Université de Montréal

Analyse de la structure logique des inférences légales et modélisation du discours juridique

par

Clayton Peterson

Département de philosophie
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en philosophie

Mai, 2014

Résumé

La présente thèse fait état des avancées en logique déontique et propose des outils formels pertinents à l'analyse de la validité des inférences légales. D'emblée, la logique vise l'abstraction de différentes structures. Lorsqu'appliquée en argumentation, la logique permet de déterminer les conditions de validité des inférences, fournissant ainsi un critère afin de distinguer entre les bons et les mauvais raisonnements. Comme le montre la multitude de paradoxes en logique déontique, la modélisation des inférences normatives fait cependant face à divers problèmes. D'un point de vue historique, ces difficultés ont donné lieu à différents courants au sein de la littérature, dont les plus importants à ce jour sont ceux qui traitent de l'action et ceux qui visent la modélisation des obligations conditionnelles. La présente thèse de doctorat, qui a été rédigée par articles, vise le développement d'outils formels pertinents à l'analyse du discours juridique. En première partie, nous proposons une revue de la littérature complémentaire à ce qui a été entamé dans Peterson (2011). La seconde partie comprend la contribution théorique proposée. Dans un premier temps, il s'agit d'introduire une logique déontique alternative au système standard. Sans prétendre aller au-delà de ses limites, le système standard de logique déontique possède plusieurs lacunes. La première contribution de cette thèse est d'offrir un système comparable répondant au différentes objections pouvant être formulées contre ce dernier. Cela fait l'objet de deux articles, dont le premier introduit le formalisme nécessaire et le second vulgarise les résultats et les adapte aux fins de l'étude des raisonnements normatifs. En second lieu, les différents problèmes auxquels la logique déontique fait face sont abordés selon la perspective de la théorie des catégories. En analysant la syntaxe des différents systèmes à l'aide des catégories monoïdales, il est possible de lier certains de ces problèmes avec des propriétés structurelles spécifiques des logiques utilisées. Ainsi, une lecture catégorique de la logique déontique permet de motiver l'introduction d'une nouvelle approche syntaxique, définie dans le cadre des catégories monoïdales, de façon à pallier les problèmes relatifs à la modélisation des inférences normatives. En plus de proposer une analyse des différentes logiques de l'action selon la théorie des catégories, la présente thèse étudie les problèmes relatifs aux inférences normatives conditionnelles et propose un système déductif typé.

Mots clés: Logique déontique, Logique catégorique, Système déductif, Logique modale, Logique linéaire, Logique non-monotone, Action, Algèbre, Théorie des catégories, Catégorie monoïdale

Abstract

The present thesis develops formal tools relevant to the analysis of legal discourse. When applied to legal reasoning, logic can be used to model the structure of legal inferences and, as such, it provides a criterion to discriminate between good and bad reasonings. But using logic to model normative reasoning comes with some problems, as shown by the various paradoxes one finds within the literature. From a historical point of view, these paradoxes lead to the introduction of different approaches, such as the ones that emphasize the notion of action and those that try to model conditional normative reasoning. In the first part of this thesis, we provide a review of the literature, which is complementary to the one we did in Peterson (2011). The second part of the thesis concerns our theoretical contribution. First, we propose a monadic deontic logic as an alternative to the standard system, answering many objections that can be made against it. This system is then adapted to model unconditional normative inferences and test their validity. Second, we propose to look at deontic logic from the proof-theoretical perspective of category theory. We begin by proposing a categorical analysis of action logics and then we show that many problems that arise when trying to model conditional normative reasoning come from the structural properties of the logic we use. As such, we show that modeling normative reasoning within the framework of monoidal categories enables us to answer many objections in favour of dyadic and non-monotonic foundations for deontic logic. Finally, we propose a proper typed deontic system to model legal inferences.

Keywords: Deontic logic, Categorical logic, Deductive system, Modal logic, Linear logic, Non-monotonic logic, Action logic, Algebra, Category theory, Monoidal category

Table des matières

Liste des tableaux

Liste des figures

Notation

$\{,\}$	accolades
$\mathcal{AL}$	*Action Logic*
$\mathcal{BA}$	algèbre de Boole
$\mathcal{HA}$	algèbre de Heyting
ATL	*alternating-time temporal logic*
$\in$	appartenance
$*$	astérisque
$\bar{}$	barre de surlignement
$\|$	barre double
$/$	barre oblique
$\backslash$	barre oblique inversée
$\mid$	barre simple
$\Box$	boîte
$\blacksquare$	boîte noire
$\circ$	cercle
CL, PL	*classical propositional logic*
$-$	complément ensembliste
CTL	*computational tree logic*
$\mathcal{CNR}$	*conditional normative reasoning*
$\wedge$	conjonction
$\models$	conséquence sémantique
$\vdash$	conséquence syntaxique
$\oslash$	co-tenseur
[,]	crochets
$\langle,\rangle$	crochets à angle
def.	définition
$\mathcal{DDS}$	*deontic deductive system*
:	deux-points
$\neq$	différent de
$\vee$	disjonction
$\oplus$	disjonction additive
DN	double négation
DDL, PD_eL	*dynamic deontic logic*
$=$	égalité
EBF	expression bien formée

$\mathbb{N}$	ensemble des nombres naturels
$\mathbb{R}$	ensemble des nombres réels
$\wp$	ensemble des parties
$\varnothing$	ensemble vide
$\equiv$	équivalence
$\perp$	*falsum*
FOL	*first order logic*
$\Uparrow$	flèche (double) vers le haut
$\uparrow$	flèche vers le haut
$\vec{F}$	futur
g.l.b.	*greatest lower bound*
H	hypothèse
$\supset$	implication
$\rightarrow$	implication
$\leftarrow$	implication inversée
$\multimap$	implication linéaire
$\multimapinv$	implication linéaire inversée
$\subset$	inclusion
$\subseteq$	inclusion propre
$<$	inférieur
$\leq$	inférieur ou égal
I/O	input/output
F	interdiction
$\cap$	intersection
IL	*intuitionistic logic*
$^{-1}$	inverse
$\cong$	isomorphisme
$\vec{U}$	jusqu'à
l.u.b.	*least upper bound*
LL	*linear logic*
$\Diamond$	lozenge
$\blacklozenge$	lozenge noir
MP	*modus ponens*
$\longrightarrow$	morphisme, preuve
MALL	*multiplicative and additive linear logic*
MLL	*multiplicative linear logic*
MNLL	*multiplicative non-commutative linear logic*
NL	*natural language*
$\neg$	négation
$\sim$	négation
$\notin$	non-appartenance
NILL	*non-commutative intuitionistic linear logic*
$\not\models$	non-conséquence sémantique
$\nvdash$	non-conséquence syntaxique
NMAS	*normative multi-agent system*

O	obligation
$\odot$	obligation
$O(/)$	obligation dyadique
$\mathcal{OL}$	*obligation logic*
(,)	parenthèses
$\parr$	parr
poset	*partially ordered set*
$\vec{P}$	passé
&	perluète
P	permission
$P(/)$	permission dyadique
P_w	permission faible
P_s	permission forte
$\preceq$	plus petit ou égal
;	point-virgule
!	point d'exclamation
$\times$	produit
$\mathcal{PAL}$	*propositional action logic*
PDL	*propositional dynamic logic*
$\bullet$	puce
$\exists$	quantificateur existentiel
$\forall$	quantificateur universel
RAA	*reductio ad absurdum*
Reit	réitération
$\curvearrowright$	séquence
$\Rightarrow$	si, alors
$\Leftrightarrow$	si et seulement si
$\frown$	*small frown*
$\smile$	*small smile*
$+$	somme
$\sqcap$	*square cap*
$\sqcup$	*square cup*
$\sqsubseteq$	*square subset or equal*
$>$	supérieur
$\geq$	supérieur ou égal
SDL	système standard de logique déontique
t.q.	tel que
$\otimes$	tenseur
$\therefore$	*therefore*
$\triangleright$	triangle
$\blacktriangleright$	triangle noir
$\cup$	union
$\top$	*verum*
WFF	*well-formed formulas*
$\ominus$	*without*

À ma mère, Alberte.

Remerciements

Je tiens en tout premier lieu à remercier mon directeur de recherche, Jean-Pierre Marquis, qui a su m'orienter et me guider. Merci de m'avoir « corrompu » dès le début de ma formation en m'initiant aux catégories. Cela aura définitivement modelé ma compréhension de la logique. Merci à François Lepage pour l'encadrement offert et pour les opportunités données, et merci à Yvon Gauthier de m'avoir permis de poursuivre mes intérêts de recherche dans un milieu de travail intellectuellement stimulant. Merci aussi à Andrew Irvine pour ses commentaires et ses remarques quant à la thèse.

Un merci spécial à ma famille pour le soutien et les encouragements. À mon père, qui se figure probablement que je fais de la « mécanique » des raisonnements et qui m'encourage à persévérer. À ma mère, pour son optimisme éternel et pour m'avoir fourni le nécessaire à l'accomplissement de mes objectifs. Finalement, à mes beaux-parents, qui savourent et soulignent chacune de nos réussites.

Merci à ma conjointe, Sarah-Geneviève, pour sa présence, autant physique, morale qu'intellectuelle. Un parcours de doctorat n'est pas une mince affaire, même lorsque l'on sait où l'on s'en va. On frappe des embûches, parfois des échecs, et par dessus tout on vit du stress, à la fois positif et négatif. Merci d'avoir été présente dans les moments difficiles et de m'avoir aidé à porter le fardeau. Merci pour les conseils de carrière et les nombreuses corrections. Merci pour les discussions et les échanges. Merci pour l'écoute et l'attention. Les choses se réalisent toujours mieux à deux.

Je tiens aussi à remercier plusieurs évaluateurs anonymes pour leurs commentaires, qui m'ont aidé à bonifier les articles de cette thèse. Cet ouvrage a été financé par le Conseil de Recherches en Sciences Humaines du Canada et par les Fonds de Recherche du Québec - Société et Culture.

Chapitre 1

Introduction

Un peu d'histoire...

La logique déontique se veut d'emblée une application de la logique aux discours éthique et juridique. Le terme « déontique ($\delta\epsilon\acute{o}\nu\tau\iota\kappa$) » provient de la racine grecque $\delta\epsilon\acute{o}\nu$ et du suffixe $\iota\kappa$, qui signifient respectivement « ce qu'il convient de faire » et « qui porte sur ». Littéralement, l'expression « logique déontique » signifie donc « la logique qui porte sur ce qu'il convient de faire (McNamara 2010, note 1) ». Par conséquent, la logique déontique est souvent présentée comme la logique qui traite des obligations, des permissions et des interdictions. Même si l'étude des modalités déontiques se retrace jusqu'à l'époque médiévale (cf. Knuuttila 1981), la première tentative de formalisation est apparue avec Ernst Mally (1926).[1] Cela dit, c'est suite aux travaux de von Wright (1951) que la logique déontique est devenue un domaine d'études en tant que tel. Initialement, l'idée de von Wright était de mettre en parallèle les notions d'obligation et de permission avec les modalités aléthiques de necessité et de possiblité. Ainsi, depuis l'article séminal de von Wright, la logique déontique est souvent conçue comme une logique modale.

Malgré qu'un tel mode de présentation soit réducteur, la logique est parfois présentée comme l'étude de la transmission des valeurs de vérité entre les énoncés. Lorsque la logique est utilisée à des fins d'analyse d'arguments et de raisonnements, les connecteurs sont considérés comme vérifonctionnels et les propositions étudiées sont présupposées être *déclarative*, c'est-à-dire qu'elles doivent minimalement avoir le potentiel d'être vraies ou fausses.[2] Or, une modalité $\Box$ est un opérateur qui transforme un énoncé déclaratif en un autre énoncé déclaratif. En ce sens, une modalité $\Box$ est quelque chose qui influence la valeur de vérité d'une proposition φ. Un exemple de modalité serait la modalité temporelle du *passé*. Au même titre que la proposition « Jean commet l'adultère » n'est pas vraie dans les mêmes conditions que « Jean a commis l'adultère », le premier énoncé n'est pas vrai dans les mêmes conditions que « Jean ne devrait pas commettre l'adultère ». De fait, lorsque considérée en tant que logique modale, la logique déontique met en jeu des opérateurs qui

[1] Voir Føllesdal et Hilpinen (1970) et McNamara (2010) pour une introduction à la logique déontique. Quant à son histoire, voir Hilpinen et McNamara (2013).

[2] À ce sujet, voir Peterson (2013b), chapitre 3.

modifient la valeur de vérité des énoncés déclaratifs.

En logique déontique, l'obligation, la permission et l'interdiction sont respectivement représentées par les opérateurs O, P et F.[3] Il s'agit d'*opérateurs* plutôt que de *prédicats* dans la mesure où les modalités déontiques peuvent être appliquées à des propositions afin de former de nouvelles propositions (possiblement d'un autre type). Chez von Wright, les opérateurs déontiques étaient appliquées à des *actions* plutôt qu'à des *propositions*. En ce sens, l'approche initiale de von Wright proposait un parallèle entre la logique déontique et la logique modale aléthique sans pour autant présenter l'opérateur O comme une modalité, c'est-à-dire comme un opérateur qui influence la valeur de vérité d'une proposition déclarative.

Ainsi, les opérateurs déontiques peuvent être interprétés de deux manières distinctes et mutuellement exclusives (cf. von Wright 1999b), dépendamment de si l'on considère que les opérateurs portent sur des actions ou des propositions (déclaratives). D'un côté, l'interprétation de type *Ought-to-do* est telle que les énoncés qui se trouvent dans la portée des opérateurs déontiques réfèrent à des *actions*. Selon cette interprétation, les opérateurs déontiques expriment « ce qui doit être fait ». Considérant que ce sont les *actions* qui sont faites, et non les propositions, l'interprétation *Ought-to-do* contraint la définition des expressions bien formées et seule une action peut se trouver dans la portée d'un opérateur déontique. Par conséquent l'itération des opérateurs déontiques n'est pas possible. À l'opposé, il y a l'interprétation de type *Ought-to-be*, où les propositions dans la portée des opérateurs déontiques sont déclaratives. Il s'agit là de l'interprétation *modale* de la logique déontique, où les opérateurs sont vus comme des modalités qui influencent la valeur de vérité des énoncés. Selon cette interprétation, une modalité déontique indique ce qui doit *être* plutôt que ce qui doit être fait. Souvent, l'interprétation de type *Ought-to-be* prendra place dans le cadre d'une sémantique des mondes possibles. Par exemple, alors que selon la première interprétation l'énoncé « il est obligatoire de respecter la loi » signifie que l'action *respecter la loi* doit être accomplie, cela signifie plutôt, selon la seconde, que la description « la loi est respectée » doit être vraie.

Plusieurs objections ont été formulées contre l'approche initiale de von Wright. La première a été avancée par Prior (1954) et visait à montrer que la notion d'*engagement* n'était pas adéquate dans les contextes mettant en jeu des obligations dérivées. Cette objection fut prise au sérieux par von Wright (1956), qui présenta la base de ce que l'on nommera plus tard la logique déontique dyadique. Même si la logique déontique dyadique a intéressé plusieurs philosophes à l'époque, notamment Rescher (1958), le système initial de von Wright est demeuré l'objet de plusieurs critiques. En 1958, Prior présenta le paradoxe du bon samaritain, précurseur du paradoxe du voleur (cf. Nozick et Routley 1962), qui visait à montrer que la relation de conséquence déontique menait à des résultats indésirables. Cela dit, l'objection la plus dommageable contre l'approche initiale de von Wright fut formulée par Chisholm (1963), qui montra que la logique déontique monadique n'est pas assez puissante pour rendre compte des inférences normatives conditionnelles et

[3] La logique déontique est majoritairement entendue comme une logique modale. Cela dit, soulignons qu'il y a néanmoins plusieurs approches qui ne la traitent pas ainsi. Outre ceux qui traitent les notions déontiques comme des prédicats, on trouve aussi la logique input/output.

des obligations contraires au devoir. Suite à ces objections, von Wright (1967) introduisit plusieurs systèmes de logique déontique dyadique afin de résoudre les objections touchant aux obligations conditionnelles. Malgré les objections contre l'approche initiale de von Wright, celle-ci s'est néanmoins vue reformulée suite aux avancées faites en logique modale et à l'introduction de la sémantique des mondes possibles, ce qui a donné lieu au fameux *système standard de logique déontique* (SDL), à savoir le système modal KD.[4] Par la suite, Forrester (1984) présenta un argument contre le système standard afin de montrer que la relation de conséquence ne permettait pas de représenter adéquatement la conséquence déontique que l'on trouve au sein de nos inférences dans la langue naturelle.

Structure de la thèse et objectifs

Bien que cela ne soit pas toujours explicite dans la littérature, la logique déontique peut être utilisée à différentes fins. Un premier objectif possible est l'utilisation de la logique déontique afin de déterminer les conditions de validité des inférences normatives. En ce sens, la logique déontique peut être vue comme l'analyse de la structure des inférences normatives, autant éthiques que légales. Cette perspective s'ancre autant dans la logique philosophique que dans la philosophie formelle et la pensée critique. Il s'agit de cette perspective que nous allons adopter. En second lieu, la logique déontique peut être utilisée afin de modéliser les situations où le comportement de plusieurs agents est régi par des normes et des contraintes. Selon cet angle d'approche, la logique déontique est un outil visant à étudier l'évolution des systèmes normatifs. Cette perspective est pertinente notamment en informatique et en droit. Dans un troisième temps, la logique déontique peut être utilisée afin d'étudier la structure d'un ensemble de normes, comme dans les cas où l'on cherche à étudier la structure des lois. Ce genre d'analyse permet la représentation de la connaissance légale et la construction de bases de données. Finalement, la logique déontique peut servir à la programmation, autant par la modélisation des contraintes (obligations, permissions et interdictions d'un programme informatique) que dans le développement de l'intelligence artificielle.

Évidemment, la visée d'un objectif plutôt que d'un autre déterminera en partie les caractéristiques de la logique déontique que l'on doit utiliser. Pareillement, le domaine d'application visé orientera les critiques et les objections que l'on peut faire envers les différents systèmes. Or, l'objectif de cette thèse est d'utiliser la logique déontique afin d'étudier la structure des inférences normatives. Plus précisément, notre objectif est de modéliser la structure des inférences légales. Le but n'est pas d'examiner ou de décrire la structure des inférences normatives comme elles se font au quotidien: il s'agit plutôt de déterminer les conditions dans lesquelles les inférences normatives sont valides. En ce sens, l'objectif est de fournir un cadre de travail permettant de déterminer la structure que *devrait* avoir une inférence normative. Notre tâche n'est donc pas descriptive mais prescriptive.

[4] Voir Åqvist (2002). Soulignons que le système initial de von Wright n'est pas équivalent au système standard.

Parmi les discours normatifs, on trouve notamment les discours éthique et juridique. Contrairement au discours éthique, les règles gouvernant les conditions dans lesquelles les inférences juridiques sont acceptables sont bien définies, ou du moins sont souvent explicites. Lorsque le juriste raisonne et intepréte la loi, ses inférences sont guidées par des principes d'interpréation bien établis, et souvent même ancrés dans la jurisprudence. Considérant cet avantage du discours légal, nous avons choisi de concentrer notre analyse sur ce dernier afin de pouvoir apporter des fondements philosophiques solides à notre approche. En un mot, l'idée derrière ce choix est fort simple: pour que la logique soit pertinente au discours légal, le système formel doit être construit sur la base de principes fondamentaux au droit. Ainsi, notre analyse sera orientée sur le discours légal de façon à pouvoir fonder notre approche sur des principes bien établis en droit. Cela dit, malgré que l'objectif principal de la thèse soit de développer des outils formels pertinents à l'analyse des inférences légales, la possibilité que ces outils soient utilisés afin d'analyser les inférences normatives en général reste ouverte.

À la lumière de ces remarques, il est possible de mieux comprendre le titre du présent ouvrage: *Analyse de la structure logique des inférences légales et modélisation du discours juridique*. Notre objectif est de développer des outils formels capables d'expliciter la structure des inférences légales, et ce de façon à déterminer leurs conditions de validité. En ce sens, le cadre de travail proposé permet l'analyse des inférences légales du point de vue de leur *structure*, ce qui ultimement permet la modélisation de la structure du discours jurique. Ainsi, l'objet de notre analyse est la structure du *discours* juridique, et non pas, par exemple, l'analyse de la structure de la législation canadienne: notre objectif est de déterminer certaines conditions qui marquent une utilisation *correcte* d'une inférence au sein du discours juridique.

La présente thèse se divise en deux parties et s'insère dans la continuité de ce qui a déjà été entamé dans Peterson (2011). La première partie se veut une revue de la littérature et complète nos travaux précédents. L'objectif principal étant de définir un cadre de travail permettant l'analyse des inférences légales, la revue de la littérature ainsi que les critiques qui sont proposées seront faites à la lumière de cet objectif. Même si plusieurs approches peuvent être pertinentes selon différents objectifs, nous tâcherons d'expliciter en quoi ce que l'on trouve actuellement au sein de la littérature ne permet pas de déterminer un cadre de travail adéquat à l'analyse des raisonnements juridiques. Ayant déjà proposé une revue critique de la littérature en abordant les paradoxes de la logique déontique ainsi que les systèmes monadiques, dyadiques et temporels, nous ne répéterons pas ce qui a déjà été exposé dans Peterson (2011), auquel le lecteur est référé pour une plus ample analyse. Néanmoins, afin que cette thèse comprenne tout le matériel nécessaire à sa compréhension, nous proposerons d'abord au chapitre 2 un résumé concis des paradoxes, des systèmes standards ainsi que de la logique déontique dyadique. Les logiques de l'action sont abordées aux chapitres 3 et 4, et une brève analyse des logiques multi-modales sans logique d'action est présentée au chapitre 5. Cela fait, les logiques input-output, qui sont des logiques non-modales, sont présentées au chapitre 6. Enfin, différentes approches qui utilisent plutôt des structures algébriques sont présentées au chapitre 7, et la logique non-monotone est abordée au chapitre 8. Le chapitre 9 clôt la première partie, où les fondements philosophiques de notre analyse ainsi qu'une critique de la littérature sont exposés.

La seconde partie présente la contribution théorique de cette thèse, qui a été rédigée par articles. Le chapitre 10 fait le pont avec le chapitre précédent et introduit les différents articles en spécifiant comment ceux-ci répondent aux critiques avancées. Notre objectif, qui est de définir un cadre de travail adéquat pour l'analyse de la structure des inférences légales, est abordé en deux temps.

D'abord, cette thèse a été rédigée avec le souci pédagogique de rendre accessible aux étudiants en philosophie ainsi qu'aux juristes une méthode leur permettant de faire l'analyse des inférences normatives. Notre objectif général est donc abordé dans un premier temps du point de vue de la pensée critique et de l'argumentation. Alors que les étudiants dans les cours de pensée critique sont généralement introduits à la logique propositionnelle classique et au calcul des prédicat, ces derniers ont rarement le matériel requis pour faire l'analyse de la validité des inférences normatives. Ce constat et la nécessité d'une nouvelle approche afin de répondre à ce problème sont exposés au chapitre 12, qui se veut un chapitre de vulgarisation s'adressant à un lecteur ayant peu de bagage en logique. Le système formel à partir duquel la méthode présentée au chapitre 12 est construite est exposé au chapitre 11, où les fondements de notre analyse sont présentés et où les paradoxes de la logique déontique sont discutés.

En plus d'aborder la question de la validité des inférences légales du point de vue de la pensée critique et de l'argumentation, notre objectif général a aussi été traité à partir d'un cadre plus formel, qui s'enracine au sein de la philosophie formelle et de la logique. Malgré que le système proposé au chapitre 11 ait des avantages face à la logique propositionnelle, au calcul des prédicats ainsi qu'aux systèmes que l'on retrouve dans la littérature quant à l'étude de la validité des inférences légales, il n'en demeure pas moins que ce dernier possède ses limites. Afin d'y répondre, un cadre de travail général et unificateur a été développé à partir d'une analyse de la logique selon la théorie des catégories. La théorie des types de Lambek en relation avec la théorie des catégories et la logique déontique est présentée au chapitre 13, et une analyse des logiques de l'action selon la théorie des catégories est proposée au chapitre 14. Le chapitre 15 aborde les obligations conditionnelles et les conflits normatifs à partir du même cadre de travail, et le chapitre 16 propose la synthèse des autres articles en introduisant un cadre unificateur permettant l'analyse des inférences légales. Enfin, une récapitulation de nos principaux résultats est présentée au chapitre 17. En un mot, la contribution de cette thèse consiste à fournir un nouveau regard sur les problèmes auxquels la logique déontique fait face, montrant qu'une interprétation de la logique selon la théorie des catégories permet de mettre en relation ces problèmes avec certaines propriétés structurelles des logiques utilisées.

Partie I

Revue de la littérature

Chapitre 2

Systèmes standards et paradoxes

Le présent chapitre se veut essentiellement une synthèse des résultats présentés dans Peterson (2011). La logique déontique standard et les paradoxes étant des points de référence au sein de la littérature, nous avons jugé nécessaire de les exposer brièvement, principalement de façon à ce que le lecteur trouve dans le présent ouvrage tout le matériel nécessaire à la compréhension du propos qui y est véhiculé. Cela dit, ces sujets ont déjà été couverts précédemment, et de fait, le présent chapitre a été rédigé de manière extrêmement concise, prenant pour acquis que le lecteur est au fait de nos travaux antérieurs. Afin que cela ne soit pas redondant, nous allons seulement exposer les notions principales et le lecteur est référé à Peterson (2011) pour une analyse approfondie.

Logique déontique standard

Tel que présenté dans Peterson (2011, p.32), les travaux de von Wright (1951) ont donné lieu aux *systèmes standards de logique déontique* (SDL), dont le représentant est le système modal KD. Ici, une précision s'impose concernant une petite ambiguïté que l'on trouve dans la littérature: même si SDL réfère à la classe des systèmes syntaxiquement équivalents à KD, l'expression *système standard de logique déontique* est parfois utilisée afin de référer à des systèmes qui ne sont pas équivalents à KD. Conformément à la présentation de Åqvist (2002, p.155), la classe des systèmes dits *standards* inclut les systèmes *normaux* et *fortement normaux*. Alors que les systèmes *normaux* de logique déontique admettent minimalement les théorèmes du système modal K, les systèmes *fortement normaux* contiennent au moins ceux de KD.

Une logique peut être vue comme un ensemble de propositions (bien formées) contenant certains axiomes et étant fermé sous certaines règles. Ainsi, un système Δ de logique déontique est normal si $K \subseteq \Delta$ et fortement normal si $KD \subseteq \Delta$. Notons que la définition proposée par Åqvist pour les systèmes normaux est équivalente à celle proposée par Chellas (1980, p.114). Soulignons aussi que le système présenté par von Wright (1951) n'est pas équivalent au système standard, notamment en raison du fait que von Wright rejette les proposition de la forme $O\top$, où $\top$ est une tautologie.[1]

La logique déontique standard est un incontournable dans la littérature considérant que lorsqu'elle n'est pas utilisée comme système de base à partir duquel on créer une extension, elle est utilisée comme point de référence afin de préciser quelles règles sont rejetées et pourquoi. Suivant la présentation faite dans Peterson (2011, p.30), le système modal KD, extension de K, est le représentant de la classe des systèmes standards. Lorsque présenté en tant que système de déduction naturelle, le système K est une extension de la logique propositionnelle classique (**CL**), à laquelle on ajoute l'opérateur $\Box$ ainsi que deux règles qui gouvernent l'introduction ($\Box in$) et l'élimination ($\Box out$) de l'opérateur (cf. Garson 2006, chapitre 1).

$$\begin{array}{l|l l} 1 & \quad \Box & \\ 2 & \quad \vdots & \\ 3 & \quad A & \\ 4 & \Box A & \Box in \text{ 1.1-3} \end{array} \qquad\qquad \begin{array}{l|l l} 1 & \vdots & \\ 2 & \Box A & \\ 3 & \quad \Box & \\ 4 & \quad A & \Box out \text{ 1.2,3} \end{array}$$

Le langage $\mathcal{L} = \{(,), \supset, \Box, Prop\}$ de K, où $Prop = \{\bot, p_1, ..., p_n, ...\}$, contient un connecteur logique primitif $\supset$, à l'aide duquel il est possible de définir (classiquement) les autres connecteurs logiques lorsque le symbole $\bot$ est considéré.

$$\neg A =_{def} A \supset \bot \qquad \text{(def. } \neg)$$
$$A \wedge B =_{def} \neg(A \supset \neg B) \qquad \text{(def. } \wedge)$$
$$A \vee B =_{def} \neg A \supset B \qquad \text{(def. } \vee)$$
$$A \equiv B =_{def} ((A \supset B) \wedge (B \supset A)) \qquad \text{(def. } \equiv)$$

L'opérateur dual de $\Box$ est défini par (def.$\Diamond$). Dans le cas de la logique déontique, les opérateur $\Box$ et $\Diamond$ sont remplacés respectivement par O et P. L'opérateur F, qui représente l'interdiction, est défini par (def. F).

$$\Diamond A =_{def} \neg\Box\neg A \qquad \text{(def.}\Diamond)$$
$$FA =_{def} O\neg A \qquad \text{(def. F)}$$

L'ensemble des expressions bien formées (EBF) est défini récursivement de manière à permettre l'itération des opérateurs déontiques.

1. si $p \in Prop$, alors $p \in EBF$;
2. si $A, B \in EBF$, alors $A \supset B \in EBF$;
3. si $A \in EBF$, alors $\Box A \in EBF$.

[1] Il s'agit du principe de contingence normative. Voir Peterson (2011, p.29) pour l'approche monadique de von Wright, lequel acceptera éventuellement ce type d'énoncé en adoptant le schéma d'axiome $O(A \vee \neg A)$ (cf. von Wright 1983, p.120).

La présentation de K en tant que système de déduction naturelle est équivalente à sa présentation axiomatique usuelle (cf. Garson 2006, p.37), où $K = \mathbf{CL}+$ (Nec) + (Dist).[2]

$$\frac{\vdash A}{\vdash \Box A} \qquad \text{(Nec)}$$

$$\vdash \Box(A \supset B) \supset (\Box A \supset \Box B) \qquad \text{(Dist)}$$

Le système KD est obtenu avec l'ajout de l'axiome (**D**).

$$\Box A \supset \Diamond A \qquad (\mathbf{D})$$

Suivant Peterson (2011, p.31 et p.35), les systèmes standards utilisent la sémantique habituelle pour les logiques modales (cf. Kripke 1963). Un modèle $\mathcal{M} = \langle U, R, a\rangle$ est construit à partir d'une structure $\mathcal{S} = \langle U, R\rangle$, où U est un univers (non vide) et R un ensemble de relation(s) au sein de U. Formellement, $U \neq \varnothing$ et $R \subseteq U \times U$. La fonction $a : U \longrightarrow \{\top, \bot\}$ attribue des valeurs de vérité aux propositions dans U. Une proposition A est dite *vraie* pour un scénario $w \in U$ à condition qu'elle soit un élément de w:

$$a_w(A) = \top \Leftrightarrow\ A \in w$$

Par convention, un scénario qui ne contient pas A contient $\neg A$. La clause sémantique quant à l'opérateur $\Box$ est que $\Box A$ est vrai pour un scénario w si et seulement si A est vrai pour tout scénario v en relation R avec w. Formellement, nous avons:

$$a_w(\Box A) = \top \Leftrightarrow\ a_v(A) = \top\ \forall v \text{ t.q. } wRv$$

Un énoncé est dit *valide* lorsqu'il est vrai pour toute interprétation $\mathcal{M}$. Dans le cas du système KD, l'axiome (**D**) induit une relation sérielle sur le modèle sémantique, et donc la relation R est restreinte par la condition suivante.

$$\forall w, \exists v \text{ t.q. } wRv$$

La logique déontique standard est une logique *monadique*, c'est-à-dire une logique pour laquelle l'opérateur $\Box$ est unaire. Autrement dit, une logique déontique monadique met en jeu un opérateur déontique qui ne contient qu'une seule proposition dans sa portée.[3]

Notons que dépendamment de l'approche préférée, la sémantique peut aussi être définie récursivement en faisant référence aux ensembles qui contiennent les scénarios où les propositions sont vraies. En prenant par exemple l'ensemble $\|V_A\| = \{w :\models_w A\}$, on peut reformuler la clause sémantique par:

$$a_w(A) = \top \Leftrightarrow\ w \in \|V_A\|$$

[2] (Nec) est une règle de fermeture alors que (Dist) est un schéma d'axiome. La différence peut s'apercevoir par le fait que $\nvdash p \supset \Box p$ mais $\vdash \Box(p \supset q) \supset (\Box p \supset \Box q)$.

[3] Le lecteur est référé à Peterson (2011, chapitre 3).

Similairement, on peut définir récursivement $\|V_A\|$ à partir des atomes et des connecteurs logiques:

$$\begin{aligned}
\|V_p\| &= \{w : \models_w p\} \\
\|V_{\neg A}\| &= \{w : w \in U - \|V_A\|\} \\
\|V_{A \wedge B}\| &= \{w : w \in \|V_A\| \cap \|V_B\|\} \\
\|V_{A \vee B}\| &= \{w : w \in \|V_A\| \cup \|V_B\|\} \\
\|V_{A \supset B}\| &= \{w : w \in \|V_{\neg A}\| \cup \|V_B\|\}
\end{aligned}$$

Logique déontique dyadique

Suivant Peterson (2011, chapitre 4), la logique déontique dyadique s'est développée en réaction aux paradoxe de Chisholm (1963), lequel met en évidence la nécessité d'avoir un opérateur dyadique $O(/)$ primitif afin de pouvoir rendre compte de la notion d'obligation contraire au devoir. L'opérateur dyadique $O(A/C)$, qui se lit « A est obligatoire dans les conditions C », vise à rendre compte des obligations conditionnelles. De manière générale, l'opérateur dual de $O(/)$ est défini de façon similaire à ce que l'on retrouve au sein de la littérature sur la logique déontique monadique. En effet, la permission conditionnelle est souvent définie comme *ce dont la négation n'est pas obligatoire* dans un certain contexte.

$$P(A/C) =_{def} \neg O(\neg A/C) \qquad \text{(def. } P(/))$$

Afin de faciliter la lecture, l'opérateur $O(A/C)$ est parfois remplacé par $O_C A$, lequel se lit plus naturellement de la gauche vers la droite.

Tel que mentionné par Chellas (1980, p.201), la définition de l'opérateur dyadique $O(/)$ en termes monadiques pose problème. Par exemple, définir $O(A/C)$ par $C \supset OA$ nous donne une version du paradoxe de Prior (1954) puisque $\neg C \supset (C \supset OA)$. Autrement dit, si Pierre ne paie pas ses taxes, alors s'il paie ses taxes, il doit commettre l'adultère! En ce sens, la définition de l'obligation conditionnelle en termes d'obligation inconditionnelle pose certaines difficultés, ce qui est souvent résolu par le fait de prendre l'opérateur dyadique comme primitif.[4] La littérature sur la logique déontique dyadique est très vaste, et le lecteur est référé à Peterson (2011, chapitre 4) pour de plus amples informations concernant son origine, son évolution et les problèmes auxquels elle fait face.

Cela dit, le problème le plus important de l'obligation conditionnelle est certainement celui du détachement (Peterson 2011, pp.60-1). Le problème du détachement se résume à la question de savoir dans quelle mesure l'obligation conditionnelle $O(A/C)$ devient une obligation inconditionnelle OA. Ce problème se résume au fait que les deux propositions suivantes sont plausibles mais incompatibles:

1. si A est une obligation conditionnelle au contexte C et que le contexte C se présente, alors il est possible de conclure que A (dans le contexte C) est une obligation actuelle;

[4] Ce n'est toutefois pas la solution proposée par Chellas (1974), lequel introduit un nouvel opérateur $\Rightarrow$ afin de rendre compte des obligations conditionnelles (cf. Peterson 2011, section 4.4).

2. même si A est une obligation conditionnelle au contexte C, il est tout à fait possible que des circonstances C' s'ajoutent au contexte C et que A ne soit pas obligatoire dans le contexte $C \wedge C'$.

En ce sens, le problème du détachement consiste à rendre compte adéquatement du fait que dans certains cas le détachement peut se faire, alors que dans d'autres cas le détachement peut être empêché. Les problèmes concernant la définition de l'opérateur dyadique, le détachement, et l'obligation conditionnelle sont analysés en détails au chapitre 15.

Paradoxes

Les paradoxes de la logique déontique se trouvent au cœur de la littérature considérant que la plupart des approches se sont développées de manière à répondre à l'un ou l'autre de ces problèmes. Dans la présente section, les paradoxes sont seulement énoncés, sans plus, principalement en raison du fait que ceux-ci ont déjà été discutés antérieurement (cf. Peterson 2011, chapitre 2).

Suivant Åqvist (2002, pp.161-173), les paradoxes de la logique déontique résultent d'une inadéquation entre les systèmes formels et la langue naturelle (normative). Soit Δ un système de logique déontique et NDL (pour *natural deontic logic*) un ensemble de propositions formant une représentation intuitive de la langue naturelle. Un paradoxe émerge lorsque Δ n'est pas une représentation adéquate de NDL. Soit $t : NDL \longrightarrow \Delta$ une fonction de traduction de la langue naturelle vers une logique déontique, et soit $t^{-1} : \Delta \longrightarrow NDL$ l'inverse de t (qui associe une traduction possible dans NDL à chaque proposition de Δ). La fonction t permet de traduire les énoncés de la langue naturelle dans le langage du système formel Δ. Par exemple, pour

si Pierre vole un vélo, alors Pierre commet un crime

un énoncé de la langue naturelle,

$$t(\text{si Pierre vole un vélo, alors Pierre commet un crime}) = p \supset q$$

et

$$t^{-1}(p \supset q) = \text{si Pierre vole un vélo, alors Pierre commet un crime.}$$

Deux situations peuvent donner lieu à un paradoxe:

1. il y a une proposition $t(\varphi) \in \Delta$ telle que $\varphi \notin NDL$ parce que intuitivement φ ne devrait pas être valide;

2. il y a une proposition $t(\varphi) \notin \Delta$ pour laquelle $\varphi \in NDL$ puisque φ est intuitivement valide.

Åqvist (2002, p.173) nomme respectivement ces deux formes d'inadéquation comme une inadéquation de la droite vers la gauche et une inadéquation de la gauche vers la droite. Le sens de l'inadéquation va de la non-appartenance $\notin$ vers l'appartenance $\in$. Dans la présentation qui suit, Δ sera utilisé afin de dénoter un système fortement normal (i.e., pour lequel $KD \subseteq \Delta$).

Le dilemme de Jørgensen

Le dilemme formulé par Jørgensen (1937) remet en doute la possibilité d'appliquer la logique au discours normatif. À strictement parler et conformément à la présentation susmentionnée, le dilemme de Jørgensen n'est pas un paradoxe, mais plutôt un problème pour la logique déontique. Le dilemme se résume aux trois propositions suivantes (Peterson 2011, pp.8-9).

1. les propositions normatives n'ont pas de condition de vérité;
2. la logique traite de la transmission des valeurs de vérité entre les propositions;
3. les raisonnements normatifs semblent néanmoins valides, notamment dans le cas des arguments légaux et moraux.

On trouvera une discussion du dilemme au chapitre 12.

Le paradoxe de l'obligation dérivée

Suivant Peterson (2011, p.11), le paradoxe de l'obligation dérivée, avancé par Prior (1954), visait principalement la traduction formelle de la notion d'*engagement* (*commitment*) telle que proposée par von Wright (1951, p.5): si une action A est obligatoire et que poser l'action A nous engage à poser l'action B, alors B est aussi obligatoire. La notion d'engagement est formellement traduite par l'implication $O(A \supset B)$, et de fait le raisonnement susmentionné se traduit par (E).

$$(OA \wedge O(A \supset B)) \supset OB \tag{E}$$

Le paradoxe se résume au fait que:

$$\vdash_{\Delta} \neg PA \supset O(A \supset B)$$

De fait, l'énoncé $\neg PA \supset O(A \supset B)$ est un (méta-) théorème des systèmes standards.[5] Considérant la lecture que fait von Wright de la notion d'engagement, cet énoncé signifie dans le langage naturel qu'enfreindre une action interdite nous engage à poser n'importe quelle autre action. Dans une perspective plus générale et en laissant de côté le traitement

[5] Il s'agit d'un méta-théorème considérant que A et B sont des méta-variables.

que von Wright fait de l'engagement, cet énoncé est paradoxal pour les systèmes standards considérant que son interprétation dans la langue naturelle est contre intuitive: s'il n'est pas permis à Paul de voler, alors il est obligatoire que si Paul vole, alors Paul commet l'adultère. D'un point de vue normatif, ce n'est pas parce qu'il n'est pas permis de voler que cela permet de conclure qu'il est obligatoire que voler implique commettre l'adultère! Le paradoxe de Prior montre une inadéquation de la droite vers la gauche:

$$\neg PA \supset O(A \supset B) \in \Delta$$
$$t^{-1}(\neg PA \supset O(A \supset B)) \notin NDL$$

Le paradoxe de Chisholm

Le paradoxe proposé par Chisholm (1963) a non seulement fait couler le plus d'encre, mais est surtout le plus important dans la littérature. En effet, ce paradoxe met en évidence l'incapacité des systèmes monadiques à rendre compte des obligations contraires au devoir (*contrary to duty obligations*, CTD). Une obligation contraire au devoir survient dans un contexte de violation. L'objection avancée par Chisholm vise à montrer que les systèmes monadiques ne sont pas capables de rendre compte des obligations qui émergent dans des contextes où d'autres obligations ont été enfreintes. Le paradoxe se résume à ce que l'on nomme souvent *Chisholm's set*, soit l'ensemble des propositions suivantes (Peterson 2011, pp.15-6).

$$O\neg p$$
$$O(\neg p \supset \neg q)$$
$$p \supset Oq$$
$$p$$

Dans la langue naturelle, le paradoxe s'énonce par

1. Paul ne doit pas voler;
2. il devrait être le cas que si Paul ne vole pas, alors il ne remet pas l'argent volé;
3. si Paul vole, alors il doit remettre l'argent volé;
4. Paul vole.

Suivant Åqvist (2002, p.190), ces quatre propositions sont indépendantes dans la langue naturelle, d'où les traductions différentes de 2 et 3 respectivement par $O(\neg p \supset \neg q)$ et $p \supset Oq$ afin de préserver l'indépendance des énoncés au sein du système formel.

Le paradoxe de Chisholm montre une inadéquation de la droite vers la gauche. En effet, ce dernier met en évidence que l'ensemble des énoncés susmentionnés, qui est intuitivement cohérent dans la langue naturelle, est inconsistant lorsque traduit dans le langage d'une logique déontique standard. Soit φ la conjonction des énoncés traduits dans le langage de Δ.

$$\varphi = O\neg p \wedge (O(\neg p \supset \neg q) \wedge ((p \supset Oq) \wedge p))$$

Le paradoxe de Chisholm est tel que:

$$\varphi \supset \perp \in \Delta$$
$$t^{-1}(\varphi \supset \perp) \notin NDL.$$

Le paradoxe de Ross

Le paradoxe de Ross (1941) montre lui aussi une inadéquation de la droite vers la gauche.

$$Op \supset O(p \vee q) \in \Delta$$
$$t^{-1}(Op \supset O(p \vee q)) \notin NDL.$$

D'une part, $Op \supset O(p \vee q)$ est un théorème des systèmes standards en vertu des règles (Nec) et (Dist). Par ailleurs, la traduction de cet énoncé dans la langue naturelle semble contre intuitive: si Paul doit poster une lettre, alors Paul doit soit poster la lettre ou la bruler!

Le paradoxe du bon Samaritain

Le paradoxe du bon Samaritain a d'abord été formulé par Prior (1958) mais a ensuite été reformulé par Åqvist (1967).[6] À l'instar des paradoxes susmentionnés, le paradoxe du bon Samaritain montre une inadéquation de la droite vers la gauche. Considérons les propositions suivantes:

1. Pierre ne doit pas voler Paul;

2. le bon Samaritain doit aider Paul, lequel a été volé par Pierre.

Dans le langage de Δ, 1 et 2 sont respectivement traduits par $O\neg p$ et $O(q \wedge p)$. Alors que les propositions 1 et 2 sont cohérentes dans la langue naturelle, $O\neg p$ et $O(q \wedge p)$ sont inconsistantes dans le langage d'un système standard Δ, principalement en vertu du fait que:

$$\vdash_{\Delta} O(q \wedge p) \supset Op$$

[6] La version initiale du paradoxe met en jeu une utilisation injustifiée de la règle dérivée (ROM) des systèmes standards (Peterson 2011, pp.25-6). Cette première version est similaire au paradoxe du voleur proposé par Nozick et Routley (1962, p.378) et au paradoxe du meurtre gentil avancé par Forrester (1984). Voir Peterson et Marquis (2012) pour une analyse détaillée.

Par conséquent, cela entraine une contradiction avec l'axiome (**D**). Nous avons donc:

$$(O\neg p \wedge O(q \wedge p)) \supset \bot \in \Delta$$
$$t^{-1}((O\neg p \wedge O(q \wedge p)) \supset \bot) \notin NDL.$$

La solution est de traduire la seconde proposition par $p \wedge Oq$ plutôt que $O(q \wedge p)$, ce qui permet d'éviter le paradoxe.

* * *

Ceci conclut notre résumé des principaux résultats présentés dans Peterson (2011). Ce résumé était évidemment très succinct et plusieurs détails ont été laissés de côté. Le présent ouvrage étant la continuation de ce que nous avons entamé précédemment, il était inévitable de mentionner certains points auxquels nous réfèrerons dans le présent texte. Les points importants ont été mentionnés extrêmement brièvement, l'objectif étant que le lecteur ait à sa disposition tout le matériel nécessaire à la compréhension de notre propos. Pour une analyse détaillée des paradoxes, le lecteur est invité à consulter Peterson (2011, chapitre 2). Pour une discussion des paradoxes susmentionnés ainsi que d'autres paradoxes de la logique déontique, le lecteur est invité à consulter les chapitres 11, 13 et 15 du présent ouvrage.

Chapitre 3

Logique stit

Les développements récents en logique déontique qui proposent une analyse du discours légal sont souvent basés sur une *logique de l'action*. Suivant Segerberg et al. (2009, p.1), la bannière *logique de l'action* regroupe deux principaux types bien distincts en philosophie, à savoir les logiques STIT et les logiques dynamiques. Alors que le présent chapitre porte sur les logiques de l'action de type STIT, les logiques déontiques dynamiques seront considérées au chapitre 4. L'intérêt d'intégrer une logique de l'action à la logique déontique se comprend en fonction des notions de *système normatif* (NS, pour *normative system*) et de *système normatif avec plusieurs agents* (NMAS, pour *multiagent normative system*).

Malgré que l'on retrouve une certaine ambiguïté au sein de la littérature quant à l'utilisation de ces deux termes, la distinction entre un système normatif et un NMAS se comprend en fonction de l'emphase mise sur la notion d'agent. Cette ambiguïté vient principalement du fait que le concept de NMAS est parfois subsumé sous le concept de système normatif. En effet, alors que le concept de système normatif est actuellement majoritairement utilisé conformément à la définition proposée par Carmo et Jones (2002), ce dernier est aussi parfois utilisé selon sa signification originelle, que l'on retrouve chez Alchourrón et Bulygin (1971).

D'entrée de jeu, Alchourrón et Bulygin (1971, p.54) considèrent un *système normatif* comme un ensemble de normes qui contient toutes ses conséquences (i.e., un ensemble de normes déductivement fermé). Cette définition est similaire à celle de Føllesdal et Hilpinen (1970, p.16): « *by a 'normative system' we understand simply any set of normative sentences closed under deduction* ». Elle a été reprise par plusieurs auteurs, trop nombreux pour tous les énumérer. À l'inverse, Carmo et Jones (2002, p.265) proposent la définition suivante[1]:

> Deontic logic is one of the formal tools needed in the design and specification of *normative systems*, where the latter are understood to be sets of agents (human or artificial) whose interactions can fruitfully be regarded as norm-governed; the norms prescribe how the agents ideally should and should not behave, what they are permitted to do, and what they have a right to do.

[1] Voir aussi Jones et Sergot (1993).

D'autres auteurs ont aussi utilisé cette définition de système normatif, par exemple Pacheco et Santos (2004, p.210), pour qui un système normatif est un « *set of interacting agents whose behaviour is ruled by norms* ». Évidemment, ces dernières conceptions d'un *système normatif* correspondent plutôt à un NMAS. Afin d'éviter toute ambiguïté, nous proposons d'utiliser la définition d'Alchourrón et Bulygin (1971) pour les systèmes normatifs et celle de Carmo et Jones (2002) pour les NMAS.[2]

Considérant que la logique déontique peut servir à la modélisation des NMAS, l'intérêt d'y intrégrer une logique de l'action devient évident. Les logiques STIT, dont l'acronyme signifie *seeing to it that*, sont des logiques qui visent à rendre compte de la réalisation de certaines actions par des agents.[3] Ces logiques visent la modélisation des NMAS dans la mesure où l'on observe l'impact des normes sur le comportement des agents et des groupes d'agents.

Bien que l'origine de la logique STIT soit généralement attribuée à Belnap et Perloff (1988), qui ont introduit l'opérateur $[i\ stit : \varphi]$, l'origine des logiques de type STIT peut être retracée aux travaux de Kanger (1957)[4], qui a été le premier à traiter des notions déontiques à l'aide d'une logique de l'action (cf. Kanger 1972, p.111). Quoique Chellas (1969) ait été le premier à proposer une sémantique bien définie pour les logiques de type STIT (cf. Segerberg 1992, p.369), l'opérateur E_i indexé à un agent i est apparu avec Pörn (1970) (cf. Carmo et Pacheco 2001, p.131), où l'auteur cherchait à rendre compte des relations d'influences et des relations normatives qui se trouvent entre certains agents. L'opérateur E_i, qui en français peut se lire 'i réalise ...', se comprend mieux en fonction de sa formulation anglaise, où l'opérateur signifie « *i sees to it that ...* » ou encore « *i brings it about that...* ». C'est d'ailleurs en raison de cette première formulation que Belnap et Perloff (1988) ont utilisé l'acronyme *stit* pour définir l'opérateur $[i\ stit : \varphi]$. À strictement parler, la logique STIT telle que développée Belnap et Perloff (1988) inclue une dimension temporelle, conformément aux travaux de Chellas (1969). Néanmoins, le présent chapitre aborde ce que nous nommons les logiques de *type* STIT, qui incluent les approches où l'on utilise un opérateur du type « i réalise... ».

Étant donné que les logiques de types STIT utilisent un formalisme similaire à celui développé par Pörn (1970), nous avons jugé bon de présenter son approche afin de donner un aperçu de l'origine de ce genre de logique. Nous avons aussi choisi de présenter l'approche de Horty (2001) considérant qu'il s'agit, à notre avis, de l'ouvrage le plus complet proposant une analyse de la logique déontique dans le cadre d'une logique STIT. Considérant que l'objectif du présent ouvrage est d'utiliser la logique déontique afin d'analyser la structure des inférences légales, les deux autres approches présentées ont été choisies en fonction du fait qu'elles utilisent la logique déontique afin d'éclaircir certaines notions

[2] Certains auteurs, comme Boella et al. (2006), refusent la définition de Carmo et Jones (2002) pour les NMAS et proposent de les définir comme la jonction entre un système multi-agents et un système normatif au sens de Alchourrón et Bulygin (1971).

[3] Le lecteur intéressé par les logiques de l'action est invité à consulter Segerberg (1992) et Peterson (2014a).

[4] Le texte de Kanger (1957) a été réimprimé dans Kanger (1971) et ses travaux se sont développés dans Kanger et Kanger (1966) et Kanger (1972).

juridiques. Nous présenterons d'abord l'approche de Pacheco et Carmo (2003), qui s'insère plutôt dans la tradition de Pörn (1970), et ensuite celle de Broersen (2011a), qui propose d'augmenter la logique STIT à l'aide d'une modalité épistémique.

Évidemment, plusieurs auteurs utilisent des logiques de type STIT, et faire une revue complète de cette littérature dépasserait la portée de notre propos. Le lecteur intéressé est invité à consulter Xu (1995) pour le développement formel de la logique STIT, et Horty et Belnap (1995) pour une introduction détaillée à la logique déontique de type STIT.

Ingmar Pörn

Dans son ouvrage intitulé *The logic of power*, Pörn (1970) visait à développer une logique capable de rendre compte des relations de pouvoir coercitif et de pouvoir d'influence entre certains agents.[5] Plus précisément, son objectif était de développer un langage formel capable de rendre compte de plusieurs nuances de la langue naturelle (normative), telles les notions de *droit*, de *devoir*, d'*obligation*, de *tolérance* ou encore de *contrôle*. Suivant les travaux de Hohfeld (1917), le langage développé a pour but de rendre compte de certaines paires de concepts normatifs, comme *droit/devoir*, *pouvoir/imputabilité* et *immunité/imputabilité* (Pörn 1970, p.46). Les travaux de Pörn visaient à rendre compte de ces concepts de manière générale, en développant une logique qui facilite leur analyse.

L'ouvrage de Pörn est général dans la mesure où il s'intéresse aux concepts de *droit* ou d'*imputabilité*, par exemple, sans s'intéresser à la logique sous-jacente aux *droits* ou à l'*imputabilité*. En ce sens, l'approche de Pörn possède ses lacunes: son système ne permet pas de rendre compte des notions d'*intention* ou de *responsabilité* du point de vue de l'action (Pörn 1970, p.77), ce qui est crucial dans la mesure où l'on cherche à formaliser une logique de l'*imputabilité*.

La logique de l'action développée par Pörn (1970) est basée sur un langage similaire au calcul de premier ordre augmenté à l'aide de deux opérateurs (modaux) visant à rendre compte de l'action.[6] Du côté de la syntaxe, le langage

$$\mathcal{L} = \{(,), \neg, \supset, E_i, C_i, \forall, =, B, \mathbb{P}, \mathbb{CV}\}$$

contient les connecteurs $\neg$, $\supset$ et le quantificateur $\forall$, à partir desquels les autres connecteurs sont définis de façon habituelle. Les opérateurs $E_i\varphi$ et $C_i\varphi$ signifient respectivement « l'agent i réalise φ » et « φ est compatible avec les actions posées par i ». Les symboles $=$ et B sont des prédicats qui représentent respectivement l'identité et le fait qu'un agent

[5] Pörn définit les relations (binaires) de pouvoir entre les agents comme des ensembles qui contiennent des formules d'un certain type. Voir Pörn (1970, p.18 et p.32) pour les relations d'influence et les relations normatives.

[6] Le symbolisme utilisé par Pörn a été modifié de façon à ce que la présentation de son approche soit plus concise et cohérente avec la notation utilisée dans le présent ouvrage. Nous avons d'ailleurs formulé la sémantique qu'il propose dans le cadre d'une sémantique de Kripke plutôt que la sémantique développée par Hintikka. Voir Peterson (2011, p.35) pour un aperçu historique de l'utilisation de la sémantique des mondes possibles pour la logique déontique.

subisse quelque chose de mauvais (B est utilisé pour *bad*). Autrement dit, Bi signifie que « i subit quelque chose de mauvais » (Pörn 1970, p.31), ce qui sera utilisé afin de définir l'obligation. Comme l'auteur le souligne (Pörn 1970, p.32), cette définition ressemble à celle proposée par Anderson (1958a), mais l'approche n'y est pas équivalente.

Les ensembles $\mathbb{P} = \{P_1^1, ..., P_m^n, ...\}$, $\mathbb{C} = \{Ag_c, Obj_c\}$ et $\mathbb{V} = \{Ag_v, Obj_v\}$ sont respectivement des ensembles de prédicats (où P^1 est un prédicat unaire, P^2 binaire, etc.), de constantes d'agents et d'objets ainsi que de variables d'agents et d'objets. L'ensemble des expressions bien formées (EBF) est défini récursivement (cf. Pörn 1970, pp.1-9):

1. si $s_1, s_2, ..., s_n \in \mathbb{C} \cup \mathbb{V}$ et $P_i^n \in \mathbb{P}$, alors $s_1 = s_2, P_i^n(s_1, ..., s_n) \in EBF$;
2. si $\varphi, \psi \in EBF$, $x \in \mathbb{V}$, alors $\neg\varphi, \varphi \supset \psi, (\forall x)\varphi \in EBF$;
3. si $\varphi \in EBF$ et $i \in Ag_c \cup Ag_v$, alors $E_i\varphi, C_i\varphi, Bi \in EBF$.

La sémantique est définie suivant les travaux de Hintikka, lequel distingue entre un *ensemble modèle* (*model set*) et un *système de modèle* (*model system*). Cela dit, ces notions peuvent être comprises en termes de *scénario* et de *modèle* dans le cadre d'une sémantique des mondes possibles. Un scénario w doit satisfaire les conditions suivantes (Pörn 1970, pp.9-10).

$$\varphi \in w \Rightarrow \neg\varphi \notin w \tag{1}$$
$$\neg\neg\varphi \in w \Rightarrow \varphi \in w \tag{2}$$
$$\varphi \supset \psi \in w \Rightarrow \neg\varphi \in w \text{ ou } \psi \in w \tag{3}$$
$$\neg(\varphi \supset \psi) \in w \Rightarrow \varphi \in w \text{ et } \neg\psi \in w \tag{4}$$
$$(\forall x)\varphi \in w \Rightarrow \varphi(s/x) \in w \text{ pour tout } s \in \mathbb{C} \tag{5}$$
$$(\forall x)\varphi \in w \Rightarrow \varphi(s/x) \in w \text{ pour au moins un } s \in \mathbb{C} \tag{6}$$
$$\neg(\forall x)\varphi \in w \Rightarrow \neg\varphi(s/x) \in w \text{ pour au moins un } s \in \mathbb{C} \tag{7}$$
$$\varphi, s_1 = s_2 \in w \text{ et } \varphi(x/s_1) = \psi(x/s_2) \Rightarrow \psi \in w \tag{8}$$
$$s_1 \neq s_1 \notin w \text{ pour tout } s_1 \in \mathbb{C} \cup \mathbb{V} \tag{9}$$

Les conditions 1 et 2 assurent la consistance de w ainsi que le tiers exclu. Les clauses 3 et 4 représentent les conditions sémantiques habituelles pour l'implication matérielle. En un mot, les conditions 1–4 assurent que les scénarios, qui sont des ensembles de propositions, sont fermés d'un point de vue syntaxique sous les règles et axiomes de la logique propositionnelle classique. La condition 5 exprime que la quantification se fait sur l'ensemble du domaine de w, représenté par les constantes de $\mathbb{C}$, où $\varphi(s/x)$ est la proposition obtenue après avoir substitué x par s au sein de φ (avec s élément du domaine). La clause 6 indique qu'une quantification se fait sur un domaine non vide, au même titre que 7 exprime que la négation d'une quantification sur φ est vraie à condition qu'il y ait au moins un objet du domaine pour lequel φ est faux. La clause 8 signifie que pour deux objets s_1 et s_2 identiques, si φ et ψ sont similaires, c'est-à-dire que lorsque toute occurrence de s_1 ou de s_2 est remplacée par x dans chaque énoncé entraîne que $\varphi(x/s_1) = \psi(x/s_2)$, alors la proposition ψ est aussi vraie. Autrement dit, les énoncés obtenus après substitution d'objets identiques dans la portée des prédicats se trouvent aussi dans le scénario. Finalement, 9 assure que tout énoncé d'identité est vrai pour un ensemble modèle, et donc que dans le scénario aucun objet (ou aucune constante) n'est différent de lui-même.

Le modèle est défini comme une paire ordonnée (U, R), où U est un ensemble de scénarios et $R \subseteq U \times U$ une relation binaire définie pour tout agent i (Pörn 1970, p.10).[7] Cette paire ordonnée doit satisfaire les conditions suivantes.

$$E_i\varphi \in w \text{ et } wR_iv \Rightarrow \varphi \in v \tag{10}$$
$$\neg E_i\varphi \in w \Rightarrow C_i\neg\varphi \in w \tag{11}$$
$$C_i\varphi \in w \Rightarrow \text{ il y a au moins un } v \text{ t.q. } wR_iv \text{ et } \varphi \in v \tag{12}$$
$$\neg C_i\varphi \in w \Rightarrow E_i\neg\varphi \in w \tag{13}$$
$$wR_iw \text{ pour tout } w \in U \tag{14}$$

La clause 10 correspond à la condition usuelle pour la valeur de vérité de $\Box$ dans le système modal K et 11 signifie que s'il est faux que i réalise φ, alors les actions de i sont compatibles avec l'état de choses décrit par $\neg\varphi$. La condition 12 exprime que si les actions de i sont compatibles avec φ pour un scénario w, alors φ est vrai dans au moins une alternative v en relation R_i avec w. Autrement dit, φ n'est pas contradictoire. La condition 13 signifie que s'il est faux que les actions de i sont compatibles avec φ, alors i réalise $\neg\varphi$. Notons que cela implique que tout agent réalise toute tautologie. Finalement, la dernière condition exprime que la relation R_i est réflexive.

Malgré que l'on ne trouve pas d'axiomatisation du système proposé dans l'ouvrage de Pörn, les restrictions du modèle sémantique nous donnent une bonne idée de ce à quoi le système axiomatique ressemblerait. Les clauses 10 et 14 donnent à l'opérateur E_i le schéma d'axiome (**M**).[8] Les clauses 11 et 13 font que d'un point de vue syntaxique C_i est l'opérateur dual de E_i, et donc le système aurait pu être axiomatisé à l'aide de l'équivalence $C_i\varphi \equiv \neg E_i\neg\varphi$. La raison pour avoir pris chaque opérateur comme primitif est probablement que l'interprétation de « φ est compatible avec les actions de i » en termes de « il est faux que i réalise $\neg\varphi$ » ne rendait pas compte adéquatement de la signification accordée par Pörn à C_i. Bref, notons simplement que la modalité d'action aurait pu être axiomatisée comme une modalité de type KM (cf. Pörn 1970, p.11 pour la comparaison entre son approche et les logiques modales).[9]

Pörn (1970, p.31) utilise le prédicat Bi afin de définir les énoncés normatifs. Ce dernier propose d'abréger la formule $E_i(E_j\varphi \supset Bj)$ par:

$$B_iE_j\varphi =_{def} E_i(E_j\varphi \supset Bj)$$

[7] Notons que, à strictement parler, il s'agit d'une structure et non d'un modèle (un modèle étant une structure à laquelle on ajoute une fonction qui assigne des valeurs de vérité).

[8] Certains auteurs réfèrent parfois à cet axiome par (T).

[9] Le lecteur familier avec les logiques modales reconnaîtra d'ailleurs plusieurs théorèmes et axiomes de KM dans la liste des énoncés valides présentée par Pörn (1970, p.14)

Cela signifie « i s'assure que si j réalise φ, alors j subit quelque chose de mauvais ».[10] D'emblée, cette définition permet de rendre compte de la relation de pouvoir coercitif considérant que l'autorité i contraint j à agir d'une certaine manière (à réaliser φ), sans quoi j fera face à des conséquences négatives. Pörn (1970, p.32) se sert de ce type d'énoncé afin de définir plusieurs concepts normatifs. Pour être précis, l'auteur définit des *relations* normatives. Soit $R \subseteq Ag_c \times (Ag_c \times EBF)$ une relation binaire qui associe un agent avec une paire ordonnée qui contient un agent et une proposition. Considérant que $i, j \in Ag_c$ et $\varphi \in EBF$, la relation normative est définie comme un ensemble de paires ordonnées $R = \{< i, < j, \varphi >>: ...\}$ qui répondent à certaines conditions « ... ». L'ensemble

$$\{< i, < j, \varphi >>: B_i \neg E_j \varphi\}$$

représente une relation normative d'*obligation* (Pörn 1970, p.32) et contient les paires ordonnées qui satisfont l'énoncé de type $B_i \neg E_j \varphi$. Autrement dit, *il est obligatoire que j fasse φ* ou encore *j doit faire φ* signifie qu'une autorité i s'assure que si j ne réalise pas φ, alors j subit quelque chose de mauvais. En ce sens, s'il est faux que φ est une description vraie du monde suite aux actions posées par j, alors j subira une conséquence négative.[11] Dans le même ordre d'idées, l'ensemble

$$\{< i, < j, \varphi >>: B_i E_j \varphi\}$$

représente une relation normative d'*interdiction*. L'énoncé *il est interdit que j fasse φ* signifie que i s'assure que si j réalise φ, alors j subit une conséquence négative. Finalement, l'ensemble

$$\{< i, < j, \varphi >>: \neg B_i E_j \varphi\}$$

est une relation de *permission*. L'énoncé *il est permis que j réalise φ* signifie qu'il est faux que i s'assure que j subisse une conséquence négative s'il réalise φ. En bref, une obligation, une interdiction et une permission s'expriment respectivement par:

$$E_i(\neg E_j \varphi \supset Bj) \tag{O}$$
$$E_i(E_j \varphi \supset Bj) \tag{F}$$
$$\neg E_i(E_j \varphi \supset Bj) \tag{P}$$

Il est possible de montrer que si φ est obligatoire, alors $\neg\varphi$ est interdit. Voici la

[10] La formulation anglaise rend mieux compte du sens de l'énoncé: *i sees to it that if j brings it about that φ, then something bad happens to j.*

[11] Soulignons que cette conception de l'obligation est similaire, comme nous le verrons, à la définition sévère de l'obligation proposée par Broersen.

preuve avec les règles de KM en déduction naturelle (restreintes à un opérateur indexé).

1	$E_i(\neg E_j\varphi \supset Bj)$	H
2	E_i	H
3	$\neg E_j\varphi \supset Bj$	$E_i$out 1,2
4	$E_j\neg\varphi$	H
5	$E_j\varphi$	H
6	$\neg\varphi$	(**M**) 4
7	φ	(**M**) 5
8	$\bot$	$\bot$in 6,7
9	$\neg E_j\varphi$	RAA 5-6
10	Bj	MP 3,9
11	$E_j\neg\varphi \supset Bj$	$\supset$in 4-10
12	$E_i(E_j\neg\varphi \supset Bj)$	$E_i$in 2-11

Cela dit, considérant que $\neg E_j\varphi$ n'implique pas $E_j\neg\varphi$, il est faux que l'interdiction implique l'obligation. De fait, nous n'avons pas l'équivalence habituelle $O\varphi \equiv F\neg\varphi$. Dans la langue naturelle, cela signifie qu'il est vrai que si φ est obligatoire, alors $\neg\varphi$ est interdit, mais qu'il est faux que si $\neg\varphi$ est interdit, alors φ est obligatoire. Par exemple, cela implique que (1) s'il est obligatoire de respecter la loi, alors il est interdit de ne pas respecter la loi, mais que (2) il est faux que s'il est interdit de voler, alors il est obligatoire de ne pas voler. Le second exemple peut être reformulé de la manière suivante: il est logiquement possible qu'il soit interdit de voler mais qu'il ne soit pas obligatoire de ne pas voler. Notons que cela implique directement que $P\varphi$ n'est pas équivalent à $\neg O\neg\varphi$. L'équivalence entre $\neg F\varphi$ et $P\varphi$ est triviale puisque l'interdiction et la permission se traduisent par la même formule:

$$\neg E_i(E_j\varphi \supset Bj) \qquad (\neg F\varphi)$$
$$\neg E_i(E_j\varphi \supset Bj) \qquad (P\varphi)$$

Cette équivalence exemplifie la relation entre O et P. Considérant que $O\varphi$ implique $F\neg\varphi$ et que $F\neg\varphi$ implique $\neg P\neg\varphi$, cela nous donne $O\varphi$ implique $\neg P\neg\varphi$. La réciproque n'est cependant pas vraie. En ce sens, la permission ne se résume pas à ce dont la négation n'est pas obligatoire. En fait, la permission se résume à ce qui n'est pas explicitement interdit.

Du côté de la sémantique, Pörn (1970, pp.34,35,37) ajoute trois conditions afin de rendre compte de la cohérence d'un système normatif (entendu comme un ensemble de relations normatives entre des agents).[12]

[12] La définition proposée par Pörn utilise en fait une méta-méta-variable $\phi(j)$ pour les clauses C1 et C2 (cf. Pörn 1970, p.31), laquelle réfère à un énoncé de la forme $E_j\psi$, où ψ est un énoncé de premier ordre (atomique ou complexe, et donc qui ne contient pas d'opérateur E_i), ou encore à un énoncé complexe composé à partir d'énoncés de cette forme. En ce qui nous concerne, nous allons nous restreindre à la présentation de la clause pour les énoncés de la forme $E_j\psi$.

$$E_i(\neg E_j\varphi \supset Bj) \in w \Rightarrow\ E_i(E_j\varphi \supset Bj) \notin w \tag{C1}$$
$$E_i(\neg E_j\varphi \supset Bj) \in w \Rightarrow\ E_k(E_j\varphi \supset Bj) \notin w \tag{C2}$$
$$E_i(\neg E_j\varphi \supset Bj) \in w \Rightarrow\ E_i(E_k\neg\varphi \supset Bk) \in w \tag{C3}$$

Autrement dit, (C1) stipule que si φ est obligatoire selon une autorité i, alors φ n'est pas interdit par cette même autorité. Plus encore, (C2) indique que si φ est obligatoire en vertu d'une autorité i, alors φ n'est pas interdit par une autre autorité k. Alors que la première clause vise à assurer une cohérence horizontale, la seconde assure une cohérence verticale (cf. chapitre 9). Finalement, la clause (C3) signifie que si i rend φ obligatoire pour j, alors i interdit à tout autre agent k de réaliser $\neg\varphi$.[13] Cette clause vise à assurer que j est en mesure de remplir ses obligations.

Soulignons que le rejet de l'équivalence entre les concepts d'*obligation* et d'*interdiction* est discutable d'un point de vue philosophique. À ce sujet, le lecteur est référé au chapitre 16. Pour une analyse approfondie de la distinction entre *action* et *omission*, voir le chapitre 14.

Somme toute, Pörn propose d'analyser les énoncés normatifs comme des relations entre des agents et des propositions, un système normatif étant un ensemble d'agents liés par des relations normatives (cf. Pörn 1970, p.32). Un énoncé d'obligation, par exemple, marque une relation normative entre une autorité i, un subordonné j et une proposition φ, où l'obligation pour j de faire φ signifie que i s'assure de punir j dans l'éventualité où il ne réalise pas son obligation φ. Comme le mentionne Pörn (1970, p.9), la logique de l'action à elle seule ne permet pas de rendre compte adéquatement de l'imputabilité et de l'intention puisque celle-ci n'inclut pas de modalités épistémiques et temporelles. Il s'agit donc d'une limite des logiques de type STIT, lesquelles ne parviennent pas à distinguer entre le résultat de la performance de l'action, un état de choses qui prend place suite à l'action ou encore la conséquence de l'action (Pörn 1970, p.5). Les approches présentées dans les sections qui suivent tentent de pallier ces limites. Pacheco et Carmo cherchent à rendre compte de la notion de responsabilité en introduisant le concept d'*agir dans un rôle*, et Broersen introduit des modalités épistémiques et temporelles afin de rendre compte de l'intentionnalité. Avant d'aborder ces approches, voyons en premier lieu celle de Horty afin de nous familiariser avec la logique STIT.

John Horty

Le livre de Horty (2001), intitulé *Agency and deontic logic*, est à ce jour probablement l'ouvrage le plus complet qui propose une analyse de la logique déontique dans le cadre d'une logique STIT. Bien que le propos de Horty ne soit pas axé sur l'analyse du discours

[13] Nous avons modifié (C3) avec une formulation équivalente afin d'exprimer explicitement l'explication que l'on trouve à la page 38: *if i makes it punishable for j not to bring it about that φ, then i makes it punishable for an arbitrary k to bring it about that $\neg\varphi$* (Pörn 1970, p.38).

légal (ce dernier propose plutôt une analyse de la notion d'agentivité (*agency*) dans un cadre utilitariste), nous avons jugé bon de présenter sa position considérant que celle-ci offre une excellente vue d'ensemble concernant l'intérêt d'utiliser une logique STIT pour l'analyse du discours normatif. Malgré que l'approche de Horty vise l'analyse de ce que les agents *doivent faire* (*ought-to-do*), l'analyse qu'il propose prend place dans le cadre plus large de ce qui *doit être* (*ought-to-be*). Cela se justifie par le fait que ce qu'un agent *doit faire* peut être identifié par ce qui *doit être* en vertu des actions de cet agent (Horty 2001, p.4). En ce sens, ce qu'un agent *doit faire* est subsumé par ce qui *doit être* dans le monde:

> The study of what agents ought to do, it is thought, can naturally be subsumed under the broader study of what ough to be, for the simple reason that among the various things that ought or ought not to be are the things that agents do, the actions they perform or refrain from (Horty 2001, p.3).

Ainsi, l'objectif de Horty est de « formuler précisément l'idée à savoir que ce qu'un agent *doit faire* peut s'identifier par ce qui *doit être* suite aux actions d'un agent (traduction libre, mes italiques, Horty 2001, p.5) ».[14]

Malgré que le système syntaxique ne soit pas explicitement défini dans l'ouvrage, nous allons tenter de présenter la syntaxe implicite à l'approche de Horty.[15] Le langage

$$\mathcal{L} = \{(,), Prop, \neg, \wedge, \Box, [a\ stit], O, \odot\}$$

comprend un ensemble dénombrable de propositions $Prop = \{p_1, ..., p_n, ...\}$, les connecteurs logiques $\neg$ et $\wedge$, à partir desquels les autres connecteurs $\vee$, $\supset$ et $\equiv$ sont définis usuellement (Horty 2001, p.9), et l'opérateur modal $\Box$ représente la nécessité historique. Considérant qu'un même scénario peut faire partie de plusieurs histoires différentes, la nécessité historique est à comprendre en termes de ce qui est *fixé pour toute histoire*:

> The formula $\Box A$ is taken to mean that A is settled, or historically necessary; $\Diamond A$, that A is still open as a possibility. The intuitive idea is that $\Box A$ should be true at some moment if A is true at that moment no matter how the future turns out, and that $\Diamond A$ should be true if there is still some way in which the future might evolve that would lead to the truth of A (Horty 2001, p.9).

Autrement dit, ce qui est historiquement nécessairement vrai pour un scénario est ce qui est vrai pour toute histoire qui passe par ce scénario. Reformulons cette idée: soit une *histoire* une suite de scénarios $w_i, \ldots, w_j, \ldots, w_k$. La notion de nécessité historique signifie que ce qui est nécessairement vrai pour w_j correspond à ce qui est vrai pour toute histoire incluant w_j.

---

[14] Malgré cette remarque, Horty assumera néanmoins deux opérateurs primitifs de types *Ought-to-be* et *Ought-to-do*.

[15] Horty (2001, p.18) laisse de côté les considérations syntaxiques relatives à la théorie de la preuve de son approche.

L'opérateur $[a\ stit]A$, plus précisément $[\{a\}\ stit]A$ comme nous le verrons bientôt, signifie *l'agent a réalise A*, où A est un état du monde qui advient suite aux actions ou aux choix de l'agent a.[16]

> A statement of the form $[a\ stit]A$, expressing the idea that the agent a sees to it that A, is defined as true at an index (w,h) just in case the action performed by a at that index garantees the truth of A; the action might result in a variety of possible outcomes, but the statement A must be true in each of them (Horty 2001, p.15).

L'opérateur O représente *ce qui doit être*, alors que $\odot$ dénote *ce qui doit être fait*, où $\odot$ ne s'applique qu'aux propositions de la forme $[a\ stit]A$ et signifie *a doit réaliser A* (Horty 2001, p.79).

En vertu de la clause sémantique (4) définie plus bas, $\Box$ peut être axiomatisé comme une modalité de type $S5$.[17] En effet, le schéma (**M**) est trivial en vertu de la condition (4) et le schéma (**5**) peut aisément être démontré.

Théorème 3.1. *$\Box$ est une modalité de type $S5$.*

Démonstration: Supposons que $\models_{\mathcal{M}_{w,h}} \Diamond A$ mais que $\models_{\mathcal{M}_{w,h}} \neg\Box\Diamond A$. D'une part, cela signifie que $\not\models_{\mathcal{M}_{w,h}} \Box\neg A$, et donc qu'il y a au moins une histoire $h' \in H_w$ pour laquelle $\models_{\mathcal{M}_{w,h'}} A$. D'autre part, cela signifie que $\not\models_{\mathcal{M}_{w,h}} \Box\Diamond A$, et donc qu'il y a un $k \in H_w$ tel que $\models_{\mathcal{M}_{w,k}} \neg\Diamond A$. Or, cela implique que $\models_{\mathcal{M}_{w,k}} \Box\neg A$, et donc que $\models_{\mathcal{M}_{w,h'}} \neg A$ pour tout $h' \in H_w$, ce qui contredit l'hypothèse de départ. Le schéma d'axiome (**5**) $\Diamond A \supset \Box\Diamond A$ est donc valide, ce qui conjointement à (**M**) signifie que $\Box$ est une modalité axiomatisable par $S5$.[18]

$\Box$

Par ailleurs, $[a\ stit]$ est aussi un opérateur de type $S5$, alors que O est de type $KD45$ (Horty 2001, p.17 et pp.35-6,43).[19] L'opérateur $\odot$ quant à lui est une modalité de type KD (cf. Horty 2001, p.79). Notons aussi que $\odot$ n'est pas une modalité de type $KD45$ considérant que l'itération de l'opérateur $\odot$ n'est pas permise, $\odot$ étant une modalité de type *Ougth-to-do* (cf. Horty 2001, p.74). Prenant tout cela en compte, l'approche de Horty peut être axiomatisée à l'aide des schémas d'axiomes suivants, assumant les schémas d'axiomes ainsi que les règles d'inférence du système modal K.

[16] À l'instar de Broersen, nous utilisons $[a\ stit]A$ plutôt que $[a\ stit : A]$ afin d'insister sur le fait que l'opérateur $[a\ stit]$ est une modalité.

[17] Voir le tableau 3.2 pour les types de modalité.

[18] Une manière plus triviale de voir la chose est de considérer les schémas (**M**) , (**4**) et (**B**), ce qui est aussi équivalent à $S5$ (cf. Garson 2006, p.222).

[19] Soulignons qu'une modalité de type $KD45$ est plus faible qu'une modalité de type $S5$.

$$\Box A \supset A \tag{$\mathbf{M}_\Box$}$$
$$\Diamond A \supset \Box\Diamond A \tag{$\mathbf{5}_\Box$}$$
$$[a\ stit]A \supset A \tag{$\mathbf{M}_{\text{STIT}}$}$$
$$\neg[a\ stit]\neg A \supset [a\ stit]\neg[a\ stit]\neg A \tag{$\mathbf{5}_{\text{STIT}}$}$$
$$\neg(OA \wedge O\neg A) \tag{$\mathbf{D}_O$}$$
$$\neg O\neg A \supset O\neg O\neg A \tag{$\mathbf{5}_O$}$$
$$OA \supset OOA \tag{$\mathbf{4}_O$}$$
$$\neg(\odot[a\ stit]A \wedge \odot[a\ stit]\neg A) \tag{$\mathbf{D}_\odot$}$$

Le principe *devoir implique pouvoir* est représenté par les axiomes ($\Diamond O$) et ($\Diamond\odot$) pour les modalités de types *Ought-to-be* et *Ought-to-do*.

$$OA \supset \Diamond A \tag{$\Diamond O$}$$
$$\odot\ [a\ stit]A \supset \Diamond[a\ stit]A \tag{$\Diamond\odot$}$$

Finalement, la cohérence entre ce qui *doit être* et ce qui *doit être fait* est exprimée par le schéma d'axiome suivant (Horty 2001, p.80):

$$\neg(\odot[a\ stit]A \wedge O[a\ stit]\neg A) \tag{$\mathbf{D}_{O\odot}$}$$

L'ensemble des expressions bien formées EBF peut être défini récursivement par:

1. si $p \in Prop$, alors $p \in EBF$;
2. si $A, B \in EBF$, alors $\neg A, A \wedge B \in EBF$;
3. si $A \in EBF$, alors $\Box A, [a\ stit]A, OA \in EBF$;
4. si $A \in EBF$, alors $\odot[a\ stit]A \in EBF$.

Comme Horty (2001, p.6) le mentionne, son approche repose sur les logiques temporelles (*theory of branching time*) proposées par Prior (1967) et Thomason (1970, 2002). Ces logiques, qui reposent sur la notion d'*arbre temporel*, conçoivent le passé comme *déterminé* mais l'avenir comme *incertain* (*indéterminé*). Une structure temporelle $\mathcal{S} = \langle U, < \rangle$ est composée d'un univers U non vide de *moments*, lesquels peuvent aussi être considérés comme des scénarios, à l'intérieur duquel se trouve une relation d'ordre transitive et antiréflexive $<$. Dans une structure temporelle, une histoire h est une suite temporelle maximale de moments liés par la relation $<$, et H_w représente l'ensemble des histoires qui contiennent le moment w (Horty 2001, pp.6,8).

$$H_w = \{h : w \in h\}$$

En ce sens, H_w représente l'ensemble des suites d'évènements possibles pouvant passer par w. Un modèle temporel $\mathcal{M} = \langle U, <, a\rangle$ est défini à partir de la structure $\mathcal{S}$, à laquelle on ajoute une fonction a qui détermine les paires moment/histoire auxquelles les propositions dans U appartiennent (Horty 2001, p.8).[20] La vérité dans le modèle est définie récursivement (Horty 2001, pp.9-10)[21]:

$$\models_{\mathcal{M}_{(w,h)}} p \Leftrightarrow\ (w,h) \in a(p) \tag{1}$$

$$\models_{\mathcal{M}_{(w,h)}} \neg A \Leftrightarrow \not\models_{\mathcal{M}_{(w,h)}} A \tag{2}$$

$$\models_{\mathcal{M}_{(w,h)}} A \wedge B \Leftrightarrow \models_{\mathcal{M}_{(w,h)}} A \text{ et } \models_{\mathcal{M}_{(w,h)}} B \tag{3}$$

$$\models_{\mathcal{M}_{(w,h)}} \Box A \Leftrightarrow \models_{\mathcal{M}_{(w,h')}} A \text{ pour toute histoire } h' \in H_w \tag{4}$$

La vérité et la validité sont définies comme à l'habitude, où *valide* signifie vrai pour toute interprétation $\mathcal{M}_{(w,h)}$, c'est-à-dire pour tout modèle $\mathcal{M}$ indexé à une paire moment/histoire (Horty 2001, p.9).

Jusqu'à présent, seul un modèle *temporel* a été défini. Afin de pouvoir définir un modèle STIT, Horty (2001, p.11) introduit l'ensemble:

$$\|A\|_w^{\mathcal{M}} = \{h \in H_w : \models_{\mathcal{M}_{(w,h)}} A\}$$

Cet ensemble contient toutes les histoires h où la proposition A est vraie selon le modèle $\mathcal{M}$ au moment w. Clairement, nous avons $\|A\|_w^{\mathcal{M}} \subseteq H_w$, c'est-à-dire que l'ensemble qui contient les histoires où A est vrai pour w est un sous-ensemble de l'ensemble qui contient toutes les histoires qui passent par le moment w. Il est aussi évident que $\|\neg A\|_w^{\mathcal{M}} = H_w - \|A\|_w^{\mathcal{M}}$, à savoir que l'ensemble qui contient les histoires où A est faux est le complément relatif à H_w de l'ensemble qui contient les histoires où A est vrai. Un modèle STIT est obtenu en ajoutant un ensemble non vide d'agent(s) Ag et une fonction de choix Ch au modèle temporel (Horty 2001, p.14).

$$\mathcal{M} = \langle U, <, Ag, Ch, a\rangle$$

La fonction de choix Ch_a^w est indexée à un moment w et à un agent (ou groupe d'agents) a. Informellement, la fonction Ch_a^w isole les histoires possibles qui s'offrent à a au moment w (Horty 2001, p.13). En ce sens, Ch_a^w est un ensemble d'histoires.

L'auteur utilise l'expression $Ch_a^w(h)$ afin de référer à l'action que a pose au moment w afin d'obtenir l'histoire h (Horty 2001, p.13). Autrement dit, alors que Ch_a^w offre l'ensemble des alternatives possibles, $Ch_a^w(h)$ dénote le choix effectué par a au moment w. Il s'agit d'un ensemble d'histoires sélectionnées par les actions de a. La clause sémantique ajoutée pour obtenir un modèle STIT est la suivante (Horty 2001, p.15).

$$\models_{\mathcal{M}_{(w,h)}} [a\ stit]A \Leftrightarrow\ Ch_a^w(h) \subseteq \|A\|_w^{\mathcal{M}} \tag{5}$$

[20] C'est-à-dire que $a(A) = \{(w,h) :\ A \in w \text{ et } w \in h\}$.

[21] Le modèle tel que défini dans le livre inclut des clauses pour les opérateurs représentant le *passé* et le *futur*. Cela dit, ces opérateurs ne seront pas réutilisés dans l'ouvrage, et de fait nous les avons laissés de côté.

Cette clause est justifiée par le fait que l'on assume que si un agent a réalise A (où A est une description du monde), alors A est vrai suite aux actions posées par a.

> The idea that an agent a sees to it that A is taken to mean that the truth of the proposition A is guaranteed by an action, or choice, of a (Horty 2001, p.12).

En ce sens, la clause signifie que s'il est vrai que a réalise A, alors l'ensemble des choix possibles que a peut faire au moment w et qui mènent à l'histoire h est un sous-ensemble de l'ensemble qui contient les histoires où la proposition A est vraie au moment w.

Notons au passage que cela marque la principale distinction entre l'approche de Broersen, basée sur une logique XSTIT, et la logique STIT. Dans le modèle proposé par Horty, l'effet de l'action est immédiat et prend place dans le même moment (dans le même scénario), alors que chez Broersen l'effet de l'action prend place dans le scénario suivant. La principale distinction entre XSTIT et STIT est que XSTIT différencie entre les scénarios et les moments alors que pour STIT il s'agit de la même chose.

Soulignons aussi que Horty utilise en fait l'opérateur [a *cstit*], qui représente le « Chellas STIT » (Horty 2001, p.14), inspiré des travaux de Chellas (1969) et auquel nous avons référé par [a *stit*]. Horty (2001, p.16) met en opposition l'opérateur [a *cstit*] avec un opérateur *délibératif* [a *dstit*], lequel diffère du premier de par l'ajout d'une condition négative sur la clause sémantique (Horty 2001, p.16).

$$\models_{\mathcal{M}_{(w,h)}} [a\ dstit]A \Leftrightarrow\ Ch_a^w(h) \subseteq \|A\|_w^{\mathcal{M}} \text{ et } \|A\|_w^{\mathcal{M}} \neq H_w$$

Cette condition négative vise à rendre compte du fait que les états du monde réalisés par un agent après délibération sont des états qui n'étaient pas *inévitables* (Horty 2001, p.16). Autrement dit, l'opérateur [a *dstit*] est caractérisé par le fait que si a réalise A après délibération, alors A n'est ni une tautologie ni une nécessité historique.[22] En effet, si $\|A\|_w^{\mathcal{M}} = H_w$, alors A est vrai pour toute histoire passant par le moment w, et donc par la clause (4) A est une nécessité historique. Cela dit, après une discussion sur les vertus de chacun des opérateurs (cf. Horty 2001, pp.16-19), Horty (2001, p.19) en vient à montrer que les opérateurs [a *cstit*] et [a *dstit*] sont interdéfinissables, et donc qu'il suffit de prendre le « Chellas STIT » comme primitif (Horty 2001, p.29).

$$[a\ dstit]A =_{def} [a\ cstit]A \wedge \neg\Box A \qquad \text{(def. } dstit\text{)}$$

En plus de développer un modèle qui permet de rendre compte de l'action d'un individu, Horty (2001, p.32) adapte le modèle afin rendre compte des actions faites en groupe (*group agency*). Pour ce faire, Horty (2001, p.31) introduit dans le modèle un ensemble Se de fonctions de sélection $s : Ag \longrightarrow H_w$ qui associent une histoire à un agent.

$$Se_w = \{s : s(a) \in Ch_a^w\}$$

[22] Notons qu'une tautologie est une nécessité historique en vertu de la définition de cette dernière.

Informellement, les fonctions de sélection isolent, lorsqu'elles sont regroupées, les histoires compatibles avec les choix de tous les agents. En effet, cette fonction doit satisfaire la condition (I), qui stipule que l'intersection de tous les $s(a)$ pour tout $a \in Ag$ n'est pas vide.[23]

$$\bigcap_{a \in Ag} s(a) \neq \varnothing \tag{I}$$

Cette condition assure qu'il y a au moins une histoire h dans laquelle chaque agent peut réaliser son choix indépendamment du choix des autres agents (Horty 2001, p.31). À l'aide de cet ensemble, Horty redéfinit la fonction Ch^w_Γ relativement à un groupe d'agents $\Gamma \subseteq Ag$.[24]

$$Ch^w_\Gamma = \{\bigcap_{a \in \Gamma} s(a) : s \in Se_w\}$$

Cela fait, Horty (2001, p.32) modifie la clause sémantique pour l'opérateur $[\Gamma\ stit]$ par:

$$\models_{\mathcal{M}_{(w,h)}} [\Gamma\ stit]A \Leftrightarrow\ Ch^w_\Gamma(h) \subseteq \|A\|^{\mathcal{M}}_w \tag{5'}$$

Il pose comme définition:

$$[a\ stit]A =_{def} [\{a\}\ stit]A \tag{def. [a stit]}$$

En mots, (5') exprime qu'un groupe d'agent réalise A à condition que l'ensemble des histoires accessibles au groupe soit sous-ensemble de l'ensemble qui contient les histoires où A est vrai.

La conception que Horty se fait de l'ensemble des choix d'un groupe s'explique par le fait qu'il considère qu'un groupe d'agents se résume aux agents qui le composent.

> As our example suggests, group actions can be defined as patterns of individual actions: an action available to a group of agents can be defined as an intersection of the actions available to the individual agents belonging to that group, one action for each agent (Horty 2001, p.30).

Notons que cette conception serait critiquée par Carmo et Pacheco (2001, p.137), lesquels soutiennent qu'un groupe ne se réduit pas aux agents individuels qui le composent.

Ayant exposé sa logique de l'action, l'auteur en vient à considérer les opérateurs déontiques.[25] L'auteur en vient à développer un modèle utilitariste en introduisant une

[23] Soulignons que la notation devrait plutôt être $\bigcap_{a \in Ag}\{s(a)\} \neq \varnothing$. La condition ne vise pas à dire que chaque histoire $s(a)$ contient au moins un même scénario, ce qui est trivial considérant que chaque $h \in H_w$, mais bien que l'ensemble des histoires accessibles lorsque l'on considère chaque agent n'est pas vide.

[24] Idem: nous devrions plutôt lire $Ch^w_\Gamma = \bigcap_{a \in \Gamma}\{s(a)\} : s \in Se_w$.

[25] Par souci d'économie, nous ne considérerons que les opérateurs de type *Ought-to-be* et *Ought-to-do*, laissant de côté les obligations conditionnelles analysées à l'aide d'une relation de préférence.

fonction de valeur V ainsi qu'un ordre partiel $\sqsubseteq$. Le modèle utilitariste général est défini à partir d'un modèle STIT.

$$\mathcal{M} = \langle U, <, Ag, Ch, a, V, \sqsubseteq \rangle$$

Soulignons que le modèle utilitariste ne présuppose aucune théorie morale autre qu'une forme d'utilitarisme en son sens large (Horty 2001, p.38). En effet, l'auteur ne précise pas quels critères permettent de déterminer l'ordre partiel, mais ne fait que présupposer une sémantique des préférences qui assume que certains scénarios sont meilleurs que d'autres, nonobstant ce que signifie la relation *être meilleur que*. Informellement, la fonction $V_w(h)$ détermine la valeur d'un moment w relativement à une histoire h. La relation d'ordre partiel $\sqsubseteq$ est une relation réflexive, transitive et antisymétrique et représente informellement la *préférence*. Le modèle doit satisfaire deux conditions (Horty 2001, p.41):

1. $V : U \times \wp(U) \longrightarrow \mathbb{R}$ est une fonction qui associe à chaque paire (w, h) une valeur dans les nombres réels;

2. l'assignation de valeurs pour une histoire ne varie pas d'un moment à l'autre, c'est-à-dire que $V_w(h) = V_{w'}(h)$ pour tout moment $w \in h$ et $w' \in h$.

La première condition vise simplement à fournir un moyen de quantifier la valeur accordée à une histoire alors que la seconde signifie que la valeur d'une histoire est basée sur l'histoire dans son ensemble. La clause sémantique pour O est représentée par (6).

$$\begin{aligned} \models_{\mathcal{M}_{(w,h)}} OA \Leftrightarrow\ & \text{il y a } h' \in H_w \text{ t.q.} \\ & 1.\ \models_{\mathcal{M}_{(w,h')}} A \text{ et} \\ & 2.\ \models_{\mathcal{M}_{(w,h'')}} A \text{ pour tout } h'' \in H_w \text{ t.q. } V_w(h') \sqsubseteq V_w(h'') \end{aligned} \tag{6}$$

Autrement dit, il est vrai que A *devrait être* pour un moment w dans la mesure où A est vrai dans au moins une histoire h' qui passe par w et où A est vrai pour toute autre histoire qui a une plus grande valeur que h' (Horty 2001, p.37). La première sous-condition assure qu'il est possible de réaliser A alors que la seconde implique que toute histoire où A est vrai a plus de valeur qu'une histoire où A est faux.

Bien que Horty (2001, p.45) propose une analyse de ce qui *doit être fait* en termes de ce qui *doit être*, ce dernier en vient à définir un opérateur primitif $\odot$ pour représenter ce qui *doit être fait* plutôt que de le définir en fonction de O. La principale raison pour cette manœuvre est que la réduction de ce qui *doit être fait* à ce qui *doit être* pose problème dans un cadre utilitariste (cf. Horty 2001, p.58 et le *gambling problem*). Horty (2001, p.74) introduit l'opérateur $\odot$, lequel ne porte que sur des expressions bien formées de type $[a\ stit]A$ et signifie que l'agent a doit réaliser A. Informellement, la clause sémantique pour cet opérateur est que a doit réaliser A à condition que pour toute action accessible à a pour un moment w qui ne garantit pas la vérité de A, il y a une autre action préférable qui la garantit (Horty 2001, p.77). Afin de bien saisir cette clause, il faut cependant introduire quelques notions. Considérons d'abord le cas de la préférence entre les propositions (Horty 2001, p.60). Une proposition A est dite *préférable* à une proposition B pour un moment w à condition que toute histoire où A est vrai pour w soit préférable à toute histoire où B est vrai pour w. Formellement:

$$\begin{aligned} B \sqsubseteq A \Leftrightarrow\ & \|B\|_w^{\mathcal{M}} \sqsubseteq \|A\|_w^{\mathcal{M}} \\ \Leftrightarrow\ & V(h) \sqsubseteq V(h') \text{ pour tout } h \in \|B\|_w^{\mathcal{M}} \text{ et } h' \in \|A\|_w^{\mathcal{M}} \end{aligned}$$

Soit $S_a^w = Ch_{Ag-\{a\}}^w$ l'ensemble qui comprend les choix accessibles à tous les agents excluant ceux accessibles à l'agent a (Horty 2001, p.67). Chaque $s \in S_a^w$ représente une histoire accessible à un agent autre que a. L'expression $k \cap s$ représente l'intersection entre deux histoires à un moment w. Autrement dit, $k \cap s$ représente ce qui est vrai conjointement de k et de s au moment w. En ce sens, $k \cap s$ peut être vue comme une proposition complexe qui décrit ce qui est vrai de w à la fois dans l'histoire k et l'histoire s. Ayant cela en tête, Horty (2001, p.68) introduit une relation $\preceq$ de *dominance* entre les choix (actions) des agents. Informellement, cette relation vise à exprimer que l'histoire réalisée par un agent (en vertu de ses actions ou de ses choix) est *préférable* à une autre dans la mesure où les actions de l'agent mènent à une histoire qui a une plus grande valeur peu importe les actions des autres agents. Si l'on considère que $k \cap s$ représente la suite des évènements qui prend place après les actions de a avec celles d'un autre agent b, alors la relation de dominance exprime que k est préférable à toute autre histoire k' accessible à a peu importe le choix de n'importe quel agent b.[26] Formellement, nous avons:

$$k \preceq k' \Leftrightarrow\ k \cap s \sqsubseteq k' \cap s \text{ pour tout } s \in S_a^w$$

Autrement dit, k domine k' à condition que k soit préférable à k' nonobstant les actions des autres agents. Ayant tout cela en tête, la clause sémantique pour $\odot$ s'exprime formellement par la condition suivante (Horty 2001, p.77)[27]:

$$\begin{aligned} &\models_{\mathcal{M}_{(w,h)}} \odot[a\ stit]A \Leftrightarrow \text{ pour tout } k \in Ch_a^w \text{ t.q. } k \notin \|A\|_w^{\mathcal{M}} \qquad (7) \\ &\text{il y a } k' \in Ch_a^w \text{ t.q.} \\ &\quad 1.\ k \prec k' \\ &\quad 2.\ k' \in \|A\|_w^{\mathcal{M}} \\ &\quad 3.\ k'' \in \|A\|_w^{\mathcal{M}} \text{ pour tout } k'' \in Ch_a^w \text{ t.q. } k' \preceq k'' \end{aligned}$$

En mots, cela signifie qu'il est vrai que a doit réaliser A au moment w à condition que pour toute histoire k où A est faux, il y a une histoire k' où A est vrai qui est strictement

[26] On trouve une coquille au niveau de la notation dans l'approche de Horty relativement à la relation de dominance. En effet, relativement à la condition (7), ce dernier écrit $k \subseteq \|A\|_w^{\mathcal{M}}$ plutôt que $k \in \|A\|_w^{\mathcal{M}}$. Or, k est une histoire et $\|A\|_w^{\mathcal{M}}$ est un ensemble d'histoire, et de fait k peut être un membre de $\|A\|_w^{\mathcal{M}}$ mais n'est pas un sous-ensemble.

[27] Notons que Ch_a^w doit être fini afin de s'assurer qu'il y ait au moins une suite d'évènements préférable. Voir Horty (2001, p.73) pour l'argument en faveur de cette condition.

préférable à k et pour laquelle toute autre histoire k'' où A est vrai est préférable à k'. La première condition assure que k' est préférable (considérée conjointement aux suites d'évènements possibles pouvant être engendrées par les autres agents) à tout autre histoire k où A est faux. La seconde signifie que A est vrai dans cette histoire, et la troisième implique que A est vrai dans toute autre histoire qui y est préférable.

Tout compte fait, notre présentation de Horty (2001) s'est limité aux chapitres 2, 3 et 4 de son ouvrage, principalement en raison du fait que cela constitue le cœur de son approche. L'analyse qu'il propose dans les chapitres 5 et 6 vise certains problèmes spécifiques à la logique déontique, comme l'introduction d'un opérateur dyadique pour la formalisation des obligations contraires et la notion d'obligation de groupe. Le chapitre 7 introduit des considérations relatives aux stratégies employées en vue de certaines fins, prenant en compte la capacité que les agents ont à accomplir certaines actions.[28] L'ouvrage de Horty occupe une place considérable dans la littérature portant sur la logique déontique, et le langage qu'il fournit est riche et permet de rendre compte de la complexité du discours normatif quotidien. Outre l'analyse de Horty, qui s'intéresse plutôt au discours éthique, d'autres approches en logique déontique de type STIT se sont penchées explicitement sur la formalisation du discours légal. Voyons maintenant deux exemples de telles approches.

Olga Pacheco et José Carmo

Parmi les approches qui utilisent les logiques de type STIT, certains cherchent à modéliser la structure d'un NMAS par le biais de la notion de *rôle*. Pacheco et Carmo (2003, p.148) introduisent cette notion en se basant sur l'analyse de la loi. D'emblée, le premier point qu'ils mettent en évidence est qu'un agent n'est pas nécessairement une personne *physique* et qu'il existe des agents *artificiels* (Pacheco et Carmo 2003, p.149). Cela est mis en évidence par le fait que d'un point de vue légal, certaines institutions (ou organisations) possèdent la *personnalité juridique* (p. ex., une entreprise), et sont donc sujettes à des obligations, des droits, des permissions, etc. Notons que cette conception de l'*agent* n'est pas surprenante d'un point de vue informatique, perspective dans laquelle s'insère l'approche de Pacheco et Carmo: un agent peut très bien être un programme, lequel pose des actions et dont le comportement est contraint par des règles (des normes, des instructions). Cela dit, un agent peut être non seulement artificiel mais peut aussi être *complexe*, voire *collectif*. Au même titre qu'un programme informatique peut être composé de sous-programmes, une organisation qui possède la personnalité juridique est composée de membres, et possiblement de sous-organisations. L'analyse du discours légal que font Pacheco et Carmo (2003, p.149) met en évidence qu'une organisation, en tant qu'agent collectif, possède certaines caractéristiques:

1. l'organisation est un agent artificiel qui n'agit pas *directement* mais agit plutôt indirectement par le biais des actions posées par les agents individuels qui le composent;

[28] Voir aussi Kooi et Tamminga (2006) pour l'analyse des actions de groupe.

2. elle possède un ensemble de rôles déterminés, un ensemble de normes qui caractérisent ces rôles et une structure stable (indépendante des agents qui remplissent les rôles);

3. certains rôles peuvent être délégués, c'est-à-dire qu'un agent peut agir en tant que représentant pour un autre agent.

Le point fondamental que soulignent les auteurs et sur lequel repose leur approche est qu'un agent agit toujours dans un certain *rôle* (Pacheco et Carmo 2003, p.152). Cela se comprend notamment en raison du fait que le comportement d'un agent est évalué en fonction du *rôle* dans lequel il a posé une action (Pacheco et Santos 2004, p.213). Par exemple, un policier qui utilise son arme dans une situation de légitime défense sera jugé en tant que policier, et non en tant que simple citoyen. En un mot, ce ne sont pas tous les rôles qui jouissent des mêmes privilèges, des mêmes droits, et qui sont contraints par les mêmes restrictions.

Le caractère innovateur de Pacheco et Carmo (2003, p.159) est de modifier l'opérateur de type STIT afin d'indexer l'action d'un agent à un rôle. L'opérateur $E_{a:r}B$ se lit « l'agent a, qui agit dans le rôle r, réalise B ». Dans le même ordre d'idées, ils indexent les opérateurs $O_{a:r}$, $P_{a:r}$ et $F_{a:r}$ à des agents qui agissent dans certains rôles. Notons que les opérateurs sont une abréviation de $OE_{a:r}$, $PE_{a:r}$ et $FE_{a:r}$, lesquels se lisent respectivement « il est obligatoire que a dans le rôle r réalise... », « il est permis que a dans le rôle r réalise... » et « il est interdit que a dans le rôle r réalise... ». Contrairement aux systèmes standards, l'opérateur $P_{a:r}$ est pris en tant que primitif et l'opérateur $F_{a:r}$ est défini par:

$$F_{a:r} =_{def} \neg P_{a:r} \qquad \text{(def. } F_{a:r}\text{)}$$

L'argument en faveur de ne pas définir l'interdiction et la permission en termes d'obligation repose sur le fait que les énoncés $OE_{a:r}\neg B$ et $O\neg E_{a:r}B$ ne sont pas équivalents (cf. Carmo et Pacheco 2000, p.100). D'une part, le premier énoncé signifie que l'agent a dans le rôle r a l'obligation de réaliser l'état de choses décrit par la proposition $\neg B$. Autrement dit, $\neg B$ est le résultat des actions de a. De l'autre côté, $O\neg E_{a:r}B$ signifie qu'il est obligatoire que a dans le rôle r ne réalise pas l'état de choses décrit par B, ce qui peut résulter autant d'une omission que d'une abstention. De fait, $O_{a:r}$ et $P_{a:r}$ sont tous deux pris en tant que primitifs afin d'avoir $O_{a:r}\neg B \equiv OE_{a:r}\neg B$, $O\neg E_{a:r}B \equiv F_{a:r}B$ mais $O_{a:r}\neg B$ non équivalent à $F_{a:r}B$. Cela découle du fait que la combinaison des opérateurs déontiques à un opérateur de type STIT engendre que l'obligation et l'interdiction ne signifient pas la même chose.

La syntaxe de leur approche est définie de la manière suivante (Pacheco et Carmo 2003, p.160). Le langage $\mathcal{L}_{\mathcal{DA}}$ contient un nombre fini $(s_1, ..., s_n)$ de types de variables et de constantes, incluant Ag (les variables/constantes d'agents), R (les variables/constantes de rôle) et AgR (les variables/constantes d'agents dans des rôles). Ces types contiennent quant à eux un nombre dénombrable de variables et au moins une constante. De plus, $\mathcal{L}_{\mathcal{DA}}$ contient un ensemble P de prédicats, incluant rg (un générateur de rôle), $is - rg$ un prédicat de qualification, où $is - rg(a, t_1, ..., t_n) = qual(a : rg(t_1, ..., t_n))$ et signifie que a

est qualifié pour le rôle $rg(t_1, ..., t_n)$, ou de façon équivalente que le rôle de a est décrit par $r(t_1, ..., t_n)$. Par exemple, $is - membre(k, i)$ pourrait signifier que k est un membre de l'organisation i. En ce sens, un générateur de rôle associe un prédicat à un rôle.

Afin de définir l'ensemble des expressions bien formées (EBF), les auteurs doivent d'abord définir ce qu'est un *terme*.

1. Si rg est de type $(s_1, ..., s_n, R)$ et que $t_1, ..., t_n$ sont des variables de type $(s_1, ..., s_n)$, alors $rg(t_1, ..., t_n)$ est un terme (de type R).

2. Si t est un terme de type Ag et $rg(t_1, ..., t_n)$ un terme de type R, alors $t : rg(t_1, ..., t_n)$ est un terme de type AgR.

En un mot, un terme d'un certain type est construit à partir de variables ou de constantes du même type. De manière plus générale, les constantes et les variables sont des termes. Cela nous amène à la définition de l'ensemble des expressions bien formées (Pacheco et Carmo 2003, p.161):

1. si $p \in P$ et $t_a, ..., t_n$ sont des termes du même type que p, alors $p(t_1, ..., t_n) \in EBF$;

2. si $B, \in EBF$, alors $\neg B, B \wedge C \in EBF$;

3. si $B \in EBF$ et x^s est une variable de type s, alors $(\forall x^s)B \in EBF$;

4. si $B \in EBF$ et $a : r$ est un terme de type AgR, alors $E_{a:r}B, O_{a:r}B, P_{a:r}B \in EBF$.

Les autres connecteurs logiques $\vee$, $\rightarrow$, $\leftrightarrow$ et $\exists$ sont définis usuellement à partir de $\neg$, $\wedge$ et $\forall$. Considérant qu'un agent agit à l'intérieur d'un rôle donné, une obligation B dans un rôle r signifie que tout agent a qualifié au rôle r doit réaliser B, c'est-à-dire que O_rB signifie:

$$(\forall x)(qual(x : r) \rightarrow O_{x:r}B)$$

Similairement, si B est permis pour un rôle R, alors B est permis à tout agent qualifié pour le rôle r, et donc P_rB signifie:

$$(\forall x)(qual(x : r) \rightarrow P_{x:r}B)$$

Finalement, les auteurs présentent leur système sous forme axiomatique, dont les axiomes et règles d'inférence se divisent en trois groupes (Pacheco et Carmo 2003, p.162)[29], sans compter les axiomes de la logique propositionnelle classique:

1. Les axiomes et règles d'inférence de premier ordre:

$$\vdash B \Rightarrow \vdash (\forall x)B \tag{Gen}$$
$$(\forall x)(B \rightarrow C) \rightarrow ((\forall x)B \rightarrow (\forall x)C) \tag{A1}$$
$$B \rightarrow (\forall x)B \qquad (x \text{ n'est pas libre pour } B) \tag{A2}$$
$$(\forall x)B \rightarrow B(x/t) \tag{A3}$$

[29] Voir Pacheco et Santos (2004, p.212) pour une présentation plus concise.

où $B(x/t)$ est obtenu en remplaçant les occurrences libres de x par t.

Notons que (Gen) est la règle d'inférence que l'on retrouve dans la présentation classique du calcul de premier ordre.[30] (A1) est un axiome de distributivité et (A3) est aussi un axiome du calcul des prédicats. (A1) et (A2) permettent de dériver l'axiome usuel

$$(\forall x)(B \rightarrow C) \rightarrow (B \rightarrow (\forall x)C)$$

du calcul des prédicats (où B ne contient pas d'occurence libre de x).

En plus des axiomes de premier ordre, on trouve les axiomes propres à la logique de l'action, qui forment le second groupe.

2. Les axiomes et règles d'inférence STIT:

$$E_{a:r}B \rightarrow B \tag{A4}$$
$$(E_{a:r}B \wedge E_{a:r}C) \rightarrow E_{a:r}(B \wedge C) \tag{A5}$$
$$E_{a:r}B \rightarrow qual(x : r) \tag{A6}$$
$$(\forall x)qual(x : itself) \tag{A7}$$
$$\vdash B \leftrightarrow C \Rightarrow \vdash E_{a:r}B \leftrightarrow E_{a:r}C \tag{Sub$_E$}$$

L'axiome (A4) correspond à l'axiome (**M**) qui représente la nécessité en logique aléthique (cf. Garson 2006, p.38). En un mot, cet axiome signifie que s'il est vrai que a dans le rôle r réalise B, alors B est vrai. L'axiome (A5) est un axiome d'agrégation qui stipule que si a réalise B et réalise C (dans le rôle r), alors a réalise la conjonction B et C.[31] L'axiome (A6) quant à lui spécifie que si un agent réalise une action dans le cadre d'un certain rôle, alors cet agent est qualifié pour le rôle. Cet axiome est, en quelque sorte, un axiome de légitimité: un agent ne peut pas réaliser une action dans le cadre d'un rôle qu'il ne possède pas. En ce sens, un agent ne peut pas agir de manière illégitime, voire frauduleuse: il est impossible que Pierre donne une contravention à Paul dans le cadre de son rôle de policier, mais que Pierre ne soit pas un policier qualifié. Autrement dit, un agent ne peut pas faire semblant d'agir dans un rôle.[32] L'axiome (A7) signifie que tout agent est qualifié à agir dans son propre rôle, où $itself$ est une forme de rôle sans convention.[33] Finalement, la règle d'inférence (Sub$_E$) est une règle de substitution qui implique que si B et C sont des propositions équivalentes dans $\mathcal{L}_{\mathcal{DA}}$, alors $E_{a:r}B$ et $E_{a:r}C$ le sont aussi.

[30] Voir par exemple Kleene (1952, p.82) ou encore Mendelson (2009, p.62).

[31] L'auteur souligne en conclusion que ce système gagnerait à être augmenté à l'aide d'une logique temporelle. Soulignons que, en effet, sans temporalité, cet axiome peut amener à des conséquences étranges. Supposons par exemple qu'à un temps t_1 a ne chante pas ($E_{a:r}\neg c$) et qu'à un temps t_2 a chante ($E_{a:r}c$). Dans une telle situation, l'agrégation permet de dire que a chante et ne chante pas!

[32] Il s'agit là d'une limite de leur approche, qui ne peut rendre compte des agents qui agissent de manière illégitime.

[33] Voir Pacheco et Carmo (2003, p.155) pour l'analyse des relations entre le rôle $itself$ et les autres rôles.

En dernier lieu, l'axiomatisation comprend les règles qui gouvernent les opérateurs déontiques et leurs relations avec l'opérateur STIT.

3. Les axiomes et règles d'inférence déontiques:

$$(O_{a:r}B \wedge O_{a:r}C) \rightarrow O_{a:r}(B \wedge C) \tag{A8}$$

$$O_{a:r}B \rightarrow P_{a:r}B \tag{A9}$$

$$O_{a:r}B \rightarrow \neg P_{a:r}\neg B \tag{A10}$$

$$(O_{a:r}B \wedge P_{a:r}C) \rightarrow P_{a:r}(B \wedge C) \tag{A11}$$

$$\vdash B \leftrightarrow C \Rightarrow \vdash O_{a:r}B \leftrightarrow O_{a:r}C \tag{Sub$_O$}$$

$$\vdash B \rightarrow C \Rightarrow \vdash P_{a:r}B \rightarrow P_{a:r}C \tag{RM$_P$}$$

$$\vdash E_{a_1:r_1}B \rightarrow E_{a_2:r_2}C \Rightarrow \vdash P_{a_1:r_1}B \rightarrow P_{a_2:r_2}C \tag{RM$_{EP}$}$$

L'axiome (A8) est, au même titre que (A5), un axiome d'agrégation pour les obligations. L'axiome (A9) correspond à l'axiome (**D**) du système standard et stipule que ce qui est obligatoire est aussi permis. L'axiome (A10) quant à lui vise à définir la relation entre $O_{a:r}$ et $P_{a:r}$. Considérant que $P_{a:r}$ est pris en tant que primitif (et donc contrairement à la définition usuelle $\neg O_{a:r}\neg \neq_{def} P_{a:r}$), l'axiome indique que si a doit, dans le rôle r, faire B, alors il est faux qu'il est permis à a (dans le rôle r) de ne pas faire B. Le dernier axiome (A11) assure une certaine cohésion entre les obligations et les permissions qui caractérisent un rôle. En effet, il assure que ce qui est permis puisse être fait en conjonction avec ce qui est obligatoire. En ce qui concerne les règles d'inférence, (Sub$_O$) est une règle de substitution à l'instar de (Sub$_E$), alors que (RM$_P$) et (RM$_{EP}$) sont des analogues de la règle (RM) en logique modale (cf. Chellas 1980, p.17):

$$\frac{\vdash A \rightarrow B}{\vdash \Box A \rightarrow \Box B} \qquad \text{(RM)}$$

Définis ainsi, les opérateurs déontiques et l'opérateur d'action ne sont pas *normaux* au sens de Chellas (1980, p.114).

Du côté de la sémantique, Carmo et Pacheco (2001, p.148) définissent un modèle

$$\mathcal{M} = \langle W, I, f_e, f_o \rangle$$

à l'aide de la sémantique des mondes possibles, où $W \neq \varnothing$ est un ensemble de scénarios possibles.[34] I est une interprétation du langage de premier ordre $\mathcal{L}_{\mathcal{DA}}$ pour chaque $w \in W$ telle que:

[34] Le système syntaxique utilisé par Pacheco et Santos (2004) est celui développé par Pacheco et Carmo (2003). Cela dit, dans les deux cas les auteurs renvoient à la sémantique développée dans Carmo et Pacheco (2000, 2001). Ici, nous proposons une synthèse de l'approche proposée dans l'article de 2003, utilisant la sémantique des articles de 2000 et 2001. Cela dit, la position des auteurs a évoluée depuis et la sémantique qu'ils ont proposée ne rend pas pleinement compte de la syntaxe telle que développée par la suite. Cette dichotomie vient principalement du fait que leur système a évolué en fonction de leurs objectifs. Nous renvoyons le lecteur au dernier paragraphe de cette section.

1. le domaine associé à chaque type s est le même pour tout scénario et est dénoté par $D_s = I(w, s)$;
2. si g est une fonction de type $(s_1, ..., s_n, R)$, alors $I(s, g) = D_{s_1} \times ... \times D_{s_n} \rightarrow D_R$;
3. si p est un prédicat de type $(s_1, ..., s_n)$, alors $I(w, s) \subseteq D_{s_1} \times ... \times D_{s_n}$;
4. si c est une constante, alors $I(w, c) = I(w_1, c)$;
5. $I(w, itself) = D_{Ag}$.

Autrement dit, la première condition stipule que le domaine est stable, et donc qu'il ne varie pas d'un scénario à l'autre. Cette contrainte assure que les scénarios possèdent tous les mêmes types de prédicats avec des extensions équivalentes. La seconde condition ne fait qu'assurer que le domaine d'un rôle est du même type que les parties dont il est composé. La troisième implique que le domaine qui contient un prédicat d'un certain type est inclus dans l'ensemble qui contient les domaines de chacun des types. La quatrième clause signifie que les individus restent les mêmes d'un scénario à l'autre, et finalement la cinquième condition stipule que le domaine des objets qui peuvent remplir le rôle *itself* pour un scénario donné est l'ensemble de tous les agents.

À cela s'ajoute une définition récursive des domaines de type R et AgR:

1. si r est un rôle de type $(s_1, ..., s_n, R)$, que $x_j \in D_{s_j}$(où $j = 1, ..., n$) et que $a \in D_{Ag}$, alors $r(x_1, ..., x_n) \in D_R$ et $a : r(x_1, ..., x_n) \in D_{AgR}$;
2. si r est un rôle de type $(AgR, s_1, ..., s_n, R)$, $x_j \in D_{s_j}$(où $j = 1, ..., n$), $a \in D_{Ag}$ et $b \in D_{AgR}$, alors $r(b, x_1, ..., x_n) \in D_R$ et $a : r(b, x_1, ..., x_n) \in D_{AgR}$;

La première condition implique que si r est un rôle dont le type est s_j, que le prédicat x_j est du même type et que a est un agent, alors le rôle caractérisé par le prédicat est élément du domaine, de même que la combinaison *agent : rôle*. La seconde stipule que les rôles peuvent se combiner. En ce sens, si r est un rôle d'un type s_j et qu'un prédicat x_j est du même type, alors si b est un rôle que peut prendre un agent, alors le rôle caractérisé par b, x_j est aussi élément du domaine.

Les deux autres fonctions servent à isoler les scénarios dans lesquels les propositions $E_{a:r}B$ et $O_{a:r}B$ sont vraies. D'une part,

$$f_e : (D_{Ag} \cup D_{AgR}) \times 2^W \longrightarrow 2^W$$

est définie de façon à ce que si $a \in D_{Ag}$ et Z est un sous-ensemble de W, alors $f_e(a : r, Z)$ dénote l'ensemble qui contient tous les scénarios où a réalise Z dans le rôle r, Z étant une description du scénario qui est vraie suite aux actions posées par a. D'autre part,

$$f_o : 2^W \longrightarrow 2^W$$

isole les scénarios où un état de choses Z est obligatoire: si $Z \subseteq W$, alors $f_o(Z)$ est l'ensemble qui contient tous les scénarios où Z est obligatoire. Cela est similaire aux clauses sémantiques proposées par Chellas (1974) et Jones (1991) afin de rendre compte des obligations conditionnelles (cf. Peterson 2011, p.58 note 10). Notons au passage qu'en vertu de cette définition, cette approche s'insère dans le cadre des logiques de type *Ought-to-be*: l'agent n'a pas l'obligation de *faire* une action, mais doit plutôt faire advenir un *état de choses* (dans le monde).

La vérité dans $\mathcal{M}$ est définie de manière usuelle. Carmo et Pacheco (2001, p.149) introduisent une fonction d'évaluation v qui assigne pour tout scénario et à tout terme t de type s un objet membre du domaine D_s. La fonction est définie de manière récursive et attribue ultimement aux constantes l'objet qui leur correspond à l'intérieur du domaine. Les conditions de vérité sont les suivantes.

$$\mathcal{M} \models_{w,v} p(t_1, ..., t_n) \Leftrightarrow\ v(w, t_1), ..., v(w, t_n) \in I(w, p) \tag{1}$$

$$\mathcal{M} \models_{w,v} \neg B \Leftrightarrow\ \mathcal{M} \not\models_{w,v} B \tag{2}$$

$$\mathcal{M} \models_{w,v} B \wedge C \Leftrightarrow\ \mathcal{M} \models_{w,v} B \text{ et } \models_{w,v} C \tag{3}$$

$$\mathcal{M} \models_{w,v} (\forall x^s)B \Leftrightarrow\ \mathcal{M} \models_{w,v_1} B \text{ pour tout } v_1 \text{ qui assigne les mêmes valeurs que } v \text{ aux variables d'autres types} \tag{4}$$

$$\mathcal{M} \models_{w,v} OB \Leftrightarrow\ w \in f_o(\|B\|_v) \text{ où } \|B\|_v = \{w : \mathcal{M} \models_{w,v} B\} \tag{5}$$

Finalement, si t est un terme de type Ag ou AgR, alors

$$\mathcal{M} \models_{w,v} E_t B \Leftrightarrow\ w \in f_{e_{v(t)}}(\|B\|_v) \tag{6}$$

La clause (1) signifie qu'une proposition de premier ordre est vraie dans la mesure où les objets dans la portée du prédicat sont dans l'extension du concept pour un scénario donné. Les conditions (2)–(3) sont les conditions habituelles pour la négation (classique) et la conjonction. (4) assure que si une quantification universelle est vraie, alors la proposition dans la portée du quantificateur est vraie pour tout autre évaluation qui assigne les mêmes valeurs aux variables des différents types. En d'autres termes, (4) assure une certaine stabilité, c'est-à-dire que la clause assure que la quantification tient dans les scénarios qui ont le même domaine et dont les prédicats identiques ont des extensions identiques. Notons que la quantification est indexée à un type, et donc ne porte par sur le domaine en entier, mais seulement sur la portion D_s. La condition (5) stipule qu'une proposition de la forme « obligatoirement B » est vraie pour un scénario w à condition que ce scénario soit membre de l'ensemble qui contient tous les scénarios où l'ensemble qui contient tous les scénarios pour lesquels B est vrai selon l'évaluation v est obligatoire. Cette dernière condition mérite d'être reformulée: l'ensemble $\|B\|_v$ dénote l'ensemble qui contient tous les scénarios où la proposition B est vraie pour l'évaluation v.[35] L'énoncé OB est vrai pour

[35] La combinaison de l'indexation scénario/évaluation permet de faire varier les descriptions des scénarios de manières différentes. Par exemple, en fixant w et en faisant varier v, on pose que les propositions vraies restent les mêmes mais que les extensions des prédicats varient. De l'autre côté, en posant que v mais en faisant varier w, on garde le même domaine mais on regarde ce qui peut être vrai considérant ce domaine.

w et l'évaluation v à condition que tout scénario x où B est vrai pour la même évaluation est obligatoire. Autrement dit, $f_o(\|B\|_v)$ dénote les scénarios obligatoires pour lesquels B est vrai (selon l'évaluation v), et OB est vrai pour w à condition que w fasse partie de cet ensemble. Notons que cette clause présume l'idéalité du modèle, liant la vérité de la proposition normative OB à celle de la proposition descriptive B dans la portée de l'opérateur déontique. En dernier lieu, la clause (6) rend compte de l'axiome (A4) et signifie que si la proposition « x réalise B dans le rôle r » est vraie pour w, alors w est membre de l'ensemble qui contient tous les scénarios où B est vrai (et où l'état de choses B est le résultat des actions de x dans le rôle r).

Une proposition est dite vraie pour un modèle $\mathcal{M}$ à condition qu'elle soit vraie pour tout scénario et toute interprétation dans $\mathcal{M}$, et elle est dite valide lorsqu'elle est vraie pour tout modèle (Carmo et Pacheco 2001, p.150).

Afin de rendre compte de la structure exprimée par les axiomes, les auteurs ajoutent des clauses supplémentaires au modèle sémantique. D'abord, Carmo et Pacheco (2001, p.151) ajoutent la clause

$$f_{e_{a:r}}(X) \subseteq f_{e_a}(X) \subseteq X$$

où $X \subseteq W$. En conjonction avec la clause (6) susmentionnée, cela vise à assurer la validité de l'axiome (A4). En mots, cela signifie que l'ensemble qui contient les scénarios où il est vrai qu'un agent réalise X dans un rôle est sous-ensemble de l'ensemble où a réalise X, lequel est sous-ensemble de X (présupposé vrai). Autrement dit, s'il est vrai qu'un agent (dans un rôle ou non) réalise X, alors X est vrai: mais ce n'est pas parce que X est vrai que l'agent a réalisé X. Les deux conditions suivantes visent à valider respectivement les axiomes (A5) et (A8).[36]

$$f_{e_b}(X) \cap f_{e_b}(Y) \subseteq f_{e_b}(X \cap Y)$$
$$f_{o_b}(X) \cap f_{o_b}(Y) \subseteq f_o(X \cap Y)$$

Dans chacun des cas, la clause est conforme à la conception ensembliste de la conjonction: l'ensemble qui contient les scénarios où A est vrai et B est vrai est inclus dans l'ensemble qui contient les scénarios où A et B est vrai. Cela dit, considérant que $O_{a:r}$ est une abréviation pour $OE_{a:r}$, la clause sémantique pour l'axiome (A8) est mieux représentée par:

$$f_o(f_{e_b}(X)) \cap f_o(f_{e_b}(Y)) \subseteq f_o(f_{e_b}(X \cap Y))$$

L'axiome (A9), équivalent à (**D**), va de paire avec la clause:

$$f_{o_b}(X) \cap f_{o_b}(W - X) = \varnothing$$

Cela signifie que l'ensemble qui contient les scénarios où X est obligatoire ne partage aucun membre avec celui qui contient les scénarios où le complément de X dans W est

[36] Notons que la seconde clause n'est pas retenue dans le texte de 2001 considérant qu'à ce moment les auteurs rejetaient l'agrégation pour l'obligation (cf. Carmo et Pacheco 2001, p.151). De plus, soulignons aussi que nous avons ajouté l'indexation de f_o à b afin d'être conforme au texte de 2003.

obligatoire. En ce sens, les scénarios où X est vrai qui sont obligatoires ne peuvent pas se retrouver à l'intérieur d'autres scénarios obligatoires où X est faux. L'axiome (A11) quant à lui est représenté sémantiquement par la clause (Carmo et Pacheco 2001, p.152):

$$(f_o(f_{e_b}(X)) - f_o(W - f_{e_b}(Y))) \cap f_o(W - f_{e_b}(X \cap Y)) = \varnothing$$

Cette clause mérite quelques explications. Allons-y étape par étape:

1. $W - f_{e_b}(Y)$ est le complément de l'ensemble qui contient les scénarios où Y (réalisé par b) est vrai par rapport à W. Il s'agit donc de l'ensemble qui contient les scénarios où Y (réalisé par b) est faux;
2. $f_o(W - f_{e_b}(Y))$ est l'ensemble qui contient les scénarios où « Y (réalisé par b) est faux » est obligatoire;
3. $(f_o(f_{e_b}(X)) - f_o(W - f_{e_b}(Y))$ est le complément de l'ensemble qui contient les scénarios où « Y (réalisé par b) est faux » est obligatoire par rapport à l'ensemble qui contient les scénarios où « X (réalisé par b) est vrai » est obligatoire. De fait, l'ensemble contient « X (réalisé par b) est vrai » est obligatoire et « Y (réalisé par b) est vrai » est obligatoire;
4. $f_o(W - f_{e_b}(X \cap Y))$ est l'ensemble qui contient les scénarios obligatoires où soit X (réalisé par b) est faux ou Y (réalisé par b) est faux.

En ce sens, l'ensemble qui contient « X (réalisé par b) est vrai » est obligatoire et « Y (réalisé par b) est vrai » est obligatoire ne partage aucun membre avec celui qui contient les scénarios où « X et Y » (réalisé par b) est faux est obligatoire.

Le lecteur attentif aura remarqué que nous n'avons pas encore discuté de tous les axiomes proposés. Le cas des axiomes et règles d'inférence de premier ordre est trivial, et de fait nous n'en discuterons pas davantage. En ce qui concerne les axiomes STIT, nous n'avons pas encore discuté de (A6), (A7) et (Sub_E). Carmo et Pacheco (2001, p.153) mentionnent au passage (en une ligne) que l'axiome (A7) est d'emblée valide dans leur sémantique, la raison étant que sinon il faudrait un scénario w et une évaluation v tels que $is - itself(x)$ est faux, et donc que x n'est pas membre de l'extension du concept *être* x.[37]

L'axiome (A6) quant à lui est validé par l'ajout de la condition:

$$w \in f_{e_{a:r}}(X) \Rightarrow \ qual(w, a : r)$$

Si w est un scénario membre de l'ensemble qui contient les scénarios où a réalise X dans le rôle r, alors a est qualifié pour le rôle r dans w.[38] La validité de la règle (Sub_E) est triviale, à supposer la complétude (dont la preuve n'est pas fournie par les auteurs), considérant la clause de vérité (6).

[37] On ne trouve d'ailleurs pas plus d'explication dans Carmo et Pacheco (2000, p.119).
[38] Le lecteur est renvoyé au texte pour la définition récursive de *qual*.

Cela nous amène aux axiomes et règles d'inférence déontiques. Jusqu'à présent, nous avons traité des axiomes (A8), (A9) et (A11). La règle Sub_O se comprend dans la même mesure que Sub_E. Les autres cas sont moins évidents et quelques remarques s'imposent. Le système syntaxique est présenté dans l'article de Pacheco et Carmo (2003), où les auteurs justifient le choix des axiomes et des règles d'inférence et analysent en profondeur les notions d'*action commune*, de *contrat*, d'*organisation* et de *rôle*. L'article de Pacheco et Santos (2004) reprend la syntaxe proposée par Pacheco et Carmo (2003) afin d'analyser la délégation de rôle, d'obligation et de responsabilité au sein d'une organisation (un NMAS). De l'autre côté, la sémantique est présentée dans les articles de Carmo et Pacheco (2000, 2001), où le texte de 2001 est en fait une reformulation de celui de 2000. Cela dit, Pacheco et Santos (2004, pp.210-1) renvoient à Pacheco et Carmo (2003) et Carmo et Pacheco (2001) pour la discussion sur les opérateurs déontiques primitifs, où Pacheco et Carmo (2003, p.159) ne font que résumer l'analyse proposée dans le texte de 2001. Par surcroît, Pacheco et Carmo (2003, p.159) renvoient aux textes de Carmo et Pacheco (2000, 2001) pour les détails relatifs à la sémantique. Or, le fait est que la syntaxe proposée dans les textes de 2004 et 2003 ne se trouve pas complètement dans ceux de 2001 et 2000. Autrement dit, les considérations sémantiques proposées dans le texte de 2001 ne couvrent pas la totalité du système syntaxique proposé en 2003. Le lecteur ne trouvera donc pas la clause sémantique de (A10) et la discussion proposée dans Carmo et Pacheco (2001, p.152) ne traite pas des règles (RM_P) et (RM_{EP}).

Somme toute, Pacheco et Carmo proposent une façon intéressante de combiner les logiques de premier ordre, STIT et déontique. Cette approche vise à rendre compte de phénomènes complexes comme la délégation d'obligations au sein d'une organisation, l'interaction entre les obligations, les agents et les rôles au sein d'une organisation et, de façon plus générale, des obligations collectives. Un aspect original de leur approche est qu'ils ne considèrent pas que l'*institution* se réduit aux agents individuels qui la composent (cf. Carmo et Pacheco 2001, p.137).[39] L'institution va au-delà de ses membres. Conformément à leur analyse de la loi, ils introduisent la notion d'agent artificiel afin de rendre compte des organisations en tant qu'agents complexes, caractérisés par les rôles qui les composent et par les normes qui régissent ces rôles.

Jan Broersen

La logique STIT vise la modélisation des NMAS, et de fait, conformément à la tradition engendrée par Belnap et Perloff (1988), elle inclut une dimension temporelle qui sert à modéliser l'évolution des obligations. En plus de cette dimension temporelle, la logique STIT peut aussi être augmentée de façon à rendre compte de phénomènes plus complexes, comme l'interaction entre les modalités temporelles, déontiques et épistémiques.[40] Broersen (2011a), par exemple, propose une logique multi-modale fondée sur une logique

[39] Voir aussi Pacheco et Carmo (2003, p.146) et Pacheco et Santos (2004, p.210).

[40] Pour une introduction à la logique épistémique et à l'épistémologie formelle, le lecteur est invité à consulter Hendricks (2006).

déontique STIT et augmentée par une modalité épistémique K_a.[41]

D'emblée, l'approche de Broersen (2011a) vise la représentation de l'action intentionnelle. En effet, ce dernier propose une nouvelle interprétation de l'opérateur K_a lorsque celui-ci est combiné avec l'opérateur $[A\ stit]$: la proposition $K_a[a\ stit]\varphi$ signifie « l'agent a réalise φ *sciemment* (Broersen 2011a, p.145) ».[42] La motivation principale de l'auteur est de formaliser certaines notions de droit criminel, notamment l'imputabilité d'un accusé. En *Common law*, le fait qu'un accusé soit imputable ou non d'un crime dépend de deux facteurs, l'*actus reus* et la *mens rea*, qui renvoient à l'acte et à l'état mental de l'individu au moment du crime. Afin qu'un crime soit imputable, il faut établir d'une part que le crime a effectivement été commis (l'*actus reus*), et d'autre part que l'individu était en contrôle de ses capacités intellectuelles.[43] Broersen (2011a, p.137) souligne que plusieurs systèmes de droit criminel distinguent entre différents *modes* de *mens rea*, comme *faire en vue d'un but*, *faire en connaissance de cause* ou encore *faire négligeamment*. De fait, le système proposé vise la représentation formelle de ces différents modes de *mens rea*, de l'*actus reus*, et, de manière générale, de la culpabilité (Broersen 2011a, p.138).

Même si le système de Broersen s'insère dans le cadre des logiques de type STIT, celui-ci utilise en fait une logique STIT augmentée. En effet, Broersen (2011a, p.139) utilise la logique XSTIT, qui se distingue de la logique STIT dans la mesure où les actions n'ont pas d'effet *immédiat*: l'effet d'une action prend place dans un état subséquent.[44] Usuellement, l'opérateur STIT satisfait le schéma d'axiome (**M**): s'il est vrai que a réalise φ dans le scénario w, alors φ est vrai pour w. Or, avec XSTIT, φ sera vrai dans l'état subséquent. À supposer que, dans une conception linéaire du temps, le monde se trouve dans un certain *état* à chaque moment, l'effet d'une action posée au temps t_i prend place au moment t_{i+1}, où t_{i+1} est le successeur immédiat de t_i.

Le langage

$$\mathcal{L}_{\text{XSTIT}} = \{Prop, Ags, (,), \neg, \wedge, \Box, [\ xstit], X, K_a\}$$

contient un ensemble dénombrable de propositions atomiques $Prop$, un ensemble fini $Ags = \{a_1, ..., a_j\}$ d'agents (où $A \subseteq Ags$ est un ensemble d'agents), les connecteurs pour la négation et la conjonction ainsi que les opérateurs pour la nécessité, l'action, l'état subséquent et la connaissance.[45]

L'ensemble des expressions bien formées (EBF) est défini récursivement (Broersen 2011a, p.139)[46]:

1. si $p \in Prop$, alors $p \in EBF$;

[41] L'article de Broersen (2011a) provient d'un acte du colloque DEON'08 (cf. Broersen 2008). Il a aussi traité (différemment) de ce sujet dans Broersen (2009).

[42] Traduction libre de « *the agent a knowingly sees to it that* φ ».

[43] Pour une analyse détaillée de ces notions, le lecteur peut consulter Parent (2007, 2008).

[44] Nous utilisons l'expression « état subséquent » comme traduction de « *next state* ».

[45] Notons que l'auteur écrit $[A\ xstit]\varphi$ plutôt que $[A\ xstit\ \varphi]$, comme on le retrouve chez Belnap et Perloff (1988), afin d'insister sur le fait qu'il s'agit d'une modalité. De plus, $[a\ xstit]$ est une abréviation de $[\{a\}\ xstit]$ (cf. Broersen 2011a, p.139 et p.145).

[46] Cette définition permet l'itération des opérateurs.

2. si $\varphi, \psi \in EBF$, alors $\neg\varphi, \varphi \wedge \psi, \Box\varphi, [A\ xstit]\varphi, X\varphi, K_a\varphi \in EBF$.

L'opérateur $\Box$ représente la nécessité historique et $X\varphi$ exprime la transition d'un état (statique) à un autre (son successeur). Les connecteurs $\rightarrow$, $\leftrightarrow$, $\vee$ et l'opérateur dual $\Diamond$ sont définis de manière usuelle et $\langle A\ xstit\rangle$ est l'abréviation de $\neg[A\ xstit]\neg$.

La notion de nécessité *historique* vise la représentation de ce qui est vrai pour tout scénario accessible (cf. Thomason 2002). Considérons par exemple le diagramme suivant (figure 3.1).

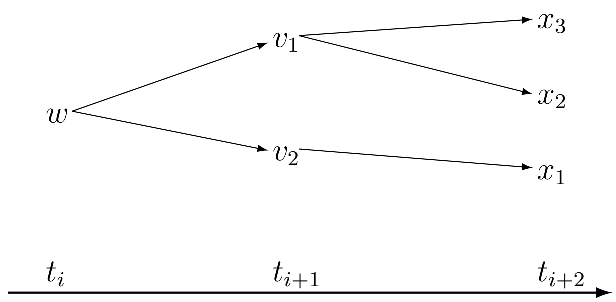


Figure 3.1: La nécessité historique

Les scénarios v_1 et v_2 sont accessibles au scénario w au temps t_i, où v_1 et v_2 sont les deux états subséquents possibles à w. De la même manière, x_3 et x_2 sont les deux scénarios subséquents possibles à v_1 au temps t_{i+1}. Une *histoire* peut informellement être considérée comme un « chemin » qui passe par différents scénarios, c'est-à-dire une suite. En ce sens, nous avons dans cet exemple trois histoires (ou suite d'évènements) possibles: (w, v_1, x_3), (w, v_1, x_2) et (w, v_2, x_1). La nécessité historique est ponctuelle et est déterminée à un certain moment t_j. Ce qui est nécessairement (historiquement) vrai à t_{i+2} par exemple est ce qui est vrai pour x_1, x_2 et x_3. Autrement dit, la nécessité historique à t_{i+2} correspond à l'ensemble de propositions (vraies) $\Lambda = \{\Gamma : \Gamma \subseteq x_1, \Gamma \subseteq x_2 \text{ et } \Gamma \subseteq x_3\}$. Cela dit, ce qui est historiquement nécessairement vrai peut changer dans le futur.

Le système axiomatique est construit à partir d'une axiomatisation de la logique propositionnelle classique ainsi que des règles d'inférence usuelles pour les opérateurs modaux (cf. Garson 2006, p.37). Autrement dit, pour $\star$ une modalité de XSTIT, le (schéma de) règle d'inférence et le schéma d'axiome suivants s'appliquent.

$$\frac{\vdash A}{\vdash \star A} \qquad \text{(Nec}\star\text{)}$$

$$\vdash \star(A \rightarrow B) \rightarrow (\star A \rightarrow \star B) \qquad \text{(Dist}\star\text{)}$$

L'axiomatisation se fait par les schémas d'axiomes suivants:

$$\Box\varphi \to \varphi \tag{\textbf{M}}$$
$$\Diamond\varphi \to \Box\Diamond\varphi \tag{\textbf{5}}$$
$$[A\ xstit]\varphi \to \langle A\ xstit\rangle\varphi \tag{\textbf{D}}$$
$$\neg X\neg\varphi \to X\varphi \tag{A1}$$
$$[\varnothing\ xstit]\varphi \leftrightarrow \Box X\varphi \tag{A2}$$
$$[Ags\ xstit]\varphi \leftrightarrow X\Box\varphi \tag{A3}$$
$$[A\ xstit]\varphi \to [A \cup B\ xstit]\varphi \tag{A4}$$
$$(\Diamond[A\ xstit]\varphi \wedge \Diamond[B\ xstit]\psi) \to \Diamond([A\ xstit]\varphi \wedge [B\ xstit]\psi) \tag{A5}$$

Les axiomes (**M**) et (**5**) donnent à la nécessité historique une structure de type $S5$ et l'axiome (**D**) donne à l'opérateur $[A\ xstit]$ une structure de type KD (cf. Garson 2006, p.115). (**M**) signifie que ce qui est nécessairement vrai pour un scénario est vrai. L'axiome (**5**) exprime que s'il est possible que φ soit vrai pour un scénario v au moment t_i, qui est un état subséquent d'un scénario w à t_{i-1}, alors φ est aussi vrai pour les autres alternatives v' de w. (**D**) stipule que si A réalise φ, alors il est faux que A réalise $\neg\varphi$. L'axiome (A1) signifie que s'il est faux que $\neg\varphi$ est le cas à un état subséquent, alors φ est le cas pour cet état. Il s'agit d'une instance de l'axiome (**CD**) et indique que l'état subséquent est unique (cf. tableau 3.1). (A2) stipule que si φ advient « naturellement », c'est-à-dire que φ n'est réalisé par aucun agent, alors c'est une nécessité historique que φ prenne place dans l'état subséquent. Autrement dit, un changement qui n'est causé par aucun agent est une nécessité historique pour l'état subséquent. (A3) signifie que si φ est réalisé par tous les agents, alors dans l'état subséquent φ est une nécessité historique. En ce sens, les propositions qui ne sont pas des nécessités historiques dans les états subséquents représentent des états de choses qui peuvent être réalisés suite aux actions et aux choix des agents. L'axiome (A4) exprime que si φ est réalisé par A, alors il est aussi réalisé conjointement par un ensemble qui contient A.[47] Soulignons que cet axiome implique que les actions d'un groupe se réduisent aux actions de chaque membre. L'axiome (A5) quant à lui est restreint par la clause $A \cap B = \varnothing$ et signifie que s'il est possible que A réalise φ et qu'il est possible que B réalise ψ, alors il est possible que A réalise φ et que B réalise ψ.

Broersen (2011a, p.145) propose ensuite d'ajouter trois axiomes épistémiques afin de modéliser adéquatement la notion de « réaliser sciemment ». Les axiomes sont:

$$K_aX\varphi \to K_a[a\ xstit]\varphi \tag{A6}$$
$$K_a[a\ xstit]\varphi \to XK_a\varphi \tag{A7}$$
$$\Diamond K_a\varphi \to K_a\Diamond\varphi \tag{A8}$$

(A6) exprime que si a sait que φ sera le cas dans l'état subséquent, alors a réalise

[47] Cela ne veut toutefois pas dire que $[B - A\ xstit]\varphi$.

sciemment φ.[48] (A7) signifie que si a réalise sciemment φ, alors dans l'état subséquent a sait que φ. Finalement, (A8) stipule que s'il est possible que a sache que φ, alors a sait que φ est possible.[49]

Cette axiomatisation se traduit littéralement sur le plan de la sémantique. Suivant les résultats de Sahlqvist (1975), qui a montré que les axiomes modaux induisent des relations précises sur les modèles sémantiques (cf. Garson 2006, p.211), Broersen (2011a, p.143) présente le modèle sémantique en explicitant avec quelle condition est lié chaque axiome. Conformément au tableau présenté dans Garson (2006, p.115), les principaux résultats peuvent être résumés par le tableau 3.1, où $\Box$ est une modalité quelconque et $\Diamond =_{def} \neg\Box\neg$ (l'ajout de ($\Box$**B**) est dû à Åqvist 2002, p.209).

	Axiome	Condition sur $\mathcal{M}$	Relation
(**D**)	$\Box A \supset \Diamond A$	$\forall w\ \exists v$ t.q. wRv	Sérielle
(**M**)	$\Box A \supset A$	$\forall w,\ wRw$	Réflexive
(**4**)	$\Box A \supset \Box\Box A$	$\forall w, v, u$ (wRv et vRu) $\Rightarrow wRu$	Transitive
(**5**)	$\Diamond A \supset \Box\Diamond A$	$\forall w, v, u$ (wRv et wRu) $\Rightarrow vRu$	Euclidienne
(**B**)	$A \supset \Box\Diamond A$	$\forall w, v\ wRv \Rightarrow vRv$	Symétrique
S5	(**M**) + (**5**)	$\forall w, v\ wRv$	Équivalence
($\Box$**M**)	$\Box(\Box A \supset A)$	$\forall w, v\ wRv \Rightarrow vRv$	Quasi-réflexive
($\Box$**B**)	$\Box(A \supset \Box\Diamond A)$	$\forall w, v, u$ (wRv et vRu) $\Rightarrow uRv$	Quasi-symétrique
(**CD**)	$\Diamond A \supset \Box A$	$\forall w, v, u\ wRv$ et $wRu \Rightarrow v = u$	Unique
(**C4**)	$\Box\Box A \supset \Box A$	$\forall w, v\ wRv \Rightarrow\ \exists u(wRu$ et $uRv)$	Dense
(**C**)	$\Diamond\Box A \supset \Box\Diamond A$	$\forall w, v,$ (wRv et wRu) $\Rightarrow$ $\exists x$ (vRx et uRx)	Convergent
(**L**)	$\Box(\Box A \supset B) \vee$ $\Box((B \wedge \Box B) \supset A)$	$\forall w, v, u$ (wRv et wRu) $\Rightarrow$ vRu, uRv ou $v = u$	Antisymétrique

Tableau 3.1: Les relations induites sur les modèles sémantiques

Le modèle sémantique $\mathcal{M} = \langle \mathcal{F}, \pi \rangle$ est défini à partir d'une structure

$$\mathcal{F} = \langle S, H, R_X, R_\Box, R_A, \sim_a \rangle$$

et d'une fonction $\pi : P \to 2^{S \times H}$ qui attribue à chaque proposition l'ensemble des états dynamiques dans lesquels la proposition est vraie (cf. Broersen 2011a, p.140).

Considérons d'abord la structure $\mathcal{F}$. S est un ensemble dénombrable d'états statiques et $H \neq \varnothing$ est un ensemble non dénombrable d'histoires h, où h est un ensemble dénombrable ordonné de sous-ensembles de S. Autrement dit, H contient toutes les combinaisons dénombrables possibles que l'on peut faire à partir des membres de S. Une histoire h est ordonnée par R_X, une relation d'état subséquent. Cette relation permet d'ordonner les états dynamiques $\langle s, h \rangle$, c'est-à-dire est une paire ordonnée qui contient un

[48] Cet axiome est discutable. Supposons que nous somme le 27 août 2013, et que Paul sait que le lendemain la situation en Égypte ne se sera pas améliorée. Donc Paul réalise sciemment que la situation en Égypte ne s'améliore pas?

[49] Cet axiome aussi est discutable. Il est possible que Simon, qui a quatre ans, sache les principes de la mécanique quantique, sans nécessairement que Simon sache qu'il est possible qu'il connaisse les principes de la mécanique quantique.

état (statique) et une histoire. Un état s' est subséquent à s dans une histoire h si et seulement si $\langle s, h\rangle R_X \langle s', h\rangle$. La relation R_X est sérielle et déterministe, ce qui est exprimé respectivement par les axiomes (**D**) et (A1). Par surcroît, la relation R_X est contextuelle à une histoire. En effet, si $\langle s, h\rangle R_X \langle s', h'\rangle$, alors $h = h'$, et donc dans l'état subséquent s', les agents demeurent dans la même histoire. $R_\Box$ est une relation de nécessité historique telle que $\langle s, h\rangle R_\Box \langle s', h'\rangle$ si et seulement si $s = s'$. En ce sens, la relation de nécessité historique est ponctuelle et s'évalue à un moment donné et pour un scénario donné (elle évalue ce qui est vrai pour toute histoire à un moment donné). Considérons la figure 3.2.

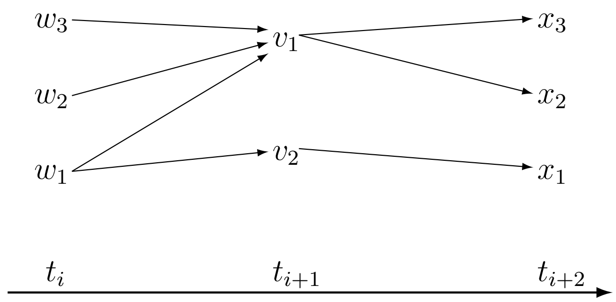


Figure 3.2: La nécessité historique (2)

Les différentes histoires visent à représenter les choix que les agents peuvent faire afin de faire advenir des états subséquents spécifiques. À l'état w_1 par exemple, un agent peut soit faire advenir v_1 ou v_2. Cela dit, la nécessité historique évalue ce qui est commun à chaque histoire pour un moment et un état donné. Par exemple, si l'on fixe le moment à t_{i+2} et l'état à x_3, la nécessité historique évalue ce qui est commun à (w_3, v_1, x_3), (w_2, v_1, x_3) et (w_1, v_1, x_3). Quant aux relations R_A, où $A \subseteq Ags$, ce sont des relations *efficaces*[50] telles que:

1. $R_\varnothing = R_\Box \circ R_X$;
2. $R_{Ags} = R_X \circ R_\Box$;
3. si $B \subset A$, alors $R_A \subseteq R_B$;
4. si $A \cap B = \varnothing$, $\langle s_1, h_1\rangle R_\Box \langle s_2, h_2\rangle$ et $\langle s_1, h_1\rangle R_\Box \langle s_3, h_3\rangle$, alors
   (a) il existe $\langle s_4, h_4\rangle$ tel que $\langle s_1, h_1\rangle R_\Box \langle s_4, h_4\rangle$;
   (b) si $\langle s_4, h_4\rangle R_A \langle s_5, h_5\rangle$, alors $\langle s_2, h_2\rangle R_A \langle s_5, h_5\rangle$;
   (c) si $\langle s_4, h_4\rangle R_B \langle s_6, h_6\rangle$, alors $\langle s_3, h_3\rangle R_B \langle s_6, h_6\rangle$.

Avant d'expliquer en mots ce que ces conditions signifient, notons que dans chacun des cas (R_X, $R_\Box$ et R_A), les relations sont des sous-ensembles de $(S \times H) \times (S \times H)$ qui

[50] La relation *effectivity* est utilisée en théorie du choix social et signifie que les choix que les individus font et que les actions qu'ils posent en fonction de ces choix réalisent l'effet désiré (cf. Pauly 2002, p.151).

respectent certaines conditions. De fait, ces relations peuvent être composées à l'aide de $\circ$. Considérant que nous avons

$$S \times H \xrightarrow{R_\Box} S \times H$$

et

$$S \times H \xrightarrow{R_X} S \times H$$

il est possible d'obtenir les compositions

$$R_\Box \circ R_X : S \times H \xrightarrow{R_X} S \times H \xrightarrow{R_\Box} S \times H$$

et

$$R_X \circ R_\Box : S \times H \xrightarrow{R_\Box} S \times H \xrightarrow{R_X} S \times H$$

La première condition signifie que les « actions » du groupe vide d'agent $\varnothing$ déterminent ce qui est nécessairement historiquement vrai pour tout scénario subséquent (R_X isole les scénarios subséquents possibles, puis $R_\Box$ détermine ce qui est vrai dans chaque scénario subséquent possible). Cette condition va de paire avec (A2). La seconde condition accompagne (A3) et signifie que les actions du groupe Ags déterminent les états subséquents possibles et fixent, pour chaque état, ce qui y sera historiquement nécessairement vrai (en vertu des actions de chaque agent). (A4) signifie que si B est un sous-groupe de A, alors A est minimalement autant efficace que B (ou encore que les choix possibles d'un agent sont inclus dans les choix possibles pour la collectivité). La dernière condition représente l'axiome (A5). Conformément à la clause susmentionnée pour la relation $R_\Box$, deux états dynamiques sont en relation de nécessité historique à condition que l'état statique soit le même. En ce sens, dans la condition 4 nous pouvons conclure qu'il s'agit du même état $s = s_1 = s_2 = s_3$. À supposer que A et B sont deux groupes d'agents qui ne partagent aucun membre et que $\langle s, h_1\rangle$, $\langle s, h_2\rangle$ et $\langle s, h_3\rangle$ sont en relation de nécessité historique, 4(a) signifie qu'il y a un autre état dynamique $\langle s_4, h_4\rangle$ en relation de nécessité historique avec $\langle s, h_1\rangle$ (et donc $s_4 = s_1 = s$). Informellement, h_1 peut être considéré comme l'histoire où A et B n'agissent pas, h_2 comme celle où A agit, h_3 celle où B agit et h_4 celle où les deux agissent. Dans l'éventualité où les deux agissent, les conditions 4(b) et 4(c) marquent l'indépendance des actions de A et de B: 4(b) signifie que si l'état dynamique $\langle s_5, h_5\rangle$ prend place après $\langle s, h_4\rangle$ et est réalisé par A, alors il prend place après $\langle s, h_2\rangle$, et (c) signifie que si l'état dynamique $\langle s_6, h_6\rangle$ prend place après $\langle s, h_4\rangle$ et est réalisé par B, alors il prend place après $\langle s, h_3\rangle$. De fait, il est possible que $h_5 = h_6$ mais que néanmoins A et B soient responsables de leurs actions respectives, dans la mesure où l'histoire dans laquelle on se trouve résulte de deux choix différents. Finalement, $\sim_a$ (avec $a \in Ags$) sont des relations d'équivalences épistémiques par rapport aux états dynamiques, où deux états sont en relation $\sim_a$ d'équivalence épistémique lorsque les connaissances de a sont les mêmes dans chaque état.

La vérité dans un modèle $\mathcal{M}$ relativement à un état dynamique $\langle s, h\rangle$, notée $\models_{\mathcal{M}_{\langle s,h\rangle}}$, est définie récursivement (cf. Broersen 2011a, p.142):

$$\models_{\mathcal{M}_{\langle s,h\rangle}} p \Leftrightarrow \ \langle s,h\rangle \in \pi(p) \tag{1}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} \neg\varphi \Leftrightarrow \not\models_{\mathcal{M}_{\langle s,h\rangle}} \varphi \tag{2}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} \varphi \wedge \psi \Leftrightarrow \models_{\mathcal{M}_{\langle s,h\rangle}} \varphi \text{ et } \models_{\mathcal{M}_{\langle s,h\rangle}} \psi \tag{3}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} \Box\varphi \Leftrightarrow \ \langle s,h\rangle R_{\Box}\langle s',h'\rangle \Rightarrow \models_{\mathcal{M}_{\langle s',h'\rangle}} \varphi \tag{4}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} [A\ xstit]\varphi \Leftrightarrow \ \langle s,h\rangle R_{A}\langle s',h'\rangle \Rightarrow \models_{\mathcal{M}_{\langle s',h'\rangle}} \varphi \tag{5}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} X\varphi \Leftrightarrow \ \langle s,h\rangle R_{X}\langle s',h'\rangle \Rightarrow \models_{\mathcal{M}_{\langle s',h'\rangle}} \varphi \tag{6}$$
$$\models_{\mathcal{M}_{\langle s,h\rangle}} K_a\varphi \Leftrightarrow \ \langle s,h\rangle \sim_a \langle s',h'\rangle \Rightarrow \models_{\mathcal{M}_{\langle s',h'\rangle}} \varphi \tag{7}$$

La clause (1) stipule qu'un atome propositionnel p est vrai dans un état dynamique à condition que cet état soit membre de l'ensemble qui contient tous les états dynamiques où p est vrai et la condition (2) est la bivalence classique, à savoir que ce qui n'est pas vrai est faux, et vice versa. (3) est la condition sémantique usuelle pour la conjonction. La clause (4) signifie qu'une proposition φ est une nécessité historique pour un état dynamique à condition que cette proposition soit vraie pour tout autre état dynamique en relation de nécessité historique avec lui. (5) implique que A réalise φ est vrai pour un état dynamique dans la mesure où φ est vrai pour tout état dynamique subséquent réalisé par A. La condition (6) exprime que φ est vrai dans le prochain état à condition que φ soit vrai pour tout état dynamique subséquent possible. Finalement, (7) signifie qu'il est vrai que a sait φ pour un état dynamique $\langle s,h\rangle$ à condition que φ soit vrai pour tout autre état dynamique en relation d'équivalence épistémique avec $\langle s,h\rangle$.

En ce qui concerne les trois axiomes proposés pour rendre compte des interactions entre les modalités $[A\ xstit]$ et K_a, les restrictions respectives sur le modèle sémantique sont:

$$\sim_a \circ R_a \subseteq \sim_a \circ R_X \tag{8}$$
$$R_X \circ \sim_a \subseteq \sim_a \circ R_a \tag{9}$$
$$\begin{aligned}&\langle s_1,h_1\rangle R_{\Box}\langle s_2,h_2\rangle \text{ et } \langle s_1,h_1\rangle \sim_a \langle s_3,h_3\rangle \Rightarrow \\ &\quad \exists s_4,h_4 \text{ t.q. } \langle s_2,h_2\rangle \sim_a \langle s_4,h_4\rangle \text{ et } \langle s_3,h_3\rangle R_{\Box}\langle s_4,h_4\rangle\end{aligned} \tag{10}$$

Cela nous amène aux considérations quant aux opérateurs déontiques. À l'instar de la tendance en logique dynamique, inspirée par Anderson (1958a), Broersen (2011a, p.147) ne prend pas l'opérateur déontique comme primitif mais le définit plutôt à l'aide de l'opérateur $[A\ xstit]$ et d'une constante de violation V. L'auteur propose deux définitions possibles pour l'obligation.

$$O[a\ xstit]\varphi =_{def} \Box(\neg[a\ xstit]\varphi \to [a\ xstit]V) \tag{def. O}$$
$$O'[a\ xstit]\varphi =_{def} \Box([a\ xstit]\neg\varphi \to [a\ xstit]V) \tag{def. O'}$$

Ces deux définitions jouent sur une nuance que le langage de XSTIT est capable d'exprimer.[51] Considérant que l'axiomatisation de l'opérateur $[A\ xstit]$ en fait un opérateur de type KD, il s'ensuit que nous obtenons d'une part

$$[A\ xstit]\neg\varphi \rightarrow \neg[A\ xstit]\varphi \tag{3.1}$$

mais d'autre part nous n'avons *pas*

$$\neg[A\ xstit]\varphi \rightarrow [A\ xstit]\neg\varphi \tag{3.2}$$

Autrement dit, il est vrai que si A réalise $\neg\varphi$ alors il est faux que A réalise φ, mais il n'est pas vrai que s'il est faux que A réalise φ (i.e., A ne réalise pas φ), alors A réalise $\neg\varphi$: ce n'est pas parce que Paul ne chante pas qu'il fait exprès de ne pas chanter. Ainsi, « ne pas réaliser φ » n'est pas équivalent à « réaliser la négation de φ ».[52]

Conformément à cette distinction, l'opérateur O est plus sévère que O'. En effet, dans le premier cas « il est obligatoire que a réalise φ » signifie que pour toute histoire (au moment t_i et pour l'état statique s), si a ne réalise pas φ, alors a réalise une violation V. Dans le second cas, la définition stipule que « il est obligatoire que a réalise φ » exprime que pour toute histoire (au moment t_i et pour l'état statique s), si a réalise $\neg\varphi$, alors a réalise une violation V. La première définition est plus sévère dans la mesure où l'absence d'action (la non réalisation) d'un agent peut mener au fait qu'il réalise néanmoins une violation, alors que dans le second cas la violation est réalisée seulement lorsque $\neg\varphi$ est réalisé. Autrement dit, la première définition implique que pour ne pas qu'il y ait de violation, un agent doit s'assurer que ses obligations sont remplies. En ce sens, cela exclut la possibilité de respecter ses obligations en omettant d'agir.

Outre ces définitions, l'auteur propose aussi d'autres définitions en conjonction avec l'opérateur K_a afin de pouvoir exprimer différents modes de *mens rea* (cf. Broersen 2011a, p.148). Éventuellement, Broersen (2011b, p.509) introduira un opérateur d'intention afin de rendre compte de certains modes de *mens rea*, où un agent agit non seulement en connaissance de cause, mais agit aussi en vue d'accomplir un certain but.

* * *

Ceci conclut le chapitre sur les logiques de type STIT. La caractéristique principale de ce genre d'approches est que celles-ci rendent compte de l'action en modélisant l'évolution d'un scénario, où l'action est associée à une description du monde. En plus des logiques STIT, on trouve aussi les logiques dynamiques, qui assument une ontologie de l'action plutôt que de considérer les actions en tant que propositions déclaratives. Passons maintenant à ces approches.

[51] Étant donné que KM est une extension de KD, cette nuance peut aussi être exprimée dans un langage STIT.

[52] À ce sujet, voir le chapitre 14.

Chapitre 4

Logique dynamique

Tel que mentionné au chapitre précédent, les logiques de l'action en philosophie se divisent en deux principales catégories, à savoir les logiques STIT et les logiques dynamiques. Bien que les logiques STIT permettent de rendre compte de certaines nuances du point de vue de l'action, comme la différence entre omettre d'agir $\neg E_i A$ et s'abstenir d'agir $E_i \neg A$, ces dernières ne sont toutefois pas en mesure de rendre compte du fait que ce sont les *actions* qui sont obligatoires, et non des descriptions du monde. Le principal avantage des logiques dynamiques par rapport aux logiques STIT est leur capacité à distinguer entre les *actions* et les *propositions*, ce qui est nécessaire afin d'exprimer que ce ne sont pas les propositions qui sont obligatoires, mais bien les actions. En ce sens, les logiques dynamiques sont de type *Ought-to-do*.

Une logique déontique est dite *réductible* lorsqu'aucun opérateur déontique n'est pris en tant que primitif, et donc que toute proposition qui contient un opérateur déontique est équivalente à une proposition qui n'en contient pas (Anglberger 2009, pp.179-80). De manière générale, les logiques déontiques dynamiques sont des réductions similaires à celle proposée par Anderson (1958a), chez qui une proposition est obligatoire lorsque la fausseté de cette proposition entraîne nécessairement une sanction S.[1] Dans le même ordre d'idées, l'obligation est définie en logique déontique dynamique à l'aide d'une constante de violation V, une action étant interdite lorsque sa performance entraîne un état où la constante de violation est vraie.

Malgré que la principale utilité de la logique déontique dynamique soit en informatique théorique et en programmation, plusieurs auteurs voient un intérêt philosophique à la dichotomie entre *actions* et *propositions*. Dans ce qui suit, nous aborderons d'abord le texte de Meyer, qui a été le premier a introduire la logique dynamique au sein de la littérature afin de résoudre certains paradoxes. Cela sera suivit de l'approche de Royakkers, qui utilise explicitement la logique déontique dynamique afin de modéliser les obligations légales en indexant les obligations à des agents et des groupes d'agents. Ensuite, il sera question de l'approche de Segerberg, qui va au-delà du système de Meyer en ajoutant des modalités temporelles à la logique déontique dynamique. Enfin, nous aborderons le système de van der Meyden, qui insiste sur la permission plutôt que sur l'obligation, et nous

[1] Voir aussi Anderson (1958b, 1967).

terminerons le chapitre avec l'analyse proposée par Broersen quant aux actions négatives en logique dynamique.

John-Jules Meyer

La logique déontique dynamique est d'abord apparue avec les travaux de Meyer (1987, 1988), qui cherchait à répondre aux paradoxes du bon Samaritain et de Forrester. La principale distinction entre la logique STIT et la logique dynamique est que cette dernière distingue entre les *propositions* et les *actions*. Contrairement à la logique STIT, la logique dynamique permet de distinguer entre les *effets* d'une action et l'*action* elle-même.[2] Le langage

$$\mathcal{L} = \{Prop, Act, (,), [\,], ;, \&, \rightarrow /, \underline{\cup}, \bar{}, \vee, \wedge, \supset, \equiv, \neg\}$$

est construit à partir des connecteurs habituels, auxquels on ajoute un opérateur [ ] ainsi que cinq connecteurs pour les actions. Tout comme pour l'opérateur *stit*, l'utilisation des crochets a pour objectif de marquer la similarité avec l'opérateur $\Box$. Soit

$$Act = \{\underline{\varnothing}, \underline{U}, a_1, ..., a_n, ...\}$$

un ensemble dénombrable d'actions atomiques et

$$Prop = \{V, p_1, .., p_n, ..\}$$

un ensemble dénombrable d'atomes propositionnels. La constante V représente une violation dans un scénario et les constantes d'action $\underline{\varnothing}$ et $\underline{U}$ représentent respectivement l'action qui *échoue* et l'action *universelle*. Les connecteurs pour les actions se lisent de la manière suivante (Meyer 1988, p.111):

1. $\bar{\alpha}$: non α (l'action négative, niée);
2. $\alpha\&\beta$: α et β (la conjonction d'actions);
3. $\alpha;\beta$: α et ensuite β (la séquence d'actions);
4. $\alpha\underline{\cup}\beta$: α ou β (le choix inclusif);
5. $A \rightarrow \alpha/\beta$: si A, alors α, sinon β (l'action conditionnelle).
6. $[\alpha]A$: si α est réalisée, alors A est une description vraie du monde.[3]

À l'instar des logiques modales, l'opérateur [ ] possède un opérateur dual (Meyer 1988, p.114).

$$\langle\alpha\rangle\varphi =_{def} \neg[\alpha]\neg\varphi \qquad \text{(def. } \langle\,\rangle)$$

[2] Pour une introduction à la logique dynamique, voir Harel et al. (2000).
[3] Notons que dans les deux derniers cas, A est une proposition déclarative, et non une action.

Considérons maintenant l'ensemble d'actions complexes $Act*$, défini récursivement par:

1. si $\alpha \in Act$, alors $\alpha \in Act*$;
2. si $\alpha, \beta \in Act*$, alors $\bar{\alpha}, \alpha; \beta, \alpha \& \beta, \alpha \cup \beta \in Act*$.

L'ensemble des expressions bien formées EBF est défini récursivement de la manière suivante:

1. si $p \in Prop$, alors $p \in EBF$;
2. si $A, B \in EBF$, alors $\neg A, A \supset B, A \vee B, A \wedge B, A \equiv B \in EBF$;
3. si $A \in EBF$ et $\alpha, \beta \in Act*$, alors $A \rightarrow \alpha/\beta \in Act*$;
4. si $\alpha, \beta \in Act*$ et $A \in EBF$, alors $[\alpha]A \in EBF$.

Tel que susmentionné, une logique déontique est dite *réductible* lorsque les opérateurs déontiques se définissent à l'aide d'énoncés qui ne contiennent pas d'opérateur déontique, et donc qu'aucun opérateur déontique n'est pris en tant que primitif (Anglberger 2009, p.180). Dans le cas de la logique déontique dynamique, Meyer offre une réduction à l'image de celle présentée par Anderson (1958a), lequel définissait Op par $\Box(\neg p \supset S)$, c'est-à-dire « nécessairement, si p est faux, alors il y aura une sanction S ». Inspiré de ce genre de réduction, Meyer (1988, p.111) propose les définitions suivantes.

$$F\alpha =_{def} [\alpha]V \qquad \text{(def. F)}$$
$$O\alpha =_{def} F\bar{\alpha} \qquad \text{(def. O)}$$
$$P\alpha =_{def} \neg F\alpha \qquad \text{(def. P)}$$

Une action α est interdite si sa performance implique une violation V, elle est obligatoire si sa négation est interdite, et elle est permise lorsqu'elle n'est pas interdite. Dans le langage de la logique dynamique, O et P signifient:

$$O\alpha \equiv [\bar{\alpha}]V$$
$$P\alpha \equiv \neg[\alpha]V$$

L'axiomatisation de la logique déontique dynamique propositionnelle PD_eL est similaire à celle du système modal K, auquel on ajoute d'autres axiomes afin de rendre compte des relations entre les connecteurs pour actions (Meyer 1988, pp.115-6). À l'instar de K, PD_eL présuppose les axiomes du calcul des propositions, l'axiome de distribution (Dist) pour [ ], le modus ponens et la règle de généralisation (Nec).

$$[\alpha](A \supset B) \supset ([\alpha]A \supset [\alpha]B) \qquad \text{(Dist)}$$

$$\frac{\vdash A}{\vdash [\alpha]A} \qquad \text{(Nec)}$$

$$\frac{\vdash A \supset B, \vdash A}{\vdash B} \qquad \text{(MP)}$$

À cela s'ajoutent les axiomes pour les actions.

$$[\overline{\overline{\alpha}}]A \equiv [\alpha]A \tag{A1}$$
$$[\alpha;\beta]A \equiv [\alpha]([\beta]A) \tag{A2}$$
$$[\overline{\alpha;\beta}]A \equiv [\overline{\alpha}]A \wedge [\alpha]([\overline{\beta}]A) \tag{A3}$$
$$[\alpha \underline{\cup} \beta]A \equiv [\alpha]A \wedge [\beta]A \tag{A4}$$
$$[\overline{\alpha}]A \vee [\overline{\beta}]A \supset [\overline{\alpha \underline{\cup} \beta}]A \tag{A5}$$
$$[\alpha]A \vee [\beta]A \supset [\alpha \& \beta]A \tag{A6}$$
$$[\overline{\alpha}]A \wedge [\overline{\beta}]A \equiv [\overline{\alpha \& \beta}]A \tag{A7}$$
$$[A \rightarrow \alpha/\beta]B \equiv (A \supset [\alpha]B) \wedge (\neg A \supset [\beta]B) \tag{A8}$$
$$[\overline{A \rightarrow \alpha/\beta}]B \equiv (A \supset [\overline{\alpha}]B) \wedge (\neg A \supset [\overline{\beta}]B) \tag{A9}$$
$$[\underline{\varnothing}]A \tag{A10}$$

L'axiome (A1) exprime le tiers exclu du point de vue de l'action, à savoir qu'une action est équivalente à la négation de sa négation. (A2) signifie que l'expression « l'action α suivie de l'action β entraîne A » est équivalente à « si l'action α est réalisée, alors si l'action β est réalisée alors A est une description vraie du monde ». Soulignons que cet axiome implique que le connecteur pour la séquence d'actions n'est pas primitif, et que par conséquent celui-ci aurait pu être omis.

L'axiome (A3) exprime que si la négation de l'action « α suivie de β » entraîne A, alors la négation de α entraîne A et α entraîne que la négation de β entraîne A. Autrement dit, si A est la conséquence de la négation de l'action combinée $\alpha;\beta$, alors A est réalisé suite à non α, et, suite à α, A est réalisé suite à non β.

(A4) signifie que si A est vrai après un choix entre α ou β, alors A est vrai après que α soit exécuté et après que β le soit. (A5) exprime que si la négation de α mène à A ou la négation de β mène à A, alors la négation de l'action disjointe $\alpha \underline{\cup} \beta$ mène aussi à A. De la même manière, (A6) signifie que si α mène à A ou β mène à A, alors l'action conjointe de α et β y mène aussi. Les axiomes (A5) et (A6) doivent cependant mettre en jeu des actions qui ont la même durée dans le temps.

(A7) stipule que si la négation de α mène à A, de même que la négation de β, alors la négation de l'action conjointe de α et β y mènera aussi (et vice versa). L'action conditionnelle, quant à elle, n'a pas à être comprise dans les même termes que pour une implication matérielle. En termes simples, $A \rightarrow \alpha/\beta$ signifie que A entraîne α et que $\neg A$ entraîne β. En ce sens, si cela mène à une description B, alors A entraîne que α mène à B et $\neg A$ entraîne que β y mène aussi. Il en va de même pour la négation de l'action conditionnelle et l'interprétation de (A9). L'expression $\overline{A \to \alpha/\beta}$ signifie que A entraîne non α et que $\neg A$ entraîne non β. Notons que ce connecteur ne sera pas utilisé par la suite. Finalement, l'axiome (A10) signifie qu'une action impossible mène à n'importe quel état.

Du côté de la sémantique, Meyer (1988, p.127) introduit les notions d'*ensemble de simultanéité* (*synchronicity set*) et de *trace de simultanéité*. Les ensembles de simultanéité $S_1, ..., S_n, ...$, sont tels que S_i est fini, $S_i \neq \varnothing$ et $S_i \subset Act$, et une trace de simultanéité est une suite t_i d'ensembles de simultanéité (laquelle peut être finie ou non). Ces notions sont introduites afin de pouvoir rendre compte formellement de la sémantique qui gouverne les opérateurs d'action. Informellement, un ensemble de simultanéité correspond à un scénario où les actions (atomiques) sont faites au même moment (ou dans la même période de temps) et une trace de simultanéité correspond à une histoire (voire à l'évolution d'un scénario).

L'idée derrière ces notions est qu'une action réfère aux scénarios dans lesquels elle est posée (Meyer 1988, p.128). Afin de pouvoir isoler certaines parties des différentes traces de simultanéité, Meyer (1988, p.128) introduit une fonction de pertinence, qui attribue 1 à un ensemble de simultanéité lorsque celui-ci est pertinent et 0 lorsqu'il ne l'est pas, ce qui est noté $S^{(i)}$.

Soit $T_1, ..., T_n$ des ensembles de traces de simultanéité. Meyer (1988, p.128) définit le domaine du modèle comme la collection $\mathcal{C}$ d'ensembles qui contiennent des traces infinies, lesquelles contiennent une partie finie pertinente suivie d'une partie infinie non pertinente. Par exemple, T_2 peut contenir $t_1 = S_1^{(1)}, S_4^{(1)}, S_5^{(0)}...$, où S_1 et S_4 forment la partie pertinente de t, suivie d'une partie non pertinente infinie. Afin de rendre compte des connecteurs pour l'action conjointe et l'action niée, Meyer définit un opérateur $\Cap$ pour l'intersection entre les ensembles de simultanéité. $S_1^{(i_1)} \Cap S_2^{(i_2)} = S^{(j)}$ si $S_1 = S_2 = S$ (où $j = i_1$ si $i_2 \leq i_1$, sinon $j = i_2$), sinon $S_1^{(i_1)} \Cap S_2^{(i_2)} = \varnothing$ lorsque $S_1 \neq S_2$.

Comme Meyer (1988, p.129) le mentionne, l'intersection définie par $T_1 \Cap T_2$ est similaire à l'intersection $T_1 \cap T_2$ à la différence que la première prend en compte les traces pertinentes. L'intersection entre deux traces isole les ensembles de simultanéité qui appartiennent aux deux traces en plus d'y associer le plus haut niveau de pertinence. En plus de cette nouvelle opération, Meyer introduit aussi une opération pour rendre compte sémantiquement de l'action niée: $\sim S^{(i)} = (\wp(Act) - S)^{(i)}$, où $\wp(Act) - S$ est le complément de S par rapport à l'ensemble des parties de Act.

Cela fait, Meyer (1988, p.129) introduit les conditions sémantiques pour les opérateurs d'action. Soit $\|\alpha\|$ l'ensemble qui contient toutes les traces possibles où α peut avoir lieu. Cet ensemble est construit dans le même esprit que $\|V_A\|$ au chapitre 2: il isole tous les scénarios où α peut être posée (et donc où « α est réalisée » est vrai). Les conditions sémantiques sont définies récursivement, où $S_1 \circ S_2$ est la composition d'ensembles de simultanéité et où $cut()$ isole la sous-trace pertinente d'une trace infinie.

$$\|a\| = \{S_i : a \in S_i\}^{(1)} \circ (\wp(Act)^{(0)})^{\omega} \tag{1}$$
$$\|\overline{\alpha}\| = \sim \|\alpha\| \tag{2}$$
$$\|\alpha;\beta\| = cut(\|\alpha\|) \circ \|\beta\| \tag{3}$$
$$\|\alpha \underline{\cup} \beta\| = \|\alpha\| \cup \|\beta\| \tag{4}$$
$$\|\alpha \& \beta\| = \|\alpha\| \mathbin{\dot{\cap}} \|\beta\| \tag{5}$$
$$\|\underline{\varnothing}\| = \varnothing \tag{6}$$
$$\|\underline{U}\| = \wp(Act)^{(1)} \circ (\wp(Act)^{(0)})^{\omega} \tag{7}$$

En mots, (1) signifie que l'ensemble qui contient toutes les traces possibles où une action atomique a est réalisée correspond à un ensemble de traces qui est composé de tous les ensembles de simultanéité pertinents qui contiennent a suivi de toutes les combinaisons non pertinentes possibles. (2) signifie que l'ensemble qui contient les traces possibles où l'action niée $\overline{\alpha}$ est réalisée équivaut au complément (relativement à $\wp(Act)$) de l'ensemble qui contient toutes les traces possibles où α est réalisée.

La troisième condition indique que l'ensemble qui contient toutes les traces possibles où l'action α suivie de β est réalisée est la composition de l'ensemble qui contient les traces pertinentes possibles où α est réalisée avec celui qui contient les traces possibles où β l'est. (4) exprime que l'ensemble qui contient les traces possibles où l'action disjointe $\alpha \underline{\cup} \beta$ est réalisée correspond à l'union des ensembles qui contiennent respectivement les traces possibles où α et β sont réalisées.

La condition (5) stipule que l'ensemble qui contient les traces possibles où l'action conjointe $\alpha \& \beta$ est réalisée correspond à l'intersection qui contient les traces pertinentes possibles où respectivement α et β sont réalisées. (6) indique que l'ensemble qui contient les traces possibles où l'action impossible est réalisée est vide, et finalement (7) stipule que l'ensemble qui contient les traces possibles où l'action universelle est réalisée équivaut à l'ensemble des parties pertinentes de Act composé avec toute les combinaisons non pertinentes possibles.

Ayant maintenant à sa disposition une sémantique pour les actions, Meyer (1988, p.133) doit faire le pont entre les traces d'actions possibles et les scénarios possibles qui peuvent en résulter. Soit $W \neq \varnothing$ l'univers du discours qui contient des scénarios possibles et $r : \wp(Act) \times W \longrightarrow W$ une fonction qui identifie l'état $r(S_i, w)$ subséquent à w lorsque les actions de S_i sont accomplies. La fonction $\mathcal{R}$ est définie récursivement à partir de r (Meyer 1988, p.133):

$$\mathcal{R}(S_i, w) = r(S_i, w)$$
$$\mathcal{R}(t_i \circ t_j, w) = \mathcal{R}(t_j, \mathcal{R}(t_i, w))$$

Considérant qu'une trace t_i est une suite de S_j, la seconde condition stipule que l'état subséquent à w lorsque les actions membres de la combinaison de traces t_i et t_j

sont réalisées équivaut à l'état subséquent de « l'état subséquent de w lorsque t_i est réalisé » lorsque t_j est réalisé. Autrement dit, l'état subséquent pour une combinaison $t_i \circ t_j$ se comprend de la manière suivante: soit v l'état subséquent de w lorsque t_i est réalisé, i.e. $\mathcal{R}(t_i, w) = v$, alors l'état subséquent de $t_i \circ t_j$ pour w est l'état subséquent de t_j pour v, i.e. $\mathcal{R}(t_i \circ t_j, w) = \mathcal{R}(t_j, v)$.

Étant donné T_i un ensemble de traces, $\mathcal{R}(T_i, w)$ correspond à l'ensemble qui contient les scénarios v pour lesquels v est l'état subséquent de w pour une trace élément de T_i.

$$\mathcal{R}(T_i, w) = \{v : v = \mathcal{R}(t, w) \text{ pour un } t \in T_i\}.$$

À l'aide de cette fonction, Meyer (1988, p.134) définit la fonction

$$\|\alpha\|_R(w) = \mathcal{R}(cut(\|\alpha\|), w) \tag{def. $\|\|_R$}$$

qui permet d'isoler l'état subséquent à w par rapport à l'ensemble qui contient les traces pertinentes possibles qui contiennent α. Autrement dit, $cut(\|\alpha\|)$ isole les traces pertinentes à l'intérieur de $\|\alpha\|$ et $\|\alpha\|_R(w)$ isole les états subséquents possibles relativement aux traces pertinentes de $\|\alpha\|$.

Cela fait, il est maintenant possible de redéfinir $\|\alpha\|$ de façon à prendre en compte l'état w dans lequel on se trouve.

$$\|a\|(w) = \{S_i : a \in S_i\}^{(1)} \circ (\wp(Act)^{(0)})^{\omega} \tag{1'}$$
$$\|\overline{\alpha}\|(w) = \sim \|\alpha\|(w) \tag{2'}$$
$$\|\alpha;\beta\|(w) = cut(\|\alpha\|(w)) \circ \|\beta\|(\|\alpha\|_R(w)) \tag{3'}$$
$$\|\alpha \underline{\cup} \beta\|(w) = \|\alpha\|(w) \cup \|\beta\|(w) \tag{4'}$$
$$\|\alpha \& \beta\|(w) = \|\alpha\|(w) \mathbin{\dot\cap} \|\beta\|(w) \tag{5'}$$
$$\|\underline{\varnothing}\|(w) = \varnothing \tag{6'}$$
$$\|\underline{U}\|(w) = \wp(Act)^{(1)} \circ (\wp(Act)^{(0)})^{\omega} \tag{7'}$$

Seule la signification de la clause (3) change: l'ensemble qui contient toutes les traces possibles où l'action α suivie de β est réalisée dans w est la composition de l'ensemble qui contient les traces pertinentes possibles où α est réalisée dans w avec celui qui contient les traces possibles où β est réalisée dans l'état subséquent à w relativement à l'ensemble qui contient les traces pertinentes qui contiennent α. Considérant que l'action $\alpha;\beta$ marque la réalisation de deux actions consécutives, l'état dans lequel β est réalisée est l'état subséquent de celui dans lequel α est réalisée. Dans les autres cas, il suffit de rajouter un « relativement à l'état w » à la lecture que nous avons proposé des clauses (1)–(7).

Ayant fait le pont entre la sémantique relative aux actions et celle des propositions, Meyer (1988, p.134) est en mesure d'exprimer la condition pour l'opérateur d'action conditionnelle $\rightarrow /$.

$$\|A \longrightarrow \alpha/\beta\|(w) = \begin{cases} \|\alpha\|(w) \text{ si } \models_w A \\ \|\beta\|(w) \text{ si } \not\models_w A \end{cases} \tag{8'}$$

À cela s'ajoute la clause (9'), qui est une légère modification de (def. $\|\|_R$) où l'on indexe $\|\alpha\|$ à w.

$$\|\alpha\|_R(w) = \mathcal{R}(cut(\|\alpha\|(w)), w) \tag{9'}$$

Tout ce matériel nous permet d'en arriver à la condition de vérité pour $[\alpha]A$.

$$\models_w [\alpha]A \Leftrightarrow \models_v A \text{ pour tout } v \in \|\alpha\|_R(w)$$

En mots, cela signifie qu'il est vrai que réaliser l'action α entraîne A pour un scénario w à condition que A soit vrai pour toute alternative subséquente v résultant de l'action α. Malgré que le modèle ne soit pas explicitement défini chez Meyer, il est possible de représenter la sémantique à l'aide d'un modèle $\mathcal{M} = \langle W, Act, \|\|(), r, V\rangle$, où W est l'univers (non vide) du discours, Act est un ensemble d'actions atomiques, $\|\|()$ est une fonction qui détermine l'ensemble des traces pertinentes possibles relativement à un scénario, r une fonction qui détermine les états subséquents possibles et $V(A)$ une fonction qui détermine l'ensemble qui contient tous les scénarios où une proposition A est vraie. Les clauses pour les autres connecteurs sont définies de manière habituelle.

Meyer proposera d'augmenter PD_eL avec un axiome de cohérence pour les obligations similaire à l'axiome (**D**), et il proposera l'introduction de plusieurs constantes de violation afin de rendre compte des conflits normatifs.

En somme, le principal attrait de la logique déontique dynamique est que celle-ci permet une distinction nette entre les *actions* et les *propositions*, ce qui permet une représentation plus adéquate des propositions normatives dans la mesure où O porte sur des actions plutôt que sur des propositions. Par surcroît, la logique dynamique possède un avantage sur la logique STIT considérant qu'elle permet de distinguer entre l'*action* et son *effet* dans le monde. Le lecteur trouvera une analyse approfondie de l'approche de Meyer au chapitre 14.

Lambèr Royakkers

Malgré que les logiques déontiques dynamiques aient l'avantage face aux logiques déontiques standards d'être en mesure de distinguer entre les les propositions *descriptives* et les *actions* (sur lesquelles portent les opérateurs déontiques), il n'en demeure pas moins que la logique déontique dynamique développée par Meyer (1988) ne permet pas de distinguer entre les obligations de groupe et les obligations individuelles. De fait, PD_eL ne permet pas de spécifier *à qui* la norme est adressée. Puisque les normes légales sont adressées à des individus spécifiques, Royakkers (1998, p.65) est d'avis qu'une logique déontique visant à formaliser le discours légal se doit d'être en mesure de spécifier les agents à qui les normes s'adressent.

L'ouvrage de Royakkers (1998), intitulé *Extending deontic logic for the formalisation of legal rules*, est une version retravaillée de la thèse de doctorat de l'auteur, où ce dernier cherchait à formaliser les contradictions inhérentes au discours légal. La consistance *nor-*

mative, par opposition avec la consistance *logique*[4], est telle que l'on ne trouve pas d'énoncé B pour lequel à la fois OB et $O\neg B$ sont dans l'ensemble des conséquences logiques d'un ensemble de propositions Γ (Royakkers 1998, p.41). La consistance normative s'exprime généralement à l'aide du schéma d'axiome (**D**) de la logique déontique standard, lequel est propositionnellement équivalent à l'énoncé suivant:

$$\neg(OA \wedge O\neg A)$$

Soulignons que si une logique admet cet axiome, alors un ensemble d'énoncé(s) normativement inconsistant sera aussi par le fait-même logiquement inconsistant. En effet, si Γ est normativement inconsistant, alors il y a B tel que OB et $O\neg B$ se trouvent dans l'ensemble de ses conséquences, et donc $OB \wedge O\neg B$ et $\neg(OB \wedge O\neg B)$ se trouveront aussi dans l'ensemble des conséquences de Γ. Par conséquent, l'ajout de (**D**) permet de réduire l'inconsistance normative à l'inconsistance logique.

La principale innovation de Royakkers est d'augmenter la logique déontique afin de rendre compte des individus à qui les normes s'adressent. Considérant que la logique déontique standard et que la logique déontique dynamique permettent respectivement de représenter les énoncés du type *Ougth-to-be* et *Ought-to-do*, Royakkers propose d'augmenter chacun de ces systèmes. Dans ce qui suit, nous avons choisi de présenter brièvement son analyse de la logique déontique standard étant donné que cela nous amènera à expliciter certaines considérations relatives aux différentes interprétations de l'opérateur O. Notre présentation sera toutefois restreinte aux points innovateurs de sa proposition, principalement afin que nous puissions nous concentrer sur le système de logique déontique dynamique qu'il avance ainsi que sur les applications qu'il en fait.

Dans le chapitre 4 de son livre, Royakkers (1998, p.67) propose d'indexer l'opérateur O de la logique déontique standard à un agent $_i$, lequel peut être considéré autant comme une constante qu'une variable sur laquelle on quantifie. Du côté de la sémantique, cela se traduit par l'indexation de la relation R sur le modèle sémantique à un agent i et à l'augmentation du modèle par un ensemble d'agents Ag.

Cet opérateur représente l'obligation personnelle et permet de définir quatre autres opérateurs (Royakkers 1998, p.78), ayant chacun une interprétation bien spécifique.

$$O^{+}(A) =_{def} \forall_{i \in Ag}\, O_i(A) \qquad \text{(def. } O^{+}\text{)}$$
$$P^{+}(A) =_{def} \exists_{i \in Ag}\, P_i(A) \qquad \text{(def. } P^{+}\text{)}$$
$$O^{-}(A) =_{def} \exists_{i \in Ag}\, O_i(A) \qquad \text{(def. } O^{-}\text{)}$$
$$P^{-}(A) =_{def} \forall_{i \in Ag}\, P_i(A) \qquad \text{(def. } P^{-}\text{)}$$

Ces opérateurs représentent respectivement l'obligation générale, la permission non spécifique, l'obligation non spécifique et la permission générale. L'obligation (ou la permission) générale est valable pour tous les agents, alors que l'obligation (ou la permission)

[4] Un ensemble de propositions Γ est logiquement consistant s'il n'y a pas d'énoncé B pour lequel B et $\neg B$ se trouvent à la fois dans l'ensemble des conséquences logiques de Γ.

non spécifique ne s'adresse qu'à au moins un individu (cf. Royakkers 1998, p.79 pour les relations entre ces opérateurs). Du côté sémantique, chaque opérateur sera représenté par une restriction. Par exemple, alors qu'usuellement nous avons OA vrai pour w si et seulement si A est vrai pour tout v t.q. wRv, nous aurons O_iA vrai pour w si et seulement si A est vrai pour tout v t.q. wR_iv. Dans le cas de O^+, nous obtenons que $O^+(A)$ est vrai pour w si et seulement si pour tout i, A est vrai pour tout v t.q. wR_iv. Pour O^-, il suffira d'avoir au moins un i qui satisfait la condition.

L'intérêt de l'approche de Royakkers est l'introduction de *groupes* d'agents. Celui-ci propose deux interprétations de la notion d'obligation collective, à savoir l'interprétation *stricte* et l'interprétation *faible* (Royakkers 1998, p.85). L'interprétation stricte est telle que l'obligation s'adresse à tout le groupe, alors que pour l'interprétation faible elle ne s'adresse qu'à un sous-groupe. Pour des raisons d'économie, nous n'allons présenter que l'interprétation stricte. L'augmentation du modèle standard, afin de rendre compte des groupes d'agents, se fait de manière similaire à ce qui a été susmentionné. En un mot, il suffit d'indexer la relation R à un groupe d'agent $_X$.

Alors que dans le cas de l'obligation personnelle O_i l'agent i est membre de l'ensemble d'agents Ag, l'obligation collective O_X met en jeu un groupe qui est membre de l'ensemble des parties de Ag, c'est-à-dire que $X \in \wp(Ag)$ et donc $X \subseteq Ag$. Le modèle sémantique $\mathcal{M} = \langle W, Ag, R_{Ag}, a\rangle$ est défini à l'aide d'un ensemble d'agents $Ag \neq \varnothing$ et d'un ensemble $R_{Ag} = \{R_X : X \in \wp(Ag)\}$ de relations indexées à des groupes d'agents (cf. Royakkers 1998, p.84). W est un ensemble non vide de scénarios possibles et a est une fonction qui attribue des valeurs de vérité aux propositions dans W. La clause sémantique pour O_X est simplement que O_XA est vrai pour w si et seulement si A est vrai pour tout v t.q. wR_Xv. Dans le cas de l'obligation collective stricte, Royakkers admet le schéma d'axiome (**D**), et donc R_X est sérielle.

Cela fait, Royakkers (1998, p.88) introduit les obligations collectives *fortes* ($O^{\oplus}$) et *faibles* ($O^{\ominus}$). Alors que l'obligation $O^+(A)$ signifie que tous les agents ont l'obligation A, $O^{\oplus}(A)$ exprime que tout groupe X possède une obligation *collective forte*, où l'obligation s'applique à tous les groupes. À partir de l'opérateur O indexé à un groupe d'agents X, l'auteur définit quatre opérateurs d'obligations collectives, à l'instar des obligations personnelles générales et non spécifiques.

$$O^{\oplus}(A) =_{def} \forall_{X \in \wp(Ag)}\, O_X(A) \qquad \text{(def. } O^{\oplus}\text{)}$$
$$O^{\ominus}(A) =_{def} \exists_{X \in \wp(Ag)}\, O_X(A) \qquad \text{(def. } O^{\ominus}\text{)}$$
$$P^{\ominus}(A) =_{def} \forall_{X \in \wp(Ag)}\, P_X(A) \qquad \text{(def. } P^{\ominus}\text{)}$$
$$P^{\oplus}(A) =_{def} \exists_{X \in \wp(Ag)}\, P_X(A) \qquad \text{(def. } P^{\oplus}\text{)}$$

Les clauses sémantiques sont définies similairement. Par exemple, $O^{\oplus}(A)$ est vrai pour w si et seulement si pour tout X, A est vrai pour tout v tel que wR_Xv. Le lecteur est invité à consulter Royakkers (1998, p.94) concernant les relations entre les différents opérateurs déontiques et les deux interprétations de l'obligation collective.

Passons maintenant à l'analyse de la logique déontique dynamique. Nous allons suivre la présentation de l'auteur, introduisant d'abord la logique déontique dynamique afin de la modifier par la suite pour rendre compte des agents et des groupes d'agents. Royakkers utilise essentiellement l'approche proposée par Meyer (1988), qu'il modifiera afin de mieux servir son propos. Soit le langage suivant:

$$\mathcal{L} = \{(,), Act, Prop, \bar{}, \&, \cup, \neg, \supset, \wedge, \vee, [\,], \textbf{any}, \textbf{fail}, \textbf{skip}, \textbf{change}\}$$

D'emblée, on remarque que Royakkers (1998, p.52) exclut les séquences d'actions et les actions conditionnelles, représentées chez Meyer par les connecteurs ; et $\rightarrow$ /. L'ensemble $Act = \{\textbf{any}, \textbf{change}, \alpha_1, ..., \alpha_n, ...\}$ contient un nombre dénombrable d'actions atomiques ainsi que les actions **any** (l'action universelle, « peu importe laquelle ») et **change** (« peu importe laquelle sauf **skip** »). L'ensemble $Act*$ d'actions complexes est défini récursivement à l'instar de chez Meyer et la lecture des connecteurs est similaire:

1. si $\alpha \in Act$, alors $\alpha \in Act*$;
2. $\textbf{fail}, \textbf{skip} \in Act*$;
3. si $\alpha, \beta \in Act*$, alors $\overline{\alpha}, \alpha\&\beta, \alpha \cup \beta \in Act*$.

Alors que l'action **skip** n'apporte aucun changement au scénario, **fail** représente l'action qui échoue et après laquelle aucun scénario n'est accessible (Royakkers 1998, p.53). $Prop = \{p_1, ..., p_n, ...\}$ est un ensemble dénombrable d'atomes propositionnels et l'ensemble des expressions bien formées (EBF) est défini récursivement par:

1. si $p \in Prop$, alors $p \in EBF$;
2. si $A, B \in EBF$ et $\alpha \in Act*$, alors $\neg A, A \supset B, A \vee B, A \wedge B, [\alpha]A \in EBF$.

L'axiomatisation de la logique dynamique chez Royakkers se fait à l'instar de celle de PD_eL, ou encore de la logique modale K. On prend les règles de généralisation (Nec), de détachement (MP) ainsi qu'une règle de substitution d'équivalences pour les actions (Subst).[5]

$$\frac{\vdash A}{\vdash [\alpha]A} \qquad \text{(Nec)}$$

$$\frac{\vdash A \supset B \qquad \vdash A}{\vdash B} \qquad \text{(MP)}$$

$$\frac{\alpha = \beta}{\vdash [\alpha]A \equiv [\beta]A} \qquad \text{(Subst)}$$

[5] Comme le souligne Royakkers (1998, p.62), la règle (Subts) est formellement inadéquate puisqu'elle met en jeu une expression qui n'est pas bien formée. Cela dit, il serait aisé de définir l'identité sur les actions, et de fait il s'agit d'une lacune mineure. L'identité est à entendre en termes de classe d'équivalence.

À cela s'ajoutent les axiomes du système, prenant pour acquis ceux de la logique propositionnelle classique (Royakkers 1998, p.63).

$$[\alpha](A \supset B) \supset ([\alpha]A \supset [\alpha]B) \tag{A1}$$
$$[\alpha \cup \beta]A \equiv ([\alpha]A \wedge [\beta]A) \tag{A2}$$
$$([\alpha]A \vee [\beta]A) \supset [\alpha \& \beta]A \tag{A3}$$
$$[\mathbf{skip}]A \equiv A \tag{A4}$$
$$[\mathbf{fail}]A \tag{A5}$$

L'axiome (A1) est l'axiome de distribution habituel, (A2) exprime que si un choix entre deux actions mène à une description A, alors chaque action individuellement y mène. Le troisième axiome signifie que si deux actions mènent individuellement à une description A, alors les deux actions faites conjointement y mènent aussi. L'axiome (A4) stipule que faire l'action **skip** équivaut à ne rien faire, et finalement (A5) signifie que n'importe quoi est vrai lorsque le système échoue.

Du côté de la sémantique, Royakkers (1998, p.53) utilise, à l'instar de Meyer, la notion d'ensemble de *simultanéité* (*synchroniticity set*). D'entrée de jeu, $\{\mathbf{fail}\}$, $\{\mathbf{skip}\}$ et tout sous-ensemble non vide de Act sont des ensembles de simultanéité. Soit $S_1, ..., S_n, ...$ des ensembles de simultanéité et $T_1, ..., T_n, ...$ des ensembles d'ensembles de simultanéité. Maintenant, soit $T^{\mathbf{fail}}$ une opération sur les ensembles telle que (cf. Royakkers 1998, p.54):

$$T^{\mathbf{fail}} = \begin{cases} T - \{\{\mathbf{fail}\}\} \text{ s'il y a un } S_i \in T \text{ t.q. } S \neq \{\mathbf{fail}\} \\ \{\{\mathbf{fail}\}\} \text{ sinon} \end{cases}$$

Autrement dit, $T^{\mathbf{fail}}$ est l'ensemble qui contient les ensembles de simultanéité qui ne contiennent pas l'action **fail** lorsqu'il y a au moins un ensemble de simultanéité différent de l'ensemble qui contient **fail**, et si aucun ensemble de simultanéité dans T est différent de celui qui contient **fail**, alors T est l'ensemble qui contient l'ensemble qui contient **fail** ($T - \{\{\mathbf{fail}\}\}$ étant le complément de $\{\{\mathbf{fail}\}\}$ relativement à T). Les opérations sémantiques ensemblistes représentant respectivement les actions négatives, disjointes et conjointes sont données par $\sim, \sqcup$ et $\sqcap$.

$$\begin{aligned} \sim S &= (\wp(Act) \cup \{skip\}) - S \\ \sim T &= \sqcap\{\sim S : S \in T\} \\ T_1 \sqcup T_2 &= (T_1 \cup T_2)^{\mathbf{fail}} \\ T_1 \sqcap T_2 &= \begin{cases} T_1 \cap T_2 \text{ si } T_1 \cap T_2 \neq \varnothing \\ \{\{\mathbf{fail}\}\} \text{ sinon} \end{cases} \end{aligned}$$

En mots, l'opération $\sqcap$ est similaire à l'intersection ensembliste usuelle dans la mesure où l'intersection n'est pas vide, sans quoi $T_1 \sqcap T_2$ se résume à l'ensemble qui contient l'ensemble de simultanéité qui contient **fail**. L'opération $\sqcup$ est similaire à l'union ensembliste à l'exception qu'elle s'assure d'exclure l'ensemble de simultanéité **fail** lorsqu'au moins un ensemble de simultanéité de T_1 ou T_2 y est différent. Finalement, $\sim S$ pour un ensemble de simultanéité S est le complément de S relativement à l'ensemble des parties de *Act* joint avec l'ensemble qui contient **skip**, et $\sim T$ pour un ensemble d'ensembles de simultanéité équivaut à l'intersection $\sqcap$ de tous les $\sim S$ pour lesquels S est élément de T.

Ayant ces opérations en main, Royakkers (1998, p.55) définit récursivement la sémantique pour les connecteurs d'actions.

$$\begin{aligned}
\|\alpha\| &= \{S_i : \alpha \in S_i\} \text{ pour un atome } \alpha \\
\|\overline{\alpha}\| &= \sim \|\alpha\| \\
\|\alpha \cup \beta\| &= \|\alpha\| \sqcup \|\beta\| \\
\|\alpha \& \beta\| &= \|\alpha\| \sqcap \|\beta\| \\
\|\mathbf{skip}\| &= \{\{\mathbf{skip}\}\} \\
\|\mathbf{fail}\| &= \{\{\mathbf{fail}\}\} \\
\|\mathbf{any}\| &= \wp(Act) \cup \{\{\mathbf{skip}\}\} \\
\|\mathbf{change}\| &= \wp(Act)
\end{aligned}$$

Cela fait, on ajoute une fonction qui permet de déterminer les scénarios accessibles suite à l'accomplissement de certaines actions. Soit l'ensemble:

$$\Omega = \wp(Act) \cup \{\{\mathbf{fail}\}\} \cup \{\{\mathbf{skip}\}\}$$

Soit W un ensemble de scénarios et $W*$ l'ensemble qui contient les sous-ensembles de W qui eux contiennent au plus un élément. La définition de $W*$ a pour objectif de pouvoir spécifier qu'aucun scénario n'est accessible après avoir posé l'action qui échoue. Royakkers (1998, p.60) introduit une fonction $\rho : (\Omega \times W) \longrightarrow W*$ telle que:

$$\begin{aligned}
\rho(\{\mathbf{fail}\})(w) &= \varnothing \\
\rho(\{\mathbf{skip}\})(w) &= \{w\}
\end{aligned}$$

Autrement dit, il n'y a aucun monde possible accessible à w lorsque l'action **fail** est posée et le monde accessible à w après l'action **skip** est w lui-même.[6] À supposer que les ensembles de scénarios ne sont pas vide, l'auteur utilise ρ afin de définir la fonction $R : \wp(W) \longrightarrow \wp(W)$ telle que pour $W_0 \subseteq W$ et $S_i \in \Omega$:

$$\begin{aligned}
R(S_i)(W_0) &= \bigcup_{w \in W_0} \rho(S_i)(w) \\
R(T_i)(W_0) &= \bigcup_{S_i \in T_i} R(S_i)(W_0)
\end{aligned}$$

[6] Notons qu'il y a une coquille dans la notation utilisée par Royakkers. Ce dernier écrit $\rho(\{\mathbf{skip}\})(w) = w$ plutôt que $\rho(\{\mathbf{skip}\})(w) = \{w\}$. Or, ρ a pour image $W*$, qui est un ensemble de singletons (et sans lequel l'auteur ne pourrait pas écrire $\rho(\{\mathbf{fail}\})(w) = \varnothing$).

En ce sens, la fonction R permet de déterminer pour un ensemble de simultanéité et un ensemble de mondes possibles l'ensemble de tous les scénarios accessibles suite aux actions posées dans S_i, de même que pour un ensemble d'ensembles de simultanéité T_i. À l'aide de cette fonction, Royakkers (1998, p.61) définit récursivement une fonction sémantique $\|\|_R$:

$$\|\alpha\|_R(w) = R(\|\alpha\|_R)(w)$$
$$\|\alpha\|_R(W_0) = \bigcup_{w \in W_0} \|\alpha\|_R(w)$$

Le lecteur aura certainement remarqué que cette approche est fort similaire à celle de Meyer. L'idée de base est la même: déterminer les ensembles qui indiquent quelles actions sont posées et définir une fonction qui permet d'isoler l'ensemble des alternatives possibles à un scénario (ou un ensemble de scénarios) relativement à un ensemble d'actions posées. Le modèle sémantique $\mathcal{M} = \langle W, Act, \|\|_R, a \rangle$ est défini usuellement, avec W un ensemble de scénarios, Act un ensemble d'actions atomiques, $\|\|_R$ une fonction qui isole les scénarios accessibles et a une fonction qui attribue des valeurs de vérité aux propositions.[7] La clause sémantique pour l'opérateur dynamique est la même que chez Meyer, à savoir que $[\alpha]A$ est vrai pour w si et seulement si A est vrai pour tout scénario v accessible à w lorsque α est posée. Les autres clauses sémantiques sont définies de manière usuelle.

$$\models_w [\alpha]A \Leftrightarrow \models_v A \text{ pour tout } v \in \|\alpha\|_R(w)$$

L'originalité de Royakkers est d'augmenter la logique déontique dynamique afin de rendre compte des agents et groupes d'agents. Pour ce faire, il modifie la sémantique en introduisant de nouvelles définitions pour les ensembles de simultanéité. D'abord, il augmente le langage $\mathcal{L}$ à l'aide des connecteurs $\cup^\star$ et $\&^\star$ pour les actions individuelles et $\cup'$ et $\&'$ pour les actions collectives. Soit Ag un ensemble d'agent. Royakkers (1998, p.98,111) définit récursivement un ensemble Ev d'évènements par la condition suivante:

si $i \in Ag$, alors

(a) si $\alpha \in Act*$, alors $[i : \alpha] \in Ev$;

(b) si $\alpha, \beta \in Ev$, alors $\overline{\alpha}, \alpha \cup^\star \beta, \alpha \&^\star \beta \in Ev$.

Notons une certaine ambiguïté quant à la notion d'*évènement nié* $\overline{\alpha}$. En effet, Royakkers (1998, p.104) considère que $\overline{[i : \beta]} = [i : \overline{\beta}]$, et donc la négation se comporte de la même manière pour les évènements et les actions, mais les autres connecteurs diffèrent. Par surcroît, cela signifie que « il est faux que i fait l'action β » est équivalent à « i fait l'action $\overline{\beta}$ », ce qui est contestable sur le plan philosophique (cf. chapitres 3 et 14). Cela

[7] Autre coquille: défini ainsi, **fail** et **skip** ne font pas partie du modèle sémantique. Il faudrait plutôt $Act \cup \{\mathbf{fail}\} \cup \{\mathbf{skip}\}$.

va à l'encontre de la distinction entre *ne pas faire* (omettre) et *ne pas faire délibérément* (s'abstenir).

La même augmentation (ainsi que la même critique) s'opère pour les évènements collectifs Ev' (où $X \in \wp(Act)$ et Act est un ensemble non vide d'agents):

1. si $\alpha \in Act*$, alors $[X : \alpha] \in Ev'$;
2. si $\alpha, \beta \in Ev'$, alors $\overline{\alpha}, \alpha \cup' \beta, \alpha \&' \beta \in Ev'$.

Dans chacun des cas, les ensembles $EBF^\star$ et EBF' sont définis respectivement par l'ajout des clauses suivantes:

1. si $\alpha \in Ev$, alors $[\alpha]A \in EBF^\star$;
2. si $\alpha \in Ev'$, alors $[\alpha]A \in EBF'$.

Au niveau axiomatique, Royakkers (1998, p.107 et p.120) ne fait qu'ajouter les axiomes (A2*) et (A3*) pour les évènements individuels et (A2') et (A3') pour les évènements collectifs. Les autres schémas d'axiomes et règles s'appliquent aux modalités $[i : \alpha]$ et $[X : \alpha]$.

$$[\alpha \cup^\star \beta]A \equiv ([\alpha]A \wedge [\beta]A) \tag{A2*}$$
$$([\alpha]A \vee [\beta]A) \supset [\alpha \&^\star \beta]A \tag{A3*}$$
$$[\alpha \cup' \beta]A \equiv ([\alpha]A \wedge [\beta]A) \tag{A2'}$$
$$([\alpha]A \vee [\beta]A) \supset [\alpha \&' \beta]A \tag{A3'}$$

Ces différents opérateurs d'action permettent de construire trois logiques dynamiques distinctes DDL, $DDL^\star$ et DDL'.

À l'instar des opérations $\sim$, $\sqcap$ et $\sqcup$, Royakkers (1998, p.101) définit les opérations $\sim^\star$, $\sqcap^\star$ et $\sqcup^\star$ pour les opérateurs individuels et $\sim'$, $\sqcap'$ et $\sqcup'$ pour les opérateurs collectifs. Dans le cas des opérateurs individuels, il faut d'abord reconsidérer les ensembles de simultanéité. Soit S_i^j un ensemble de simultanéité pour un agent i qui contient des évènements (relatifs à i). Soit S_i^* l'ensemble qui contient tous les ensembles de simultanéité pour l'agent i (incluant $[i : \mathbf{skip}]$). Soit Σ l'ensemble qui contient le produit cartésien de tous les S_i^* (où $i \in Ag$). Une *trace* t est un élément de Σ, c'est-à-dire une suite d'ensembles de simultanéité (qui peuvent mettre en jeu différents agents).

Tout cela peut informellement être considéré de la manière suivante. On considère d'abord tous les ensembles de simultanéité possibles pour un agent i. En faisant le produit cartésien de tous les ensembles de simultanéité pour tous les agents i, on obtient un ensemble de n-uplets (où Ag contient n agents) qui propose toutes les combinaisons d'actions possibles pour tous les agents. En ce sens, une trace t correspond à un n-uplet et représente les actions posées par chacun des agents. Cela permettra d'évaluer les scénarios accessibles suite aux actions de tous les agents.

Maintenant, soit T_i un ensemble de traces.[8] Les opérations sont définies à l'instar des opérations susmentionnées pour la logique déontique dynamique, à l'exception que l'on considère des traces plutôt que des ensembles de simultanéité.

$$T_i^{\mathbf{fail}\star} = \begin{cases} T_i - \{\{\mathbf{fail}\}\} \text{ s'il y a un } t \in T_i \text{ t.q. } t \neq \{\mathbf{fail}\} \\ \{\{\mathbf{fail}\}\} \text{ sinon} \end{cases}$$

$$\begin{aligned}
\sim^\star t &= \Sigma - \{t\} \\
\sim^\star T_i &= \sqcap^\star \{\sim^\star t : t \in T_i\} \\
T_1 \sqcup^\star T_2 &= (T_1 \cup T_2)^{\mathbf{fail}} \\
T_1 \sqcap^\star T_2 &= \begin{cases} T_1 \cap T_2 \text{ si } T_1 \cap T_2 \neq \varnothing \\ \{\{\mathbf{fail}\}\} \text{ sinon} \end{cases}
\end{aligned}$$

Royakkers (1998, p.101) définit récursivement la sémantique pour les connecteurs d'actions et d'évènements individuels.

$$\begin{aligned}
\|i : \alpha\| &= \{t \in \Sigma : \alpha \in S_i^j \text{ et } S_i^j \text{ est dans } t\} \text{ pour un atome } \alpha \\
\|\overline{\alpha}\| &= \sim^\star \|\alpha\| \\
\|i : \alpha \cup \beta\| &= \|i : \alpha\| \sqcup^\star \|i : \beta\| \\
\|i : \alpha \& \beta\| &= \|i : \alpha\| \sqcap^\star \|i : \beta\| \\
\|i : \overline{\alpha}\| &= \sim^\star \|i : \alpha\| \\
\|\alpha \cup^\star \beta\| &= \|\alpha\| \sqcup^\star \|\beta\| \\
\|\alpha \&^\star \beta\| &= \|\alpha\| \sqcap^\star \|\beta\| \\
\|i : \mathbf{skip}\| &= \{t \in \Sigma : \mathbf{skip} \in S_i^j \text{ et } S_i^j \text{ est dans } t\} \\
\|\mathbf{fail}\| &= \{\{\mathbf{fail}\}\} \\
\|\mathbf{any}\| &= \Sigma \\
\|\mathbf{change}\| &= \Sigma - \|i : \mathbf{skip}\|
\end{aligned}$$

Encore une fois, la signification de ces conditions est similaire à ce que l'on trouve en logique dynamique sans agent. La première clause exprime que l'ensemble qui contient les évènements où il est vrai que i pose l'action atomique α est l'ensemble qui contient les traces dans lesquelles se trouvent les ensembles de simultanéité qui contiennent α pour i. L'évènement nié $\overline{\alpha}$ est représenté par le complément de α relativement à Σ. Les évènements qui mettent en jeu des actions disjointes et conjointes s'interprètent de la même manière, à l'exception que l'on parle de trace d'ensembles de simultanéité plutôt que d'ensembles de simultanéité. Les évènements disjoints et conjoints se représentent sémantiquement

[8] Les autres conditions pour les ensembles de simultanéité restent les mêmes, c'est-à-dire que $\{\mathbf{fail}\}$, $\{\mathbf{skip}\}$ et $[i : A]$ pour tout sous-ensemble non vide A de $Act*$ sont des ensembles de simultanéité.

par les opérations ensemblistes définies plus haut, et finalement les constantes d'actions s'interprètent trivialement selon leur signification respective.

En somme, du point de vue syntaxique, la notation change mais l'interprétation sémantique requiert simplement que l'on prennent en compte des combinaisons d'actions possibles pour tous les agents, par opposition avec de simples ensembles de simultanéité. En un certain sens, nous aurions pu parler d'*ensembles collectifs de simultanéité*, lesquels représenteraient toutes les actions posées par tous les agents à un moment donné et équivaudrait à la notion de trace.

On ajoute au modèle sémantique $\mathcal{M} = \langle W, Act, Ag, \|\|_R, a\rangle$ un ensemble non vide d'agents, où $\|\|_R$ est redéfinie en termes de traces.[9]

$$\begin{aligned}
\rho(\{\mathbf{fail}\})(w) &= \varnothing \\
R(t)(w) &= \rho(t)(w) \\
R(T_i)(w) &= \{v : v = R(t)(w) \text{ pour tous les } t \in T_i\} \\
\|\alpha\|_R(w) &= R(\|\alpha\|)(w)
\end{aligned}$$

La fonction ρ identifie *le* scénario accessible à w lorsque toutes les actions (atomiques) d'une trace t sont réalisées par les agents. Évidemment, l'action qui échoue ne mène à aucun scénario possible. Lorsqu'une trace met en jeu des évènements complexes, l'ensemble des scénarios accessibles à w est donné par $R(T_i)(w)$, qui contient les scénarios accessibles à chaque trace t de T_i. La clause sémantique pour $[i : \alpha]A$ est identique à celle de $[\alpha]A$, à l'exception des remarques susmentionnées par rapport à $\|\alpha\|_R(w)$. Il est vrai pour w que i pose l'action α, ce qui rend la description A vraie, à condition que A soit vraie pour tout scénario accessible à w après que i ait posé l'action α.

Dans le cas de la logique dynamique mettant en jeu des groupes d'agent, la sémantique de DDL' est identique à celle de $DDL^\star$, à l'exception du fait que l'on redéfinit les opérations sur des traces collectives plutôt que sur des traces (cf. Royakkers 1998, pp.111-121). Que ce soit pour DDL, $DDL^\star$ ou DDL', la définition de l'obligation se fait à l'instar de chez Meyer, où V est une constante de violation et $[\,]$ la modalité appropriée.

$$F(\alpha) =_{def} [\alpha]V, \qquad \text{(def. F)}$$

Notons au passage que Royakkers propose la même analyse pour la logique déontique dynamique que pour la logique déontique standard concernant les obligations générales, non spécifiques, collectives fortes et collectives faibles. À ce sujet, le lecteur peut consulter le chapitre 5 de l'ouvrage de Royakkers.

En guise de conclusion, notons que les travaux de Royakkers sont fortement inspirés de ceux de Dignum et al. (1996), qui utilisent une logique dynamique augmentée de certaines modalités temporelles, mais ne formalisent pas les agents et groupes d'agents.

[9] Soulignons que la définition de ρ pose encore problème puisque $\{\mathbf{fail}\}$ n'est pas une trace mais bien un ensemble de simultanéité. En ce sens, il faudrait plutôt indiquer que $\rho(t)(w) = \varnothing$ pour une trace t dans laquelle $\{\mathbf{fail}\}$ est un ensemble de simultanéité.

Par ailleurs, Hughes et Royakkers (2008) utilisent la logique déontique dynamique afin de rendre compte des obligations qui perdurent dans le temps et n'ont pas de « date d'expiration ». Les auteurs introduisent une constante I de *dette* et définissent $O\alpha$ comme $[\overline{\alpha}]I$, c'est-à-dire que tant que α n'est pas réalisée, l'agent reste en *dette* (Hughes et Royakkers 2008, p.72). Le lecteur peut consulter l'article de Royakkers (2000) pour une critique de la logique STIT en faveur de la logique dynamique.[10]

Kirster Segerberg

Krister Segerberg est un logicien qui a surtout travaillé en logique modale, principalement en logique dynamique épistémique ainsi qu'en logique de l'action. Ce dernier offre une logique déontique dynamique enrichie permettant de rendre compte de beaucoup plus de subtilité que celle proposée par Meyer. L'originalité de l'approche de Segerberg est d'augmenter la logique dynamique avec des modalités temporelles, l'objectif étant de fournir une représentation formelle adéquate de l'*action*. Dans ce qui suit, nous présenterons la sémantique telle que présentée dans Segerberg (2009). Le lecteur peut aussi consulter Segerberg (2007, 2012) pour l'essentiel de sa position en logique déontique dynamique.

Afin de bien comprendre la sémantique proposée, il faut d'abord expliquer certains concepts. Le modèle proposé par Segerberg (2009, p.389) comprend un univers U (non vide), lequel contient des *points*. Les suites non vides (finies ou infinies) de points sont nommées des *traces* (ou chemins, en anglais *path*), auxquelles l'auteur réfère par les lettres p, q, r, etc.[11] Deux traces p et q peuvent être combinées en une trace pq à condition que le dernier point de p soit le premier point de q. L'ensemble Ev est un ensemble d'*évènements*[12], lequel est aussi considéré comme un ensemble d'actions (i.e., un ensemble de choses qui sont réalisées). Un évènement $a \in Ev$, où $Ev \subseteq \wp(U)$ est tel que les évènements de Ev ne contiennent que des traces finies, est un ensemble de traces finies. En ce sens, U contient des points, $\wp(U)$ contient des traces et $\wp(\wp(U))$ contient des évènements.

Informellement, une action a peut être réalisée (advenir) de plusieurs manières. Par exemple, si la porte est ouverte, alors Paul peut ouvrir la porte, la garder ouverte avec son pied, mettre un objet pour empêcher que la porte se ferme, etc. Plusieurs traces différentes peuvent donc faire advenir le même évènement (la même action). L'évènement est formellement conçu comme un ensemble qui contient toutes les traces (finies, sans quoi l'évènement n'adviendrait jamais) qui peuvent y mener.

Les évènements *trivial* et *impossible* correspondent respectivement à **any** et **fail** chez Royakkers, ou encore à U et $\varnothing$ chez Meyer.

L'ensemble H est un ensemble d'*histoires* maximales (à comprendre de la même

[10] Dans cet article, l'auteur intègre le connecteur $\Rightarrow_s$ de Jones et Sergot (1996) à la logique dynamique pour actions collectives.

[11] Attention à ne pas confondre les traces avec les variables propositionnelles.

[12] À ne pas confondre avec la notion d'évènement chez Royakkers, qui associe un agent ou groupe d'agents à une action.

manière que dans le cas des logiques temporelles), où une histoire $hg = (h, g)$ est une trace infinie pour laquelle le dernier point de h est le premier point de g et représente le présent (et donc h et g représentent respectivement le passé et le futur). L'ensemble $cont(h) = \{g : hg \in H\}$ dénote les continuations possibles (les futurs possibles) de h.

Le langage contient un ensemble dénombrable de propositions atomiques $Prop = \{P_1, ..., Pn, ...\}$ (en majuscule pour ne pas confondre avec les traces) et un ensemble dénombrable de termes $T = \{e_1, ..., e_n, ...\}$ qui représentent des évènements atomiques. Le modèle $\mathcal{M} = \langle U, Ev, H, V\rangle$ contient les ensembles susmentionnés (l'univers U, un ensemble d'évènements Ev et un ensemble d'histoires H) ainsi qu'une fonction d'évaluation V qui assigne des valeurs de vérité aux propositions.

Formellement, $V : (Prop \cup T) \longrightarrow \wp(U)$ détermine les ensembles à l'intérieur desquels se trouvent (respectivement) les atomes propositionnels et les termes. Informellement, V détermine les scénarios où les atomes propositionnels sont vrais et les termes sont vrais, ce qui est dénoté par $V(A) = \|A\|$. La fonction est définie récursivement de façon habituelle pour les propositions (où $\varphi, \psi \in EBF_{prop}$, les expressions bien formées étant définis usuellement):

$$\begin{aligned} V(P_i) &= \|P_i\| \\ \|\neg\varphi\| &= U - \|\varphi\| \\ \|\varphi \vee \psi\| &= \|\varphi\| \cup \|\psi\| \\ \|\varphi \wedge \psi\| &= \|\varphi\| \cap \|\psi\| \end{aligned}$$

Quant aux actions (où $\alpha, \beta \in EBF_T$ pour un ensemble de termes bien formés)[13]:

$$\begin{aligned} V(e_i) &= \|e_i\| \\ \|\alpha + \beta\| &= \|\alpha\| \cup \|\beta\| \\ \|\alpha; \beta\| &= \|\alpha\| \uparrow \|\beta\| \\ \|\alpha; ; \beta\| &= \|\alpha\| \Uparrow \|\beta\| \end{aligned}$$

Cette dernière définition (cf. Segerberg 2009, p.390) utilise les ensembles définis par (où A et B contiennent des traces finies):

$$\begin{aligned} &A \uparrow B = \{pq : p \in A, q \in B \text{ et } pq \text{ est une combinaison}\} \\ &A \Uparrow B = \{prq : p, r \in A, q \in B \text{ et } pr \text{ et } rq \text{ sont des combinaisons}\} \end{aligned}$$

Le connecteur + permet d'ouvrir des choix possibles pour les suites d'évènements et les connecteurs ; et ; ; permettent de concaténer respectivement des traces finies et infinies. En ce sens, + représente l'action disjonctive et ; la séquence d'actions.

[13] La définition des expressions bien formées n'est pas explicite chez Segerberg. Néanmoins, considérant que celui-ci introduira des opérateurs dynamiques, on peut en comprendre que ceux-ci sont définis de manière usuelle en logique dynamique.

Cela fait, Segerberg (2009, p.390) introduit des opérateurs (temporels) modaux afin de représenter le passé $\vec{P}$, le futur $\vec{F}$ et la nécessité historique $\Box$.[14] Un scénario peut être représenté par une histoire (h, g). Soit w le point (présent) qui constitue le dernier point de h et le premier point de g. Les clauses sémantiques pour ces opérateurs sont:

$$\models_{(h,g)} \varphi \Leftrightarrow w \in \|\varphi\| \tag{1}$$

$$\models_{(h,g)} \vec{F}\varphi \Leftrightarrow \models_{(h',g')} \varphi \text{ pour tout } v, h', g' \text{ t.q. } hv = h' \text{ et } vg' = g \tag{2}$$

$$\models_{(h,g)} \vec{P}\varphi \Leftrightarrow \models_{(h',g')} \varphi \text{ pour tout } v, h', g' \text{ t.q. } h'v = h \text{ et } g' = vg \tag{3}$$

$$\models_{(h,g)} \Box\varphi \Leftrightarrow \models_{(h',g')} \varphi \text{ pour tout } g' \in \mathit{cont}(h) \tag{4}$$

Les clauses (1) et (4) vont de soi: une proposition est vraie pour un scénario lorsque ce scénario fait partie de l'ensemble qui contient les scénarios où la proposition est vraie et une proposition est historiquement nécessairement vraie pour un scénario w relativement à un passé h à condition que la proposition soit vraie pour tout futur possible accessible à w relativement à h. Les clauses (3) et (4) quant à elles se comprendront mieux à la lumière de la figure 4.1.

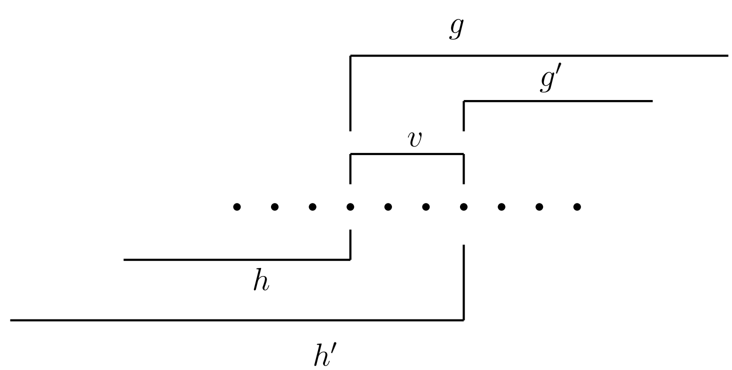


Figure 4.1: Futur

En un mot, l'idée est de dire qu'une proposition $\vec{F}\varphi$ est vraie pour un scénario w d'une histoire hg à condition que φ soit vrai pour tout scénario v de la même histoire qui vient après w. Soulignons qu'un scénario est un ensemble de points, et qu'une histoire peut être vue comme la représentation de l'évolution d'un scénario (global) où chaque point représente un état statique, voire une description, du scénario à un moment donné.

Dans la figure 4.1, on voit clairement que $hv = h'$ et que $g = vg'$. Ainsi, une proposition $\vec{F}\varphi$ est vraie lorsque φ est vrai dans toutes les évolutions possibles d'une histoire. Notons qu'une proposition est vraie pour un futur relativement à un passé spécifique. La proposition $\vec{F}\varphi$ est vraie lorsque φ est vrai pour tous les déploiements possibles de w.

La clause pour le passé se comprend dans le même ordre d'idées. Il est vrai à un moment w pour une histoire hg que φ était vrai à condition que φ soit vrai pour tout moment qui vient avant w sur hg.

[14] Ici, nous avons modifié la notation de Segerberg afin de ne pas confondre avec les opérateurs dynamiques.

En plus de ces opérateurs temporels, Segerberg définit un opérateur binaire $\vec{U}$ qui signifie *jusqu'à* (*until*). L'énoncé $\vec{U}_\varphi\psi$ signifie ψ *jusqu'à ce que* φ. L'interprétation sémantique de cet énoncé est similaire à celle du *à moins que* en logique propositionnelle.

$$\begin{aligned} &\models_{(h,g)} \vec{U}_\varphi\psi \Leftrightarrow \text{ soit} \\ &\quad 1.\ \not\models_{(h',g')} \varphi \text{ pour tout } h', g' \text{ t.q. } h'g' = hg \text{ ou} \\ &\quad 2.\ \models_{(h',g')} \varphi \text{ pour un } v \text{ t.q. } h' = hv \text{ et } g = vg' \text{ et} \\ &\quad\quad \models_{(h'',g'')} \psi \text{ pour tout } x \text{ t.q. } h'' = hx \text{ et } g = xg'' \end{aligned} \tag{5}$$

Autrement dit, il est vrai pour w relativement à hg que ψ *jusqu'à ce que* φ à condition que soit 1. φ soit faux pour tout scénario v de hg ou 2. φ est vrai pour un v de hg qui suit w et que ψ est vrai pour tout x de hg qui vient avant w.

Cela nous amène aux opérateurs dynamiques. Alors qu'en logique dynamique propositionnelle on peut simplement exprimer qu'une description est vraie après qu'une action soit réalisée, l'approche de Segerberg permet de rendre compte de beaucoup plus de subtilités. En effet, ce dernier introduit les opérateurs **does**, **realises**, **occurs**, **done**, **realised** et **occured**.[15] Pour des raisons d'économie (et parce que leur interprétation sémantique est triviale), nous ne présenterons pas en détails la sémantique de ces opérateurs. Soulignons simplement que leur signification est univoque et se traduit directement à l'aide des modalités temporelles développées jusqu'à présent. La différence entre **does** et **realises** est que le premier s'applique aux actions alors que le second s'applique aux propositions, au même titre que **occurs** et **occured** s'appliquent seulement à des actions.

En ce qui concerne les notions déontiques, Segerberg (2009, p.392) mentionne que son approche, étant basée sur une logique dynamique, s'insère dans le cadre des approches de type *Ougth-to-do*. D'abord, Segerberg (2009, p.393) considère les normes *simples*, c'est-à-dire les normes qui assignent à un passé plusieurs futurs idéaux possibles. Étant donné que $cont(h)$ correspond à l'ensemble des futurs possibles suite au passé h, $N(h) \subseteq cont(h)$ isole les futurs accessibles idéaux. Cet ensemble doit satisfaire à la condition suivante, qui exprime la cohérence normative.

$$g = vg' \Rightarrow [g' \in N(hv) \Rightarrow g \in N(h)]$$

En d'autres termes, si g' est un futur idéal accessible à l'histoire hv, alors g est un futur accessible à l'histoire h lorsque g est vg'. Segerberg adopte aussi la condition de sérialité (i.e., $N(h) \neq \varnothing$) mais souligne que celle-ci pourrait très bien ne pas être incluse. De fait, le schéma d'axiome (**D**) est valide. Cela fait, l'auteur est en mesure d'introduire quatre modalités déontiques *locales* et quatre modalités déontiques *normales* (Segerberg 2009, pp.293-5). Les quatre modalités sont l'*obligation*, la *permission*, l'*interdiction* et l'*omission*. Commençons d'abord par les modalités locales. Soit les quatre conditions sémantiques suivantes:

[15] À ce sujet, soulignons que la signification de ces opérateurs est beaucoup plus clair dans Segerberg (2012, p.13) que dans Segerberg (2009, p.391).

$$\models_{(h,g)} O\alpha \Leftrightarrow \tag{6}$$

1. pour tout $g' \in N(h)$ il y a $v \in \|\alpha\|$ (v sous-trace de g')
2. pour tout $v \in \|\alpha\|$ il y a $g' \in N(h)$ (v sous-trace de g')

$$\models_{(h,g)} P\alpha \Leftrightarrow \tag{7}$$

2. pour tout $v \in \|\alpha\|$ il y a $g' \in N(h)$ (v sous-trace de g')

$$\models_{(h,g)} F\alpha \Leftrightarrow \tag{8}$$

3. pour tout $g' \in N(h)$ et pour tout $v \in \|\alpha\|$ (v n'est pas sous-trace de g')

$$\models_{(h,g)} OMMI\alpha \Leftrightarrow \tag{9}$$

4. il y a un $g' \in N(h)$ t.q. pour tout $v \in \|\alpha\|$ (v n'est pas sous-trace de g')

En mots, il est vrai que α est obligatoire relativement à (h, g) si et seulement si 1. tout futur accessible à partir de h contient au moins une partie v membre de l'ensemble qui contient les scénarios où α est réalisée et 2. pour tout scénario v qui est membre de l'ensemble qui contient les scénarios où α est réalisée il y a au moins une trace g' membre de l'ensemble des futurs idéaux à partir de h et dont v fait parti. Il est vrai que α est permise si et seulement si pour tout scénario v qui est membre de l'ensemble qui contient les scénarios où α est réalisée il y a au moins une trace g' membre de l'ensemble des futurs idéaux à partir de h et dont v fait partie. Il est vrai que α est interdite si et seulement si pour tout futur idéal accessible à h et pour tout scénario v membre de l'ensemble qui contient les scénarios où α est réalisée, v ne fait pas partie de g'. Autrement dit, si α est interdite, alors chaque futur idéal accessible ne contient pas de traces qui font partis de l'ensemble qui contient les traces où α est réalisée. Finalement, α peut être omise si et seulement s'il y a au moins un futur accessible idéal qui ne contient pas de trace qui fait partie de l'ensemble qui contient les traces où α est réalisée.

Dans le cas des modalités normales, nous avons besoin de la condition suivante, à laquelle nous réfèrerons par (C): « pour toute trace finie v telle que hv est une combinaison (et donc le dernier point de h est le premier point de v), s'il n'existe pas de trace x élément de $\|\alpha\|$ telle que x est une sous-trace de v, alors ... »:

$$\models_{(h,g)} O^{\star}\alpha \Leftrightarrow (C) \tag{6'}$$

1. pour tout $g' \in N(hv)$ il y a $x \in \|\alpha\|$ (x sous-trace de g')
2. pour tout $x \in \|\alpha\|$ il y a $g' \in N(hv)$ (x sous-trace de g')

$$\models_{(h,g)} P^{\star}\alpha \Leftrightarrow (C) \tag{7'}$$

2. pour tout $x \in \|\alpha\|$ il y a $g' \in N(hv)$ (x sous-trace de g')

$$\models_{(h,g)} F^{\star}\alpha \Leftrightarrow (C) \tag{8'}$$

3. pour tout $g' \in N(hv)$ et pour tout $x \in \|\alpha\|$ (x n'est pas sous-trace de g')

$$\models_{(h,g)} OMMI^{\star}\alpha \Leftrightarrow (C) \tag{9'}$$

4. il y a un $g' \in N(hv)$ t.q. pour tout $x \in \|\alpha\|$ (x n'est pas sous-trace de g')

Chacune des clauses utilise la condition (C), laquelle exprime que s'il n'y a pas de sous-trace x incluse dans v pour hv (pour toute trace v), alors les conditions 1–4 doivent être remplies dépendamment des cas. Autrement dit, s'il n'y a pas de v tel que x est une sous-trace de v et x est membre de l'ensemble des traces où α est réalisée, alors: dans le cas de l'obligation normale, cela implique que tout futur accessible à hv contient au moins une trace qui fait partie de l'ensemble des traces où α est réalisée et que toute trace qui fait partie de l'ensemble des traces où α est réalisée fait partie d'au moins un futur accessible à hv. Pour la permission, cela implique que toute trace qui fait partie de l'ensemble des traces où α est réalisée fait partie d'au moins un futur accessible à hv. Pour l'interdiction, cela signifie que les traces qui font parties de l'ensemble qui contient les traces où α est réalisée ne sont pas des sous-traces des futurs accessibles à hv. Finalement, dans le cas de l'omission, cela implique qu'il y a au moins un futur accessible à hv qui ne contient aucune trace dans laquelle α est réalisée.

Après avoir considéré les normes *simples*, Segerberg (2009, p.395) en vient à considérer les normes *complètes*, la principale distinction étant que la dernière ne sélectionne pas l'ensemble des futurs idéaux accessibles, mais plutôt un sous-ensemble (préférable) de l'ensemble des futurs idéaux accessibles. Le formalisme est similaire à celui des normes simples, à l'exception du fait que l'on considère une histoire (h, g) relativement à un $S \subseteq N(h)$ où tout futur accessible membre de S est préférable aux futurs accessibles qui ne sont pas membres de S. Cela se fait notamment en introduisant un opérateur qui permet de sélectionner un ensemble de scénarios où une certaine proposition est vraie.

Tout compte fait, l'approche de Segerberg concernant la logique déontique dynamique a ses mérites, incluant le fait d'introduire plusieurs modalités dynamiques afin de rendre compte des modes d'actions. En plus de la richesse de son langage, un aspect intéressant de l'approche de Segerberg est que celui-ci propose une sémantique pour les logiques dynamiques différente de celle de Meyer. Plutôt que de définir des ensembles de simultanéité et des mondes possibles, Segerberg fait d'une pierre deux coups en considérant les mondes possibles comme des suites d'évènements.

Outre ses travaux en logique dynamique, Segerberg a aussi proposé une formalisation de la logique déontique reposant sur une logique de l'action basée sur une algèbre de Boole. Nous reviendrons sur ses travaux dans le chapitre 7. Les travaux de Segerberg ont notamment été repris par Demolombe (2014), lequel supplémente le cadre de travail de Segerberg afin de rendre comte des obligations qui ont une date de tombée.[16]

Ron van der Meyden

Alors que la plupart des logiciens se concentrent sur la formalisation de l'*obligation*, certains sont d'emblée attirés par la *permission*. Van der Meyden (1996), par exemple, a proposé un cadre de travail différent de celui de Meyer afin de pouvoir rendre compte autant des permissions de *choix* (*free-choice permission*) que des permissions qui résultent de ce qui

[16] Concernant les obligations sujettes à certains *deadlines*, voir Broersen et al. (2004).

n'est pas interdit (van der Meyden 1996, p.467).[17]

Une première particularité de l'approche de van der Meyden est d'utiliser l'opérateur dynamique $\langle\alpha\rangle$ comme primitif plutôt que $[\alpha]$. Cela se comprend en fonction du fait que ce dernier met l'emphase sur la permission plutôt que sur l'obligation. Rappelons que la permission est usuellement définie comme l'opérateur dual de l'obligation. Par ailleurs, cette approche provient plutôt de la tradition de l'informatique théorique que de la logique philosophique (cf. van der Meyden 1996, p.466). En effet, l'*action* est plutôt considérée comme un *programme* qui, suite à son exécution, fait que le système est dans un certain état φ. Ainsi, φ est la description de l'état dans lequel le système se trouve suite à l'exécution du programme. Le langage

$$\mathcal{L} = \{(,), Act, Prop, \cup, ;, ^*, \vee, \neg, \langle\,\rangle, \diamond, \pi\}$$

contient un ensemble dénombrable d'actions atomiques $Act = \{a_1, ..., a_n, ...\}$, un ensemble dénombrable de propositions atomiques $Prop = \{p_1, ..., p_n, ...\}$ et les connecteurs d'actions représentant respectivement le choix, la séquence d'actions et l'itération d'action. Soulignons que ces connecteurs correspondent à ceux de la logique dynamique telle que développée initialement par Pratt (1976, 1980) et qui est souvent utilisée en programmation (cf. chapitre 14). Le langage proposé par van der Meyden diffère donc considérablement de celui proposé par Meyer: même si $\cup$ et ; s'interprètent (informellement) de la même manière, on ne trouve pas d'action négative chez van der Meyden, mais plutôt une action *répétée*, où a^* est l'action obtenue en répétant a un nombre fini de fois (van der Meyden 1996, p.467).

Outre les connecteurs pour les actions, on trouve les connecteurs propositionnels usuels, l'opérateur dynamique $\langle\alpha\rangle\varphi$, qui se lit « φ est vrai après certaines exécutions de α » et les opérateurs déontiques $\diamond$ et π, où $\diamond(\alpha, \varphi)$ et $\pi(\alpha, \varphi)$ signifient respectivement que φ est vrai suite à certaine exécutions de α qui ne sont pas interdites et que toute exécution de α qui mène φ est permise (van der Meyden 1996, p.469). L'opérateur dynamique $[\alpha]\varphi$ est défini par $\neg\langle\alpha\rangle\neg\varphi$ et signifie que φ est vrai dans tous les scénarios possibles qui prennent place suite à l'exécution de α. Finalement, $O(\alpha, \varphi)$ est utilisé comme abréviation de $\neg \diamond (\alpha, \neg\varphi)$ et signifie que toute exécution permise (non interdite) de α mène à un état décrit par φ (van der Meyden 1996, p.467,469). Les connecteurs propositionnels sont définis normalement.

L'ensemble des expressions bien formées (EBF) est défini récursivement par:

1. si $\alpha \in Act$, alors $\alpha \in Act*$;
2. si $\alpha, \beta \in Act*$, alors $\alpha \cup \beta, \alpha; \beta, \alpha^* \in Act*$
3. si $p \in Prop$, alors $p \in EBF$;
4. si $\varphi, \psi \in EBF$, alors $\neg\varphi, \varphi \vee \psi \in EBF$;

[17] La seconde notion renvoit à la permission faible, voire implicite.

5. si $\varphi \in EBF$ et $\alpha \in Act*$, alors $\langle\alpha\rangle\varphi, \diamond(\alpha,\varphi), \pi(\alpha,\varphi) \in EBF$.

L'axiomatisation se fait en quatre temps. D'abord, on introduit les axiomes pour la modalité dynamique. Pour ce faire, on utilise le symbole $\perp$ afin de référer à une contradiction arbitraire (p. ex., $p \wedge \neg p$).

$$\langle\alpha\rangle\perp \equiv \perp \tag{1}$$
$$\langle\alpha;\beta\rangle\varphi \equiv \langle\alpha\rangle\langle\beta\rangle\varphi \tag{2}$$
$$\langle\alpha \cup \beta\rangle\varphi \equiv \langle\alpha\rangle\varphi \vee \langle\beta\rangle\varphi \tag{3}$$
$$\langle\alpha^*\rangle\varphi \equiv (\varphi \vee \langle\alpha;\alpha^*\rangle\varphi) \tag{4}$$
$$\langle\alpha\rangle(\varphi \vee \psi) \equiv (\langle\alpha\rangle\varphi \vee \langle\alpha\rangle\psi) \tag{5}$$
$$(\varphi \wedge [\alpha^*](\varphi \supset [\alpha]\varphi)) \supset [\alpha^*]\varphi \tag{6}$$

Le premier axiome signifie qu'une contradiction prend place suite à α si et seulement s'il y a déjà une contradiction dans le système. Comme van der Meyden (1996, p.471) le souligne, les axiomes (2)–(5) permettent d'éliminer les actions complexes de la portée de l'opérateur dynamique. L'axiome (2), comme chez Meyer, signifie que φ est vrai suite à l'action complexe « α suivie de β » si et seulement si « φ est vrai suite à β » est vrai suite à α. (3) stipule qu'un choix entre α ou β mène à φ à condition que chacune des actions y mène individuellement. Le quatrième axiome signifie que si φ est vrai suite à une séquence finie d'action α, alors soit φ est vrai ou φ est le résultat de α suivie d'une séquence finie de α. (5) exprime que α mène à φ ou ψ si et seulement si α mène à φ ou α mène à ψ. Finalement, (6) signifie que si φ est vrai et que « φ implique que φ est le résultat de toute exécution de α » est le résultat d'une séquence finie d'exécutions de α, alors φ est le résultat d'une séquence finie d'exécutions de α.

Dans un deuxième temps, van der Meyden (1996, p.472) propose d'axiomatiser la modalité *non interdit*, c'est-à-dire la permission faible, de la même manière que l'opérateur $\langle\,\rangle$.

$$\diamond(\alpha,\varphi) \supset \langle\alpha\rangle\varphi \tag{7}$$
$$\diamond(\alpha;\beta,\varphi) \equiv \diamond(\alpha,\diamond(\beta,\varphi)) \tag{8}$$
$$\diamond(\alpha \cup \beta,\varphi) \equiv \diamond(\alpha,\varphi) \vee \diamond(\beta,\varphi) \tag{9}$$
$$\diamond(\alpha^*,\varphi) \equiv \varphi \vee \diamond(\alpha;\alpha^*,\varphi) \tag{10}$$
$$\diamond(\alpha,\varphi \vee \psi) \equiv \diamond(\alpha,\varphi) \vee \diamond(\alpha,\psi) \tag{11}$$
$$(\varphi \wedge O(\alpha^*,\varphi \supset O(\alpha,\varphi))) \supset O(\alpha^*,\varphi) \tag{12}$$

(7) signifie que s'il n'est pas interdit que φ soit le résultat de (d'une certaine exécution de) α, alors φ est le résultat d'une certaine exécution de α. L'axiome (8) exprime que s'il n'est pas interdit que φ soit le résultat de l'exécution de la séquence « α suivie de β », alors il n'est pas interdit que « il n'est pas interdit que φ soit le résultat de β » soit le résultat de α. (9) stipule que s'il n'est pas interdit que φ soit le résultat du choix entre α ou β, alors soit il n'est pas interdit que φ soit le résultat de α ou il n'est pas interdit que φ soit

le résultat de β. Le dixième axiome signifie que s'il n'est pas interdit que φ soit le résultat d'une suite d'exécutions de α, alors soit φ est d'emblée vrai ou il n'est pas interdit que φ soit le résultat de α suivie d'une suite d'exécutions de α. L'axiome (11) exprime que s'il n'est pas interdit que φ ou ψ soit vrai suite à α, alors soit il n'est pas interdit que φ soit le résultat de α ou il n'est pas interdit que ψ soit le résultat de α. En dernier lieu, (12) signifie que si φ est vrai et que toute séquence d'exécutions non interdite de α mène à un scénario où la proposition « si φ est vrai, alors toute exécution non interdite de α mène à φ » est vraie, alors toute séquence d'exécutions non interdite de α mène à un scénario où φ est vrai.

En troisième lieu, l'axiomatisation de la permission se fait aussi de manière similaire à celle de l'opérateur dynamique.

$$[\alpha]\neg\varphi \supset \pi(\alpha,\varphi) \tag{13}$$
$$\pi(\alpha;\beta,\varphi) \equiv \pi(\alpha,\langle\beta\rangle\varphi) \wedge [\alpha]\pi(\beta,\varphi) \tag{14}$$
$$\pi(\alpha\cup\beta,\varphi) \equiv \pi(\alpha,\varphi) \wedge \pi(\beta,\varphi) \tag{15}$$
$$\pi(\alpha^*,\varphi) \equiv \pi(\alpha;\alpha^*,\varphi) \tag{16}$$
$$\pi(\alpha,\varphi\vee\psi) \equiv \pi(\alpha,\varphi) \wedge \pi(\alpha,\psi) \tag{17}$$
$$[\alpha^*]\pi(\alpha,\langle\alpha^*\rangle\varphi) \supset \pi(\alpha^*,\varphi) \tag{18}$$

L'axiome (13) signifie que si toute exécution de α mène à $\neg\varphi$, alors toute exécution de α qui mène à φ est permise, ce qui ne pose pas problème puisque par hypothèse aucune exécution de α ne mènera à φ. (14) exprime que si toute exécution de « α suivie de β » qui mène à φ est permise, alors toute exécution de α qui mène à « certaines exécutions de β mènent à φ » est permise et toute exécution de α mène à un scénario où toute exécution où β mène à φ est permise. (15) signifie que si toute exécution de α ou β qui mène à φ est permise, alors toute exécution de α qui mène à φ est permise et toute exécution de β qui mène à φ est permise. (16) stipule que si toute séquence d'exécutions de α qui mène à φ est permise, alors toute action α suivie d'une séquence d'exécutions de α qui mène à φ est permise. L'axiome (17) signifie que si toute exécution de α qui mène à φ ou ψ est permise, alors toute exécution de α qui mène à φ est permise et toute exécution de α qui mène à ψ est permise. Finalement, (18) signifie que si après toute exécution d'une séquence d'exécutions de α il en résulte que toute exécution de « α qui mène à un scénario où certaines exécutions de séquences d'exécutions de α mènent à φ » est permise, alors toute séquence d'exécutions de α qui mène à φ est permise.

La quatrième étape consiste à établir des liens entre les diverses modalités.

$$(\pi(\alpha,\varphi) \wedge \langle\alpha\rangle\varphi) \supset \diamond(\alpha,\varphi) \tag{19}$$
$$(\diamond(\alpha,\varphi) \wedge [\alpha](\varphi \supset \psi)) \supset \diamond(\alpha,\psi) \tag{20}$$
$$(\pi(\alpha,\psi) \wedge [\alpha](\varphi \supset \psi)) \supset \pi(\alpha,\varphi) \tag{21}$$

L'axiome (19) signifie que si toute exécution de α qui mène à φ est permise et que certaines exécutions de α mènent à φ, alors certaines exécutions de α qui mènent à φ

ne sont pas interdites. (20) signifie que si certaines exécutions de α qui mènent à φ ne sont pas interdites et que toute exécution de α mène à un scénario où φ implique ψ, alors certaines exécutions de α qui mènent à ψ ne sont pas interdites. L'axiome (21) quant à lui signifie que si toute exécution de α qui mène à ψ est permise et que toute exécution de α mène à un scénario où φ implique ψ, alors toute exécution de α qui mène à φ est permise.[18]

À ces axiomes s'ajoutent les règles usuelles de *modus ponens* et de généralisation pour l'opérateur dynamique dual $[\,]$. Soulignons qu'en vertu de l'axiome (1) la règle de généralisation semble superfétatoire. En effet, cela implique que $\vdash [\alpha]\top \equiv \top$, et donc $\vdash \top \supset [\alpha]\top$, ce qui permet de conclure par *modus ponens* que $\vdash [\alpha]\varphi$ pour tout φ tel que $\vdash \varphi$ (puisque $\top$ représente n'importe qu'elle tautologie).

La sémantique proposée par van der Meyden est dans la lignée de celle de Segerberg plutôt que celle de Meyer. En effet, ce dernier propose d'interpréter une action comme référant à une suite d'états statiques. Par exemple, si l'on suppose que w est un scénario dynamique, alors w_1, w_2 peuvent être utilisés afin de décrire certains états statiques du scénario w. Au même titre que Segerberg considère un univers rempli de *points*, van der Meyden considère un ensemble d'*états* et une action réfère à un ensemble de suites d'états. Autrement dit, une action réfère à un ensemble qui contient les suites d'états $e_1, ..., e_n$, où e_1 est l'état actuel, e_n est l'état final (le résultat de l'action) et « ... » est ce qui mène à l'état final. En ce sens, considérant que plusieurs suites d'évènements peuvent mener au même résultat, une action est considérée comme l'ensemble des suites d'évènements possibles pouvant mener à ce résultat. Soit $U \neq \varnothing$ un univers d'états et $U^+ \subset \wp(U)$ l'ensemble de toutes les suites d'états de longueur finie $n \geq 1$.

Avant d'aller plus loin, voyons d'abord un exemple informel afin de bien mettre en évidence la conception de l'action que se fait van der Meyden. Son approche est basée sur une logique *dynamique*, et cet aspect est crucial pour quiconque veut bien saisir l'effet d'une action sur le monde.[19] Mettons-la en opposition avec Meyer et Royakkers. Chez ces derniers, l'exécution de certaines actions restreint l'ensemble des scénarios qui sont accessibles au scénario actuel: il y a un ensemble d'actions, un ensemble de scénarios et une fonction qui permet de déterminer quels scénarios sont accessibles suite à l'exécution de certaines actions. Bien que l'idée soit similaire chez van der Meyden (et Segerberg, l'approche de van der Meyden étant basée sur une logique dynamique pour programmation telle qu'exposée dans Segerberg (1982a)), il y a une différence relativement à la conception qu'il se fait de l'action. En un certain sens, une action chez van der Meyden réfère à une (un ensemble de) transition(s) dans le monde. Autrement dit, une action correspond à un (ensemble d') état(s) dynamique(s). Soit w le monde actuel. Chez van der Meyden, l'univers U du modèle sémantique n'est pas un ensemble de *scénarios*, mais plutôt un ensemble de *points*, voire de *tranches de scénarios*. En effet, chaque point est considéré comme la description d'un scénario à un moment précis. D'un point de vue formel, la différence est mineure, mais au niveau de l'interprétation elle est considérable. L'état

[18] Réitérons que dans les explications précédentes, l'opérateur dynamique *après* doit être compris au sens de *certaines exécutions de ... mènent à ...*.

[19] Van der Meyden n'aborde toutefois pas la question de la causalité.

dynamique w, qui est soumis au changement, peut être vue comme un ensemble d'états statiques qui décrivent l'état dynamique à chaque moment.

Dans un telle perspective, l'action est considérée comme une suite d'états statiques ou, plus précisément, comme un ensemble de suite d'états statiques. L'action « Paul vole une bicyclette », par exemple, peut être vue comme la suite s_1, s_2, s_3 où à s_1 Paul met la main sur la bicyclette, à s_2 il embarque sur la bicyclette et à s_3 il part avec la bicyclette. Évidemment, cette conception pose problème si l'on postule la densité de l'action et du temps, où s_1, s_2 et s_3 peuvent eux-mêmes être considérés comme des ensembles d'états (i.e., mettre la main sur le vélo peut être considéré aux $\frac{1}{10^{10}}$ ième de seconde et ainsi de suite).[20] Néanmoins, l'intuition est assez claire: le vol du vélo peut être vu comme la suite s_1, s_2, s_3. Plus encore, considérant qu'il y a plusieurs manières de voler la bicyclette, l'action « Paul vole une bicyclette » peut être considérée comme l'ensemble de toutes les suites d'états possibles qui peuvent être interprétées comme le vol d'une bicyclette par Paul. Dans une certaine mesure, l'action réfère à toutes les manières de la réaliser.[21]

Le modèle $\mathcal{M} = \langle U, P, \tau, a \rangle$ est constitué d'un univers $U \neq \varnothing$ d'états, d'un ensemble $P \subseteq U \times U$ de paires d'états, d'une fonction $\tau : Act \longrightarrow \wp(U^+)$ qui associe à chaque action atomique un ensemble de suites d'états et $a : Prop \times U \longrightarrow \{\bot, \top\}$ une fonction qui attribue des valeurs de vérité aux propositions relativement à certains états (van der Meyden 1996, p.468). Informellement, l'ensemble P est un ensemble de transitions permises. Pour bien comprendre la sémantique de van der Meyden, il faut voir que ce n'est pas un *état* qui est permis ou non, mais bien la *transition d'un état à un autre*. En ce sens, si $< w_1, w_2 >\in P$, alors la transition de w_1 à w_2 est permise.

Afin de rendre compte sémantiquement des actions, van der Meyden doit définir de nouvelles opérations sur les ensembles pour modéliser les connecteurs ; et *. Soit $\sigma_j, \sigma_k \in U^+$ des suites de longueur respective n et m. Autrement dit, $\sigma_j = (s_1, ..., s_n)$ et $\sigma_k = (t_1, ..., t_m)$. Laissons $\sigma[i]$ et $\sigma[f]$ dénoter respectivement l'élément *initial* et l'élément *final* de la suite σ. La composition de suite $\sigma_k; \sigma_j$ n'est définie que lorsque $\sigma_j[i] = \sigma_k[f]$. Ayant cela en main, van der Meyden (1996, p.468) définit les opérations suivantes pour $A, B \in \wp(U^+)$ des ensembles de suites.

$$A; B = \{(\sigma_k; \sigma_j) : \sigma_k \in A, \sigma_j \in B \text{ et } \sigma_j[i] = \sigma_k[f]\}$$
$$A^* = \{w : w \in U\} \cup A \cup A; A\ \cup ...$$

Alors que la première définition indique que $A; B$ est l'ensemble de toutes les combinaisons possibles (définies) des suites de A et B, la seconde dépend du nombre d'itération(s). En enlevant les '...' de la définition précédente, cela donnerait par exemple A itéré 3 fois. Dépendamment du nombre d'itération(s), l'ensemble A^* contiendra tous les scénarios possibles (en effet, $\{w : w \in U\}$ est égal à U!) ainsi que toutes les suites de A, toutes les combinaisons des suites de A, toutes les combinaisons de séquences de suites de A avec les

[20] Il s'agit là selon nous d'une objection contre la réduction d'une action à des suites de descriptions.

[21] Cela fait place à une seconde objection: le mérite de la logique dynamique est d'opérer une distinction entre les propositions déclaratives et les actions. Toutefois, en réduisant l'action à un ensemble de descriptions, la distinction devient floue.

suites de A, etc. De fait, pour les connecteurs d'actions, nous avons:

$$\tau(\alpha;\beta) = \tau(\alpha);\tau(\beta)$$
$$\tau(\alpha \cup \beta) = \tau(\alpha) \cup \tau(\beta)$$
$$\tau(\alpha^*) = \tau(\alpha)^*$$

En ce sens, la séquence d'action $\alpha;\beta$ dénote l'ensemble qui contient toutes les combinaisons possibles de α et de β (n'oublions pas qu'une action est considérée comme référant à un ensemble qui contient des suites possibles). Le choix entre α et β est équivalent à l'union qui contient les suites auxquelles réfèrent α et β, alors que l'action itérée réfère à l'ensemble itéré.

Maintenant, soit $\tau_w(\alpha)$ l'ensemble qui contient toutes les séquences σ_j de $\tau(\alpha)$ pour lesquelles $\sigma_j[i] = w$, c'est-à-dire pour lesquelles l'état (initial) w est le point de départ. La vérité est définie récursivement.[22] Soit $\|p\|$ l'ensemble qui contient les états pour lesquels p est vrai.

$$\models_w p \Leftrightarrow\ w \in \|p\|$$
$$\models_w \neg\varphi \Leftrightarrow \not\models_w \varphi$$
$$\models_w \varphi \vee \psi \Leftrightarrow \models_w \varphi \text{ ou } \models_w \psi$$
$$\models_w \langle\alpha\rangle\varphi \Leftrightarrow \models_{\sigma_j[f]} \varphi \text{ pour un } \sigma_j \in \tau(\alpha)$$
$$\models_w \diamond(\alpha,\varphi) \Leftrightarrow \models_{\sigma_j[f]} \varphi \text{ pour un } \sigma_j \in \tau(\alpha) \text{ vert}$$
$$\models_w \pi(\alpha,\varphi) \Leftrightarrow \text{ tout } \sigma_j \in \tau(\alpha) \text{ t.q. } \models_{\sigma_j[f]} \varphi \text{ est vert}$$

Les quatre premières clauses sont usuelles pour la logique dynamique, les trois premières étant les clauses propositionnelles classiques et la quatrième étant celle pour l'opérateur *après*, laquelle stipule qu'il est vrai que φ est la conséquence de α pour un état w à condition que φ soit vrai pour l'état final d'une suite d'états membre de l'ensemble qui contient les suites auxquelles réfère l'action α. Autrement dit, considérant que α réfère à un ensemble de suites, il est vrai que φ est le cas après l'action α à condition que φ soit vrai pour au moins un état qui est l'état final d'une suite à laquelle réfère l'action α. Les cinquième et sixième clauses se comprennent à l'aide de la notion de suite *verte*. Cela n'a rien d'écologique, mais est plutôt à comprendre en analogie avec le code de la route! Soit $\sigma_j = (s_1, ..., s_n)$ une suite où chaque s_i, s_{i+1} représente un état transitoire (dynamique entre deux états statiques). On dit que σ_j est une suite *verte* lorsque pour tout état transitoire s_i, s_{i+1} de σ_j, $< s_i, s_{i+1} > \in P$. La cinquième clause exprime qu'il est vrai qu'il n'est pas interdit que l'action α entraîne φ pour l'état w à condition que φ soit vrai pour l'état final d'au moins une suite verte membre de l'ensemble qui contient les suites auxquelles l'action α réfère. En d'autres termes, soit σ_j une suite verte d'états à laquelle réfère l'action α. Il est vrai qu'il n'est pas interdit d'exécuter α et d'obtenir φ à condition

[22] L'auteur définit la sémantique avec une implication de la droite vers la gauche. Nous avons cependant défini la sémantique avec un bi-conditionnel conformément à ce que l'on trouve usuellement en logique dynamique.

que φ soit vrai pour l'état final de cette suite verte (pour au moins une suite verte). Quant à la sixième clause, cela signifie qu'il est vrai qu'il est permis que l'exécution de α mène à φ à condition que toute les suites auxquelles réfère l'action α et pour lesquelles φ est vrai pour le dernier état soient vertes. Autrement dit, l'action α réfère seulement à des suites vertes et φ est vrai pour l'état final de chaque suite.

L'intérêt de l'approche de van der Meyden vient du fait que ce dernier met l'emphase sur la permission plutôt que sur l'obligation. Soulignons que van der Meyden (1996, p.467) propose cette approche sous prétexte que celle de Meyer ne rend pas compte adéquatement de la notion de permission *légale*. Par ailleurs, comme van der Meyden (1996, pp.477-8) le mentionne, même si son approche permet de rendre compte de différents types de permissions, celle-ci n'est toutefois pas en mesure de rendre compte de la notion d'*obligation*, laquelle est fondamentale au discours légal. Plus encore, van der Meyden ne traite pas explicitement de la notion d'*interdiction*, ce qui se comprend en fonction du fait qu'il ne présente pas de sémantique pour l'action *niée*. Cela est d'ailleurs plutôt problématique: même si la formalisation de l'action négative pose problème, il n'en demeure pas moins que celle-ci est nécessaire dans la mesure où l'on veut rendre compte de l'abstention. L'action négative a notamment été étudiée par Broersen (2004), ce que nous aborderons dans la prochaine section (voir aussi le chapitre 14). Outre les suites *vertes* introduites par van der Meyden, Csirmaz (1990) a aussi introduit d'autres *couleurs* (et d'autres ensembles de paires d'états) afin de rendre compte de différents degrés de permission. Enfin, van der Meyden (1991) a aussi travaillé sur des systèmes où les actions réalisées changent l'ensemble des suites vertes.

Jan Broersen

Dans un article où il cherche à fournir une formalisation adéquate du concept d'*action négative*, Broersen (2004) reprend une partie des travaux qu'il a fait pour sa thèse de doctorat. (cf. Broersen 2003). Étant donné que nous avons jusqu'à présent donné un bon aperçu de ce que sont les logiques déontiques dynamiques, nous n'exposerons pas l'approche de Broersen en détails mais présenterons seulement son analyse de l'action négative dans le cadre de la logique dynamique ainsi que les caractéristiques originales de son approche.[23]

D'emblée, l'objectif de Broersen (2004, p.155) est de proposer un formalisme adéquat à la représentation de l'action négative $-\alpha$, interprétée comme une abstention (délibérée), par opposition avec une simple omission. Autrement dit, $-\alpha$ (lorsque considérée avec un opérateur dynamique) signifie *s'abstenir de faire* α (*to refrain from doing* α). Selon l'auteur, le formalisme doit être en mesure de rendre compte de trois caractéristiques:

1. l'interprétation doit être plausible lorsque la négation est combinée avec les autres connecteurs pour actions;

[23] Notons que Broersen (2004, p.154) utilise une logique dynamique plus complexe que celle proposée par Meyer, où l'itération, l'action inverse et l'action *test* sont considérées.

2. cela n'introduit pas de restriction *ad hoc* sur la séquence ou l'itération;

3. l'interprétation doit être plausible dans un contexte normatif.

Broersen (2004, p.165) refuse de réduire les définitions de O, F ou P à l'un ou l'autre de ces opérateurs et ne fait que poser quelques conditions de cohérence normative en tant qu'axiomes.[24] Plutôt, il propose d'introduire trois constantes de violations V_P, V_O et V_F afin de rendre compte individuellement de chaque opérateur déontique.

$$P\alpha =_{def} [\alpha]\neg V_P \qquad \text{(def. P)}$$

$$O\alpha =_{def} P\alpha \wedge [\wr^I \alpha]\neg V_O \qquad \text{(def. O)}$$

$$F\alpha =_{def} \langle\alpha\rangle V_F \qquad \text{(def. F)}$$

La définition de la permission signifie que α est permise pour w lorsque tous les scénarios accessibles à w suite à l'exécution de α ne contiennent pas de violation V_P. L'action α est interdite pour w lorsque son exécution mène à une violation V_F dans certains scénarios accessibles à w suite à l'exécution de α. La définition de O se comprend à la lumière de l'interprétation que Broersen fait de l'action négative. Comme il le mentionne (Broersen 2004, p.160), le connecteur pour l'action négative s'interprète différemment dépendamment des autres connecteurs avec lesquels il peut se jumeler. En ce sens, la logique du connecteur $\wr^I$ dépend de l'ensemble de connecteurs pour actions qui sont considérés.[25]

Brièvement, Broersen identifie les ensembles de connecteurs suivants, où l'interprétation sémantique de la négation varie d'un ensemble à l'autre.

$$(\cup, \wr^K)$$
$$(^{\leftarrow}, \cup, \wr^B)$$
$$(;, \cup, \wr^4)$$
$$(^*, ;, \cup, \wr^{S4})$$
$$(^{\leftarrow}, ;, \cup, \wr^{B4})$$
$$(^{\leftarrow}, ^*, ;, \cup, \wr^{S5})$$

Autrement dit, lorsque considéré avec l'action disjointe, la relation sémantique sur le modèle pour l'action négative est celle de K, elle est symétrique lorsque l'on ajoute le connecteur pour l'action inverse, elle est transitive lorsque l'on considère l'action disjointe et la séquence d'actions, elle est transitive et réflexive lorsque l'on ajoute a cela l'itération d'actions, elle est symétrique et transitive lorsque l'on considère l'action inverse, la séquence et l'action disjointe, et finalement elle est universelle lorsque l'on ajoute l'itération. Prenant cela en considération, la définition de l'obligation signifie que α est obligatoire pour w lorsque α est permis et que dans tout scénario accessible suite à l'exécution de $\wr^I \alpha$ (où $\wr^I$ est modélisé dépendamment de l'ensemble des connecteurs pour actions de la logique dynamique utilisée) il n'y a pas de violation V_O.

[24] Voir Broersen (2004, pp.164-5) pour les conditions.

[25] Voir Broersen (2004, pp.158-9) pour une brève revue des différentes formalisations de l'action négative.

* * *

Pour conclure, le lecteur intéressé par la logique déontique dynamique est aussi invité à consulter Anglberger (2008), où l'auteur montre que la logique déontique dynamique laisse place à de nouveaux paradoxes, et ce malgré qu'elle soit en mesure d'éviter la majorité des paradoxes de la logique déontique standard. Par ailleurs, le lecteur peut aussi consulter Anglberger (2009) pour un bref survol des différentes réductions proposées pour la logique dynamique (réduction d'un opérateur déontique à une proposition qui ne contient pas d'opérateur déontique) ainsi que pour une réduction alternative, où plusieurs constantes de violations sont introduites et sur lesquelles on impose un ordre partiel afin de rendre compte des alternatives qui sont préférables à d'autres. Étant donné que la logique dynamique provient de l'informatique et que plusieurs informaticiens s'intéressent à la logique déontique, le lecteur ne sera pas surpris de trouver beaucoup d'applications de la logique déontique dymamique à la programmation. Herzig et al. (2011) par exemple augmente le langage de la logique dynamique afin de pouvoir programmer des NMAS. Pour une analyse approfondie des propriétés algébriques des logiques dynamiques ainsi que de leur théorie de la preuve, le lecteur est invité à consulter le chapitre 14 du présent ouvrage.

Chapitre 5

Logique multi-modale

Même si le présent chapitre s'intitule *Logique multi-modale*, celui-ci ne regroupe pas toutes les approches qui tombent sous cette banière. En effet, plusieurs approches en logique STIT (p. ex., Horty 2001, Pacheco et Carmo 2003 et Broersen 2011a) et en logique dynamique (p. ex., Segerberg 2009 et van der Meyden 1996), sont des approches multi-modales qui utilisent différentes modalités primitives. Néanmoins, nous avons préféré regrouper les logiques STIT au chapitre 3 et les logiques dynamiques au chapitre 4 dans la mesure où celles-ci possèdent d'autres caractéristiques distinctives. Le présent chapitre porte sur les approches multi-modales qui ne formalisent pas explicitement l'action. Tel que vu au chapitre 3, la logique déontique peut être considérée comme un outil pertinent à l'analyse des systèmes normatifs et des NMAS. Dans chaque cas, l'objectif est de voir l'impact des normes sur le comportements d'un agent, la différence entre les deux étant que du côté du NMAS, on considère les interactions entre les agents, alors que pour le système normatif (entendu en un sens plus général), on ne considère l'impact des normes que sur un seul agent. Malgré que l'objectif soit de raisonner par rapport à ce que les agents doivent faire, l'objectif principal n'est pas de modéliser l'inférence, mais est bien d'étudier l'influence des normes sur les comportements d'un individu. Dans ce qui suit, nous verrons d'abord l'approche de Carmo et Jones (2002), qui utilisent une myriade de modalités afin de pallier les problèmes et paradoxes de la logique déontique standard. Nous aborderons ensuite brièvement l'approche de Dellunde (2008), qui propose des fondations communes pour les différentes logiques déontiques multi-modales.

José Carmo et Andrew Jones

Suite aux nombreux paradoxes de la logique déontique standard, mais surtout en réaction au paradoxe de Chisholm (1963), les approches en logique déontique se sont complexifiées afin de modéliser certains paramètres mis en lumière par ces problèmes. Même si d'emblée la logique déontique visait l'analyse de l'inférence normative, les travaux des chercheurs ayant un intérêt en informatique et en droit ont fait que la logique déontique est parfois considérée comme un outil pertinent à la modélisation d'un système normatif. Par *système normatif*, Carmo et Jones (2002, p.265) entendent un ensemble d'agents dont le comportement et les interactions sont contraints par des normes. Bien que leur objectif soit de formaliser l'impact des normes sur les agents individuels, ceux-ci n'incluent cependant pas de logique de l'action à leur approche.

L'objectif de Carmo et Jones est de modéliser l'interaction qui se trouve entre un ensemble de normes et le comportement d'un agent. Outre la dimension temporelle implicite à l'évolution des obligations d'un agent exposée par le paradoxe de Chisholm, ce dernier met aussi en évidence qu'une logique déontique qui traite d'un système normatif doit être en mesure de rendre compte du fait que certaines obligations sont parfois violées. De telles situations donnent souvent lieu à des obligations contraires au devoir (*contrary-to-duty obligations*). Afin de rendre compte de ce phénomène mais en laissant de côté la dimension temporelle, Jones et Pörn (1985) ont introduit la notion d'*alternative sous-idéale* (*sub-ideal world*), où certaines obligations ne sont pas réalisées.

Cela dit, en plus de mettre en évidence que les obligations d'un agent évoluent dans le temps et varient lorsque certaines obligations sont violées, le paradoxe de Chisholm montre aussi que certaines obligations sont conditionnelles. Or, les obligations conditionnelles, qui sont souvent traitées à l'aide d'un opérateur dyadique, font face au problème du détachement et à celui de la définition de l'opérateur dyadique.[1] Plus encore, il y a la question de déterminer quels sont les liens qui se trouvent entre les obligations conditionnelles et les obligations actuelles, et quelles sont les règles qui gouvernent le passage d'une obligation conditionnelle à une obligation actuelle.

Afin de répondre à ces problèmes, Carmo et Jones (2002) ont développé une logique déontique multi-modale conformément à la notion d'alternative sous-idéale introduite par Jones et Pörn (1985).[2] Leur système vise à rendre compte de sept principales caractéristiques mises en évidence par le paradoxe de Chisholm. Rappelons nous brièvement le paradoxe (cf. chapitre 2):

$$OA \tag{1}$$
$$O(A \supset B) \tag{2}$$
$$\neg A \supset O\neg B \tag{3}$$
$$\neg A \tag{4}$$

Parmi les caractéristiques mises de l'avant par Carmo et Jones (2002, pp.275-7), les plus importantes sont:

1. les propositions 1–4 sont cohérentes dans la langue naturelle;
2. les propositions 1–4 sont indépendantes dans la langue naturelle;
3. il est possible de dériver des obligations actuelles;

[1] Voir le chapitre 15 du présent ouvrage.

[2] Voir aussi Jones et Pörn (1986) et Jones (1991). Voir Peterson (2011, p.45) pour une synthèse de leur approche.

4. il est possible de dériver des obligations idéales;

5. les obligations sont parfois violées, ce qui donne lieu à des alternatives sous-idéales.

Principalement en raison du fait que le paradoxe de Chisholm peut survenir même dans un contexte où il n'y a aucune considération temporelle (Carmo et Jones 2002, p.275), les auteurs proposent une logique multi-modale sans modalité temporelle. À la base, l'idée est de distinguer entre les obligations *idéales* et les obligations *actuelles* d'un agent, ce qui a une incidence sur la distinction entre les contextes qui dépendent des agents versus ceux qui n'en dépendent pas.

Du côté de la syntaxe, le langage

$$\mathcal{L} = \{Prop, (,), \neg, \supset, \wedge, \vee, \equiv, O(/), O_a, O_i, \Box, \blacksquare\}$$

contient les connecteurs logiques usuels, un ensemble dénombrable de variables propositionnelles $Prop = \{\bot, \top, p_1, ..., p_n, ...\}$ ainsi que les opérateurs $O(/)$, O_a, O_i, $\Box$ et $\blacksquare$ (Carmo et Jones 2002, p.289). Les opérateurs $\Diamond$ et $\blacklozenge$ sont respectivement les opérateurs duaux de $\Box$ et $\blacksquare$, définis par:

$$\Diamond A =_{def} \neg\Box\neg A \quad \text{(def. } \Diamond\text{)}$$
$$\blacklozenge A =_{def} \neg\blacksquare\neg A \quad \text{(def. } \blacklozenge\text{)}$$

L'ensemble des expressions bien formées (EBF) n'est pas explicitement défini mais les auteurs mentionnent que celui-ci est défini de manière habituelle. Considérant que l'axiome (A9) admet la combinaison de différentes modalités, cela suggère que l'itération et la combinaison des opérateurs est possible, et donc que l'ensemble peut être défini récursivement par:

1. si $p \in Prop$, alors $p \in EBF$;

2. si $A, B \in EBF$, alors $\neg A, A \supset B, A \wedge B, A \vee B, A \equiv B \in EBF$;

3. si $A, B \in EBF$, alors $O(A/B), O_a A, O_i A, \Box A, \blacksquare A \in EBF$.

En prenant pour acquis les axiomes de la logique propositionnelle classique, les schémas d'axiomes proposés sont (Carmo et Jones 2002, p.293):

$$
\begin{aligned}
&KM \text{ pour } \Box && \text{(A1)}\\
&KD \text{ pour } \blacksquare && \text{(A2)}\\
&\Box A \supset \blacksquare A && \text{(A3)}\\
&\neg O(\bot/C) && \text{(A4)}\\
&(O(A/C) \wedge O(B/C)) \supset O(A \wedge B/C) && \text{(A5)}\\
&O(A/C) \supset O(A/A \wedge C) && \text{(A6)}\\
&\Diamond O(A/C) \supset \Box O(A/C) && \text{(A7)}\\
&(\Diamond(A \wedge (B \wedge C)) \wedge O(A/C)) \supset O(A/C \wedge B) && \text{(A8)}\\
&(O_i A \wedge O_i B) \supset O_i(A \wedge B) && \text{(A9)}\\
&(O_a A \wedge O_a B) \supset O_a(A \wedge B) && \text{(A10)}\\
&\blacksquare A \supset (\neg O_a A \wedge \neg O_a \neg A) && \text{(A11)}\\
&\Box A \supset (\neg O_i A \wedge \neg O_i \neg A) && \text{(A12)}\\
&\blacksquare(A \equiv B) \supset (O_a A \equiv O_a B) && \text{(A13)}\\
&\Box(A \equiv B) \supset (O_i A \equiv O_i B) && \text{(A14)}\\
&[O(A/C) \wedge (\blacksquare C \wedge (\blacklozenge A \wedge \blacklozenge \neg A))] \supset O_a A && \text{(A15)}\\
&[O(A/C) \wedge (\Box C \wedge (\Diamond A \wedge \Diamond \neg A))] \supset O_i A && \text{(A16)}\\
&O(A/C) \wedge (\blacklozenge(A \wedge C) \wedge (\blacklozenge(\neg A \wedge C))) \supset O_a(C \supset A) && \text{(A17)}\\
&O(A/C) \wedge (\Diamond(A \wedge C) \wedge (\Diamond(\neg A \wedge C))) \supset O_i(C \supset A) && \text{(A18)}
\end{aligned}
$$

Les axiomes (A1) et (A2) signifient que les modalités $\Box$ et $\blacksquare$ sont axiomatisés respectivement par les systèmes modaux KM et KD. L'opérateur $\Box$ représente une forme de nécessité qu'un agent ne peut éviter (Carmo et Jones 2002, p.287). Autrement dit, $\Box A$ signifie que A est nécessairement vrai, indépendamment des choix ou des actions d'un agent. De la même manière, $\Diamond A$ signifie que A a le potentiel d'être réalisé (A n'est pas fixé actuellement). L'opérateur $\blacksquare$ représente ce qui est nécessaire relativement aux choix et aux actions posées par l'agent. Il s'agit de ce qui est nécessaire en vertu de ce qui est fixé actuellement. Quant à $\blacklozenge$, cela représente la possibilité actuelle. Par exemple, il peut être physiquement possible pour un agent de sauver une personne de la noyade ($\Diamond A$) sans pour autant qu'il soit actuellement possible de le faire ($\neg\blacklozenge A$), comme dans un cas où l'agent ne sait pas nager. Informellement, l'opérateur $\blacksquare$ restreint l'ensemble de choix possibles pour un agent relativement au contexte et à ses capacités.

L'axiome (A3) signifie que ce qui est nécessaire est aussi nécessaire actuellement. La contraposée de cet axiome permet de mieux en saisir le sens: ce qui est possible d'être fait actuellement est possible. En ce sens, la notion de possibilité exprimée par $\Diamond$ est plus large que celle exprimée par $\blacklozenge$. S'il est possible actuellement de faire A, alors il est logiquement, physiquement et temporellement possible de faire A.

Dans le même ordre d'idées, l'opérateur O_i représente ce qui est *idéalement* obligatoire, c'est-à-dire ce qui est obligatoire dans toutes les alternatives idéales à un scénario.

De l'autre côté, O_a représente ce qui est *actuellement* obligatoire, et qui peut être le fruit de la violation d'une obligation idéale. L'opérateur dyadique $O(A/C)$ se lit « dans les circonstances C, A est obligatoire », et l'axiome (A4) signifie que dans toutes circonstances C, il est faux que l'absurde (l'impossible, le faux) est obligatoire. (A5) est le principe d'agrégation pour les obligations dues à un même contexte: si A est obligatoire dans le contexte C et que B l'est aussi, alors la conjonction de A et B est aussi obligatoire dans ce contexte. (A6) stipule que si A est obligatoire conditionnellement au contexte C, alors A est aussi obligatoire dans un contexte où la conjonction $C \wedge A$ est vraie.

L'axiome (A7) gère la relation entre la nécessité et l'obligation conditionnelle. Sémantiquement, l'axiome exprime que s'il est possible que A soit obligatoire dans le contexte C pour un scénario w, alors il est vrai que A est une obligation conditionnelle à C dans les alternatives possibles à w. En d'autres termes, si $O(A/C)$ est vrai pour un scénario accessible, alors $O(A/C)$ est vrai pour tous les scénarios accessibles. Quant à l'axiome (A8), cela signifie que s'il est possible que trois évènements A, B et C se produisent conjointement, alors si A est obligatoire dans les conditions C, A est aussi obligatoire dans les conditions $C \wedge B$.

Les axiomes (A9) et (A10) représentent respectivement l'agrégation pour les obligations idéales et les obligations actuelles. Les axiomes (A11) et (A12) posent que la nécessité influence ce qui est obligatoire. (A11) indique que si A est actuellement nécessaire, alors ni A ni $\neg A$ n'est actuellement obligatoire. De même pour l'obligation idéale: si A est nécessaire, alors ni A ni $\neg A$ n'est idéalement obligatoire. Ces axiomes expriment qu'il doit être possible pour un agent d'agir à l'encontre de ses obligations. Il s'agit du principe de contingence.

(A13) et (A14) expriment que les équivalences peuvent être substituées au sein des obligations. (A13) signifie que si A et B sont équivalent dans un contexte, alors A est actuellement obligatoire si et seulement si B l'est aussi. Par exemple, si dans un contexte « Paul sauve la vie de Pierre si et seulement si Paul conduit à 150 km/h », alors si Paul dans ce contexte a l'obligation actuelle de sauver la vie de Pierre, il aura aussi l'obligation actuelle de conduire à 150 km/h. La même interprétation s'applique pour (A14): s'il est nécessaire que A est vrai si et seulement si B est vrai, alors A est une obligation idéale si et seulement si B en est aussi une. Notons que la nécessité actuelle rend *nécessairement actuellement* vraies des propositions qui ne sont pas *nécessairement* vraies. L'opérateur $\blacksquare$ restreint les interprétations où une proposition est vraie à un certain contexte.

L'axiome (A15) quant à lui signifie que si A est une obligation conditionnelle à un contexte C, alors si C est nécessairement actuellement vrai (les interprétations sont restreintes aux scénarios où C est vrai), qu'il est actuellement possible que A soit vrai et qu'il est actuellement possible que A soit faux, alors A est actuellement obligatoire. Autrement dit, si A est une obligation conditionnelle à C et que le contexte C est nécessairement vrai dans les circonstances actuelles (il est actuellement impossible que C soit faux), alors si un agent a la possibilité actuelle de faire A ou de ne pas faire A, et donc que l'agent n'est pas contraint dans le contexte à ne pas faire A, alors A est actuellement obligatoire. La condition $\blacklozenge(\neg A \wedge C)$ vise à assurer que l'agent a la possibilité d'agir à l'encontre de ses obligations. La même lecture s'applique pour (A18) quant à l'obligation idéale et à la nécessité.

La présentation axiomatique est accompagnée de quelques règles d'inférences usuelles, incluant le modus ponens et les deux règles de substitution suivantes, qui permettent de rendre compte des équivalences logiques du point de vue de l'antécédent (REA) et du conséquent (REC).

$$\frac{\vdash B \equiv C}{\vdash O(A/B) \equiv O(A/C)} \qquad \text{(REA)}$$

$$\frac{\vdash C \supset (A \equiv B)}{\vdash O(A/C) \equiv O(B/C)} \qquad \text{(REC)}$$

En ce qui à trait à la sémantique, Carmo et Jones (2002, p.289) définissent un modèle $\mathcal{M} = \langle W, av, pv, ob, V \rangle$, où $W \neq \varnothing$ est un ensemble de scénarios possibles et $V : Prop \longrightarrow \wp(W)$ une fonction qui détermine l'ensemble des scénarios pour lesquels une proposition atomique est vraie. La fonction $av : W \longrightarrow \wp(W)$, où $av \neq \varnothing$, détermine l'ensemble des alternatives actuelles à un scénario w (cf. Carmo et Jones 2002, p.287), lesquelles prennent en compte le contexte.[3] Dans le même ordre d'idées, la fonction $pv : W \longrightarrow \wp(W)$ détermine les alternatives potentielles à w (Carmo et Jones 2002, p.288). La fonction pv est restreinte par les clauses $av(w) \subseteq pv(w)$ et $w \in pv(w)$. Autrement dit, toute alternative actuelle est une alternative potentielle (cf. axiome A3) et tout scénario w est une alternative potentielle à lui même (cf. KM). Cependant, w n'est pas nécessairement une alternative actuelle à lui-même puisque $\blacksquare$ est axiomatisé par KD, et de fait la relation induite sur le modèle sémantique n'est pas réflexive pour cet opérateur. Finalement, la fonction $ob : \wp(W) \rightarrow \wp(\wp(W))$ sélectionne les scénarios dans lesquels une proposition est obligatoire relativement à un certain contexte (Carmo et Jones 2002, p.286). Elle prend un ensemble de scénarios et l'associe à un ensemble d'ensembles de sécnarios. Si $Y \in ob(X)$, alors Y est un ensemble de scénarios qui sont obligatoires dans le contexte X. La fonction est restreinte par quatre conditions (Carmo et Jones 2002, p.290), où $X, Y, Z \in \wp(W)$:

1. $\varnothing \notin ob(X)$;
2. si $Y \cap X = Z \cap X$, alors $Y \in ob(X)$ ssi $Z \in ob(X)$;
3. si $Y, Z \in ob(X)$, alors $Y \cap Z \in ob(X)$;
4. si $Y \subseteq X$, $Y \in ob(X)$ et $X \subseteq Z$, alors $((Z - X) \cup Y) \in ob(Z)$.

[3] Tel que mentionné dans Peterson (2011, p.58, note 10), il y a un fort parallèle entre l'opérateur O_δ, introduit par Jones (1991) afin de rendre compte des obligations conditionnelles et le traitement de l'obligation conditionnelle que fait Chellas (1974). On trouve ici le même parallèle en ce qui concerne les clauses sémantiques pour $\Box$ et $\blacksquare$.

Carmo et Jones (2002, p.291) mentionnent que la première condition indique qu'une contradiction ne peut pas être obligatoire (cf. axiome A4): l'absurde n'est pas obligatoire dans un contexte exprimé par un scénario membre de X. En fait, la condition exprime que la relation pour l'obligation conditionnelle est sérielle (rappelons-nous que dans KD il y a une équivalence entre **D** et $\neg O\bot$): pour n'importe quel contexte exprimé par un scénario membre de X, il y a minimalement une alternative obligatoire, c'est-à-dire que l'ensemble qui contient les ensembles qui contiennent les scénarios obligatoires dans le contexte X n'est pas vide. La seconde condition vise à assurer que si deux propositions sont équivalentes dans un contexte, alors l'une est obligatoire dans le contexte X si et seulement si l'autre l'est aussi (cf. REC et Carmo et Jones 2002, p.291). Cela se comprend par le fait qu'un scénario est un ensemble de propositions, lesquelles décrivent un état de choses dans le monde. À supposer que l'intersection des ensembles X et Y est égale à celle de X et Z, et donc que chacune des intersections *décrit la même chose*, alors la description faite par Y est membre de l'ensemble qui contient les scénarios obligatoires relativement au contexte X si et seulement si celle faite par Z l'est aussi. La troisième condition correspond au principe d'agrégation et signifie que si deux propositions sont obligatoires dans un contexte X, alors leur conjonction l'est aussi (cf. A5, A9 et A10). Finalement, dans les mots de Carmo et Jones (2002, p.292), la dernière condition signifie que si Y est sous-ensemble de X, que X est sous-ensemble d'un contexte Z et que Y est obligatoire dans le contexte X, alors ce qui est obligatoire dans le contexte Z se trouve soit dans Y ou dans la partie de Z qui ne contient pas X. En quelque sorte, l'idée est d'isoler l'ensemble de scénarios obligatoires Y relativement à un contexte plus large X, où le complément de Y par rapport à X n'est pas nécessairement obligatoire, et de dire que dans un contexte encore plus large Z, ce qui est obligatoire dans le contexte Z est soit dans Y ou dans le complément de X par rapport à Z.

La vérité dans le modèle sémantique est définie récursivement de manière usuelle (Carmo et Jones 2002, p.290). Soit $\|A\| = \{w \in W : \models_{\mathcal{M},w} A\}$ l'ensemble qui contient les scénarios où la proposition A est vraie pour le modèle $\mathcal{M}$.[4]

$$\models_{\mathcal{M},w} p \Leftrightarrow w \in V(p) \tag{1}$$
$$\models_{\mathcal{M},w} \neg A \Leftrightarrow \not\models_{\mathcal{M},w} A \tag{2}$$
$$\models_{\mathcal{M},w} A \wedge B \Leftrightarrow \models_{\mathcal{M},w} A \text{ et } \models_{\mathcal{M},w} B \tag{3}$$
$$\models_{\mathcal{M},w} A \vee B \Leftrightarrow \models_{\mathcal{M},w} A \text{ ou } \models_{\mathcal{M},w} B \tag{4}$$
$$\models_{\mathcal{M},w} A \supset B \Leftrightarrow \not\models_{\mathcal{M},w} A \text{ ou } \models_{\mathcal{M},w} B \tag{5}$$
$$\models_{\mathcal{M},w} A \equiv B \Leftrightarrow (\models_{\mathcal{M},w} A \text{ et } \models_{\mathcal{M},w} B) \text{ ou } (\not\models_{\mathcal{M},w} A \text{ et } \not\models_{\mathcal{M},w} B) \tag{6}$$
$$\models_{\mathcal{M},w} \blacksquare A \Leftrightarrow av(w) \subseteq \|A\| \tag{7}$$
$$\models_{\mathcal{M},w} \square A \Leftrightarrow pv(w) \subseteq \|A\| \tag{8}$$

[4] Nous avons explicité les conditions sémantiques des connecteurs propositionnels, laissées implicites chez Carmo et Jones.

$$\models_{\mathcal{M},w} O(A/C) \Leftrightarrow \|A\| \cap \|C\| \neq \varnothing \text{ et}$$
$$X \subseteq \|C\| \text{ et } X \cap \|A\| \neq \varnothing \Rightarrow \ \|A\| \in ob(X) \tag{9}$$
$$\models_{\mathcal{M},w} O_a A \Leftrightarrow \|A\| \in ob(av(w)) \text{ et } av(w) \cap \|\neg A\| \neq \varnothing \tag{10}$$
$$\models_{\mathcal{M},w} O_i A \Leftrightarrow \|A\| \in ob(pv(w)) \text{ et } pv(w) \cap \|\neg A\| \neq \varnothing \tag{11}$$

Comme à l'habitude, A est vrai pour un modèle $\mathcal{M}$ à condition que $\models_{\mathcal{M},w} A$ pour tout w, et A est valide lorsque A est vrai pour tout modèle.

Les six premières clauses sont les conditions habituelles pour les connecteurs classique. Les clauses (7) et (8) signifient respectivement que $\blacksquare A$ est vrai à condition que l'ensemble qui contient les alternatives *actuelles* à w soit inclus dans l'ensemble qui contient les scénarios où A est vrai, et $\Box A$ est vrai à condition que l'ensemble qui contient les alternatives *potentielles* à w est inclus dans l'ensemble qui contient les scénarios où A est vrai. La clause (9) signifie que la proposition « A est obligatoire dans les conditions C » est vraie si et seulement si il y a au moins un scénario où à la fois A est vrai et C est vrai et que si X est un ensemble de scénarios où C est vrai et qu'il y a au moins un scénario membre de X où A est vrai, alors l'ensemble qui contient les scénarios où A est vrai est obligatoire dans les conditions exprimées par les scénarios de X. Autrement dit, $O(A/C)$ est vrai pour w à condition qu'il y ait au moins un scénario v tel que $A, C \in v$ et que s'il y a au moins un scénario v' membre de $X \cap \|A\|$ (où X contient des scénarios où la condition C est vraie), alors tout scénario où A est vrai est obligatoire dans un contexte où les membres de X sont vrais. En isolant un ensemble $X \subseteq \|C\|$, on détermine des scénarios où la proposition C est vraie mais où il n'y a pas d'autres conditions C' qui viennent s'ajouter au scénario et qui pourraient faire changer la valeur de vérité de l'obligation actuelle O_a.

La première partie de la clause (10) indique que A est actuellement obligatoire pour un scénario w si tout scénario où A est vrai est obligatoire dans les conditions v, où v est une alternative actuelle à w. La seconde indique qu'il y a au moins un scénario v' qui est une alternative actuelle à w mais où A est faux pour v'. En ce sens, ce qui est obligatoire n'est pas inévitable. Enfin, la clause (11) stipule qu'il est vrai que A est idéalement obligatoire pour un scénario w si tout scénario où A est vrai est obligatoire dans les conditions v, où v est une alternative potentielle à w, et qu'il y a au moins un scénario v' tel que v' est une alternative potentielle à w et A est faux pour v'. La seconde partie de cette clause exprime, à l'instar de la clause (10), que ce qui est idéalement obligatoire n'est pas inévitable. La seconde partie des clauses (10) et (11) se comprend en fonction de l'argument de Jones et Pörn (1985, p.279): puisqu'il est impossible d'agir à l'encontre d'une tautologie, il s'ensuit que les tautologies ne sont pas obligatoires.

Somme toute, le système proposé par Carmo et Jones vise à rendre compte des obligations conditionnelles dans un cadre qui admet des situations non idéales, où certaines obligations sont enfreintes. L'idée de base est d'être en mesure de déterminer ce qu'un agent devrait faire, et ce même dans des situations où le comportement de l'agent n'est pas conforme à celui prescrit par les normes. Le lecteur intéressé par l'approche de Carmo et Jones trouvera une critique et une réduction de leur système dans l'article de Gabbay et Schlechta (2010), ainsi que la preuve de complétude et de la décidabilité du système dans Carmo et Jones (2013).

Pilar Dellunde

La majorité des approches qui visent à modéliser l'interaction entre un système normatif et un système multi-agents utilisent une variante ou une autre des logiques modales. Cela dit, trois logiques importantes utilisées afin de modéliser l'évolution d'un système dans le temps (en informatique), de par la représentation du passage d'un état à un autre, sont les logiques dynamiques (cf. chapitre 4), **ATL** (*Alternating-Time Temporal Logic*) et **CTL** *Computational Tree Logic* (cf. Dellunde 2008, p.265).

L'approche de Dellunde (2008) vise à construire une logique *générale* qui permet de rendre compte de plusieurs logiques différentes. Ainsi, plutôt que de définir une logique déontique spécifique, Dellunde présente un cadre de travail général pouvant servir de fondation pour la construction de différentes logiques multi-modales. L'idée est d'obtenir des fondements unificateurs permettant d'englober différentes approches qui partagent des caractéristiques communes. En ce sens, Dellunde définit un système qui englobe une classe de systèmes spécifiques, lesquels varieront dépendamment des axiomes (ou règles) ajoutés.

Plutôt que de présenter l'analyse proposée par Dellunde en détails, ce qui recouperait plusieurs chapitres vus jusqu'à maintenant, nous allons seulement présenter le matériel nécessaire à la compréhension du système général permettant la représentation et la comparaison de différentes logiques multi-modales. Nous avons jugé bon de présenter l'approche de Dellunde considérant, d'une part, que cela offre un cadre commun à partir duquel différentes logiques peuvent être analysées, et que d'autre part, l'auteur propose une interprétation d'un système normatif différente de ce que l'on retrouve usuellement dans la littérature.

L'objectif de Dellunde est de définir un cadre de travail commun pour les logiques qui visent à formaliser la structure d'un système normatif (Dellunde 2008, p.262). Cependant, contrairement à la définition usuelle d'un système normatif, Dellunde (2008, p.261) conçoit un système normatif η comme un *ensemble de contraintes* restreignant la transition d'un état à un autre dans un système.

Plus précisément, à supposer une structure $\langle U, R\rangle$, où U est un ensemble d'états et R un ensemble de paires ordonnées (voire une relation, $R \subseteq U \times U$), $\eta \subseteq R$ contient les transitions interdites pour un système donné. En ce sens, si $(w, v) \in \eta$ (ou encore $wR_\eta v$), alors la transition de l'état w à l'état v est interdite. (Soulignons que dans les approches qui sont plutôt axées vers l'informatique, nous parlons d'*états* (statiques) plutôt que de *scénarios* considérant qu'il y a toujours une forme ou une autre de logique temporelle en arrière plan.) En supposant une logique L visant à modéliser un système multi-agent, Dellunde (2008, p.262) considère la logique d'un système normatif L_η comme l'augmentation de L à l'aide d'un ensemble de systèmes normatifs (i.e., un ensemble d'ensembles de contraintes).

Même si le langage développé par Dellunde (2008, p.264) peut, dans une certaine mesure, être compris comme un *méta-langage* (il s'agit d'un langage général qui permet de parler de plusieurs autres langages plus spécifiques), il faut garder à l'esprit que l'objectif est de développer un cadre de travail abstrait permettant de voir sous le même angle des systèmes qui sont fondamentalement différents. Soit $\tau = \langle F, \rho \rangle$ un type qui contient un ensemble de modalités F pour lesquelles $\rho : F \longrightarrow \mathbb{N}$ assigne un nombre naturel à chaque modalité (et T l'ensemble de tous les types). Un langage $\mathcal{L}$ pour un type τ est composé des connecteurs logiques (propositionnels classiques) usuels, des symboles $\top$ et $\bot$ (qui représentent respectivement une tautologie et une contradiction arbitraire) et un nombre fini de modalités $\Box_1, ..., \Box_n$.

Rappelons-nous qu'une logique peut être définie comme « le plus petit ensemble de formules bien formées étant fermé sous les règles $r_1, ..., r_k$ et contenant les axiomes (A1),...,(A_n) ». Lorsque conçue ainsi, on considère un ensemble d'axiomes Γ avec qu'une relation de conséquence $\vdash$ (ou encore Cn) qui répond aux règles d'inférence désirées, et on définit la logique (i.e., l'ensemble des théorèmes) comme $Cn(\Gamma)$, soit l'ensemble qui contient toutes les conséquences logiques de Γ. Autrement dit, $Cn(\Gamma)$ est la logique ayant comme axiomes Γ et comme règles d'inférence les conditions sur Cn.

L'objectif de Dellunde est de proposer une logique contenant le plus petit ensemble d'axiomes (communs à toutes logiques modales) possibles, En ce sens, Dellunde (2008, p.264) définit une relation de conséquence $\vdash_\tau$ relativement à un type τ (i.e., relativement à un langage; un langage contient un nombre fini de modalités et Dellunde définit la logique pour tout langage modal). La relation de conséquence se résume de la manière suivante:

1. $\vdash_\tau$ contient les axiomes de la logique propositionnelle classique;
2. $\vdash_\tau$ est fermée sous le modus ponens;
3. $\vdash_\tau$ est compacte;
4. $\vdash_\tau$ est transitive;
5. $\vdash_\tau$ est monotone;
6. $\vdash_\tau$ est fermée sous une règle de substitution pour les propositions équivalentes;
7. $\vdash_\tau$ est fermée sous la règle (RE) pour chaque modalité $\Box_i$ du langage;
8. toute proposition membre d'un ensemble de propositions est la conséquence de cet ensemble.

Malgré que Dellunde (2008, p.265), faisant appel à l'ouvrage de Gabbay (1994), nomme les logiques qui répondent à ces critères des *logiques modales classiques*, il s'agit en fait de la définition proposée par Chellas (1980, p.231). Notons que la définition est plus succincte chez Chellas: un système *classique* est fermé sous (RE) et possède la définition de l'opérateur dual par $\Box =_{def} \neg\Diamond\neg$ (cette dernière condition est introduite par la suite chez Dellunde).

$$\frac{\vdash A \equiv B}{\vdash \Box A \equiv \Box B} \qquad \text{(RE)}$$

Notons qu'un système classique ne doit pas être confondu avec un système *normal*. En général, le système K est la logique modale utilisée comme point de référence dans la littérature pour quiconque veut modéliser un phénomène avec une logique modale. Cela dit, K est un système *normal* (cf. Chellas 1980, p.114). Un système *normal*, plutôt que d'être fermé sous (RE), est fermé sous (RK).

$$\frac{\vdash (A_1 \wedge \cdots \wedge A_n) \supset B}{\vdash (\Box A_1 \wedge \cdots \wedge \Box A_n) \supset \Box B} \qquad \text{(RK) avec } n \geq 0$$

Au même titre que KD est une extension de K, les systèmes normaux sont des extensions (conservatrices) des systèmes classiques. En ce sens, le système classique est plus *faible* que le système normal, c'est-à-dire qu'il permet de prouver moins de théorèmes.

Les conditions 2–6 représentent la notion de conséquence pour la logique propositionnelle classique, que l'on obtient en ajoutant la condition 1. La condition 3 (compacité) signifie que si une proposition est la conséquence d'un ensemble de propositions Γ, alors elle est la conséquence d'un sous-ensemble fini de Γ. Autrement dit, s'il existe une dérivation de A à partir de Γ, alors il existe une dérivation de A à partir d'un sous-ensemble fini de Γ (ou de manière similaire, si A est prouvable à partir de Γ, alors il existe une dérivation de longueur finie de A à partir de Γ). La condition 7 fait appel à la règle (RE), que l'on retrouve dans la définition de Chellas.

$$\frac{\vdash A \equiv B}{\vdash \Box_i A \equiv \Box_i B} \qquad \text{(RE)}$$

Finalement, 5 signifie que si A est la conséquence de Γ et que $\Gamma \subseteq \Delta$, alors A est aussi la conséquence de Δ.

Nous avons donc la logique $L = \{L_\tau : \tau \in T\}$ où $L_\tau = \{A : \vdash_\tau A\}$. En d'autres termes, L est la logique multi-modale qui contient toutes les logiques multi-modales de type τ.

En supposant un type τ et une structure $S = \langle U, R\rangle$ (avec $U \neq \varnothing$ et un ensemble de relations d'accessibilité pour chaque modalité $R = \{R_f : f \in F\}$), Dellunde (2008, p.265) définit un système normatif η comme un sous-ensemble de l'ensemble:

$$U* = \{(w_1, ..., w_n) : \text{pour tout } i < n \text{ il y a } f \text{ t.q. } w_i R_f w_{i+1}\}$$

Autrement dit, $U*$ est un élément de $\wp(U)$ qui contient des séquences finies d'états pour lesquelles chaque w_{i+1} est accessible à son prédécesseur w_i par le biais d'une relation R_f pour une modalité f donnée. L'ensemble $U*$ contient donc des séquences $w_1, \ldots, w_n$ où chaque état w_{i+1} est un successeur à son prédecesseur w_i en vertu d'une modalité (et de la clause sémantique qui l'accompagne) du langage.

Ainsi, étant un sous-ensemble de $U*$, un système normatif isole les séquences finies où chaque transition de w_i à w_{i+1} est interdite. Le système normatif stipule qu'une séquence d'accessibilité est interdite, peu importe le chemin par lequel on passe de w_1 à w_n. En ce sens, $U*$ contient les séquences $(w_1, ..., w_n)$ pour lesquelles toute sous-séquence binaire (w_i, w_{i+1}) est interdite. Un système normatif est un ensemble qui contient des séquences finies à l'intérieur desquelles chaque w_{i+1} est accessible à l'état qui le précède w_i de par une relation R_f quelconque, mais où le passage de w_i à w_{i+1} est interdit.

À partir d'un type τ, le langage est augmenté afin d'obtenir un type τ^η, où $\eta \in N$ est élément d'un ensemble de symboles dénotant des systèmes normatifs. Le langage est alors augmenté par les modalités $\Box_1^\eta, ..., \Box_n^\eta$, où $\Box_i^\eta A$ se lit « A est obligatoire selon le système normatif η (cf. Dellunde 2008, p.261) ». En ce sens, A appartient à tout scénario accessible à w, c'est-à-dire que les séquences qui admettent la transition d'un scénario w à un scénario où A est faux sont interdites.

En un mot, l'idée est de prendre une logique multi-modale L et de créer une extension conservatrice L^η à l'intérieur de laquelle la relation d'accessibilité $R_f^\eta \subseteq R_f$ détermine les transitions interdites, où le langage de L^η est celui de L augmenté par τ^η. Autrement dit, pour chaque $L_\tau = \{A : \vdash_\tau A\}$ on définit $L_{\tau^\eta} = \{A : \vdash_{\tau^\eta} A\}$ afin d'obtenir $L^\eta = \{L_{\tau^\eta} : \tau^\eta \in T^\eta\}$, où T^η est l'ensemble qui contient tous les types et types augmentés (Dellunde 2008, p.265). La relation de conséquence pour $\vdash_{\tau^\eta}$ est définie de la même manière, seul le langage est augmenté. De fait, nous obtenons une logique multi-modale pour les systèmes normatifs capable de rendre compte d'une diversité de logiques spécifiques qui seront obtenues en ajoutant des conditions sur Cn et des axiomes.

Dans la suite de son article, Dellunde (2008, pp.267-8) concentrera son attention sur les systèmes normatifs *élémentaires*, à savoir ceux dont la structure (et plus particulièrement les relations d'accessibilité) peut être décrite par une logique de premier ordre monadique. Le lecteur intéressé par les formules de Sahqlvist trouvera chez Dellunde (2008, p.268) une définition précise.

* * *

Tout compte fait, les logiques multi-modales offrent un cadre de travail riche permettant une analyse détaillée du discours. Bien qu'il se trouve une panoplie d'approches multi-modales en logique déontique, celles-ci utilisent toutes les différentes relations qui peuvent être définies sur un modèle sémantique afin de modéliser le langage et les raisonnements. Dans chaque cas, l'idée est de représenter sémantiquement une modalité en définissant des relations qui permettent d'isoler les scénarios qui seront accessibles suite à son instanciation. Le présent chapitre clôt pour le moment la présentation des logiques déontiques *modales*, auxquelles nous retournerons brièvement à la fin de cette partie au chapitre 8. Pour l'heure, nous allons tourner notre attention à deux autres types d'approches, notamment les logiques I/O et les approches algébriques.

Chapitre 6

Logique input/output

Le présent chapitre porte sur les logiques input-output, que nous nommons les logiques I/O. Il regroupe trois approches, incluant la présentation des logiques I/O comme on la retrouve chez Makinson et van der Torre (2000). Même si la seconde approche ne porte pas sur la logique I/O, nous avons préféré présenter le texte de Jones et Sergot (1996) au sein de ce chapitre puisque celle-ci a été incorporée à la logique I/O dans l'approche de Boella et van der Torre (2006c), que nous aborderons en troisième lieu.

David Makinson et Leendert van der Torre

Les logiques I/O offrent un cadre de travail différent de celui des logiques modales et sont apparues chez Makinson et van der Torre (2000). Plutôt que de tenter de développer une logique déontique qui soit en mesure de rendre compte de l'inférence normative dans sa totalité, ceux-ci introduisent les logiques I/O en guise d'*assistants* qui peuvent être utilisés afin d'aider un processus de transformation (Makinson et van der Torre 2000, p.384). Les logiques I/O sont à comprendre en termes d'*entrée* et de *sortie*: on cherche à déterminer certaines règles qui permettent de gouverner le processus de transformation qui part d'un *input* et donne un *output*.

D'entrée de jeu, on considère un langage propositionnel booléen (qui obéit aux règles de la logique propositionnelle classique) à l'aide duquel on définit des paires $(a, x) \in EBF \times EBF$, lesquelles sont interprétées comme des règles qui déterminent le output x relativement à l'input a (Makinson et van der Torre 2001, p.155).[1] Même si le parallèle peut être fait entre une paire (a, x) et un conditionnel $a \supset x$, la paire ne doit pas être entendue au sens de l'implication matérielle. Cela se comprend notamment en raison du fait qu'une norme (exprimée par une paire) est sans valeur de vérité (Makinson et van der Torre 2003b, p.163), sans compter que les règles qui gouvernent les paires diffèrent

[1] Par convention les premières lettres de l'alphabet sont souvent utilisées comme des constantes et les dernières comme des variables, alors que les lettres p, q, r sont utilisées comme variables propositionnelles. Ici, les premières lettres de l'alphabet sont utilisées comme variables propositionnelles pour input et les dernières comme variables propositionnelles pour output. Par ailleurs, EBF ici est l'ensemble des expressions bien formées de la logique propositionnelle classique.

de celles qui régissent l'implication.

Un *ensemble générateur* $G \subseteq EBF \times EBF$ est un ensemble qui contient des paires (a, x) et l'ensemble $G(A) = \{x : (a, x) \in G \text{ pour certains } a \in A\}$ est un ensemble qui contient le outputs x pour certains a éléments d'un ensemble de propositions A relativement à l'ensemble générateur G. Autrement dit, G est un ensemble de règles (de normes) et $G(A)$ est un ensemble de propositions contenant les outputs x pour lesquels $a \in A$ et $(a, x) \in G$.

Makinson et van der Torre (2001, p.156) introduisent l'opération $out(G, A)$, qui, à partir d'un ensemble générateur et d'un ensemble de inputs, donne un ensemble de outputs. Cette opération est utilisée afin de définir quatre opérateurs, qui à leur tour donneront quatre logiques I/O. Soit $Cn(A)$ l'ensemble des conséquences classiques de A. Soulignons que $\top \in out(G, A)$ pour toute tautologie classique. Les opérateurs (qui représentent des opérations sur des ensembles) sont:

$$out_1(G, A) =_{def} Cn(G(Cn(A)))$$

$$out_2(G, A) =_{def} \bigcap_V Cn(G(V)) \text{ t.q. } A \subseteq V \text{ et } V \text{ est complet}$$

$$out_3(G, A) =_{def} \bigcap_B Cn(G(B)) \text{ t.q. } B = Cn(B), A \subseteq B \text{ et } G(B) \subseteq Cn(B)$$

$$out_4(G, A) =_{def} \bigcap_V Cn(G(V)) \text{ t.q. } A \subseteq V, G(V) \subseteq V \text{ et } V \text{ est complet}$$

Examinons chacune de ces définitions une par une.[2] La première définition est telle que l'ensemble qui contient les outputs pour les paires de G formées à partir des inputs de A est l'ensemble qui contient les conséquences logiques de l'ensemble qui contient les paires ayant pour inputs les conséquences logiques de A. Reprenons cette définition à l'envers: $Cn(A)$ représente l'ensemble des conséquences de l'ensemble de propositions A. Autrement dit,

$$Cn(A) = \{x : a \in A \text{ et } a \vdash_{LC} x\}.$$

L'ensemble $G(Cn(A))$ correspond à l'ensemble qui contient les outputs pour lesquels l'ensemble qui contient les conséquences de A sont des inputs. Si par exemple $G = \{(a, x), (a \vee b, y)\}$, alors $G(a) = \{x\}$ et $G(Cn(a)) = \{x, y\}$.[3] Le output de type 1 est par définition l'ensemble qui contient les conséquences logiques de $G(Cn(A))$. Informellement, on entre l'ensemble A comme input, on considère ses conséquences, ensuite on considère le output des conséquences de A (relativement à G) et enfin les conséquences du output des conséquences de A.

La seconde définition requiert l'introduction de la notion d'ensemble *complet*. Un ensemble est *complet* lorsqu'il est soit *maximalement consistant* (i.e., Γ tel que pour tout

[2] À partir de ces définitions, les auteurs introduisent aussi $out_i^+(G, A) = out_i(G^+, A)$ avec $G^+ = G \cup I$ et $I = \{(a, a) : a \in EBF\}$. Nous laissons ces définitions de côté par souci d'économie. Les auteurs notent que $out_2^+ = out_4^+$ et que cela redonne la logique classique.

[3] Afin de faciliter la lecture, les auteurs écrivent $G(a)$ ou $Cn(a)$ plutôt que $G(\{a\})$ ou $Cn(\{a\})$ lorsqu'il s'agit d'un singleton.

$A \in EBF$, soit $A \in \Gamma$ ou $\neg A \in \Gamma$ et $\perp \notin \Gamma$) ou qu'il est équivalent à l'ensemble de toutes les propositions du langage (Makinson et van der Torre 2001, p.156). Conformément au lemme de Lindenbaum, à savoir que tout ensemble consistant possède une extension maximalement consistante, il est possible de construire un nombre dénombrable d'extensions maximalement consistantes pour un ensemble (consistant) de inputs A. Un ensemble complet est un ensemble auquel on ne peut rien ajouter de plus. Dans le cas d'un ensemble maximalement consistant, on ne peut rien ajouter de plus sans briser sa consistance Lorsqu'il s'agit de l'ensemble de tous les énoncés du langage, alors clairement rien ne peut y être ajouté puisque tous les énoncés s'y trouvent. Dans l'éventualité où l'ensemble correspond à tous les énoncés du langage, ce dernier est trivialement inconsistant.

La définition de out_2 considère l'intersection de tous les ensembles des conséquences des outputs qui ont V comme input, où V est une extension complète de A. Par opposition avec out_1 qui ne fait que considérer l'ensemble de inputs, out_2 permet de prendre en considération tout ce qui est commun avec toute extension complète de l'ensemble de inputs. En plus de considérer les outputs de A relativement à G (ainsi que leurs conséquences), on considère aussi les outputs des extensions maximalement consistantes de A. Par exemple, si G contient des paires pour lesquelles le input n'est pas élément de A, alors out_2 permet d'isoler les outputs qui sont consistants avec les outputs de A.

La troisième définition (qui ressemble à la deuxième) fait appel aux conditions $B = Cn(B)$ et $A \subseteq B$, ce qui peut sembler étrange à première vue. En fait, il suffit de remarquer que $EBF = Cn(EBF)$ et que $V = Cn(V)$ pour un ensemble V maximalement consistant. Par surcroît, il est trivial que $Cn(Cn(A)) = Cn(A)$ puisque la conséquence est transitive (i.e., les conséquences des conséquences de A sont des conséquences de A). En ce sens, il est possible de considérer toute extension B de A sans pour autant que B soit complète. Il suffit d'avoir $B = Cn(X)$ avec $A \subseteq X$, ce qui trivialement donne $A \subseteq B$.

La troisième condition de la définition, soit que $G(B) \subseteq Cn(B)$, signifie que l'ensemble des outputs qui ont B comme input (relativement à G) est sous-ensemble de l'ensemble des conséquences de B. De fait, les outputs peuvent être réutilisés comme inputs. Sachant que $G(B) \subseteq Cn(B)$, cela implique que $Cn(G(B)) \subseteq Cn(B)$: B contient toutes ses conséquences et contient $G(B)$, et donc contient toutes les conséquences de $G(B)$. Mais puisque $Cn(G(B)) \subseteq Cn(B)$, il s'ensuit que l'ensemble des conséquences de l'ensemble qui contient les outputs qui ont B comme input fait partie de l'ensemble des conséquences de B. Étant donné que B est considéré comme input dans $G(B)$, il s'ensuit que les outputs de $G(B)$ sont aussi d'emblée considérés en tant qu'inputs. La définition de out_3 donne l'intersection des ensembles de conséquences de outputs qui ont pour inputs B (relativement à G) lorsque B équivaut à l'ensemble de ses conséquences, A est sous-ensemble de B et l'ensemble des outputs qui ont pour input B fait partie de B. Cette dernière condition amène la possibilité de réutiliser les outputs comme inputs.

Une façon peut-être un peu plus simple de voir la chose serait de définir $out_3(G, A) = Cn(G(A*))$ où $A*$ est le plus petit ensemble fermé sous Cn et G tel que $A \subseteq A*$ (Makinson et van der Torre 2000, p.393). Autrement dit, $A*$ est le plus petit ensemble pour lequel A est un sous-ensemble et qui est fermé sous la relation de conséquence et G.

En dernier lieu, la définition de out_4 est à l'image de celle de out_2, à l'exception que l'on permet de reprendre les outputs comme inputs à l'instar de out_3. Il s'agit de prendre l'intersection des conséquences de toute extension complète V de A où l'ensemble des outputs de V est sous-ensemble de V.

Évidemment, les outputs varient en fonction de l'ensemble générateur G. Les auteurs nomment ces définitions respectivement le *simple minded output*, le *basic output*, le *reusable simple minded output* et le *reusable basic output*. Les points cruciaux à souligner sont que les inputs ne se retrouvent pas dans les outputs (sauf pour les out_i^+) et que la contraposition ne s'applique pas dans chacun des cas (Makinson et van der Torre 2000, p.407)

Le lecteur qui consulte les articles portant sur les logiques I/O ne verra peut-être pas tout de suite ce qui distingue les quatre types de output du point de vue de la sémantique, sans compter que certains exemples donnés par les auteurs contribuent à engendrer la confusion. Par exemple, Makinson et van der Torre (2001, p.159) donnent l'exemple suivant.

Exemple 6.1. Soit $G = \{(\top, \neg a), (a, x)\})$. Dans ce cas, $out_i(G, a) = Cn(\{\neg a, x\}$ pour $i \in \{1, 2, 3, 4\}$.

Considérons chacun des outputs.

1. $out_1(G, a) = Cn(G(Cn(a)))$ et puisque $\top, a \in Cn(a)$ il s'ensuit que $G(Cn(a)) = \{\neg a, x\}$ et donc que $out_1(G, a) = Cn(\{\neg a, x\})$.

2. Si l'on prend en compte tout $G(V)$ pour toute extension complète V de a, nous obtiendrons encore $G(V) = \{\neg a, x\}$ pour chaque V, et de fait $out_2(G, a) = \bigcap Cn(G(V)) = \{\neg a, x\}$.

3. $G(B)$ où $B = Cn(B)$ et $\{a\} \subseteq B$ donne aussi seulement $G(B) = \{\neg a, x\}$, donc $out_3(G, a) = \bigcap Cn(G(B_) = \{\neg a, x\}$.

4. Le même raisonnement s'applique que pour 2 et 3.

Mais alors, qu'est-ce qui différencie chacun des outputs? D'abord, il faut insister sur le fait que l'intersection dans les définitions susmentionnées ne porte pas sur V mais bien sur $Cn(G(V))$. Voyons deux exemples afin de mettre en lumière les différences qui se trouvent entre le *simple minded* et le *basic* output, d'un côté, et le *simple minded* et le *reusable simple minded* output de l'autre.

Considérons l'ensemble $G = \{(a, x), (b, x)\}$. Quel est $out_1(G, a \vee b)$ et $out_2(G, a \vee b)$? Alors que $out_1(G, a \vee b) = Cn(G(Cn(a \vee b))) = Cn(\varnothing)$, $out_2(G, a \vee b) = \bigcap Cn(G(V)) = Cn(x)$. En effet, toute extension complète de $\{a \vee b\}$ contient soit a, b ou a et b. De fait, dans chacun des cas, $G(V) = \{x\}$. En ce sens, la sémantique de out_2 permet d'aller chercher ce qui est commun à toutes les extensions complètes de $a \vee b$ du point de vue de leur output relativement à G. La distinction entre out_1 et out_2 se transpose à out_3 et out_4.

Quant à la différence entre out_1 et out_4 (ou encore out_2 et out_3), considérons l'ensemble $G = \{(a,x),(a \wedge x, y)\}$. Quels sont $out_1(G,a)$ et $out_4(G,a)$? D'un côté, $out_1(G,a) = Cn(G(Cn(a))) = Cn(x)$. De l'autre côté, $out_4(G,a) = \bigcap Cn(G(V))$. Puisque $a \in V$, il s'ensuit que $x \in G(V)$. Or, considérant que $G(V) \subseteq V$, il en résulte que $x \in V$, et donc que $y \in G(V)$. De fait, $out_4(G,a) = Cn(\{x,y\})$. Le lecteur est invité à consulter les figures dans Makinson et van der Torre (2000) afin de mieux visualiser les différents types de outputs. Voir Makinson et van der Torre (2000, p.407) pour un résumer des propriétés des logiques I/O.

Les définitions sémantiques des différents types de outputs s'accompagnent du côté de la syntaxe par des règles (Makinson et van der Torre 2001, p.157). (SI) signifie que si x est le output de a et que a est une conséquence logique de b, alors x est le output de b.

$$\text{(SI)}\,\frac{(a,x)}{(b,x)} \qquad \text{(lorsque } b \vdash_{LC} a)$$

(AND) signifie que si x et y sont les outputs de a, alors $x \wedge y$ est le output de a.

$$\text{(AND)}\,\frac{(a,x)\ (a,y)}{(a,x \wedge y)}$$

(WO) stipule que si y est la conséquence de x et que x est le output de a, alors y est le output de a.

$$\text{(WO)}\,\frac{(a,x)}{(a,y)} \qquad \text{(lorsque } x \vdash_{LC} y)$$

(OR) exprime que si x est le output de a et est le output de b, alors c'est aussi le output de $a \vee b$.

$$\text{(OR)}\,\frac{(a,x)\ (b,x)}{(a \vee b, x)}$$

(CT) signifie que si x est le output de a et que y est le output de $a \wedge x$, alors y est le output de a.

$$\text{(CT)}\,\frac{(a,x)\ (a \wedge x, y)}{(a,y)}$$

Conjointement à chaque définition de outputs se joignent des règles de dérivation spécifiques. Soit $deriv_i(G)$ le plus petit ensemble qui contient G, les paires $(\top,\top)$ où $\top$ est une tautologie, et qui est fermé sous les règles faisant parties de R_i.

$$R_1 = \{(SI),(AND),(WO)\}$$
$$R_2 = \{(SI),(AND),(WO),(OR)\}$$
$$R_3 = \{(SI),(AND),(WO),(CT)\}$$
$$R_4 = \{(SI),(AND),(WO),(OR),(CT)\}$$

Les notions de dérivabilité $deriv_1(G) - deriv_4(G)$ correspondent respectivement aux outputs $out_1 - out_4$.[4]

Les logiques I/O ont été développées par Makinson et van der Torre afin de s'appliquer à la logique déontique. Jusqu'à présent, nous avons vu que le discours normatif met en jeu deux formes d'inconsistances, à savoir l'inconsistance propositionnelle et l'inconsistance normative. Le paradoxe de Chisholm met en évidence que tout système qui veut rendre compte des obligations contraires au devoir se doit d'être capable d'exprimer (dans une certaine mesure) les inconsistances normatives. Or, les logiques I/O permettent d'emblée de faire la distinction entre les deux types d'inconsistance. Makinson et van der Torre (2001, p.159) distinguent entre l'inconsistance du output, qui ultimement est une inconsistance propositionnelle, et l'inconsistance entre les inputs et les outputs, qui est une forme d'inconsistance normative. Alors qu'un output de type i est dit inconsistant lorsque $\perp \in out_i(G, A)$, l'inconsistance entre les inputs et les outputs advient lorsque $\perp \in Cn(out_i(G, A) \cup A)$. Voici l'exemple qu'offrent les auteurs afin de mettre en évidence la différence entre les deux types d'inconsistance.

Exemple 6.2. Soit $G = \{(\top, \neg a), (a, x)\}$. Informellement, G affirme l'existence d'une obligation conditionnelle: en général le output $\neg a$ est préféré, mais s'il advient que a est donné, alors le output x sera désiré. Tel que susmentionné, $out_i(G, a) = Cn(\{\neg a, x\})$ pour $i \in \{1, 2, 3, 4\}$. Le output est consistant puisque $\perp \notin Cn(\{\neg a, x\})$. Cependant, le output est inconsistant avec le input puisque $\perp \in Cn(\{\neg a, x\} \cup \{a\})$.

Cela dit, les logiques I/O permettent d'arriver à un résultat étrange: considérant que $a \vdash_{LC} \top$, la règle (SI) nous donne dans l'exemple précédent que $(\top, \neg a)$ entraîne $(a, \neg a)$.[5] Ce résultat semble peu probable, surtout si l'on considère que les obligations contraires au devoir surviennent à un certain moment dans le temps. Par exemple, si a tient pour « Paul vole de l'argent » et x pour « Paul remet l'argent volé », alors on obtient que dans toutes circonstances idéales Paul ne vole pas, que dans une circonstance sous-idéale où Paul vole il remet l'argent volé, mais que dans une circonstance où Paul vole il ne vole pas. Il ne s'agit pas d'une contradiction à proprement parler puisqu'il s'agit d'une paire input/output, mais néanmoins cela pose problème dans un contexte déontique. Plus précisément, une fois que l'action qui ne devait pas être faite a été accomplie, on veut mettre de côté les scénarios où l'action n'était pas faite et se concentrer uniquement sur ceux où elle est accomplie.

La stratégie pour parvenir à ce résultat est d'utiliser une technique propre aux logiques épistémiques:

> Our strategy is to adapt a technique that is well-known in the logic of belief change – cut back the set of norms to just below the threshold of making the current situation contrary-to-duty. In effect, we carry out a contraction on the set G of given norms.

[4] Afin d'obtenir les outputs out_i^+, il suffit d'ajouter la règle (ID) qui permet d'ajouter (a, a) à tout moment (Makinson et van der Torre 2001, p.157).

[5] Un exemple plus complexe est donné dans Makinson et van der Torre (2003b, pp.169-70).

> Specifically, we look at the maximal subsets $G' \subseteq G$ such that $out(G', A)$ is consistent with input A. (Makinson et van der Torre 2003b, p.170)

Pour ce faire, les auteurs introduisent des *contraintes* sur les outputs (Makinson et van der Torre 2001, p.160). Soit C un ensembles de contraintes qui contient des propositions. L'ensemble $maxfamily(G, A, C)$ est défini comme l'ensemble de tous les plus grands $H \subseteq G$ pour lesquels $out(H, A)$ est consistant avec C (i.e., $\perp \notin Cn(out(H, A) \cup C)$). L'ensemble $outfamily(G, A, C)$ est l'ensemble des outputs générés par $maxfamily(G, A, C)$, à savoir que $outfamily(G, A, C) = \{out(H, A) : H \in maxfamily(G, A, C)\}$.

Par exemple (Makinson et van der Torre 2001, p.162), afin de restreindre l'ensemble des normes G en excluant celle qui stipule que Paul ne doit pas voler, on ajoute la contrainte $C = \{a\}$. Alors que dans le cas sans contrainte on obtient $maxfamily(G, a, \varnothing) = \{G\}$ et $outfamily(G, a, \varnothing) = \{Cn(\{\neg a, x\})\}$, l'ajout de la contrainte $C = \{a\}$ permet d'isoler $maxfamily(G, a, a) = \{\{(a, x)\}\}$ puisque

$$\perp \in Cn(out(G, a) \cup \{a\} = Cn(Cn(\{\neg a, x\}) \cup \{a\})$$

et $outfamily(G, a, a) = \{Cn(\{x\})\}$.[6]

À l'aide de l'imposition d'une contrainte sur les différents types de outputs, Makinson et van der Torre (2001, p.166) introduisent la notion de *contrainte* I/O (*input/output constraint*). Pour un output $out(G, A)$, l'ajout de la contrainte I/O correspond à la restriction de l'ensemble des outputs à l'ensemble de inputs $maxfamily(G, A, C)$ pour $C = A$. Ainsi, on s'assure que l'ensemble de outputs est consistant et que ce dernier est consistant avec l'ensemble des inputs.

Les logiques I/O contraintes s'obtiennent d'un point de vue syntaxique en ajoutant la notion de *dérivation contrainte*. Conformément aux quatre types de outputs ainsi qu'à leur notion conjointe de dérivation, les outputs *contraints* s'obtiennent de par leur notion respective de dérivation contrainte, où une dérivation est *contrainte* lorsque $(a, \neg a) \notin deriv_i(L)$ pour tout $L \subseteq G$ (Makinson et van der Torre 2001, p.169). Le lecteur est invité à consulter Makinson et van der Torre (2001, p.177) pour un résumé des résultats sur les logiques I/O contraintes.

Jusqu'à présent, les logiques I/O nous offrent une façon de formuler des obligations qui tiennent dans tous les contextes, p. ex., $(\top, a)$, et des obligations qui sont contraires au devoir, p. ex., $\{(\top, a), (\neg a, x)\}$. Traité de la sorte, un ensemble G est considéré comme un ensemble de normes, lesquelles dictent des obligations (des outputs) dans certains contextes (des inputs). Ainsi, la paire $(\top, a)$ signifie que dans toutes circonstances a doit être alors que $\{(\top, a), (\neg a, x)\}$ signifie que a doit être, mais que si jamais a n'est pas le cas alors x doit être. Les logiques I/O fournissent donc une manière de traiter des obligations contraires au devoir et des obligations conditionnelles.

Mais qu'en est-il de la permission? Makinson et van der Torre (2003a, pp.391-2) assument une distinction entre les permissions *explicites* et *implicites*, ce à quoi on

[6] Pour d'autres exemples, voir Makinson et van der Torre (2001, pp.162-3).

réfère aussi par le biais de permissions *forte* et *faible*. Une permission *implicite*, voire une permission *faible*, correspond à quelque chose qui n'est pas interdit par un ensemble de normes. Par exemple, il est implicitement permis de se baigner en une chaude journée d'été puisque la législation canadienne ne l'interdit pas. Cela dit, le pendant de la permission implicite est la permission *explicite*, voire la permission *forte*: une action est explicitement permise lorsqu'un ensemble de norme le mentionne. Par exemple, il est explicitement permis par la législation d'arrêter son véhicule sur la voie d'accotement d'une autoroute en cas d'urgence, notamment si le véhicule tombe en panne. Une autre façon de nommer ces types de permissions serait de les appeler les permissions *négatives* et *positives*.

D'abord, les auteurs considèrent la permission implicite conditionnelle. Une action est dite implicitement permise par un ensemble de normes dans un certain contexte lorsque cet ensemble de normes n'interdit pas cette action dans les mêmes conditions (Makinson et van der Torre 2003a, p.393). Dans le langage des logiques I/O, $x \in negperm(G, a)$ si et seulement si $\neg x \notin out_i(G, a)$ pour $i \in \{1, 2, 3, 4\}$. Considérant que G est un ensemble de normes qui exprime des obligations conditionnelles, quelque chose est implicitement permis dans un certain contexte relativement à G lorsque la négation de ce quelque chose n'est pas obligatoire dans ce contexte. Soulignons que défini ainsi, $negperm(G, a)$ n'est pas nécessairement consistant.

Outre la notion de permission implicite, les auteurs adaptent aussi leur approche afin de rendre compte de la permission explicite. Toutefois, ces derniers distinguent entre deux types de permission explicite: la permission *statique* et la permission *dynamique* (Makinson et van der Torre 2003a, p.392). Voici les trois passages où les auteurs les comparent.

> Roughly speaking, static positive permission inherits many properties from the underlying intput/output operation, and behaves like a diminished obligation. On the other hand, dynamic positive permission behaves like an amplified negative permission (Makinson et van der Torre 2003a, pp.392-3).

> Static permission seems to answer to the needs of the citizen, who needs to anticipate the deontic status of his actions. [...] On the other hand, dynamic permission corresponds to the needs of the legislator, who needs to anticipate the effect of *changing* an existing corpus of norms by adding a prohibition. If prohibiting x in condition a would commit us to forbidding something that has been positively permitted in a certain realizable situation, then adding the prohibition would give rise to a certain kind of incoherence (Makinson et van der Torre 2003a, p.398).

> Static positive permission guides the citizen in the deontic assessment of specific actions, and behaves like a weakened obligation. Dynamic positive permission guides the legislator by describing the limits on what may be prohibited without violating static permissions. It behaves like a strenghened negative permission in many respects, but differs in those in which the set G of explicit obligations is allowed to vary (Makinson et van der Torre 2003a, p.408).

Ainsi, alors que la permission statique est celle qui détermine explicitement le cadre à l'intérieur duquel un agent peut agir, la permission dynamique guide le législateur, qui doit s'assurer de ne pas légiférer à l'encontre des actions permises statiquement. Voyons maintenant les défintions formelles de ces notions.

Considérons d'abord la permission statique. Soit l'ensemble G un ensemble générateur d'obligations et P un ensemble générateur de permissions explicites. La définition de la permission statique est telle que $x \in statperm_i(P, G, a)$ si et seulement si $x \in out_i(G \cup \{(c, z)\}, a)$ pour $i \in \{1, 2, 3, 4\}$ et une paire $(c, z) \in P$ (Makinson et van der Torre 2003a, p.397). Prenons par exemple le cas de out_1, où $G*$ est la fonction qui isole les outputs potentiels de $G \cup \{(c, z)\}$.

$$\begin{aligned} x \in statperm_1(P, G, a) &\Leftrightarrow \ x \in out_1(G \cup \{(c, z)\}, a) \\ &\Leftrightarrow \ x \in Cn(G * (Cn(a))) \end{aligned}$$

Si l'on pose $G = \{(b, y), (a \vee b, u)\}$ et $P = \{(b \wedge c, w), (a, z), (a \vee c, v)\}$, alors nous obtenons (en répétant l'opération pour chaque paire de P) $G * (Cn(a)) = \{u, v, z\}$, ce qui nous donne $statperm_1(P, G, a) = Cn(\{u, v, z\})$. Le lecteur aura remarqué que tout ce qui est (explicitement) obligatoire est par le fait même explicitement statiquement permis. Soulignons aussi que dans le cas où $P = \varnothing$, l'ensemble de permissions statiques se réduit aux outputs de a relativement à G. La définition est telle que l'ensemble des permissions statiques se construit en observant une à une les paires de P et le contexte (le input) dans la permission, qui est un élément de P, peut être une conséquence de a. En ce sens, l'ensemble de permissions statiques offre à un agent toutes les permissions explicites qu'il a dans un certain contexte relativement aux ensembles de normes G et P.

Toutefois, ce type de permission pourrait très bien mener à une impasse. Par exemple, posons $G = \{(b, y), (a \vee b, u), (a, \neg v)\}$ et $P = \{(b \wedge c, w), (a, z), (a \vee c, v)\}$. En ajoutant $\neg v$ obligatoire dans le contexte a, on contrevient à une permission explicite statique, ce qui engendre $statperm_1(P, G, a) = Cn(\{u, v, \neg v, z\})$. Même si un ensemble de permissions peut être inconsistant (une action et sa négation peuvent toutes les deux être permises, voire *indifférentes*), cela pose problème dans le cas de la permission *explicite* puisque si l'on pose que v est explicitement permis dans le contexte a, alors dans le même contexte v ne doit pas être interdit. Cela serait une forme d'inconsistance normative.

C'est là qu'entre en jeu la notion de permission explicite *dynamique*. Formellement, $x \in dynperm_i(P, G, a)$ si et seulement si $\neg z \in out_i(G \cup \{(a, \neg x)\}, c)$ pour $i \in \{1, 2, 3, 4\}$ et un $z \in statperm_i(P, G, c)$ et où c n'est pas contradictoire (Makinson et van der Torre 2003a, p.398). En ce sens, l'ensemble des permissions dynamiques correspond aux actions que l'on ne peut interdire sans engendrer d'inconsistance normative. Autrement dit, il s'agit d'un ensemble de permissions qui, si elles étaient interdites, engendreraient des inconsistances normatives. Reprenons l'exemple susmentionné:

$$
\begin{aligned}
&G = \{(b, y), (a \vee b, u)\}\\
&G' = G \cup \{(a, \neg v)\}\\
&P = \{(b \wedge c, w), (a, z), (a \vee c, v)\}\\
&statperm_1(P, G, a) = Cn(\{u, v, z\})
\end{aligned}
$$

Dans ce cas, il y a un $v \in statperm_1(P, G, a)$ pour lequel $\neg v \in out_1(G', a)$, et donc $v \in dynperm(P, G, a)$. Autrement dit, dans le contexte a, v est une permission dynamique, c'est-à-dire une permission qui ne peut être interdite sans engendrer d'inconsistance normative. Évidemment, si l'on veut absolument interdire v, il y a toujours une possibilité de redéfinir P.[7]

En laissant de côté l'aspect « vérité » des propositions, les logiques I/O offrent une solution intéressante au dilemme de Jørgensen (1937), en mettant simplement de côté l'aspect vérifonctionnel des normes, similairement à ce que fait Alchourrón (1990). Cela dit, malgré que la question de savoir si les normes sont sujettes à une attribution de valeur de vérité ou non ait fait couler beaucoup d'encre, il n'est pas si évident que la solution soit simplement de mettre de côté cette hypothèse et de se concentrer seulement sur les critères syntaxiques qui gouvernent les inférences normatives. Après tout, malgré que les normes soient sans valeur de vérité, les propositions normatives demeurent déclaratives: il est tout à fait légitime de se demander si un énoncé de la forme « a est obligatoire » est vrai ou non.

Comme nous le verrons dans les sections qui suivent, les logiques I/O sont utilisées dans le domaine de l'informatique afin de rendre compte de la structure des systèmes normatifs et des NMAS. Cependant, avant de voir comment cela s'opère, il nous faut d'abord présenter l'approche de Jones et Sergot (1996).

Andrew Jones et Marek Sergot

L'approche de Jones et Sergot (1996) vise à rendre compte de la notion de *pouvoir institutionnalisé*, c'est-à-dire de pouvoir accordé à un agent par le biais d'une institution (p. ex., le pouvoir d'un policier de donner une contravention). Afin de rendre compte de cette notion, ceux-ci introduisent un opérateur $\Rightarrow_s$ en conjonction avec une logique de type STIT. La raison pour laquelle nous n'avons pas placé l'approche de Jones et Sergot dans le chapitre 3 est que celle-ci est utilisée en conjonction avec les logiques I/O par Boella et van der Torre (2003b, 2004, 2006a,b,c) afin d'analyser la structure d'un NMAS.

Du côté de la syntaxe, le langage utilisé par Jones et Sergot utilise les connecteurs habituels et permet l'itération de l'opérateur d'action E_x. Cependant, même si leur approche est de type STIT, celle-ci est cependant axiomatisée de manière différente. Plutôt

[7] Le lecteur qui désire avoir plus de détails formels concernant les propriété des différents ensembles de permissions est invité à consulter plus en détails Makinson et van der Torre (2003a).

que d'être axiomatisé comme une modalité de type K, l'opérateur E_x répond à la règle (RE) de substitution d'équivalences logiques dans la portée de l'opérateur E, au schéma d'axiome (**M**) de la logique aléthique ainsi qu'à la règle (non Nec) qui permet de représenter le fait qu'un agent ne réalise pas une tautologie (Jones et Sergot 1996, p.434).

$$E_x A \supset A \tag{\textbf{M}}$$

$$\frac{\vdash A \equiv B}{\vdash E_x A \equiv E_x B} \qquad \text{(RE)}$$

$$\frac{\vdash A}{\vdash \neg E_x A} \qquad \text{(non Nec)}$$

Les auteurs mentionnent que le schéma d'axiome pourrait être ajouté, sans pour autant l'adopter. Cependant, le schéma d'axiome est à rejeter afin de ne pas obtenir (ROM), ce qui contredirait la règle (non Nec).[8] L'opérateur E_x est donc axiomatisé à l'aide d'une logique modale classique (cf. Chellas 1980, p.231).

$$(E_x A \wedge E_x B) \supset E_x(A \wedge B) \tag{Agg}$$
$$E_x(A \wedge B) \supset (E_x A \wedge E_x B) \tag{Agg'}$$

En plus de la logique de l'action sur laquelle se base leur approche, Jones et Sergot (1996, pp.435-6) introduisent le connecteur conditionnel $\Rightarrow_s$. En assumant les règles de la logique propositionnelle classique, ce connecteur est régit par les règles et schémas d'axiomes suivants:

$$\frac{\vdash B \equiv B'}{\vdash (A \Rightarrow_s B) \equiv (A \Rightarrow_s B')} \qquad \text{(EqC)}$$

$$\frac{\vdash A \equiv A'}{\vdash (A \Rightarrow_s B) \equiv (A' \Rightarrow_s B)} \qquad \text{(EqA)}$$

$$((A \Rightarrow_s B) \wedge (A \Rightarrow_s C)) \supset (A \Rightarrow_s (B \wedge C)) \tag{Agg$_s$}$$
$$((A \Rightarrow_s B) \wedge (C \Rightarrow_s B)) \supset ((A \vee C) \Rightarrow_s B) \tag{Disj}$$

Les règles (EqC) et (EqA) permettent de substituer les équivalences logiques respectivement dans le conséquent et l'antécédent du connecteur conditionnel. Le connecteur $A \Rightarrow_s B$ s'interprète comme A *compte pour* B *relativement à* s. En ce sens, le schéma d'axiome (Agg$_s$) exprime que si A compte pour B et A compte pour C, alors A compte pour B et C. Dans le même ordre d'idées, (Disj) indique que si A compte pour B et que C compte aussi pour B, alors la disjonction A ou C compte pour B.

[8] À supposer que le système contient (Agg) et $\Box\top$ (cf. Chellas 1980, p.115).

Étant donné que les auteurs rejettent la règle (Nec) et (Agg'), la sémantique de l'opérateur E_x se traduit en termes de *modèle minimal*. Bien que Jones et Sergot (1996, p.437) renvoient le lecteur au chapitre 7 de l'ouvrage de Chellas (1980) pour la définition du modèle minimal, nous allons néanmoins l'expliciter brièvement. En un mot, la particularité d'un modèle minimal est que ce dernier permet de rejeter la règle (Nec) ainsi que les schémas (Agg) et (Agg'). Les modèles minimaux sont donc des modèles pour les logiques modales classiques, monotones et régulières.

Un modèle minimal $\mathcal{M} = \langle W, N, a\rangle$ est constitué d'un univers non vide W et d'une fonction qui assigne des valeurs de vérité aux propositions relativement aux scénarios. La principale différence est que N est une collection de sous-ensembles de W (Chellas 1980, p.280), par opposition avec la relation R dans les modèles normaux qui est un ensemble de paires ordonnées. Informellement, N_w est un ensemble de propositions qui sont *nécessaires* au moment w. Il faut insister sur la notion de nécessité dans la mesure où il ne s'agit pas d'une nécessité *logique*, comme dans le cas des modèles normaux. Plutôt, cette notion doit être entendue comme une notion arbitraire de nécessité: il s'agit de ce qui est postulé comme nécessaire dans un scénario donné. De fait, il est possible d'avoir $N_w = \varnothing$, et donc $\top \notin N_w$ pour $\top$ une tautologie du système.

Le modèle sémantique $\mathcal{M} = \langle W, E, f_s, V\rangle$ contient un univers $W \neq \varnothing$, un ensemble E contenant différentes collections de sous-ensembles E_w où certaines propositions sont jugées *nécessaires*, une fonction f_s qui détermine les scénarios dans lesquels « A compte pour B » est vrai relativement à une institution s, et finalement V un ensemble d'assignations $\|A\|_{\mathcal{M}} = \{w : A \in w\}$, où $\|A\|_{\mathcal{M}}$ est l'ensemble qui contient les scénarios où A est vrai (relativement au modèle $\mathcal{M}$).

$$\models_w p \Leftrightarrow\; w \in \|p\|_{\mathcal{M}} \tag{1}$$
$$\models_w \neg A \Leftrightarrow \not\models_w A \tag{2}$$
$$\models_w A \supset B \Leftrightarrow \not\models_w A \text{ ou } \models_w B \tag{3}$$
$$\models_w E_x A \Leftrightarrow\; \|A\|_{\mathcal{M}} \in E_w \tag{4}$$
$$\models_w A \Rightarrow_s B \Leftrightarrow\; \|B\|_{\mathcal{M}} \in f_s(w, \|A\|_{\mathcal{M}}) \tag{5}$$

Les clauses (1)–(3) sont usuelles et (4) exprime que « x réalise A » est vrai pour w si et seulement si l'ensemble qui contient les scénarios où A est vrai est élément de la collection d'ensembles qui contiennent les propositions nécessairement vraies pour w. La clause (5) signifie que « A compte pour B » est vrai relativement à l'institution s pour le scénario w si et seulement si l'ensemble qui contient les scénarios où B est vrai est élément de l'ensemble qui contient les scénarios où A est vrai. Autrement dit, f_s isole les scénarios accessibles à w où A est vrai. En ce sens, « A compte pour B » est vrai si et seulement si tout scénario où B est vrai est une alternative à w dans laquelle A est vrai.

À cela s'ajoutent les conditions (C1) et (C2) pour assurer la validité de (Agg_s) et (Disj) respectivement (Jones et Sergot 1996, p.437).

$$Y \in f_s(w, X) \text{ et } Z \in f_s(w, X) \Rightarrow\; Y \cap Z \in f_s(w, X) \tag{C1}$$
$$X \in f_s(w, Y) \text{ et } X \in f_s(w, Z) \Rightarrow\; X \in f_s(w, Y \cup Z) \tag{C2}$$

En mots, (C1) signifie que si les ensembles de propositions Y et Z sont vrais dans les scénarios accessibles à w et où X est vrai, alors l'intersection de Y et Z est aussi vraie dans les alternatives à w où X est vrai. Quant à (C2), cela signifie que si X est vrai dans les alternatives à w où Y est vrai et que X est vrai dans les alternatives à w où Z est vrai, alors X est vrai dans les alternatives à w où l'union de Y et Z est vraie (autant les scénarios où Y est vrai que ceux où Z l'est mènent à des alternatives où X est vrai).

L'introduction de l'opérateur $\Rightarrow_s$ amène Jones et Sergot (1996, p.439) à considérer la modalité (de contrainte) D_s: si A compte pour B dans l'institution s, alors A implique B est une contrainte pour l'institution s. Autrement dit, à partir du moment où l'on accepte que A compte pour B relativement à une institution s, alors cette institution est contrainte par le fait que $A \supset B$ est vrai. En ce sens, « A et non B » n'est pas acceptable pour s. Cela se traduit syntaxiquement par le schéma d'axiome suivant.

$$(A \Rightarrow_s B) \supset D_s(A \supset B) \qquad (D_s)$$

Outre ce schéma d'axiome, la modalité D_s est axiomatisée par KD. Du côté de la sémantique, le modèle $\mathcal{M} = \langle W, E, f_s, d_s, V\rangle$ est augmenté par une relation d'accessibilité d_s, laquelle détermine les scénarios accessibles à w relativement à une institution s. La relation d_s est sérielle (en vertu de KD) et doit satisfaire la condition (C3).

$$Y \in f_s(w, X) \Rightarrow\ d_s(w) \cap X \subseteq Y. \qquad \text{(C3)}$$

En d'autres termes, (C3), qui vise à assurer la validité du schéma (D_s), signifie que si Y est vrai dans les scénarios accessibles à w où X est vrai, alors les scénarios accessibles à w où X est vrai sont sous-ensembles de ceux où Y est vrai. La clause sémantique (cf. Jones et Sergot 1996, p.440) est que:

$$\models_w D_sA \Leftrightarrow\ d_s(w) \subseteq \|A\|_{\mathcal{M}}, \qquad (6)$$

S'il est vrai que A est une contrainte pour l'institution s, alors les scénarios accessibles à w sont parmi ceux où A est vrai.

À la fin de leur article, Jones et Sergot (1996, pp.442-3) soulignent l'intérêt de leur approche du point de vue du discours normatif, sans pour autant développer d'extension déontique de leur système. L'approche présentée jusqu'à présent vise à formaliser la notion de *pouvoir institutionnel*. Cela a nécessité l'introduction du connecteur $\Rightarrow_s$ afin de pouvoir représenter le fait que certaines actions *comptent pour* d'autres actions dans le cadre d'une institution. Par exemple, *Paul prend l'argent de Pierre sans sa permission* compte pour un vol du point de vue légal. Même si l'approche de Jones et Sergot ne fait pas partie des logiques I/O, celle-ci a été reprise éventuellement par des logiciens dont l'attention est portée sur la représentation informatique d'un système normatif, ce que nous allons maintenant aborder.

Guido Boella et Leendert van der Torre

Une bonne partie de la littérature portant sur la logique déontique vise ses applications en informatique et en programmation. L'approche de Boella et van der Torre (2006c), qui jumelle les logiques I/O de Makinson et van der Torre (2000, 2001, 2003a) avec l'analyse de Jones et Sergot (1996), s'insère dans cette catégorie.[9] En effet, leur objectif est de développer des outils formels pertinents à la modélisation de l'architecture d'un système normatif, facilitant ainsi l'automatisation des raisonnements. Dans ce qui suit, l'idée est de définir récursivement un système normatif par le biais d'opérations input/output. L'objectif est de débuter à partir de définitions de base (p. ex., la permission, l'obligation, les contraintes institutionnelles, etc.) et ensuite de déterminer les outputs *combinés* de ces opérations, l'objectif ultime étant d'obtenir *une* opération qui représente le système global. En ce sens, l'objectif est de définir un système normatif comme *une seule* opération input/output, où l'on « entre A et on obtient B ».

L'approche de 2006c est notamment basée sur ce qui a été présenté dans Boella et van der Torre (2006a), où les auteurs introduisent l'architecture d'un système normatif. D'entrée de jeu, Boella et van der Torre (2006c, p.27) utilisent l'approche de Jones et Sergot (1996), à laquelle ils ajoutent un axiome pour la transitivité (Tran) ainsi que la condition (C4) sur le modèle sémantique. Cette dernière signifie simplement que si Y est vrai pour les alternatives à w où X est vrai et que Z est vrai pour les alternatives à w où Y est vrai, alors Z est vrai pour les alternatives à w où X est vrai.[10]

$$((A \Rightarrow_s B) \wedge (B \Rightarrow_s C)) \supset (A \Rightarrow_s C) \qquad \text{(Tran)}$$
$$Y \in f_s(w, X) \text{ et } Z \in f_s(w, Y) \Rightarrow Z \in f_s(w, X) \qquad \text{(C4)}$$

Cela dit, Boella et van der Torre (2006c, p.27) reformulent la logique des *count-as conditionals* de Jones et Sergot en termes de logique I/O. Alors que chez Makinson et van der Torre on parlait d'ensemble générateur G, Boella et van der Torre parlent d'ensemble de paires CA (pour *count-as*), lequel s'évalue relativement à un système normatif ou à une institution s.

L'ensemble CA est fonction de l'ensemble de normes qui caractérise le système normatif. L'opération $out_{CA}(CA, s, a)$ est régie par les règles (AND), (OR) et (T), la dernière exprimant la transitivité.

$$\frac{(a, x)(x, y)}{(a, y)} \qquad \text{(T)}$$

Empruntant l'idée de *contrainte institutionnelle* exprimée par l'opérateur modal D_s de Jones et Sergot, Boella et van der Torre (2006c, p.29) introduisent l'ensemble de paires

[9] Voir aussi notamment Boella et van der Torre (2003b, 2004, 2006a,b,c).

[10] Dans leur article, Jones et Sergot (1996, p.438) acceptent l'axiome par défaut simplement parce qu'ils n'ont pas de contre-exemple en tête pour le rejeter. Nous l'avons mis de côté dans la présentation de leur approche puisqu'ils n'insistent pas sur ce dernier.

IC (*institutional constraints*), dont le output $out_{IC}(IC, s, a)$ est régi par les règles (SI), (WO) et (Id).

$$\frac{}{(a, a)} \qquad \text{(Id)}$$

La définition de out_{IC} est plutôt simple:

$$out_{IC}(IC, s, a) =_{def} Cn(\{a\} \cup \{a \supset y : (a, y) \in IC\}) \qquad \text{(def. } out_{IC}\text{)}$$

Autrement dit, le output des contraintes institutionnelles pour un input a correspond à l'ensemble des conséquences de a avec $a \supset y$ pour toute paire $(a, y) \in IC$.

L'axiome (D_s) de Jones et Sergot est remplacé par la règle suivante sur les outputs.

$$\frac{x \in out_{CA}(CA, s, a)}{x \in out_{IC+CA}(IC, CA, s, a)} \qquad (out_{IC+CA})$$

Le output out_{IC+CA} répond aux règles (AND), (OR), (T), (SI), (WO) et (Id) et est sémantiquement défini récursivement par (Boella et van der Torre 2006c, p.30):

$$\begin{aligned} &out_{IC+CA}(IC, CA, s, X) && \text{(def. } out_{IC+CA}\text{)} \\ &\quad =_{def} out_{IC}(IC \cup CA, s, X) \\ &\quad =_{def} Cn(X \cup \{x \supset y : x \in X \text{ et } (x, y) \in IC \cup CA\}) \end{aligned}$$

En mots, cela signifie que out_{IC+CA} est par définition l'ensemble des outputs de IC en union avec CA relativement à une institution s et à un input X. Tout ce qui est un output out_{CA} est un output out_{IC+CA}, mais certaines contraintes institutionnelles peuvent ne pas être des *count-as*. Le output $out_{IC+CA}(IC, CA, s, x)$ équivaut donc à l'ensemble des conséquences des tous les $\{x, x \supset y\}$ pour lesquels $x \supset y$ est soit dans IC ou dans CA.

Cela fait, les auteurs introduisent une opération pour les obligations conditionnelles (Boella et van der Torre 2006c, pp.30-1). L'introduction de l'obligation conditionnelle se fait exactement dans le même esprit que la définition de l'ensemble générateur G chez Makinson et van der Torre. En effet, Boella et van der Torre introduisent l'ensemble de paires O, où chaque paire de O donne une obligation (output) conditionnelle à un contexte (input). L'opération out^i_O est définie au même titre que les outputs 1–4 chez Makinson et van der Torre. Ensuite, les auteurs introduisent le output relatif aux contraintes institutionnelles, à l'ensemble CA ainsi qu'à l'ensemble d'obligations conditionnelles. L'opération $out_{IC+CA+O}$ est donnée par la règle (nous avons omis le i):

$$\frac{x \in out_{IC+CA}(IC, CA, s, a), y \in out_O(O, s, x)}{y \in out_{IC+CA+O}(IC, CA, O, s, a)} \qquad (out_{IC+CA+O})$$

Cette règle est accompagnée par la définition sémantique suivante.

$$\begin{aligned} & out_{IC+CA+O}(IC, CA, O, s, a) && \text{(def. } out_{IC+CA+O}\text{)} \\ & \qquad =_{def} out_O(O, s, out_{IC+CA}(IC, CA, s, a)) \\ & \qquad =_{def} out_O(O, s, out_{IC}(IC \cup CA, s, a)) \\ & \qquad =_{def} out_O(O, s, Cn(\{a\} \cup \{a \supset y : (a, y) \in IC \cup CA\})) \end{aligned}$$

En d'autres termes, le output conjoint $out_{IC+CA+O}$ ayant a comme input correspond au output out_O (relativement à l'ensemble générateur O) lorsque le input est l'ensemble des conséquences logiques de l'ensemble $\{a, a \supset y\}$ qui contient tous les $a \supset y$ pour lesquels (a, y) est soit une contrainte institutionnelle, soit un *count-as*.

Conformément à ce que l'on trouve dans Makinson et van der Torre (2001), Boella et van der Torre (2006c, pp.31-2) introduisent aussi les contraintes sur les outputs afin de mieux rendre compte des obligations conditionnelles. Toutefois, ces derniers utilisent une notion qui avait été laissée de côté chez Makinson et van der Torre, à savoir la notion de *meet*, voire de *geatest lower bound*. Tel que mentionné au début de ce chapitre, l'ensemble *outfamily* est obtenu après avoir mis comme input l'ensemble *maxfamily*, lequel permet de déterminer tous les sous-ensembles de l'ensemble générateur qui sont consistants avec le output. Cela dit, *outfamily* est une collection d'ensembles, laquelle contient tous les outputs des sous-ensembles de l'ensemble générateur qui, lorsque pris en tant que input, sont consistants avec le output.

Il y a deux opérations possibles sur *outfamily* que nous avions laissé de côté dans la présentation de Makinson et van der Torre, principalement en raison du fait que ces notions, bien que mentionnées, n'étaient pas à l'étude dans leur article. Ces deux opérations sont celles du *meet* et du *join*, voire du *greatest lower bound* et du *least upper bound*, représentées respectivement par:

$$\bigcap outfamily(G, A, C)$$
$$\bigcup outfamily(G, A, C)$$

En un mot, le *meet* identifie le plus petit ensemble commun à tous les membres de *outfamily* (l'intersection) alors que le *join* sélectionne le plus grand (l'union). À l'aide de la définition du *meet*, Boella et van der Torre définissent un nouvel output, soit $out_O^{\cap}$, où C est un ensemble de contraintes (à ne pas confondre avec l'ensemble des contraintes institutionnelles IC).

$$out_O^{\cap}(O, s, a, C) =_{def} \bigcap outfamily(O, s, a, C) \qquad \text{(def. } out_O^{\cap}\text{)}$$

Le i a encore une fois été laissé de côté afin de ne pas encombrer l'écriture. Ayant cette opération en main, Boella et van der Torre (2006c, p.32) définissent récursivement $out_{IC+CA+O}^{\cap}$ par:

$$
\begin{aligned}
&out^{\cap}_{IC+CA+O}(IC, CA, O, s, a) && \text{(def. } out^{\cap}_{IC+CA+O}\text{)}\\
&\quad =_{def}\ out^{\cap}_{O}(O, s, out_{IC+CA}(IC, CA, s, a), out_{IC+CA}(IC, CA, s, a))\\
&\quad =_{def}\ \bigcap outfamily(O, s, out_{IC+CA}(IC, CA, s, a), out_{IC+CA}(IC, CA, s, a))\\
&\quad =_{def}\ \bigcap outfamily(O, s, \Gamma, \Gamma)
\end{aligned}
$$

Ici, nous avons:

$$\Gamma = Cn(\{a\} \cup \{a \supset y : (a, y) \in IC \cup CA\})$$

Autrement dit, l'opération $out^{\cap}_{IC+CA+O}$ garantit la consistance du output des obligations avec les contraintes institutionnelles en isolant le plus petit ensemble commun à tous les membres de *outfamily*.

Ayant traité des obligations conditionnelles et de l'ajout de contraintes au système normatif afin que les obligations soient consistantes avec les contraintes institutionnelles de s, Boella et van der Torre (2006c, p.32) introduisent un ensemble de paires P représentant des permissions conditionnelles. Conjointement à cet ensemble, il définissent l'opération suivante.

$$
\begin{aligned}
&out_P(P, s, A) && (out_P)\\
&\quad =_{def} \{Cn(x) : (a_i \wedge ... \wedge a_n, x) \in P \text{ pour certains } a_i, ..., a_n \in A\}
\end{aligned}
$$

Cette définition est univoque: le output out_P relativement à l'input A correspond à l'ensemble qui contient les ensembles des conséquences de x pour lesquels x est la permission conditionnelle à la conjonction de certaines propositions membres de A. Autrement dit, on observe toutes les combinaisons possibles des propositions de A en termes de conjonction, on isole les x qui sont les permissions conditionnelles à ces combinaisons et on détermine l'ensemble qui contient les ensembles des conséquences de chacun de ces x.

En plus des permissions explicites introduites par le output out_P, les auteurs définissent l'opération *merge* afin de déterminer l'ensemble qui contient non seulement les permissions explicites, mais aussi les permissions qui découlent du fait que quelque chose est obligatoire (ce qui valide le schéma d'axiome (**D**) des systèmes standards).

$$merge(X, Y) = \{Cn(x \cup Y) : x \in X\} \qquad \text{(def. Merge)}$$

En ce sens, l'opération *merge* détermine l'ensemble qui contient tous les ensembles des conséquences de $x \cup Y$, où x est un ensemble de permissions et Y un ensemble d'obligations. Dans cette définition, X est donc un ensemble de sous-ensembles de propositions, alors que Y est un ensemble de propositions, conformément à la définition de out_P.

Les dernières étapes consistent à définir les outputs du système global. Pour ce faire, les auteurs doivent d'abord définir le output conjoint $out_{IC+CA+P}$, défini de la même manière que $out_{IC+CA+O}$ par:

$$\begin{aligned} & out_{IC+CA+P}(IC, CA, P, s, a) && \text{(def. } out_{IC+CA+P}\text{)} \\ & \quad =_{def} out_P(P, s, out_{IC+CA}(IC, CA, s, a)) \\ & \quad =_{def} out_P(P, s, out_{IC}(IC \cup CA, s, a)) \\ & \quad =_{def} out_P(P, s, Cn(\{a\} \cup \{a \supset y : (a, y) \in IC \cup CA\})) \end{aligned}$$

Ce dernier est régi par la règle suivante.

$$\frac{x \in out_{IC+CA}(IC, CA, s, a), y \in out_P(P, s, x)}{y \in out_{IC+CA+P}(IC, CA, P, s, a)} \qquad (out_{IC+CA+P})$$

Cela fait, il reste deux opérations à définir. En premier lieu, l'opération $out^{\cap}_{IC+CA+P+O}$ s'assure que le output conjoint de IC, CA, P et O est consistant avec les contraintes institutionnelles. Dans la définition qui suit, nous avons écrit out_{IC+CA} plutôt que $out_{IC+CA}(IC, CA, s, a)$ afin de faciliter la lecture.

$$\begin{aligned} & out^{\cap}_{IC+CA+P+O}(IC, CA, P, O, s, a) && \text{(def. } out^{\cap}_{IC+CA+P+O}\text{)} \\ & \quad =_{def} out^{\cap}_{O}(O, merge(out_{IC+CA+P}, out_{IC+CA}), s, out_{IC+CA}) \\ & \quad =_{def} \bigcap outfamily(O, s, merge(out_{IC+CA+P}, out_{IC+CA}), out_{IC+CA}) \end{aligned}$$

Autrement dit, on prend le plus petit ensemble commun à tous les outputs de O consistants avec les outputs de IC et CA et qui ont pour input l'ensemble qui contient toutes les combinaisons possibles entre les permissions (prises en considérations avec les contraintes institutionnelles) et les contraintes institutionnelles. En d'autres termes, on prend comme input l'ensemble de tous ce qui est permis (au sens large, et non seulement explicite) en fonction des contraintes institutionnelles et des *count-as* afin de déterminer le plus petit output consistant avec les contraintes institutionnelles.

Finalement, l'opération $out_{IC+CA+P+O+P}$ détermine le output du système normatif global. Cette opération est définie par (nous avons encore laissé de côté ce qui se trouve dans la portée des opérations):

$$\begin{aligned} & out_{IC+CA+P+O+P}(IC, CA, P, O, s, a) && \text{(def. } out_{IC+CA+P+O+P}\text{)} \\ & \quad =_{def} merge(out_{IC+CA+P}, out^{\cap}_{IC+CA+P+O}) \end{aligned}$$

Somme toute, l'idée de Boella et van der Torre est de définir récursivement un système normatif à l'aide d'opérations input/output. Ayant défini étape par étape les différents types de outputs ainsi que leurs combinaisons possibles, les auteurs parviennent à la définition d'un système normatif en termes d'entrée et de sortie. Voici une reconstruction de l'explication à partir de la dernière définition. Prenons out_1 pour les définitions de out_P et out_O, avec $P(A)$ et $O(A)$ définis dans le même esprit que $G(A)$ au début de ce chapitre. De plus, présupposons que le input de départ est a. Le output d'un système normatif est considéré comme le *merge* de $out_{IC+CA+P}$ et de $out^{\cap}_{IC+CA+P+O}$. Le *merge* est l'ensemble qui contient l'ensemble des conséquences de $X \cup out^{\cap}_{IC+CA+P+O}$ pour tout $X \in out_{IC+CA+P}$. Un $X \in out_{IC+CA+P}$ est un élément de out_P, où $out_P = Cn(P(Cn(\Gamma)))$, Γ étant l'ensemble qui contient les conséquences des unions de $\{a\}$ avec $\{a \supset y\}$ pour tous y tel que $(a, y) \in IC \cup CA$. En ce sens, un $X \in out_{IC+CA+P}$ est un ensemble qui contient les conséquences de $\{a\} \cup \{a \supset y\}$ pour un y tel que $(a, y) \in IC \cup CA$. De fait, le *merge* consiste à prendre les conséquences de $X \cup out^{\cap}_{IC+CA+P+O}$ pour chaque X qui contient les conséquences de $\{a\} \cup \{a \supset y\}$ pour un y tel que $(a, y) \in IC \cup CA$. De manière informelle, $out_{IC+CA+P}$ équivaut à considérer le out_P ayant comme input toutes les paires de $IC \cup CA$ qui partagent le même output. On prend cela, on observe les conséquences, on regarde ce que cela donne comme output relativement à P lorsqu'on met ces conséquences en input et on observe les conséquences du tout. Quant à $out^{\cap}_{IC+CA+P+O}$, il s'agit du plus petit sous-ensemble partagé par tous les outputs consistants avec out_{IC+CA} et ayant comme input des sous-ensembles de $merge(out_{IC+CA+P}, out_{IC+CA})$. En ce sens, on prend l'ensemble qui contient les conséquences de $X \cup out_{IC+CA}$ pour tout $X \in out_{IC+CA+P}$ (où out_{IC+CA} contient les conséquences de tout ce que a implique dans $IC \cup CA$) et ensuite on détermine l'intersection de $out_O(O, s, \Gamma)$ consistant avec $IC \cup CA$ pour tout $\Gamma \subseteq merge(out_{IC+CA+P}, out_{IC+CA})$.

L'approche de Boella et van der Torre est décidément très complexe et il est évident que de tenter de la résumer en quelques phrases simples nous éloignerait de la subtilité du formalisme. Néanmoins, par souci pédagogique, voici comment nous pourrions tenter de la résumer: le système normatif est considéré comme une entité qui se comprend en termes de inputs et de outputs; on entre des données dans la « boîte », on brasse et ensuite on observe ce qui sort. La « boîte » du système normatif et divisée en « sous-boîtes ». Dans le cas des obligations conditionnelles, on entre le input, on ajoute certaines contraintes pour s'assurer que le résultat est cohérent (et ainsi éviter le paradoxe de Chisholm), on brasse et on observe ce qui sort (c'est-à-dire les obligations qui s'appliquent). Il est aussi possible de regarder la « sous-boîte » de la permission de la même manière. Au total, on peut créer une « boîte » qui est constituée d'une ou plusieurs « sous-boîte(s) » afin d'observer ce qui se passe lorsqu'on augmente l'information que l'on prend en compte. Au final, le système normatif peut se comprendre ainsi (en supposant qu'à chaque étape on brasse la boîte!): on entre un input a, on regarde les permissions contextuelles à a relativement à l'ensemble de permissions explicites et l'ensemble de contraintes institutionnelles (i.e., tout ce que a implique dans le cadre du système normatif s), on observe les conséquences de cela en rapport avec toutes les obligations qui tiennent dans ce contexte, et on brasse à nouveau la boîte afin d'observer les conséquences que cela nous donne.

En un mot, l'opération input/output d'un système normatif consiste à prendre en compte tout ce que le système présuppose (les contraintes institutionnelles, les implications considérées vraies dans le système, les équivalences entre les actions et les descriptions de faits, les *count-as*, etc.), toutes les permissions explicites et obligations relatives au input et finalement à considérer les conséquences logiques du tout. Malgré l'intérêt de ce genre de définition récursive pour la programmation de systèmes normatifs, l'approche de Boella et van der Torre n'est pas destinée à servir à l'analyse du discours et des raisonnements. Par surcroît, les auteurs assument qu'un système légal se réduit à une opération inputs/outputs bien définie, ce qui est non seulement contesté par la majorité des juristes mais va aussi à l'encontre des principes qui guident l'interprétation des lois (cf. Côté 2006): au moment de la rédaction des lois, le législateur ne pouvait pas prévoir toutes les façons possibles dont un système légal pouvait évoluer, et par conséquent, si *a posteriori* il faut réévaluer cette législation à la lumière de l'information actuelle, il faut se questionner sur l'intention du législateur et sur l'esprit de la loi, chose qui n'est pas définissable d'emblée par un ensemble de contraintes institutionnelles et de *count-as*.

Pour une extension de l'article de 2006c où l'on cherche à rendre compte autant des systèmes normatifs que des NMAS, le lecteur peut consulter Boella et van der Torre (2006b).

* * *

En guise de conclusion, les logiques I/O offrent une alternative aux logiques modales, fournissant un cadre de travail permettant d'éviter le dilemme de Jørgensen. Ainsi, en les utilisant en vue de l'analyse des raisonnements, les logiques I/O permettent l'analyse des inférences normatives dans une sémantique autre que celle des mondes possibles. Cela dit, le principal intérêt de ce genre d'approche se trouve surtout en informatique, et d'ailleurs la lourdeur du formalisme inhérent aux logiques I/O rend leur applications en argumentation et en analyse du discours une tâche difficile. Passons maintenant aux approches algébriques.

Chapitre 7

Algèbre

Outre les logiques I/O ainsi que les logiques modales, on trouve aussi plusieurs approches algébriques au sein de la littérature en logique déontique. Alors que les logiques modales sont souvent de type *Ought-to-be*, les approches de type *Ought-to-do* utilisent usuellement une algèbre d'actions afin de bien distinguer entre les propositions déclaratives et les actions dans la portée des opérateurs déontiques. Néanmoins, on trouve des approches algébriques de types *Ought-to-be*, et certains auteurs préfèrent considérer les systèmes normatifs comme des ensembles sujets a certaines opérations. Tel est le cas de Lindahl et Odelstad (2000), dont il sera d'abord question. Par la suite, nous présenterons quelques approches algébriques mettant l'accent sur la structure des actions dans une perspective *Ought-to-do*, incluant l'approche de Segerberg (1982b), qui par la suite aura influencé celle de Trypuz et Kulicki (2009). Finalement, il sera question de l'approche de Castro et Maibaum (2009), qui introduisent une logique dynamique fondée sur une logique de l'action, et en dernier lieu nous considérerons les travaux de Lucas, qui propose de faire une analyse de la logique de l'action de von Wright dans le cadre de la théorie des catégories. Pour une analyse approfondie des algèbres d'action, le lecteur est référé au chapitre 14 du présent ouvrage.

Lars Lindahl et Jan Odelstad

L'article de Lindahl et Odelstad (2000), où les auteurs proposent une analyse algébrique des systèmes normatifs, est très fortement inspirée du texte de Odelstad et Lindahl (2000). Outre la similarité entre ces deux article, l'approche de Lindahl et Odelstad (2000) s'insère dans la lignée des travaux de Alchourrón et Bulygin (1971). Sans proposer de définition précise, Lindahl et Odelstad (2000, p.261) entendent un *système normatif* comme un ensemble de lois. En un mot, l'idée de Lindahl et Odelstad est de représenter un système normatif de par la jonction de deux (ou plusieurs) fragments d'une algèbre de Boole. Plus précisément, leur objectif est d'être en mesure représenter formellement la jonction d'un fragment *normatif* à un fragment *descriptif* (Lindahl et Odelstad 2000, p.265).[1]

[1] Voir aussi Odelstad et Lindahl (2000, p.161).

Dans le même esprit que celui des logiques I/O, Lindahl et Odelstad conçoivent un système normatif comme un processus input/output qui permet de déduire à partir de certains faits les conséquences légales applicables dans une situation donnée (Lindahl et Odelstad 2000, p.262).

Essentiellement, Lindahl et Odelstad (2000, p.266) considèrent un système normatif comme un pré-ordre booléen. Cette notion repose sur celle d'une *algèbre de Boole*. Pour le dire en trois mots, une algèbre de Boole est un *treillis distributif complémenté* (Goldblatt 2006, p.134). Un *treillis* est un *poset*, c'est-à-dire un ensemble partiellement ordonné (*Partially Ordered SET*) pour lequel chaque paire d'éléments possède un g.l.b. (*greatest lower bound*) et un l.u.b. (*least upper bound*) (Goldblatt 2006, p.55). Un *poset* $\mathbf{P}$ est une structure $\langle P, \sqsubseteq \rangle$, où P est un ensemble et $\sqsubseteq$ est une relation d'ordre partiel (Goldblatt 2006, p.29). Un *ordre partiel* est un pré-ordre antisymétrique, c'est-à-dire une relation réflexive, transitive et antisymétrique (Goldblatt 2006, pp.28-9). Soulignons qu'un ordre partiel n'est pas nécessairement linéaire. Un *g.l.b.*, ici dénoté par $x \sqcap y$, est tel que (cf. Goldblatt 2006, p.49):

1. $x \sqcap y \sqsubseteq x$ et $x \sqcap y \sqsubseteq x$;
2. si $z \sqsubseteq x$ et $z \sqsubseteq y$, alors $z \sqsubseteq x \sqcap y$.

Un *l.u.b.*, ici dénoté par $x \sqcup y$, est tel que (cf. Goldblatt 2006, p.55):

1. $x \sqsubseteq x \sqcup y$ et $y \sqsubseteq x \sqcup y$;
2. si $x \sqsubseteq z$ et $y \sqsubseteq z$, alors $x \sqcup y \sqsubseteq z$.

Un treillis est donc un poset pour lequel chaque paire d'éléments possède un g.l.b. et un l.u.b.. Un treillis est dit *distributif* lorsque les deux lois suivantes sont respectées pour tout x, y et z (Goldblatt 2006, p.134):

1. $x \sqcap (y \sqcup z) = (x \sqcap y) \sqcup (x \sqcap z)$;
2. $x \sqcup (y \sqcap z) = (x \sqcup y) \sqcap (x \sqcup z)$.

Finalement, un treillis est dit *complémenté* lorsque chaque élément possède un complément, c'est-à-dire que pour tout x il y a un y tel que:

1. $x \sqcup y = 1$;
2. $x \sqcap y = 0$.

Ici, 0 est un objet initial et 1 est un objet terminal, c'est-à-dire des objets pour lesquels (respectivement) $0 \sqsubseteq x$ et $x \sqsubseteq 1$ pour tout x (Goldblatt 2006, pp.43-4). Pour une analyse plus détaillée des algèbres, le lecteur est invité à consulter le chapitre 14.

Mais quel est le rapport entre un système normatif et un treillis distributif complémenté? D'emblée, il faut voir que la logique propositionnelle classique se résume à une algèbre de Boole (cf. chapitre 14). En effet, en logique classique, le g.l.b. est la conjonction, le l.u.b. la disjonction, l'objet initial est le *falsum* $\perp$ et l'objet terminal est le *verum* $\top$. De plus, nous pouvons aisément prouver que:

1. $p \wedge (q \vee r) = (p \wedge q) \vee (p \wedge r)$;
2. $p \vee (q \wedge r) = (p \vee q) \wedge (p \vee r)$;
3. $p \vee \neg p = \top$;
4. $p \wedge \neg p = \perp$.

Le symbole $=$ renvoit à l'équivalence logique et l'ordre partiel à la conséquence logique.[2] C'est dans cet esprit que Lindahl et Odelstad parlent d'un système normatif en tant que pré-ordre booléen. Leur objectif est de déterminer les propriétés de la relation R d'un système normatif qui permettent d'interpréter aRb comme $a \supset b$.

> Let $\mathcal{S}$ be a normative system and let a, b be two conditions. The subject-matter of our inquiry is a relation R such that aRb represents that it follows from $\mathcal{S}$ that a implies b (Lindahl et Odelstad 2000, p.262).

Une *condition* peut être soit une proposition atomique au sens du calcul des propositions ou une proposition atomique au sens du calcul des prédicats. Malgré que cela ne soit pas présenté ainsi, voici une façon de comprendre la notion de *condition*. Soit C_i un ensemble dénombrable de conditions qui ont i variable(s) dans leur portée. Si $i = 0$, alors nous obtenons l'ensemble des conditions qui ne portent sur aucune variable et qui sont représentées par des atomes propositionnels. Lorsque $i = 1$, on obtient les conditions représentées par des prédicats unaires (p. ex., x est majeur), $i = 2$ donne des prédicats binaires (p. ex., x est le père de y), $i = 3$ des prédicats ternaires, etc. Lorsque la condition a_i est définie, la condition $a'_i(x_1, ..., x_i)$ représente la négation de la condition $a_i(x_1, ..., x_i)$, $(a_i \wedge b_i)(x_1, ..., x_i)$ la conjonction $a_i(x_1, ..., x_i)$ et $b_i(x_1, ..., x_i)$, et la condition $(a_i \vee b_i)(x_1, ..., x_i)$ représente la disjonction $a_i(x_1, ..., x_i)$ ou $b_i(x_1, ..., x_i)$, où $a \vee b =_{def} (a' \wedge b')'$ (Lindahl et Odelstad 2000, p.263). Soit

$$C = \bigcup_{i \in \mathbb{N}} C_i$$

l'ensemble de toutes les conditions atomiques (portant sur le nombre de variables i possibles, avec $C_i = \{c_i^j : j \in \mathbb{N}\}$. L'ensemble des expressions bien formées pour les conditions est défini récursivement à partir de l'ensemble C de toutes les conditions atomiques (Lindahl et Odelstad 2000, p.264).[3]

[2] Notons cependant que lorsque l'on traite la logique classique dans un cadre algébrique, il faut considérer un ordre partiel entre des classes d'équivalences de propositions (en vertu de l'antisymétrie).

[3] Nous avons modifié la formulation de la définition récursive des expressions bien formées afin que cela soit cohérent avec la suite de leur texte ainsi qu'avec la position de Odelstad et Lindahl (2000).

1. si $a \in C$, alors $a \in EBF$;

2. si $A, B \in EBF$, alors $A', A \wedge B \in EBF$.

Ainsi, en posant $\mathbf{EBF} = \langle EBF, \sqsubseteq \rangle$, on obtient une algèbre de Boole $\langle \mathbf{EBF}, \wedge, ' \rangle$, où $\perp$ et $\top$ sont respectivement l'objet initial et l'objet terminal.

À partir de cela, les auteurs définissent un *système normatif* $\mathcal{S} = \langle \mathcal{N}, \wedge, ', R \rangle$ (où $\mathcal{N} \subseteq EBF$ est un poset) lorsque pour tout $a, b \in \mathcal{N}$, aRb si et seulement si a implique b selon $\mathcal{S}$ (Lindahl et Odelstad 2000, p.265). Un parallèle mérite d'être fait entre la conception d'un système normatif chez Lindahl et Odelstad et l'ensemble des contraintes institutionnelles chez Boella et van der Torre. Au même titre que l'ensemble des contraintes institutionnelles contient les paires (a, x) pour lesquelles a compte pour x au sein du système normatif (i.e., pour lesquelles $a \supset x$ est vrai), un système normatif est conçu chez Lindahl et Odelstad comme une structure qui détermine un ensemble de paires $< a, b > \in R$ pour lesquelles $a \supset b$ est vrai dans le système normatif. La principale différence est que chez Lindahl et Odelstad, $\mathcal{S}$ contient à la fois les contraintes institutionnelles et les ensembles (générateurs) d'obligations et de permissions. Cela dit, dans les cas où a est descriptif et que b est normatif, et donc que (a, b) est plutôt dans un ensemble générateur d'obligations ou de permissions chez Boella et van der Torre, Lindahl et Odelstad parleront de *corrélations normatives* (Lindahl et Odelstad 2000, p.265).

Afin de formaliser la notion de système normatif, Lindahl et Odelstad (2000, pp.264-5) introduisent la notion de *pré-ordre booléen*. En termes simples, un *pré-ordre booléen* est une algèbre de Boole sur laquelle on impose un pré-ordre R, c'est-à-dire une relation réflexive et transitive, qui doit satisfaire les conditions suivantes (cf. Lindahl et Odelstad 2000, p.266):

1. si aRb et aRc, alors $aR(b \wedge c)$;

2. si aRb, alors $b'Ra'$;

3. $(a \wedge b)Ra$;

4. non $\top R \perp$.

Le lecteur familier avec la logique propositionnelle aura remarqué que cela se résume aux conditions suivantes:

1. si a implique respectivement b et c, alors a implique la conjonction de b et c;

2. si a implique b, alors non a implique non b (contraposition);

3. une conjonction implique chacun de ses membres;

4. le vrai n'implique pas le faux.

Soulignons au passage une remarque technique. Lindah et Odestad utilisent un pré-ordre booléen plutôt qu'une algèbre de Boole afin de représenter l'implication matérielle par la relation R et ainsi éviter de traiter des classes d'équivalences. En utilisant un ordre partiel plutôt qu'un pré-ordre, on obtient que aRb et bRa implique que $a = b$, ce qui requiert que l'on parle d'identité sur les classes d'équivalences des propositions plutôt que sur les propositions elles-mêmes. Les auteurs préfèrent utiliser un pré-ordre afin de pouvoir distinguer entre deux propositions qui ont une signification différente malgré que aRb et bRa.

Dans le but de rendre compte des corrélations normatives, c'est-à-dire des conditionnels qui mettent en jeu un antécédent descriptif et un conséquent normatif, Lindahl et Odelstad (2000, p.267) représentent un système normatif $\mathcal{S}$ comme la jonction de deux (ou plusieurs) sous-structures $\mathcal{N}_1$ et $\mathcal{N}_2$. Dans un tel contexte, $\mathcal{S}_1 = \langle \mathcal{N}_1, \wedge, ', R_1 \rangle$ et $\mathcal{S}_2 = \langle \mathcal{N}_2, \wedge, ', R_2 \rangle$ sont des *fragments* de $\mathcal{N}$, c'est-à-dire des sous-structures pour lesquelles $\mathcal{N}_1, \mathcal{N}_2 \subset \mathcal{N}$ et $R_i \subset R$ est une sous-relation qui isole les implications pertinentes à $\mathcal{N}_i$ (Lindahl et Odelstad 2000, p.269). Deux fragments peuvent être joints par le biais de deux manières, en l'occurrence la *connexion* et l'*assemblage* (*coupling*).

À supposer deux fragments $\mathcal{S}_1$ et $\mathcal{S}_2$, une *connexion* a lieu lorsqu'il y a $a \in \mathcal{N}_1$ et $b \in \mathcal{N}_2$ tels que aRb et un *assemblage* est tel qu'il y a une et une seule connexion entre $\mathcal{S}_1$ et $\mathcal{S}_2$. En fait, l'*assemblage* se distingue de la connexion dans la mesure où $\mathcal{S}_1$ et $\mathcal{S}_2$ sont uniquement connectés de par le l.u.b. de $\mathcal{S}_1$ et le g.l.b. de $\mathcal{S}_2$. En ce sens, dans le cas d'une *connexion*, il y a des liens directs entre les conditions de $\mathcal{S}_1$ et celles de $\mathcal{S}_2$, alors que dans le cas de l'*assemblage*, toute relation est indirecte et requiert le passage par un intermédiaire (voir la représentation graphique dans Lindahl et Odelstad 2000, pp.267-8).

L'assemblage permet de mettre l'ensemble $\mathcal{S}_1$ en relation avec $\mathcal{S}_2$ alors que la connexion ne met qu'une partie de $\mathcal{S}_1$ en relation avec $\mathcal{S}_2$. Considérons la figure 7.1, où a, $a*$ et b, $b*$ représentent respectivement des membres de $\mathcal{S}_1$ et $\mathcal{S}_2$. À gauche, la flèche entre a et b représente une connexion à partir de laquelle il est seulement possible de lier $a \vee a*$ et a à b et $b*$. En ce sens, la connection ne permet de lier qu'une partie des deux fragments. Notons qu'il pourrait très bien y avoir une autre connexion de $a*$ à $b*$. Dans le cas de droite, la flèche représente plutôt un assemblage puisqu'il s'agit de la seule connexion, laquelle permet de lier chaque membre de $\mathcal{S}_1$ avec ceux de $\mathcal{S}_2$.

L'idée de Lindahl et Odelstad est donc de représenter un système normatif complexe, par exemple la législation canadienne, par le biais d'un pré-ordre booléen $\mathcal{S}$ que l'on peut diviser en fragments, comme le Code criminel, le Code civil, la Charte des droits et libertés, etc., lesquels peuvent eux aussi être divisés en fragments, et ainsi de suite. Cela dit, chaque fragment peut lui aussi être divisé en deux fragments, l'un descriptif et l'autre normatif. En ce sens, la législation canadienne consiste en la jonction d'un fragment descriptif et d'un normatif, de même pour le Code criminel, le Code civil, etc. Selon cette conception, un système normatif peut donc être divisé de plusieurs manières afin d'obtenir différents fragments dépendamment de ce que l'on cherche à étudier. Dans l'ensemble, un système normatif correspond à la jonction d'un fragment qui contient des « vérités » descriptives (ou du moins des conventions, ce qui est pris pour acquis) avec un fragment qui contient

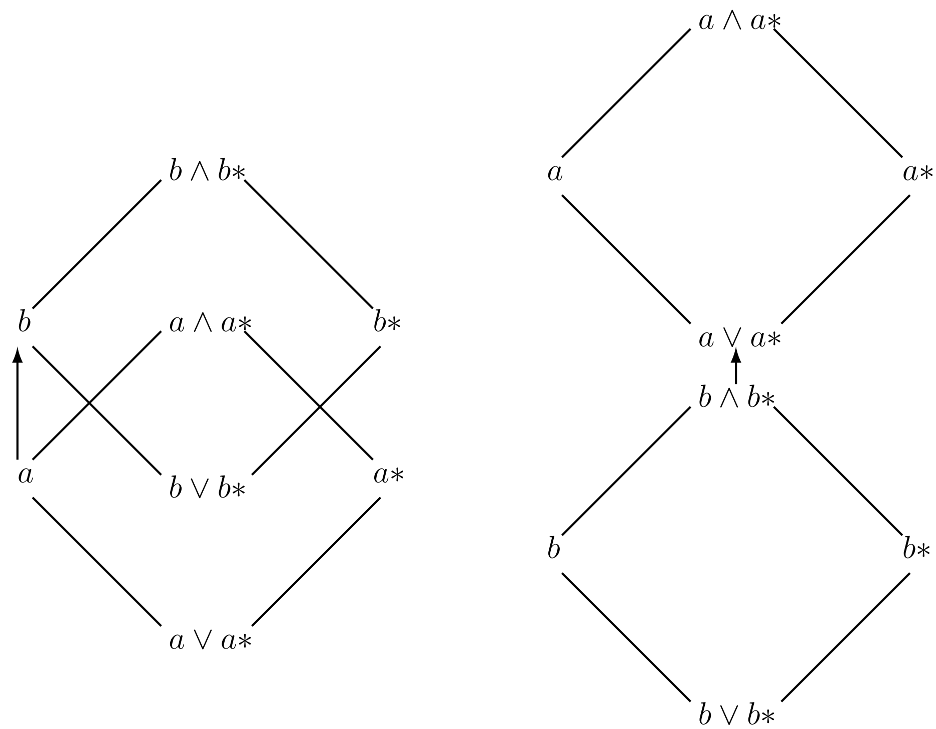


Figure 7.1: Connexion et assemblage

des normes.

Il s'agit là d'un portrait brossé à gros traits. Pour rendre justice à l'approche de Lindahl et Odelstad, il faudrait plutôt comprendre un système normatif comme la jonction d'un fragment descriptif, d'un fragment de corrélations normatives (qui font le pont entre le descriptif et le normatif) et d'un fragment purement normatif. Dans la suite de leur article, les auteurs utilisent leur conception d'un système normatif comme pré-ordre booléen afin d'étudier les notions de *concepts intermédiaires* et de *conséquence légale*, laquelle se comprend en fonction de la relation R transitive qui permet de passer d'un fragment à l'autre. La notion de concept intermédiaire se comprend en parallèle avec la notion de *count-as* chez Jones et Sergot ou chez Boella et van der Torre (cf. chapitre 6).

À titre d'exemple de concept intermédiaire, considérons l'énoncé suivant: tout citoyen canadien a le droit de vote. Autrement dit, si x est un citoyen canadien, alors x a le droit de vote. Mais qu'est-ce qui *compte pour* (*count-as*) un citoyen canadien? Notamment, le fait que x soit née au Canada de parents ayant la citoyenneté canadienne et qu'il réside au Canada compte pour « x est un citoyen canadien ». Dans ce cas, « citoyen canadien » est un concept intermédiaire entre les conditions et le fait que x ait le droit de vote.

Le lecteur intéressé par cette approche trouvera dans Lindahl et Odelstad (2004) une combinaison des pré-ordres booléens avec une forme de logique de l'action. Dans Lindahl et Odelstad (2002), les normes sont conçues comme des corrélations normatives, c'est-à-dire comme les connections entre les fragments purement descriptifs et les fragments purement normatifs. Lindahl et Odelstad (2003) utilisent les pré-ordres booléens afin d'étudier les changements de normes ainsi que la préservation de la cohérence du système normatif dans les cas d'obligations contraires au devoir. Lindahl et Odelstad (2006) reprennent ce qui a été présenté dans Lindahl et Odelstad (2000) par rapport aux concepts intermédiaires et un exemple légal plus complexe est donné dans Lindahl et Odelstad (2008b). Les travaux de Lindahl et Odelstad (2006, 2008b) sont repris dans Lindahl et Odelstad (2008a, 2011). Finalement, Lindahl et Odelstad (2004) reprennent l'analyse des concepts intermédiaires dans le cadre algébrique afin de la vulgariser et rendre le tout pertinent pour les juristes.

Krister Segerberg

En plus de ses travaux en logique déontique dynamique, Segerberg (1982b) a aussi développé la notion de *modèle urne*, basée sur une algèbre déontique d'action. À l'instar de son approche en logique déontique dynamique, une action est considérée en parallèle avec l'ensemble des séquences d'états qui mènent à un certain résultat. De même, une distinction est faite entre les *actions* et les *propositions*.

Le langage $\mathcal{L} = \{Act, \smile, \frown, ^{-}, =, P, F, \neg, \supset\}$ distingue entre les *termes*, qui réfèrent aux actions, et les *propositions* (Segerberg 1982b, p.269). Le langage contient un ensemble dénombrable d'actions atomiques (d'évènements) ainsi que les constantes $\mathbf{0}$ et $\mathbf{1}$, c'est-à-dire que $Act = \{\mathbf{0}, \mathbf{1}, \alpha_1, \ldots, \alpha_n, \ldots\}$, où $\mathbf{0}$ et $\mathbf{1}$ sont respectivement les actions impossible et universelle. Le langage contient les symboles $\smile$, $\frown$ et $^{-}$, qui représentent respectivement l'union, la conjonction et la négation d'action, le symbole $=$ pour l'identité entre les actions, les opérateurs P et F pour la permission et l'interdiction, et finalement les connecteurs usuels de la logique propositionnelle classique à l'aide desquels les autres connecteurs sont définis.

L'ensemble des expressions bien formées est défini récursivement de la manière suivante (où $Prop$ est un ensemble de propositions atomiques):

1. si $\alpha \in Act$, alors $\alpha \in Act*$;
2. si $\alpha, \beta \in Act*$, alors $\overline{\alpha}, \alpha \smile \beta, \alpha \frown \beta \in Act*$;
3. si $\alpha, \beta \in Act*$, alors $\alpha = \beta, P\alpha, F\alpha \in Prop$;
4. si $A \in Prop$, alors $A \in EBF$;
5. si $A, B \in EBF$, alors $\neg A, A \supset B \in EBF$.

Une structure $\mathcal{S} = \{\mathsf{A}, \frown, \smile, -, 0, 1, F, P\}$ ($\mathsf{A} \subseteq Act$) est qualifiée d'*algèbre déontique d'action* lorsque celle-ci respecte les trois conditions suivantes:

1. $\mathcal{S} = \{\mathsf{A}, \frown, \smile, -, 0, 1\}$ est une algèbre de Boole;
2. F et P sont des ensembles d'alternatives idéales;

3. $F \cap P = \{\mathbf{0}\}$ (i.e., rien n'est à la fois permis et interdit).

Les conditions sémantiques sont représentées par la fonction $v : Act \longrightarrow \{0, 1\}$ qui attribue des valeurs de vérité aux actions atomiques. Les clauses sémantiques sont (cf. Segerberg 1982b, p.270):

$$\begin{aligned}
v(\mathbf{0}) &= 0 \\
v(\mathbf{1}) &= 1 \\
v(\alpha \frown \beta) &= v(\alpha) \frown v(\beta) \\
v(\alpha \smile \beta) &= v(\alpha) \smile v(\beta) \\
v(\overline{\alpha}) &= -v(\alpha) \\
v(\alpha = \beta) &= \begin{cases} \top & \text{si } v(\alpha) = v(\beta) \\ \bot & \text{sinon} \end{cases} \\
v(F\alpha) &= \begin{cases} \top & \text{si } v(\alpha) \in F \\ \bot & \text{sinon} \end{cases} \\
v(P\alpha) &= \begin{cases} \top & \text{si } v(\alpha) \in P \\ \bot & \text{sinon} \end{cases}
\end{aligned}$$

Quelques explications s'imposent. Le cas des deux premières clauses est relativement simple: pour toute algèbre déontique d'action, $v(\mathbf{0})$ est l'objet initial 0 et $v(\mathbf{1})$ est l'objet terminal 1. Du point de vue de l'action, α est une action *tautologique*, c'est-à-dire *universelle*, lorsque $v(\alpha) = 1$ et α est une action *contradictoire* (*impossible*) lorsque $v(\alpha) = 0$. Dans les autres cas, il s'agit d'actions *satisfiables*, c'est-à-dire d'actions qui ont le potentiel d'être réalisées ou non. Cela dit, le lecteur aura remarqué que $v(\alpha)$ n'est pas défini pour les actions atomiques. La première chose à noter est qu'il ne faut pas y réfléchir en termes ensemblistes: dans les ensembles, nous dirions que $v(\alpha)$ correspond à l'ensemble qui contient les scénarios où l'action α est accomplie. Ici, ce n'est pas tout à fait juste. Un parallèle qui peut être fait (et qui sera explicité sous peu lorsque nous aborderons les modèles déontiques d'actions) est que l'action est considérée comme un *évènement*, et donc que celle-ci est informellement comprise comme référant à une multitude de séquences de descriptions qui peuvent être interprétées comme « l'action α ». Toutefois, dans le cas d'une algèbre déontique d'action, la clause sémantique ne signifie pas simplement que $v(\alpha) = 1$ lorsque $\alpha \in \mathsf{A}$. Plutôt, il faut concevoir (informellement) une infinité de valeurs intermédiaires: $v(\alpha)$ est égal à *quelque chose*, ce quelque chose n'étant pas réductible à 0 ou 1 dans le cas des actions atomiques. Toujours informellement, ce quelque chose peut être entendu comme un ensemble de regroupements de points (où chaque regroupement constitue une exécution de α). En ce sens, l'attribution faite par v à

1. l'action impossible est 0;

2. l'action universelle est 1;

3. la conjonction d'actions correspond aux regroupements de points conjoints de chaque action;

4. la disjonction d'actions correspond à au moins un regroupement de points d'une action;

5. la négation d'actions correspond aux regroupements de points autres que ceux qui correspondent à α.

Soulignons que cette explication en termes de regroupements de points est informelle et a une visée purement pédagogique. Il ne s'agit pas de ce qui est présenté dans l'article de Segerberg. En fait, il s'agit d'une interprétation de ce qui est présenté, interprétation principalement motivée par le fait qu'aucune explication relativement à $v(\alpha)$ pour une action atomique n'y est présentée. Cette interprétation provient notamment du parallèle qui est fait entre l'algèbre déontique d'action et le modèle déontique d'actions. Pour l'instant, poursuivons l'analogie avec les regroupements de points. Considérons la figure 7.2. Selon l'analogie présentée jusqu'à présent, une action peut être informellement représentée comme un ensemble de regroupements de points. En ce sens, une action atomique α peut être vue comme dénotant les regroupements a,b,c et d. Pour les fins de notre exemple, supposons que α correspond aux regroupements a et b et que β correspond à c et d. La conjonction d'actions $\alpha \frown \beta$ correspond à l'accomplissement simultané d'un regroupement (p. ex., a et b et c et d). La disjonction d'actions $\alpha \smile \beta$ correspond à l'accomplissement de l'un des regroupements pour l'une des actions (p. ex., a ou b, ou c ou d). Quant à la négation $\overline{\alpha}$, il s'agit de tous les regroupements autres que a et b. En ce sens, $v(\alpha)$ pour une action atomique correspond à un ensemble de regroupements de points dans A.

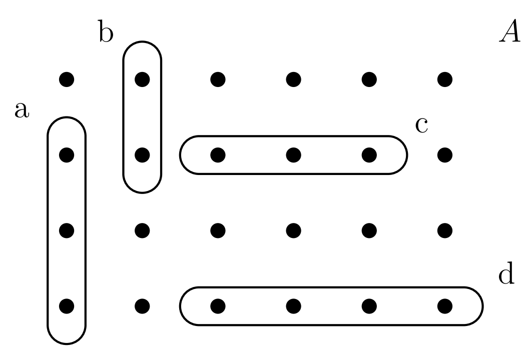


Figure 7.2: Actions et regroupements de points

Alors que dans le cas de l'attribution faite par v pour les actions on ne peut pas réellement parler de *valeur de vérité*, v assigne les valeurs $\bot$ (faux) et $\top$ (vrai) pour les propositions atomiques. La sixième clause signifie qu'il est vrai que $\alpha = \beta$ à condition que $v(\alpha) = v(\beta)$. Considérant que $v(\alpha)$ pour une action atomique réfère à l'ensemble des regroupements de points qui correspondent à l'action α, deux actions atomiques α et β sont dites identiques lorsque $v(\alpha)$ renvoit exactement aux mêmes regroupements que $v(\beta)$. Les clauses pour la permission et l'interdiction sont à comprendre de la même manière. À supposer que F et P renvoient à des ensembles d'actions interdites et permises, $F\alpha$ (respectivement, $P\alpha$) est vrai lorsque $v(\alpha)$ est membre de F (respectivement, de P). Dans ce cas, l'interprétation de v comme une fonction qui attribue des valeurs de vérité serait étrange: une action est interdite lorsque sa valeur de vérité est membre d'un ensemble de valeurs de vérité interdites? De fait, v est plutôt à comprendre comme une fonction qui, concernant des actions, détermine des ensembles de points (lesquels représentent des actions). Il est vrai que α est interdite à condition que l'ensemble qui contient les regroupements de points qui correspondent à des exécutions de α, soit $v(\alpha)$, est membre d'un ensemble qui contient des regroupements de points interdits. Dans le cas des propositions complexes formées par le biais des connecteurs propositionnels, l'attribution faite par v se fait usuellement.

Une proposition A est dite *vraie* pour $\mathcal{S}$ lorsque $v(A) = \top$ et elle est *valide* lorsqu'elle est vraie pour tout $\mathcal{S}$ et tout v (Segerberg 1982b, p.270).

Outre l'interprétation algébrique proposée par Segerberg, ce dernier construit aussi un modèle en termes ensemblistes (cf. Segerberg 1982b, p.271). Soit $\mathbb{S} = \langle U, \mathcal{I}ll, \mathcal{L}eg\rangle$ où $U \neq \varnothing$ est un ensemble de points (d'états) et $\mathcal{I}ll$ et $\mathcal{L}eg$ sont des sous-ensembles de U qui ne partagent aucun membre. Informellement, les sous-ensembles de $\mathcal{I}ll$ représentent des ensembles illégaux de points alors que ceux de $\mathcal{L}eg$ représentent des ensembles légaux de points. Soulignons que $\mathcal{L}eg$ et $\mathcal{I}ll$ ne sont pas le complément l'un de l'autre dans U. Un modèle $\mathbb{M} = \langle U, \mathcal{I}ll, \mathcal{L}eg, V\rangle$ est construit à l'aide d'une fonction d'attribution de valeurs de vérité V qui se comporte dans le même esprit que v pour les algèbres déontiques d'actions. C'est d'ailleurs ici que le parallèle susmentionné devient évident relativement à l'interprétation informelle de ce qu'est une action: une action est considérée comme un *évènement*, c'est-à-dire comme un ensemble de *résultats* (*outcome*) qui prennent place suite à une séquence d'états subséquents (Segerberg 1982b, p.271). La fonction V est définie de manière similaire à v.

$$\begin{aligned}
V(\mathbf{0}) &= \varnothing \\
V(\mathbf{1}) &= U \\
V(\alpha \frown \beta) &= V(\alpha) \cap V(\beta) \\
V(\alpha \smile \beta) &= V(\alpha) \cup V(\beta) \\
V(\overline{\alpha}) &= U - V(\alpha) \\
V(\alpha = \beta) &= \begin{cases} \top & \text{si } V(\alpha) = V(\beta) \\ \bot & \text{sinon} \end{cases} \\
V(F\alpha) &= \begin{cases} \top & \text{si } V(\alpha) \subseteq \mathcal{I}ll \\ \bot & \text{sinon} \end{cases} \\
V(P\alpha) &= \begin{cases} \top & \text{si } V(\alpha) \subseteq \mathcal{L}eg \\ \bot & \text{sinon} \end{cases}
\end{aligned}$$

Autrement dit, l'action impossible et l'action universelle sont représentées respectivement par l'ensemble vide et l'univers en entier. La conjonction d'actions se représente par le biais de l'intersection entre les résultats possibles pour chaque action. En d'autres termes, la conjonction d'actions correspond aux séquences qui mènent à des résultats communs à chacune des actions. La disjonction d'actions correspond à l'union entre les ensembles qui correspondent à la réalisation de chacune des actions. La négation d'une action équivaut au complément dans U de l'action elle-même. Les clauses pour les propositions atomiques sont similaires à celles de v: deux actions sont identiques si elles renvoient exactement aux mêmes ensembles de points, une action est interdite lorsque les ensembles de points auxquels elle renvoit sont illégaux et une action est permise lorsque les ensembles de points auxquels elle renvoit sont légaux. Encore une fois, les conditions pour les connecteurs propositionnels sont usuelles. Tout comme dans le cas de v, un énoncé est dit vrai pour un modèle lorsque $V(A) = \top$ et valide lorsque vrai pour tout modèle.

En ce qui concerne la syntaxe, le système axiomatique proposé est assez direct. Segerberg (1982b, pp.271-2) n'utilise que la règle *modus ponens*, les schémas d'axiomes de la logique propositionnelle classique, les axiomes qui correspondent à une algèbre de Boole pour les actions (distributivité et complémentation), les axiomes usuels pour l'identité.

$$\alpha = \alpha \tag{Id}$$
$$\alpha = \beta \supset (A \supset A(\alpha|\beta)) \tag{Subst}$$

Le premier axiome exprime l'identité de chaque action avec elle-même et le second indique que si deux actions sont identiques, alors elles peuvent être substituées dans une proposition. Finalement, il y a les axiomes pour les opérateurs déontiques:

$$F(\alpha \smile \beta) \equiv (F\alpha \wedge F\beta) \tag{1}$$
$$P(\alpha \smile \beta) \equiv (P\alpha \wedge P\beta) \tag{2}$$
$$a = \mathbf{0} \equiv (F\alpha \wedge P\alpha) \tag{3}$$

L'axiome (1) exprime que si une disjonction d'actions est interdite, alors chaque action est interdite individuellement. (2) représente la notion de permission libre de choix (*free choice permission*), à savoir que si une disjonction d'actions est permise, alors chaque action prise individuellement est permise. Finalement, (3) signifie que si une action est impossible, alors elle est à la fois permise et interdite.

Dans la suite de son article, Segerberg étudie les systèmes clos où $\mathcal{I}ll$ est le complément de $\mathcal{L}eg$ dans U (et donc où toute action est explicitement permise ou interdite). De plus, ce dernier propose aussi deux définitions distinctes sur la base d'un seul opérateur primitif (cf. Segerberg 1982b, p.279). Par exemple, il propose de définir une forme d'obligation et de permission sur la base de l'opérateur F, où

$$O_F\alpha =_{def} F\overline{\alpha} \tag{def. O_F}$$
$$P_F\alpha =_{def} \neg F\alpha \tag{def. P_F}$$

Lorsque définis ainsi, les opérateurs se comportent comme les opérateurs usuels en logique déontique standard. Contrairement à la permission définie par P qui exprime un choix libre entre deux actions, la permission définie par P_F est une permission faible (implicite, négative, qui dénote ce qui n'est pas interdit).

Somme toute, Segerberg propose une interprétation intéressante de la logique déontique en termes de ce qui *doit être fait*, où les opérateurs déontiques portent sur les actions plutôt que sur les propositions. Comme nous le verrons cependant au chapitre 14, il est toutefois peu plausible qu'une algèbre de Boole soit le meilleur candidat pour modéliser adéquatement l'action humaine.

Robert Trypuz et Piotr Kulicki

Les travaux de Segerberg (1982b) ont notamment influencé ceux de Trypuz et Kulicki (2009).[4] Leur objectif est de construire diverses extensions et ainsi comparer les travaux de Segerberg avec l'approche plus récente de Castro et Maibaum (2009). Dans ce qui suit, nous allons nous limiter à présenter la syntaxe ainsi que la conception que Trypuz et Kulicki (2009) se font de l'action atomique.

Conformément aux travaux de Segerberg, Trypuz et Kulicki distinguent entre les *actions* et les *propositions*. Soit le langage $\mathcal{L} = \{(,), GA, 0, 1, \bar{}, \sqcap, \sqcup, \neg, \wedge, \bot, P, F\}$. La logique de l'action est formalisée par une algèbre de Boole et est construite à partir des connecteurs usuels. Comme nous le verrons plus bas, GA est un ensemble de générateurs d'action. L'ensemble des expressions bien formées pour la logique de l'action est défini récursivement par (Trypuz et Kulicki 2009, p.255):

$$\alpha := a_i \mid 0 \mid 1 \mid \overline{\alpha} \mid \alpha \sqcap \beta \mid \alpha \sqcup \beta$$

Les connecteurs sont interprétés usuellement comme la négation (le complément), la conjonction et la disjonction d'action. L'ordre partiel est défini par:

$$\alpha \sqsubseteq \beta =_{def} (\alpha \sqcap \beta) = \alpha$$

Cette logique de l'action se combine à une logique propositionnelle construite à partir de l'ensemble d'expressions bien formées défini récursivement par:

$$\varphi := \top \mid \alpha = \beta \mid P(\alpha) \mid F(\alpha) \mid \neg\varphi \mid \varphi \wedge \psi$$

Les autres connecteurs propositionnels sont définis de manière usuelle. L'opérateur P représente la permission forte alors que F dénote l'interdiction. La permission faible est définie par $P_w(\alpha) =_{def} \neg F(\alpha)$.

Le système $\mathcal{DAL}^0$ est construit à partir de $\mathcal{L}$, des axiomes du calcul propositionnel, des axiomes d'identités (AI1) et (AI2), et des axiomes (D1), (D2) et (D3) pour les opérateurs déontiques (Trypuz et Kulicki 2009, p.257).

[4] Ce texte est repris dans Trypuz et Kulicki (2010).

$$\alpha = \alpha \tag{AI1}$$
$$(\alpha = \beta) \supset (\varphi \supset \varphi(\alpha/\beta)) \tag{AI2}$$
$$P(\alpha \cup \beta) \equiv P(\alpha) \wedge P(\beta) \tag{D1}$$
$$F(\alpha \cup \beta) \equiv F(\alpha) \wedge F(\beta) \tag{D2}$$
$$(\alpha = 0) \equiv P(\alpha) \wedge F(\beta) \tag{D3}$$

Plusieurs extensions sont construites sur la base de $\mathcal{DAL}^0$ (cf. Trypuz et Kulicki 2009, pp.262-8) et la sémantique est inspirée des travaux de Segerberg (cf. Trypuz et Kulicki 2009, pp.259-261).

L'intérêt de cette approche s'aperçoit notamment lorsque l'on considère la conception que les auteurs se font de l'action atomique. Trypuz et Kulicki (2009, p.256) distinguent enre une *action atomique* et un *générateur d'action*. Alors qu'un *générateur d'action* est membre de l'ensemble fini $GA = \{a_1, \ldots a_n\}$, une *action atomique* est composée de générateurs d'action:

> We shall think about the action generators as *the bricks from which all actions are made up*. [...] Every atom is a combination of all action generators and has the form $\delta_1 \sqcap \cdots \sqcap \delta_n$ where δ_k is a generator or its complement (Trypuz et Kulicki 2010, p.134).

Cette distinction est importante dans la mesure où les auteurs soutiennent qu'une action est la description complète de ce qu'un agent accomplit:

> In other words an atom of action algebra is *a complete description of what an agent is doing in a certain situation* (Trypuz et Kulicki 2010, p.134).

Bien que cette perspective puisse être utile en informatique (l'utilisation d'un ensemble fini de générateurs permet la description finie d'un programme), celle-ci est cependant douteuse d'un point de vue philosophique lorsque l'on considère l'action humaine. Cela peut s'apercevoir sous deux angles.

D'une part, la *description complète* de ce qu'un agent accompli actuellement est impossible, surtout considérant que les auteurs incluent les compléments d'action au sein de la description. Une telle description (finie) est impossible dans la mesure où il y a un nombre dénombrable d'actions que nous n'accomplissons pas, et ce même si l'on accepte un ensemble fini de générateurs: prenons simplement une action a_i qui n'est pas accomplie. Si a_i n'est pas accomplie, alors la conjonction de a_i avec a_i n'est pas accomplie, au même titre que la conjonction de a_i avec a_i avec a_i, etc. Par ce raisonnement, on peut aisément montrer qu'il y a un nombre dénombrable d'actions qui ne sont pas accomplies. Par conséquent, si la description complète de ce qu'un agent accomplit requiert la description complète de ce qu'il n'accomplit pas, alors une description complète (finie) de ce qu'un agent accomplit est impossible puisqu'il y a un nombre dénombrable d'actions que l'agent n'accomplit pas.[5]

D'autre part, cette restriction est *ad hoc*: malgré qu'un agent ne puisse accomplir qu'un nombre fini d'actions, pourquoi supposer qu'il n'y a qu'un nombre fini de générateurs d'action? Même si cela n'est pas explicitement mentionné chez Trypuz et Kulicki, probablement que cette hypothèse vise l'application de leur système en informatique. En supposant un nombre fini de générateurs d'action, on peut faire une description complète de chaque action atomique pouvant être accomplie par le programme, et ainsi une description complète du programme est possible, incluant ce qu'il peut faire, ce qu'il doit faire et ce qu'il ne doit pas faire. Cela dit, du point de vue de l'action humaine, cette hypothèse est superflue dans la mesure où le nombre d'actions possibles pouvant être accomplies par un agent est tellement grand qu'une description totale de ce que l'agent fait est actuellement impossible. Prenons simplement l'ensemble des endroits possibles dans l'univers où quelqu'un peut se trouver. La description complète de « Paul est dans son bureau » requiert que l'on spécifie partout où Paul ne se trouve pas, ce qui, lorsque l'on considère l'univers en entier, entraîne beaucoup de possibilités! Par surcroît, il faudra spécifier tout ce que Paul ne fait pas à chacune de ces positions, et par conséquent l'hypothèse d'un nombre fini de générateurs d'action perd son utilité considérant que la description totale ne pourra pas être actuellement accomplie.

Par ailleurs, les auteurs affirment qu'un générateur d'action s'interprète comme l'ensemble des états possibles qu'il est susceptible d'engendrer:

> From those definitions it is clear that every action generator is interpreted as a set of (its) possible outcomes (Trypuz et Kulicki 2009, p.259).

Il s'agit là de la même idée que l'on retrouve chez Segerberg: une action renvoit à l'ensemble des résultats possibles qui peuvent en découler. Bien qu'elle soit inoffensive en elle-même, cette conception peut cependant difficilement être jumelée à l'autre hypothèse de Trypuz et Kulicki, à savoir qu'une action peut être décrite dans sa totalité. En effet, si l'action se résume à l'ensemble de résultats possibles et que l'action peut être décrite complètement, alors il s'ensuit que l'ensemble des résultats possible peut, lui aussi, être décrit complètement. Or, l'ensemble des résultats possibles est une totalité, et une totalité ne peut être décrite complètement que si chacune de ses parties l'est aussi, sans quoi une partie de la totalité ne serait pas décrite complètement, et de fait la totalité ne le serait pas non plus. Par conséquent, il s'ensuit que si une action peut être décrite complètement, alors chaque état peut aussi être décrit complètement. Rappelons-nous ici qu'un *état* dans l'interprétation ensembliste correspond à un scénario et que l'action est interprétée

[5] Rappelons-nous que l'objectif de cette thèse est d'utiliser la logique déontique afin d'étudier la structure des inférences légales. La présente critique doit donc être comprise en fonction de l'application que nous visons. Il ne s'agit pas de soutenir que l'approche de Trypuz et Kulicki (2009) est inadéquate en informatique, mais plutôt de montrer que celle-ci ne peut pas simplement être importée et appliquée à l'analyse des inférences légales. Les auteurs ont en tête l'analyse des actions d'un *programme*. Or, cette analyse n'est pas simplement transférable aux actions *humaines*, qui doivent nécessairement être prise en compte lors de la modélisation du discours légal.

comme un regroupement d'états (de points). Le problème avec ce raisonnement est que, comme le souligne Hansson (1969, p.376), une description *totale* d'un scénario est au mieux problématique, sinon peu plausible.

Un état correspond à une description de l'univers à un certain moment. Il s'agit, en un certain sens, d'une *photo*: en un clique, on fige l'univers et on en obtient une description. Or, malgré que la description complète de l'univers dans le cas d'une application informatique puisse être faite, cela pose problème lorsque l'on considère des systèmes plus complexes. Prenons simplement le Canada. Est-il possible de faire une description totale du Canada au moment présent? Cela est logiquement possible, certes, mais, pour reprendre le raisonnement de Hempel (1972, p.17), cela est irréaliste d'un point de vue pratique: il faudrait notamment décrire l'emplacement de chaque grain de sable, de chaque roche, de chaque insecte, de chaque bactérie, etc. Au mieux, nous ne pouvons avoir qu'une description partielle d'un état du système. Considérant que l'utilité de l'hypothèse à savoir que les actions atomiques sont en nombre fini vise la possibilité (actuelle) de faire une description totale du système, et que, considérant l'action humaine, il est actuellement (dans les faits, mais non pas logiquement) impossible de décrire la totalité d'un système, il en résulte que cette hypothèse est inutile à la modélisation de l'action humaine.

En dernier lieu, soulignons un dernier problème avec la conception de l'action atomique que se font Trypuz et Kulicki. La notion de générateur d'action n'est qu'intuitive et mal définie: un générateur d'action est une brique à partir de laquelle une action atomique est construite. L'exemple que donne les auteurs est celui des générateurs *être assis* (*sit*) et *fumer* (*smoke*). Les quatre actions atomiques possibles construites à partir de ces générateurs sont:

$$\begin{gathered} sit \sqcap smoke \\ sit \sqcap \overline{smoke} \\ \overline{sit} \sqcap smoke \\ \overline{sit} \sqcap \overline{smoke} \end{gathered}$$

Prenons simplement le cas du générateur *sit*, qui est un « fondement à partir duquel il est possible de construire des actions atomiques ».[6] Nous allons montrer que ce fondement n'est en fait pas si fondamental et que le critère qui permet de différencier un générateur de ce qui n'est pas un générateur n'est pas si évident. En fait, nous allons montrer qu'une action peut toujours être divisée en sous-actions. L'action de boire un verre d'eau, par exemple, se résume à lever le bras, serrer les doigts autour du verre, lever le verre, ouvrir la bouche, pencher le verre afin de faire couler l'eau dans la bouche, avaler l'eau, etc.

Serait-il possible de trouver un générateur d'action qui lui-même ne soit pas réductible à des sous-générateurs d'action? L'intuition derrière ce raisonnement est que, considérant qu'une action implique un mouvement, le seul bloque indivisible possible d'une action est une description statique de chaque moment de l'action dans le temps. Pensons

[6] Soulignons d'ailleurs que l'usage de l'expression action *atomique* est curieux considérant qu'une action atomique est divisible en générateurs d'action.

à l'action de serrer les doigts autour du verre d'eau: au temps 1 les doigts sont immobiles, au temps 2 ils ont bougé un peu, au temps 3 un peu plus, et ainsi de suite. Mais si l'on considère que l'action est un mouvement, ce qui implique le passage d'un moment à un autre, alors conceptuellement il est toujours possible de penser à une subdivision de ce mouvement en intervalles plus petits.

Le problème ici est que l'action, en tant que totalité, n'est pas réductible à la somme de ses descriptions. Une action peut être décrite, certes, mais cette description ne capture pas conceptuellement ce que signifie pour une action d'être une action. Revenons au générateur *sit*. Est-ce réellement un bloque indivisible? Si l'on pense seulement au *résultat* de l'action de s'asseoir, alors *être assis*, pris en tant que description d'un état statique, est effectivement un bloque indivisible. Cela dit, une action est quelque chose de dynamique: une action est quelque chose qui est *fait*, et la notion de *faire quelque chose* requiert une évolution dans le temps. En ce sens, l'action *être assis* ne se résume pas à une description statique: l'action *être assis* est une action qui est faite par un agent à partir d'un temps t_1 jusqu'à un temps t_n. Si l'on réfléchit au générateur *sit*, on se rend rapidement compte que l'action d'*être assis* nécessite plusieurs sous-actions. Par exemple, les muscles du dos doivent travailler afin de préserver la posture assise de l'agent, les muscles des jambes travailleront afin de les garder dans une certaine position, cela sans compter toutes les autres actions vitales nécessaires à l'accomplissement de l'action *sit*.

Par ailleurs, l'action d'un muscle peut elle-même être analysée en termes d'autres actions: le processus de transformation d'énergie des cellules, la génération d'influx nerveux, les sous-mouvements nécessaires à l'accomplissement du mouvement total, etc. Bien que informellement on comprenne la conception que Trypuz et Kulicki se font d'une action atomique, il n'en demeure pas moins que leur notion de *générateur d'action* est vague et mal définie. Où trace-t-on la ligne afin de déterminer les bloques fondamentaux? L'action *sit*, prise comme bloque de départ pour construire des actions atomiques, peut elle-même être conçue comme action atomique construite à partir d'autres générateurs d'action plus précis.

Ces arguments seraient en toute probabilité rejetés par la communauté informatique étant donné que la réduction d'une action à une séquence de descriptions (une suite d'états du système) est fondamentale à la programmation. Mais cela ne fait qu'exemplifier les différences qui se trouvent entre les diverses applications de la logique déontique. Est-ce que la logique déontique vise l'analyse du discours, ou vise-t-elle plutôt la programmation?

En informatique, la logique déontique a une visée *pratique*. On veut des résultats et des applications. On veut être en mesure de déterminer le comportement d'une base de données lorsque sujette à certaines contraintes (p. ex., Carmo et al. 2001; Carmo et Jones 1996), de formaliser les ententes contractuelles (p. ex., Boella et van der Torre 2004; Brown 2005; Prisacariu et Schneider 2012), de gérer le comportement et l'évolution d'un système informatique (p. ex., Boella et van der Torre 2003a), de gérer l'évolution des obligations qui sont sujettes à des échéances (p. ex., Broersen et al. 2004; Demolombe 2014; Dignum et al. 2005; Dignum et Kuiper 1997, 1998), ou encore de construire des robots capables d'agir conformément à certaines normes (p. ex., Bringsjord et Taylor 2012; Bringsjord et al. 2011). Certains vont même jusqu'à vouloir formaliser un système légal de façon à pouvoir déterminer sans équivoque le résultat normatif d'une situation lorsque l'on entre certains faits avec certaines normes.

Le lecteur est invité à consulter Wieringa et Meyer (1993) pour un survol des applications de la logique déontique en informatique. Dans chacun des cas, la logique déontique a une visée pratique, et malgré les lacunes conceptuelles de la conception de l'action on justifie la réduction de l'action à une description (une séquence de descriptions d'états) par un argument pragmatique. En ce qui nous concerne, le principal objectif de cette thèse n'est pas *pratique* mais bien *théorique*: l'objectif est d'utiliser les outils de la logique contemporaine afin de clarifier la structure conceptuelle propre au discours juridique, et ce par l'analyse de la relation de conséquence dans les inférences normatives. L'action sera analysée plus en détails au chapitre 14. Pour le moment, passons à l'approche de Castro et Maibaum.

Pablo Castro et Tom Maibaum

L'approche de Castro et Maibaum (2009) présente une logique dynamique construite sur la base d'une logique de l'action booléenne, d'où sa présentation au sein de ce chapitre. Soit $Prop = \{p_1, \ldots, p_n, \ldots\}$ un ensemble dénombrable de propositions atomiques et $Act = \{a_1, \ldots, a_n\}$ un ensemble fini d'actions atomiques. Soit le langage

$$\mathcal{L} = \{(,), Prop, Act, 0, 1, \bar{}, \sqcap, \sqcup, =, \neg, \supset, [\,], P, P_w, \top, \bot\}$$

avec les définitions usuelles pour les autres connecteurs propositionnels. Au niveau des actions, les expressions bien formées sont définis récursivement par:

$$\alpha := a_i \mid 0 \mid 1 \mid \overline{\alpha} \mid \alpha \sqcap \beta \mid \alpha \sqcup \beta$$

Quant aux propositions, elles sont définies par (Castro et Maibaum 2009, p.442):

$$\varphi := p_i \mid \bot \mid \top \mid \alpha = \beta \mid \neg\varphi \mid \varphi \supset \psi \mid [\alpha]\varphi \mid P(\alpha) \mid P_w(\alpha)$$

L'obligation est définie par:

$$O(\alpha) =_{def} P(\alpha) \wedge \neg P_w(\overline{\alpha})$$

Ici, P_w représente la permission faible et P la permission forte. Une action est obligatoire lorsqu'elle est explicitement permise et que sa négation n'est pas faiblement permise (non interdite).

Le système axiomatique est construit à partir des axiomes du calcul propositionnel et d'une algèbre de Boole pour les actions.[7] L'opérateur [ ] est axiomatisé de manière usuelle par une modalité de type K, et les axiomes pour les autres opérateurs sont:

[7] L'axiomatisation proposée dans Castro et Maibaum (2009) est redondante. Par conséquent, nous nous sommes limité à ne présenter que ce qui est pertinent.

$$[0]\varphi \tag{A1}$$
$$[\alpha \sqcup \beta]\varphi \equiv [\alpha]\varphi \wedge [\beta]\varphi \tag{A2}$$
$$[\alpha]\varphi \supset [\alpha \sqcap \beta]\varphi \tag{A3}$$
$$P(0) \tag{A4}$$
$$P(\alpha \sqcup \beta) \equiv P(\alpha) \wedge P(\beta) \tag{A5}$$
$$P(\alpha) \vee P(\beta) \supset P(\alpha \sqcap \beta) \tag{A6}$$
$$\neg P_w(0) \tag{A7}$$
$$P_w(\alpha \sqcup \beta) \equiv P_w(\alpha) \vee P_w(\beta) \tag{A8}$$
$$P_w(\alpha \sqcap \beta) \supset P_w(\alpha) \wedge P_w(\beta) \tag{A9}$$
$$(P(\alpha) \wedge \alpha \neq 0) \supset P_w(\alpha) \tag{A10}$$

En plus de cela, on trouve des règles de substitution pour les égalités et des contraintes sur l'élément terminal (cf. Castro et Maibaum 2009, p.446).

(A1) stipule qu'après une action impossible toute proposition est vraie et (A2) assure que lorsqu'une action disjointe mène à un état alors chaque membre de l'action mène individuellement l'état. (A3) implique que si un état est le résultat d'une action, alors ce sera aussi le résultat de la combinaison de cette action avec une autre action. (A4) stipule que l'action impossible est fortement permise alors que (A7) indique qu'elle n'est pas faiblement permise (d'où l'impossibilité de l'accomplir, voir les remarques de Castro et Mailbaum sur ce point). (A5) exprime qu'une disjonction d'action est fortement permise lorsque chaque action l'est individuellement et (A6) signifie que lorsque soit α ou β est fortement permise, alors leur conjonction l'est aussi. (A8) et (A9) signifient respectivement qu'une disjonction d'action est faiblement permise lorsque l'une ou l'autre des actions l'est et que si une conjonction est faiblement permise, alors chaque action prise individuellement l'est aussi. Finalement, (A10) stipule que si une action est fortement permise et qu'elle n'est pas impossible, alors elle est aussi faiblement permise.

Le lecteur est invité à consulter Castro et Maibaum (2009, pp.443-4) pour la sémantique, qui est construite dans les ensembles de manière usuelle à l'aide du complément, de l'intersection et de l'union pour les opérateurs en prenant l'ensemble des scénarios où une proposition est vraie. Par ailleurs, notons que l'approche de Castro et Maibaum (2009) a été reprise dans Castro et Maibaum (2012).

Thierry Lucas

Thierry Lucas est l'un des rares logiciens ayant appliqué explicitement la théorie des catégories à la logique déontique. La logique qu'il propose s'inspire des travaux de von Wright et est construite à partir du langage:

$$\mathcal{L} = \{Prop, (,), \neg, \wedge, Ac(-,-,-), \Box\}$$

$Prop = \{p_1, ..., p_n, ...\}$ est un ensemble dénombrable d'atomes propositionnels descriptifs et $\neg$ et $\wedge$ sont des connecteurs booléens (Lucas 2006, p.87). Une action atomique $Ac(\varphi, \chi, \psi)$ est la description de trois états: l'état présent φ, l'état χ qui résulterait des actions d'un agent et l'état ψ qui résulterait advenant que l'agent n'agisse pas. Les énoncés dans la portée de $Ac(\varphi, \chi, \psi)$ sont uniquement propositionnels et appartiennent à l'ensemble défini récursivement de la manière suivante:

$$\varphi := \ p_i \mid \neg\varphi \mid \varphi \wedge \psi$$

L'ensemble des expressions bien formées est défini par:

$$\varphi := \ p_i \mid a \mid \neg\varphi \mid \varphi \wedge \psi \mid \Box\varphi$$

Les autres connecteurs sont définis de manière usuelle (avec a une action atomique). Les axiomes pour l'action sont:

$$Ac(\varphi_1 \vee \varphi_2, \chi, \psi) \equiv Ac(\varphi_1, \chi, \psi) \vee Ac(\varphi_2, \chi, \psi) \tag{A1}$$
$$Ac(\varphi, \chi_1 \vee \chi_2, \psi) \equiv Ac(\varphi, \chi_1, \psi) \vee Ac(\varphi, \chi_2, \psi) \tag{A2}$$
$$Ac(\varphi, \chi, \psi_1 \vee \psi_2) \equiv Ac(\varphi, \chi, \psi_1) \vee Ac(\varphi, \chi, \psi_2) \tag{A3}$$
$$Ac(\varphi_1 \wedge \varphi_2, \chi_1 \wedge \chi_2, \psi_1 \wedge \psi_2) \equiv Ac(\varphi_1, \chi_1, \psi_1) \wedge Ac(\varphi_2, \chi_2, \psi_2) \tag{A4}$$
$$\varphi \equiv Ac(\varphi, \top, \top) \tag{A5}$$
$$\neg Ac(\varphi, \bot, \psi) \tag{A6}$$
$$\neg Ac(\varphi, \chi, \bot) \tag{A7}$$

Et les règles comprennent le modus ponens ainsi que:

$$\frac{\varphi \equiv \varphi', \chi \equiv \chi', \psi \equiv \psi'}{Ac(\varphi, \chi, \psi) \equiv Ac(\varphi', \chi', \psi')}$$

L'opérateur $\Box$ est un opérateur de nécessité axiomatisé par KM.

L'axiome (A1) indique que si l'action est telle que la disjonction $\varphi_1 \vee \varphi_2$ est vraie pour l'état présent, que χ sera vrai si l'action est accomplie et que ψ sera vrai si elle ne l'est pas, alors soit $Ac(\varphi_1, \chi, \psi)$ est vrai ou $Ac(\varphi_2, \chi, \psi)$ est vrai. (A2) exprime que si l'action est telle que φ est vrai, que $\chi_1 \vee \chi_2$ serait vrai advenant que l'action soit accomplie, et que ψ le serait si elle ne l'était pas, alors soit $Ac(\varphi, \chi_1, \psi)$ est vrai ou $Ac(\varphi, \chi_2, \psi)$. Dans le même ordre d'idées, (A3) stipule que si l'action est telle que φ est vrai, χ serait vrai si l'action est accomplie et $\psi_1 \vee \psi_2$ serait vrai si elle ne l'était pas, alors soit $Ac(\varphi, \chi, \psi_1)$ est vrai ou $Ac(\varphi, \chi, \psi_2)$ l'est. Le quatrième axiome permet d'éliminer les conjonctions à l'intérieur de la portée de Ac, et le cinquième indique qu'une description φ équivaut à l'action décrite par $Ac(\varphi, \top, \top)$. Finalement, (A6) indique qu'il est faux qu'il y a une action telle que φ est vrai mais que $\bot$ serait réalisé advenant que l'action soit accomplie, et donc il est impossible de faire une action contradictoire, au même titre que (A7) exprime qu'il est faux qu'il y a une action telle que $\perp$ serait réalisé advenant que l'action ne soit pas accomplie. La règle permet la substitution d'équivalences.

Lucas (2006, p.88) définit un modèle sémantique $\mathcal{M} = \langle U, Ag, R, M\rangle$ par:

$$\begin{aligned}
&a : U \longrightarrow U\\
&n : U \longrightarrow U\\
&R \subseteq U \times U\\
&M : U \times Prop \longrightarrow \{0, 1\}
\end{aligned}$$

La fonction a représente l'action d'un agent, à savoir la transition d'un état à un autre, tandis que la fonction n représente les « actions » de la *nature*, c'est-à-dire la transition d'un état à un autre advenant l'absence d'action de la part de l'agent. La relation d'accessibilité R entre les scénarios est réflexive et M est une assignation de valeurs de vérité qui assigne à chaque paire <scénario,proposition> une valeur de vérité. Les conditions de vérité pour les énoncés propositionnels sont définies usuellement et la condition de vérité pour les actions est:

$$\models_w Ac(\varphi, \chi, \psi) \text{ ssi } \models_w \varphi \text{ et } \models_{a(w)} \chi \text{ et } \models_{n(w)} \psi$$

Autrement dit, l'action $Ac(\varphi, \chi, \psi)$ est vraie pour w si et seulement si φ est vrai pour w, χ est vrai pour les scénarios accessibles à w qui correspondent aux scénarios qui résultent des actions de l'agent, et ψ est vrai pour les scénarios accessibles à w qui correspondent aux actions qui résultent de la nature.

L'originalité de l'approche de Lucas (2006, p.101) est que ce dernier considère les actions en tant que *morphismes* dans le cadre de la théorie des catégories. En effet, il entend les actions comme des flèches allant d'un ensemble de conditions initiales à un ensemble de résultats.

Lucas (2008, p.370) définit une algèbre d'action comme un triplet $\langle B, \mathcal{ACT}, C\rangle$ avec B et C deux algèbres de Boole (possiblement différentes l'une de l'autre) et $\mathcal{ACT} = \langle A, 0, 1, \neg, \wedge, \vee, C_0, \rangle$ un ensemble partiellement ordonné se comportant comme une logique bi-intuitionniste (Lucas 2006, p.86). (La notion de co-support C_0 est introduite pour représenter les conditions dans lesquelles la négation d'une action est obtenue.) Supposant que $\mathcal{F}_1$ et $\mathcal{F}_2$ sont respectivement les ensembles d'expressions bien formées de B et C, Lucas (2006, p.102) conçoit une action comme un morphisme $a : \Sigma \rightarrow \mathcal{F}_2$, avec $\Sigma \subseteq \mathcal{F}_1$. En ce sens, une action est vue comme un morphisme d'un ensemble de conditions à un ensemble de résultats. Les actions doivent satisfaire une condition de cohérence telle que $a(\sigma) = a(\sigma')$ pour $\sigma, \sigma' \in \Sigma$, c'est-à-dire que l'action a donne le même résultat pour toute sous-condition $\sigma \in \Sigma$. En d'autres termes, les sous-conditions sont cohérentes entres-elles, c'est-à-dire que deux sous-conditions d'un ensemble Σ donnent les mêmes résultats.

Les actions sont pré-ordonnées par $\leq$ (Lucas 2008, p.369), 1 est l'action vide et 0 l'action nulle. L'action vide correspond à ne rien faire, alors qu'il est impossible d'accomplir l'action nulle.

Sur la base de $(A, 0, 1, \wedge, \vee)$, Lucas (2006, p.106) définit $(A, 0, 1, \wedge, \vee, \sim, \rightarrow)$ comme un système déductif intuitionniste avec $\rightarrow$ l'adjoint de $\wedge$ et $\sim$ la négation intuitionniste, définie par $\sim \alpha = \alpha \rightarrow 0$ (voir aussi Lucas 2008, p.383). Cela dit, Lucas ajoute aussi un adjoint à la disjonction en plus d'une négation v. L'adjoint à la disjonction est gouverné par (cl') et la négation est définie par:

$$v\alpha = 1 \backslash \alpha$$

:

$$\frac{\gamma \leq \alpha \vee \beta}{\gamma \backslash \beta \leq \alpha} \qquad \text{(cl')}$$

C'est en vertu de cette seconde négation que $\mathcal{ACT}$ est défini comme une logique bi-intuitionniste. Une logique bi-intuitionniste est entendue comme une algèbre de bi-Heyting, c'est-à-dire un treillis distributif qui comprend une algèbre de Heyting et une algèbre de co-Heyting.

Lucas (2008, p.370) utilise une troisième négation afin de dualiser les connecteurs, en introduisant les connecteurs duaux $(-)^*$ à l'aide de:

$$\begin{aligned}
&\alpha \leq^* \beta \Leftrightarrow \neg\beta \leq \neg\alpha \\
&\alpha \leq \beta \Leftrightarrow \neg\beta \leq^* \neg\alpha \\
&\alpha \wedge^* \beta = \neg(\neg\alpha \vee \neg\beta) \\
&\alpha \vee^* \beta = \neg(\neg\alpha \wedge \neg\beta) \\
&1^* = \neg 0 \\
&0^* = \neg 1 \\
&\alpha \backslash^* \beta = \neg(\neg\beta \rightarrow \neg\alpha) \\
&v^* \alpha = \neg \sim \neg\alpha \\
&\alpha \rightarrow^* \beta = \neg(\neg\beta \backslash \neg\alpha) \\
&\sim^* \alpha = \neg v \neg\alpha
\end{aligned}$$

Ainsi, $\wedge$ et $\wedge^*$ distribuent sur $\vee$. L'interprétation de ces connecteurs n'est cependant pas explicitée.

Outre les avancées faites par Lucas (2006) et reprises dans Lucas (2007), Lucas (2008) propose d'utiliser sa logique de l'action afin de construire une logique déontique. En un mot, l'idée est de définir l'obligation de α comme la nécessité que la description des conditions entraîne les résultats. Autrement dit, une action est obligatoire lorsque les conditions entraînent nécessairement les résultats. Cette conception est à entendre dans les mêmes lignes qu'en logique modale, où la relation induite sur le modèle sémantique par l'axiomatisation de l'obligation permet d'isoler les scénarios où l'obligation est réalisée. En ce sens, il s'agit d'une sémantique idéale: on met de côté les scénarios où l'obligation n'est pas réalisée. De même, chez Lucas, la conception de l'obligation est telle que la vérité des conditions entraînera nécessairement la vérité du résultat. Comme Lucas (2008, p.79) le mentionne, sa conception de l'obligation permet de dériver plusieurs conséquences du système standard, notamment l'agrégation, la distribution de O sur l'implication, (ROM) et le paradoxe de Ross.

Un point intéressant de l'approche de Lucas est que celui-ci adopte la distinction courante en logique dynamique entre la logique de l'action et la logique propositionnelle. En effet, l'idée fondamentale derrière son approche est que la structure algébrique des actions est fondamentalement différente de la structure algébrique des propositions que l'on utilise afin de parler de ces actions. Il s'agit là d'un principe auquel nous adhérons et qui, comme nous le verrons, sera au fondement de l'analyse proposée au chapitre 14.

* * *

Ceci clos notre la revue de la littérature portant sur les approches algébriques en logique déontique. Au total, outre l'approche de Lucas (qui n'a d'ailleurs malheureusement pas eu énormément d'impact au sein de la littérature), on voit aisément que toute approche en logique déontique qui repose sur une algèbre d'action utilise en fait une algèbre de Boole. Au chapitre 14, nous analyserons plus en détails d'autres logiques de l'action, qui ne sont pas nécessairement utilisées en vue de la logique déontique. Pour l'instant, passons à la logique non-monotone avant de conclure cette revue de la littérature.

Chapitre 8

Logique non-monotone

La logique déontique non-monotone constitue une branche importante de la logique déontique. Ses principales applications se trouvent en informatique théorique et en programmation, plus précisément en intelligence artificielle, où l'on cherche à modéliser le raisonnement et l'apprentissage. La logique non-monotone est pertinente en intelligence artificielle considérant qu'elle permet de modéliser les raisonnements qui se font à la lumière d'une information spécifique, où la conclusion du raisonnement peut varier lorsque l'information change. L'objectif de ce chapitre est de familiariser le lecteur avec la logique déontique non-monotone. Notre objectif n'est pas de dresser un portrait exhaustif de la littérature sur le sujet, mais plutôt de fournir au lecteur le matériel nécessaire à la compréhension des approches que l'on y trouve, ainsi que des enjeux fondamentaux qui guident cette discipline. Le lecteur intéressé par la logique déontique non-monotone est invité à consulter l'ouvrage édité par Nute (1997).

La première section expose les principales motivations qui guident l'introduction des logiques déontiques non-monotones. Les notions de monotonie et de non-monotonie y sont expliquées, et les arguments en faveur de la non-monotonie des inférences normatives sont exposés. Par la suite, l'approche de Horty (1997) est présentée à titre d'exemple d'approche non-monotone en logique déontique. Ce choix a été réalisé en fonction du fait que l'approche de Horty est très claire et est moins axée vers ses applications en informatique, ce qui rend le texte plus abordable pour le lecteur qui provient de la philosophie plutôt que de l'informatique théorique. Finalement, différentes notions de défaisabilité sont présentées à la dernière section, conformément à la proposition de van der Torre et Tan (1997).

Motivations

La logique non-monotone vise la formalisation des inférences quotidiennes (ou ordinaires), qui se font à la lumière d'une information qui est plus souvent qu'autrement incomplète. Ce type d'approche a une incidence directe en intelligence artificielle, où l'objectif est de programmer des machines capables de raisonner à l'instar de l'homme. Ainsi, la logique non-monotone permet la représentation formelle des inférences où les conclusions sont adaptées en fonction de l'information que l'on a. Elle permet à une machine de « changer d'idée ». En termes de input et de output, la logique non-monotone permet de faire varier le output dépendamment du input, ce qui n'est pas possible dans les cadres de travail monotones. Outre ses applications en informatique et en intelligence artificielle, la logique déontique non-monotone vise aussi la représentation des inférences légales, où certaines conclusions sont parfois renversées, notamment lorsque celles-ci sont prises à partir d'information incomplète.

Grossièrement, l'idée de base derrière les cadres de travail non-monotones est de fournir des systèmes qui permettent de modifier les conclusions auxquelles on arrive en fonction de l'information que l'on possède. Une relation de conséquence est dite *monotone* lorsque l'ajout de certaines prémisses à une inférence valide ne change pas la conclusion. Autrement dit, une relation de conséquence monotone respecte la règle structurale d'affaiblissement (i.e., *weakening*, cf. Gentzen 1934):

$$\frac{A \vdash B}{A, C \vdash B} \quad \text{(wk)}$$

Si B est une conséquence de A et que la relation de conséquence est monotone, alors B sera aussi la conséquence d'un ensemble de prémisses qui contient A. Une relation de conséquence est donc monotone lorsque celle-ci est stable et produit toujours le même résultat à partir d'un ensemble de prémisses. Lorsque la relation de conséquence est monotone, l'ajout de certaines prémisses ne change pas la conclusion.

Certains raisonnements quotidiens ne respectent cependant pas ce schéma d'inférence. En effet, certaines inférences peuvent être *défaites*, voire *déconstruites*. Dans certains cas, l'ajout d'information change la conclusion à laquelle on parvient. Cela peut être mis en évidence à l'aide des deux exemples suivants. Considérons d'abord les énoncés qui suivent.

$p =$ Il pleut.
$q =$ Paul va au cinéma.
$r =$ Marie appelle Paul.
$s =$ Paul va chez Marie.

Supposons que Paul planifie sa journée et qu'il en arrive à la conclusion que *s'il pleut, alors il va au cinéma, mais cependant si Marie appelle, alors il ira chez elle.* Supposons maintenant que cette information est transmise à Pierre, et qu'en plus on lui dit que cette journée là il pleuvait. On demande à Pierre ce qu'a fait Paul. Son raisonnement sera alors

$$p \supset q, r \supset s, p \vdash q$$

et ce dernier conclura que Paul est allé au cinéma. Cela dit, si par la suite on apprend à Pierre que Marie a appelé Paul, alors son raisonnement sera

$$p \supset q, r \supset s, p, r \vdash s$$

et il conclura plutôt que Paul est allé chez Marie. De plus, si on lui demande si Paul est allé au cinéma, alors Pierre répondra que non puisque Paul est allé chez Marie. De fait, l'ajout de l'information r à l'ensemble de prémisses $\{p \supset q, r \supset s, p\}$ fait changer la conclusion. En ce sens, l'inférence de Pierre n'est pas monotone considérant que

$$p \supset q, r \supset s, p \vdash q$$

mais

$$p \supset q, r \supset s, p, r \nvdash q$$

et donc que la règle (wk) n'est pas respectée.

Par ailleurs, la non-monotonie s'aperçoit aussi de par le fait que certaines inférences ne respectent pas la règle (cut), qui représente la transitivité de la relation de conséquence.[1]

$$\frac{A \vdash B \qquad B \vdash C}{A \vdash C} \quad \text{(cut)}$$

Cette règle indique que lorsque B est une conséquence de A et que C est une conséquence de B, alors nécessairement il en résulte que C est une conséquence de A. Cependant, certaines inférences quotidiennes ne respectent pas cette règle. Prenons par exemple les propositions suivantes.[2]

$p =$ Paul est québécois.
$q =$ La langue maternelle de Paul est le français.
$r =$ Paul est né en France.

Supposons que l'on transmet l'information suivante à Pierre:

1. La langue maternelle de la plupart des québécois est le français.

2. La plupart des français sont nés en France.

Pierre acceptera donc comme prémisses que $p \supset q$ et que $q \supset r$. Toutefois, ce n'est pas parce que Pierre accepte ces deux prémisses qu'il acceptera aussi nécessairement que $p \supset r$. Malgré que, de manière générale, si une personne est québécoises, alors sa langue maternelle est le français, et que dans la plupart des cas si la langue maternelle d'un individu est le français alors celui-ci est né en France, cela n'implique pas que si un individu est québécois alors celui-ci est né en France.

[1] Soulignons cependant que ce type d'argument n'est pas présenté en faveur de la non-monotonie des inférences normatives.

[2] L'exemple est adapté de Antonelli (2012).

Dans le cas de la logique déontique, l'introduction d'un cadre de travail non-monotone peut être justifiée de deux manières. Alors que certains utilisent la logique déontique non-monotone afin de pouvoir représenter les conflits d'obligations, comme Ryu et Lee (1997) et Horty (1997), d'autres l'utlisent plutôt afin de pouvoir représenter formellement les obligations conditionnelles, par exemple van der Torre et Tan (1997).

Prenons le système standard à titre d'exemple. Considérant l'axiome (**D**), qui est propositionnellement équivalent à

$$\neg(OA \wedge O\neg A)$$

il en résulte que dans une situation où il y a un conflit d'obligations, alors n'importe quelle action est obligatoire puisque:

$$\vdash_{KD} (OA \wedge O\neg A) \supset OB$$

Il s'agit là de l'explosion déontique. Si, par exemple, Paul a le devoir d'aider sa patrie en allant à la guerre, mais qu'il a aussi le devoir moral de rester à la maison pour aider sa mère qui est paraplégique, alors Paul est dans une situation où il est dans l'obligation d'aller à la guerre et de ne pas aller à la guerre. Au sein de KD, cela entraîne que Paul est dans l'obligation de voler une banque, ce qui, évidemment, n'est pas un résultat désiré.

Par ailleurs, la non-monotonie des inférences normatives s'aperçoit de par le fait que celles-ci ne respectent pas le schéma d'inférence (wk). Le lecteur est référé au chapitre 15 pour une analyse détaillée des arguments en faveur de la logique déontique non-monotone. Ayant maintenant vu les principales motivations pour l'introduction de la logique déontique non-monotone, passons maintenant à l'approche de Horty (1997), qui nous servira à titre de paradigme afin d'exemplifier la stratégie derrière les cadres de travail non-monotones.

John Horty

L'introduction des relations de conséquences non-monotones vise la représentation des inférences *ordinaires* ou *quotidiennes*, où les conclusions auxquelles on parvient varient en fonction de l'information que nous avons. En plus de viser à représenter comment les individus réfléchissent au quotidien, ces relations de conséquences sont aussi appliquées dans le domaine de l'intelligence artificielle, où l'objectif est de faire le tri dans l'information fournie afin d'obtenir les conclusions désirées. Ainsi, on s'assure que les machines soient en mesure d'atteindre certains de leurs buts même lorsque que ceux-ci sont contradictoires. Un des principaux objectifs qui a motivé l'introduction des logiques non-monotones était de développer une logique qui soit capable de rendre compte des inférences humaines afin de pouvoir servir à des fins de programmation en intelligence artificielle (Horty 1997, p.23).

Dans le contexte de la logique déontique, les logiques non-monotones sont pertinentes lorsque l'on cherche à formaliser les conflits d'obligations ou les obligations conditionnelles.

Le texte de Horty (1997) offre une introduction quant aux motivations qui guident les développements de la logique déontique non-monotone (voir aussi Horty 1994). D'emblée, ce dernier considère que la logique non-monotone est pertinente lorsque l'on cherche à formaliser les conflits d'obligations. Un conflit d'obligations survient lorsque deux normes dictent des actions incompatibles, c'est-à-dire qui ne peuvent être réalisées simultanément. Afin de mettre en évidence l'importance des fondements non-monotones pour une logique déontique, Horty (1997, p.18) insiste sur le fait que les conflits d'obligations ne sont pas possibles au sein du système standard. En vertu de l'axiome (**D**), il est impossible d'avoir une situation telle que $OA \wedge O\neg A$ est vrai, et par conséquent le système standard ne permet pas de rendre compte des conflits d'obligations. Considérant que les conflits d'obligations sont courants, Horty en conclut qu'une logique déontique digne de ce nom se doit de rendre compte de telles situations.

Comme il le mentionne (Horty 1997, p.20), on trouve déjà des approches monotones au sein de la littérature qui permettent de rendre compte des conflits d'obligations. Le système proposé par Chellas (1974), par exemple, permet la représentation des conflits d'obligations. La proposition de Chellas est de modéliser les obligations conditionnelles à l'aide d'une logique modale non-normale plus faible que le système K, construite comme une extension de la logique propositionnelle à laquelle on ajoute la règle d'inférence (ROM) et l'axiome $\neg O\bot$.[3] Les conflits d'obligations sont possibles au sein d'un tel système dans la mesure où, contrairement au système standard, il n'y a pas d'équivalence entre (**D**) et $\neg O\bot$, ce qui résulte du fait que l'agrégation $(OA \wedge OB) \supset O(A \wedge B)$ n'est pas dérivable.

Or, Horty (1997, p.22) soutient que le rejet complet de l'agrégation rend le système de Chellas *trop* faible. En effet, malgré qu'il soit souhaitable de rejeter l'énoncé $(OA \wedge O\neg A) \supset O(A \wedge \neg A)$, sans quoi nous obtiendrions $(OA \wedge O\neg A) \supset OB$ puisque $O(A \wedge \neg A) \supset OB$ en vertu de (ROM) et du fait que $(A \wedge \neg A) \supset B$, l'agrégation est néanmoins parfois désirable, notamment lorsqu'il n'y a pas de conflit.

Afin de mettre ce point en évidence, Horty (1997, p.21) propose l'argument suivant, qui pour être validé dans une approche comme celle de Chellas doit utiliser une instance du principe d'agrégation. Supposons qu'un agent est dans l'obligation de soit payer ses taxes en un seul paiement ou de les payer en deux versements. Supposons maintenant que cet agent est dans l'obligation de ne pas payer ses taxes en un seul paiement en raison de contraintes financières familiales. Dans une telle situation, nous voulons conclure que l'agent est dans l'obligation de payer ses taxes en deux versements. Cependant, afin que ce raisonnement soit valide dans le système de Chellas, il faudrait être en mesure de passer de $O(p \vee q) \wedge O\neg p$ à $O((p \vee q) \wedge \neg p)$ à l'aide de l'agrégation afin de pouvoir ensuite utiliser (ROM) et le fait que $((p \vee q) \wedge \neg p) \supset q$ pour conclure que Oq. Cependant, l'agrégation n'est pas disponible au sein du système de Chellas, et par conséquent ce dernier n'est pas en mesure de valider l'argument susmentionné.

La conclusion à laquelle arrive Horty est que le principe d'agrégation ne doit pas permettre de conclure un conflit d'obligations, mais doit néanmoins pouvoir être utilisé

[3] Pour une présentation détaillée, voir Chellas (1980) sections 6.5 et 10.2 et Peterson (2011) section 4.4.

lorsque les conflits sont prévenus. La solution proposée par Horty (1997, p.22) est inspirée de la définition de la conséquence déontique chez van Fraassen (1973, p.17). Soit Γ un ensemble de propositions qui contient des énoncés du type OA et $\mathcal{M}$ un modèle de la logique classique. Le *pointage* d'un modèle $\mathcal{M}$ relativement à un ensemble Γ est défini par:

$$score_\Gamma(\mathcal{M}) = \{OA \in \Gamma : \models_\mathcal{M} A\}$$

Autrement dit, l'ensemble $score_\Gamma(\mathcal{M})$ contient les propositions OA membres de Γ pour lesquelles A est vrai dans $\mathcal{M}$. Par exemple, si

$$\Gamma = \{Op, O\neg q, O(p \supset q)\}$$

alors il y a

$$\begin{aligned}
\mathcal{M}_1 &= \{p\} \\
\mathcal{M}_2 &= \{p, \neg q\} \\
\mathcal{M}_3 &= \{p, p \supset q\} \\
\mathcal{M}_4 &= \{\neg q\} \\
\mathcal{M}_5 &= \{\neg q, p \supset q\} \\
\mathcal{M}_6 &= \{p \supset q\}
\end{aligned}$$

tels que

$$\begin{aligned}
score_\Gamma(\mathcal{M}_1) &= \{Op\} \\
score_\Gamma(\mathcal{M}_2) &= \{Op, O\neg q\} \\
score_\Gamma(\mathcal{M}_3) &= \{Op, O(p \supset q)\} \\
score_\Gamma(\mathcal{M}_4) &= \{O\neg q\} \\
score_\Gamma(\mathcal{M}_5) &= \{O\neg q, O(p \supset q)\} \\
score_\Gamma(\mathcal{M}_6) &= \{O(p \supset q)\}
\end{aligned}$$

Cependant, il n'y a pas de modèle $\mathcal{M}$ qui satisfasse à la fois p, $\neg q$ et $p \supset q$. Maintenant, soit $|A| = \{\mathcal{M} : \models_\mathcal{M} A\}$ l'ensemble qui contient les modèles pour lesquels A est vrai. Par exemple, $|p| = \{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3\}$.

Ayant ces notions en mains, la conséquence déontique $\vdash_F$ est définie par:

$$\begin{aligned}
\Gamma \vdash_F OA \Leftrightarrow\ &\text{il y a } \mathcal{M}_1 \in |A| \text{ pour lequel il n'y a pas de} \\
&\mathcal{M}_2 \in |\neg A| \text{ t.q. } score_\Gamma(\mathcal{M}_1) \subseteq score_\Gamma(\mathcal{M}_2)
\end{aligned}$$

Conformément à l'exemple précédent, nous avons:

$$\begin{aligned}
score_\Gamma(\mathcal{M}_1) &\subseteq score_\Gamma(\mathcal{M}_2)\\
score_\Gamma(\mathcal{M}_1) &\subseteq score_\Gamma(\mathcal{M}_3)\\
score_\Gamma(\mathcal{M}_4) &\subseteq score_\Gamma(\mathcal{M}_2)\\
score_\Gamma(\mathcal{M}_4) &\subseteq score_\Gamma(\mathcal{M}_5)\\
score_\Gamma(\mathcal{M}_6) &\subseteq score_\Gamma(\mathcal{M}_3)\\
score_\Gamma(\mathcal{M}_6) &\subseteq score_\Gamma(\mathcal{M}_5)
\end{aligned}$$

et

$$\begin{aligned}
|q| &= \{\mathcal{M}_3\}\\
|\neg q| &= \{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_4, \mathcal{M}_5, \mathcal{M}_6\}
\end{aligned}$$

Selon cet exemple, on voit aisément que $\Gamma \vdash_F O\neg q$ puisque $\mathcal{M}_2 \in |\neg q|$ et $score_\Gamma(\mathcal{M}_2) \not\subseteq score_\Gamma(\mathcal{M}_3)$. De même, nous pouvons conclure que $\Gamma \vdash_F Oq$ puisqu'il n'y a pas de modèle dans $|\neg q|$ tel que $score_\Gamma(\mathcal{M}_3) \subseteq score_\Gamma(\mathcal{M}_i)$ pour $i \in \{1, 2, 4, 5, 6\}$. Cependant, nous avons $\Gamma \nvdash_F O(q \wedge \neg q)$ puisque $|q \wedge \neg q| = \varnothing$ et donc il n'y a pas de $\mathcal{M} \in |q \wedge \neg q|$ qui satisfait la condition.

L'objectif de Horty est donc de représenter ces idées dans le cadre d'une logique non-monotone. Pour ce faire, il utilise la logique des conditions de base (*default logic*) de Reiter (1980). Ici, l'idée est d'indexer une inférence de A vers B à une condition C. Autrement dit, plutôt que de représenter une inférence comme une paire (A, B), où A est une prémisse et B une conclusion, l'inférence est représentée par un triplet $(A, B|C)$ où la conclusion B ne peut être inférée à partir de A que lorsque cela est consistant avec C (Horty 1997, p.24).

Une *théorie de base* (*a default theory*) $\Delta = \langle \mathcal{W}, \mathcal{D} \rangle$ est définie comme un ensemble de propositions $\mathcal{W}$ auquel on ajoute un ensemble de règles $\mathcal{D}$ qui guident les inférences. Ayant une théorie de base à notre disposition, l'objectif est de définir ce qu'est une extension $\mathcal{E}$ pour une théorie Δ. Une extension $\mathcal{E}$ pour une théorie $\Delta = \langle \mathcal{W}, \mathcal{D} \rangle$ est définie par les trois conditions suivantes.

$$\mathcal{W} \subseteq \mathcal{E} \tag{8.1}$$

$$\mathcal{E} = Cn(\mathcal{E}) \tag{8.2}$$

$$\text{pour tout } (A, B|C) \in \mathcal{D},\ A \in \mathcal{E} \text{ et } \neg C \notin \mathcal{E} \Rightarrow\ B \in \mathcal{E} \tag{8.3}$$

Alors que la première condition stipule que $\mathcal{E}$ est une extension de $\mathcal{W}$, c'est-à-dire que toute formule de $\mathcal{W}$ est contenue dans $\mathcal{E}$, la seconde indique que l'extension contient l'ensemble de ses conséquences logiques. La troisième condition assure que l'extension contient les conclusions qui peuvent être dérivées à partir des règles de base (*default rules*), et donc pour chaque règle de la forme *B est dérivable à partir de A à condition que cela est consistant avec C*, on conclut que B est dans l'extension lorsque A s'y trouve et que rien ne permet de contredire C.

Cela fait, Horty (1997, p.28) en vient à faire le parallèle entre la relation de conséquence $\vdash_F$ et le cadre de travail non-monotone. En définissant la théorie de base Δ_Γ pour un ensemble d'obligations Γ par $\mathcal{W} = \varnothing$ et l'ensemble de règles de base par

$$\mathcal{D} = \{(\top, A|A) : OA \in \Gamma\}$$

Horty obtient que $\Gamma \vdash_F OA$ si et seulement si $A \in \mathcal{E}$ pour une extension $\mathcal{E}$ de Δ_Γ. En mots, ce résultat signifie qu'une obligation OA peut être inférée d'un ensemble d'obligations lorsque A est consistante avec les autres propositions au sein d'une extension de la théorie de base.

Afin d'illustrer cela, considérons l'exemple susmentionné. L'ensemble de règles de base de Δ_Γ est $\mathcal{D} = \{(\top, p|p), (\top, \neg q|\neg q), (\top, p \supset q|p \supset q)\}$. Cette théorie possède plusieurs extensions, dépendamment de l'ordre dans lequel elles sont construites. Par exemple, si l'on prend d'abord $(\top, p|p)$, alors nous avons $\top \in \mathcal{E}_1$ puisque $\mathcal{W} = \varnothing$ et $\top \in Cn(\varnothing)$, et puisque $\neg p \notin \mathcal{E}_1$ nous obtenons $p \in \mathcal{E}$. Ensuite, il y a deux possibilités. Si l'on considère d'abord $(\top, p \supset q|p \supset q)$, alors on obtient $p \supset q \in \mathcal{E}_1$ par le même raisonnement et puisque $\mathcal{E}_1 = Cn(\varnothing \cup \{p, p \supset q\})$, on obtient que $q \in \mathcal{E}_1$, voire que $\neg\neg q \in \mathcal{E}_1$. Dans cette optique, la condition (3) ne permet pas de conclure que $\neg q \in \mathcal{E}_1$ étant donné que $\neg\neg q \in \mathcal{E}$ (pour que ce soit le cas il faudrait que $\neg\neg q \notin \mathcal{E}_1$). Par ailleurs, si plutôt nous avions opté pour l'ajout de $(\top, \neg q|\neg q)$, alors nous aurions obtenu que $\neg q \in \mathcal{E}_2$ puisque $\top \in \mathcal{E}_2$ et $\neg\neg q \notin \mathcal{E}_2$, et par conséquent $\neg(p \supset q) \in \mathcal{E}_2$ puisque $\neg(p \supset q) \in Cn(\varnothing \cup \{p, \neg q\})$. Finalement, si l'on débute par considérer $(\top, p \supset q|p \supset q)$ et ensuite $(\top, \neg q|\neg q)$, on obtient une extension $\mathcal{E}_3 = Cn(\varnothing \cup \{p \supset q, \neg q\})$ pour laquelle $\neg p \in \mathcal{E}_3$. En ce sens, dépendamment de l'ordre dans lequel les propositions sont considérées, on peut obtenir différentes extensions pour une même théorie.

$$\begin{aligned}
\mathcal{E}_1 &= Cn(\varnothing \cup \{p, p \supset q\}) \\
\mathcal{E}_2 &= Cn(\varnothing \cup \{p, \neg q\}) \\
\mathcal{E}_3 &= Cn(\varnothing \cup \{p \supset q, \neg q\})
\end{aligned}$$

Certaines théories peuvent avoir aucune, une ou plusieurs extensions. Lorsqu'il y a plusieurs extensions disponibles, les deux stratégies possibles sont soit la stratégie *crédule*, qui consiste à prendre une extension au hasard et à accepter comme conclusion ce qui se trouve dans cette extension, ou la stratégie *sceptique*, qui consiste à n'accepter comme conclusion que ce qui fait partie de chaque extension (Horty 1997, p.26).

Le fait qu'une théorie de base puisse posséder plusieurs extensions amène Horty (1997, p.32) à définir une relation de conséquence déontique *sceptique* $\vdash_S$, pour laquelle une obligation OA peut être dérivée d'un ensemble d'obligations Γ seulement lorsque A appartient à toutes les extensions $\mathcal{E}$ de Δ_Γ. Autrement dit, la relation de conséquence est définie par $\Gamma \vdash_S OA$ si et seulement si $A \in \mathcal{E}$ pour tout $\mathcal{E}$ extension de Δ_Γ. Selon cette définition et conformément à l'exemple susmentionné, cela entraîne que même si

$$\begin{aligned}
&\Gamma \vdash_F Op \\
&\Gamma \vdash_F O\neg q \\
&\Gamma \vdash_F O(p \supset q)
\end{aligned}$$

nous avons cependant

$$\begin{aligned}
&\Gamma \nvdash_S Op \\
&\Gamma \nvdash_S O\neg q \\
&\Gamma \nvdash_S O(p \supset q)
\end{aligned}$$

puisque $\mathcal{E}_1 \cap (\mathcal{E}_2 \cap \mathcal{E}_3) = Cn(\varnothing)$. Cela dit, si nous avions plutôt $\Gamma' = \Gamma \cup \{Or\}$, alors

$$\begin{aligned}
\mathcal{E}'_1 &= Cn(\varnothing \cup \{p, p \supset q, r\}) \\
\mathcal{E}'_2 &= Cn(\varnothing \cup \{p, \neg q, r\}) \\
\mathcal{E}'_3 &= Cn(\varnothing \cup \{p \supset q, \neg q, r\})
\end{aligned}$$

et donc $\Gamma' \vdash_S Or$.

Outre la relation de conséquence déontique sceptique, Horty discute des relations qui se trouvent entre le système de Chellas, KD et la relation de conséquence $\vdash_F$. Par la suite, ce dernier aborde aussi la question des obligations conditionnelles, ce qui ne sera pas présenté considérant que les développements qu'il propose ne sont que préliminaires et sujets à plusieurs problèmes (cf. Horty 1997, p.40).

Pour conclure, soulignons que même si les approches non-monotones peuvent sembler souhaitable, celles-ci laissent néanmoins place à des résultats indésirables. Dans le cas de Horty, par exemple, autant $\vdash_F$ que $\vdash_S$ nous laissent dans une impasse.

D'un côté, l'interprétation de $\vdash_F$ nous laisse le choix entre les différentes extensions qui s'offrent à nous. Ainsi, conformément à l'exemple susmentionné, on peut autant choisir entre $\Gamma \vdash_F Op$ et $\Gamma \vdash_F O(p \supset q)$ que $\Gamma \vdash_F O\neg q$. Cependant, certaines obligations au sein d'un conflit sont prioritaires à d'autres, et par conséquent il est plausible que plutôt que d'avoir un choix à faire entre deux obligations, l'une domine l'autre. En ce sens, si l'objectif de la logique déontique non-monotone est de permettre de déterminer quelles sont les obligations en place lors d'un conflit, la stratégie crédule ne permet pas d'atteindre cet objectif puisqu'elle laisse le choix à l'agent. Dans le même ordre d'idées, la relation de conséquence déontique sceptique élimine simplement toutes les obligations qui sont liées au conflit et ne garde que les obligations qui n'y sont pas rattachées. Dans l'exemple précédent, nous n'avions ni $\Gamma \vdash_S Op$, ni $\Gamma \vdash_S O\neg q$, ni $\Gamma \vdash_S O(p \supset q)$.

Or, l'intérêt de pouvoir modéliser les conflits d'obligations est d'être en mesure de pouvoir déterminer les obligations actuelles d'un agent. De manière générale, un conflit d'obligations ne se résoud pas simplement en se débarassant des obligations qui sont en conflit les unes avec les autres. Lorsqu'un conflit d'obligations surgit, il est plutôt plausible que l'une des obligations domine l'autre. Certains, notamment Royakkers et Dignum (1997) et van der Torre et Tan (1997), répondent à cette intuition en introduisant une relation de préférence (de hiérarchie) entre les obligations. Ce type de solution est inspiré des travaux de Alchourrón et Makinson (1981).

Toutefois, il n'en demeure pas moins qu'il soit possible d'avoir une situation où les deux obligations sont incommensurables, comme demander à un parent ce qu'il préfère entre sauver l'un ou l'autre de ses enfants. Dans une telle situation, même si l'on accepte une hiérarchie entre les obligations, la relation de conséquence ne permettra pas de conclure ce qui devrait être fait.

En ce sens, même si l'approche de Horty permet de répondre à certaines objections qui peuvent être soulevées contre la logique déontique standard, son approche ne parvient pas à modéliser adéquatement les conflits d'obligations. Lorsqu'un conflit d'obligations prend place, la solution ne consiste pas simplement à *tout* rejeter, mais consiste plutôt, lorsque cela est possible, à sélectionner l'obligation qui devrait prendre place.

Différents types de défaisabilité

Ces considérations nous amènent aux différentes manières de défaire un argument.[4] Dans un article portant sur le traitement de l'obligation conditionnelle, van der Torre et Tan (1997, p.82) distinguent entre trois types de défaisabilité, notamment la *défaisabilité factuelle* (*factual defeasability*), la *défaisabilité de priorité forte* (*strong overridden defeasability*) et la *défaisabilité de priorité faible* (*weak overridden defeasability*).

La défaisabilité factuelle survient lorsqu'un fait contredit une règle de base. Prenons le fameux exemple de Tweety. Dans le raisonnement

P1	Les oiseaux volent.
P2	Tweety est un oiseau.
C	Tweety vole.

la règle de base *Les oiseaux volent* peut être défaite par l'ajout de l'information *Tweety ne vole pas.* Il s'agit donc d'un cas de défaisabilité factuelle dans la mesure où la règle de base est rejetée en vertu d'un fait.

La défaisabilité de priorité quant à elle survient lorsqu'une règle annule une autre. Dans le cas de Kant et du nazi, par exemple, la règle *Il faut toujours dire la vérité* peut être surpassée par la règle *il faut mentir si cela permet de sauver la vie de quelqu'un.* En ce sens, le raisonnement

P1	Il faut toujours dire la vérité.
C	Paul doit avouer au nazi qu'il cache une famille juive dans son grenier.

peut être défait en accordant priorité à la règle *il faut mentir si cela permet de sauver la vie de quelqu'un.* Comme le mentionnent van der Torre et Tan (1997, p.83), la défaisabilité de priorité est pertinente dans les cas de conflits d'obligations où l'une doit être retenue en faveur de l'autre.

[4] Faute d'une meilleure traduction, nous allons utiliser les termes défaire, défaisable ainsi que défaisabilité pour référer aux concepts de *defeasible reasoning* et de *defeasibility* en logique non-monotone.

La différence entre la défaisabilité de priorité *faible* et *forte* peut s'apercevoir à l'aide de la distinction entre l'*annulation* d'une obligation et son *surpassement contextuel* (*overshadowing*). Alors que la défaisabilité de priorité forte survient lorsqu'il y a nécessairement annulation (rejet), la défaisabilité de priorité est faible lorsqu'il y a possibilité de surpassement contextuel. Dans l'exemple susmentionné, l'obligation de dire la vérité n'est pas complètement rejetée, mais est seulement mise de côté en faveur d'une autre obligation qui a préséance sur celle-ci. En ce sens, il s'agit d'un cas de défaisabilité de priorité faible puisque le surpassement de l'obligation *il faut mentir pour sauver la vie de quelqu'un* sur *il faut dire la vérité* dépend du contexte dans lequel Paul se trouve.

Un cas de défaisabilité de priorité forte survient lorsqu'une règle est complètement abandonnée en faveur d'une autre. Cela peut se produire lorsque dans un certain contexte une obligation n'est simplement plus applicable.

La conclusion à laquelle parviennent les auteurs est que chaque type de défaisabilité s'applique à une situation précise. Alors que la défaisabilité factuelle permet de représenter les cas où certaines obligations sont violées, la défaisabilité de priorité forte permet de représenter les cas où certaines obligations sont plus spécifiques et la défaisabilité faible représente les obligations prima facie (van der Torre et Tan 1997, p.118).

Selon van der Torre et Tan (1997, p.98), les obligations contraires au devoir mettent en jeu des violations et doivent être modélisées à l'aide de défaisabilité factuelle. Par exemple, dans le cas du paradoxe de Chisholm, nous avons

$$O\neg p \tag{8.4}$$

$$O(\neg p \supset \neg q) \tag{8.5}$$

$$p \supset Oq \tag{8.6}$$

$$p \tag{8.7}$$

et dans le système standard cela permet de dériver $O\neg q$ et Oq. Dans cette situation, la défaisabilité de l'argument en faveur de Oq vient du *fait* que l'obligation $O\neg p$ a été violée. Ainsi, la défaisabilité du paradoxe de Chisholm en faveur de Oq vient du fait que p.

Les conflits d'obligations qui ne découlent pas d'obligations contraires au devoir, quant à eux, sont mieux représentés par la défaisabilité de priorité (van der Torre et Tan 1997, p.99). La défaisabilité de priorité faible représente les conflits d'obligations entre des obligations *actuelles* (toutes choses considérées) et des obligations prima facie dans la mesure où même si les obligations actuelles ont priorité dans un contexte donné, les obligations prima facie restent en force. Dans l'exemple du nazi, l'obligation de ne pas mentir reste en force, même si l'on opte plutôt pour mentir afin de sauver la vie de la famille juive.

Finalement, la défaisabilité de priorité forte s'applique lorsque les conflits d'obligations proviennent d'un conflit de normes. En effet, ce type de défaisabilité permet de représenter les cas où certaines règles sont rejetées en faveur d'autres règles plus spécifiques. Afin de mettre ce point en évidence, van der Torre et Tan (1997, p.115) utilisent l'exemple de la clôture, tel qu'exposé dans Prakken et Sergot (1996). Supposons un règlement municipal tel que:

1) il ne doit pas y avoir de clôture autour de la villa;

2) si la villa est au bord de la mer, alors il peut y avoir une clôture autour de la villa.

À supposer que la villa est au bord de la mer, cela est inconsistant au sein du système standard dans la mesure où nous obtenons $O\neg c$ et $\neg O\neg c$ en vertu de:

$$O\neg c \tag{8.8}$$

$$v \supset \neg O\neg c \tag{8.9}$$

$$v \tag{8.10}$$

Ce raisonnement peut cependant être défait par le biais d'une priorité forte dans la mesure où la règle (8.9) a priorité sur la règle (8.8) puisqu'elle est plus spécifique. Dans une telle situation, la règle (8.9) *annule* la règle (8.8), c'est-à-dire que, contrairement à la défaisabilité de priorité faible, l'obligation qu'il n'y ait pas de clôture n'est plus en place.

* * *

Somme toute, les arguments en faveur de fondations non-monotone pour la logique déontique opèrent de deux manières distinctes: soit on insiste sur le fait que les systèmes standards ne permettent pas de rendre compte des conflits d'obligations, soit on insiste sur le fait qu'ils ne sont pas en mesure de modéliser adéquatement les obligations conditionnelles et les conflits potentiels qui peuvent en résulter. Dans le chapitre qui suit, différents arguments contre les approches présentées jusqu'ici seront avancés. La seconde partie de cette thèse, où plusieurs articles sont présentés, vise à répondre à la majorité de ces objections. Pour une plus ample analyse des arguments en faveur de la logique non-monotone, le lecteur est invité à consulter le chapitre 15.

Chapitre 9

Conclusion

Ce chapitre clôt la revue de la littérature, et par le fait même la première partie de cette thèse. Évidemment, il va sans dire que la revue proposée ne couvre pas la totalité des publications sur le sujet et que certains auteurs n'ont pas été traités.[1] Néanmoins, lorsque combinée avec ce que nous avons fait précédemment dans Peterson (2011), la revue est exhaustive dans la mesure où elle couvre les différents types d'approches que l'on retrouve en logique déontique. Ainsi, nous avons tâché de dresser un portrait le plus complet possible, fournissant au lecteur le matériel de base nécessaire à la compréhension des approches que l'on retrouve au sein de la littérature.

Le présent chapitre vise à faire la transition entre la première et la deuxième partie de cette thèse. Dans ce qui suit, nous allons d'abord présenter les fondements philosophiques et juridiques de notre position afin d'expliciter la pertinence de notre analyse pour le discours légal. Cela sera suivi d'une liste d'arguments pouvant être objectés contre les différents types d'approches que l'on retrouve en logique déontique. Par le fait même, cela nous permettra de mettre en lumière les points importants à considérer lors de l'analyse du discours normatif.

Fondements philosophiques

Du côté des fondements, notre position s'insère dans la continuité de ce qui a été présenté dans Peterson (2011, chap. 6). Afin de bien saisir les fondements philosophiques de notre approche, il faut d'abord comprendre quel est notre objectif. Tel que mentioné au début de cette partie, l'objectif principal de cette thèse est de dévlopper des outils formels pertinents à l'analyse des inférences légales. En ce sens, notre objectif n'est pas de développer des outils pertinents à l'informatique ou à la modélisation des NMAS: l'objectif est de fournir une compréhension approfondie de la structure des inférences légales, et par le fait même de préciser la structure du discours juridique. Une conséquence directe de cet objectif est

[1] Par exemple, nous n'avons pas pris la peine de traiter explicitement des approches qui se qualifient de réductions andersonniennes, puisque l'idée de base derrière ces approches est similaire à la réduction que l'on retrouve en logique déontique dynamique.

que notre tâche n'est pas d'expliciter la structure des lois: le but est de déterminer la structure des inférences faites sur la base des lois.

D'entrée de jeu, nous sommes d'avis que si la philosophie formelle prétend être pertinente à l'analyse du discours légal, alors celle-ci doit minimalement réfléchir sur les principes de base qui sont admis en droit. Ainsi, une logique déontique pertinente à l'analyse des inférences légales doit à tout le moins être fondée sur certains principes ou caractéristiques que l'on retrouve au sein du discours juridique. En ce sens, pour qu'une logique déontique soit d'une quelconque utilité quant à l'analyse des inférences légales, celle-ci doit être fondée sur des principes admis au sein de la législation et de la jurisprudence.

En cherchant à déterminer les conditions dans lesquelles les inférences légales sont valides ou non, on vise à déterminer certains critères de rationalité permettant de distinguer entre les inférences légales dont la forme est acceptable de celle dont elle ne l'est pas. L'objectif de cette thèse n'est cependant pas de fournir une théorie de la rationalité, et par conséquent notre conception de ce qu'est un argument dont la forme est rationnellement acceptable repose un critère de rationalité plutôt minimal.

Par un argument *rationnel*, nous entendons simplement un argument qui respecte le principe de non-contradiction. Le critère de rationalité qui se trouve au fondement de notre analyse est donc la non-contradiction, voire la cohérence ou encore la consistance. À titre de précision terminologique, notons que l'usage du terme *consistance* est réservé à son utilisation en logique, où la consistance est une propriété (syntaxique) formelle d'un système. Le terme *cohérence* sera utilisé afin de référer à la cohérence de la langue naturelle, lorsqu'une définition formelle explicite n'est pas donnée. Soulignons aussi la distinction entre la cohérence propositionnelle et la cohérence normative. Alors que la première indique qu'une proposition ne peut être à la fois vraie et fausse, la seconde implique qu'une action et son contraire ne peuvent être obligatoires en même temps. Formellement, la cohérence normative est réductible à la cohérence propositionnelle que si l'on admet le schéma d'axiome (**D**) de la logique déontique standard, qui est propositionnellement équivalent à $\neg(O\varphi \wedge O\neg\varphi)$.

Or, il s'avère que la non-contradiction est non seulement un principe fondamental en droit, mais qu'il s'agit aussi d'un critère de rationalité pour le discours juridique. Évidemment, quiconque est au fait du système juridique verra aisément que les lois sont souvent contradictoires, ou du moins incompatibles. Cela dit, réitérons que notre objectif n'est pas d'étudier la structure des lois mais que le but est plutôt d'étudier la structure des inférences légales. En ce sens, notre approche se place *après* interprétation. À supposer que le sens des lois est fixé[2], notre objectif est de déterminer les conditions de validité des inférences qui peuvent être faites sur la base de cette législation.

Notre analyse prenant place après interprétation, nous avons investigué du côté des principes qui guident l'interprétation des lois afin de fournir des fondements philosophiques solides à notre approche. Malgré que les différentes législations soient souvent contradic-

[2] L'analyse de la signification de la loi se fait en fonction de l'acceptabilité des prémisses d'un argument. L'analyse de la validité d'un argument fait abstraction de la valeur de vérité actuelle des prémisses.

toires, il n'en demeure pas moins que les lois sont présupposées cohérentes *après* interprétation, et donc la présomption de cohérence est un principe qui guide l'interprétation de la loi. Cette présomption de cohérence du système juridique résulte notamment du fait que le législateur est présupposé rationnel, et donc que sa pensée est présuposée rationnelle et logique (Côté 2006, p.387). La cohérence se veut une « valeur fondamentale des systèmes juridiques, dont elle contribue à assurer l'autorité, l'accessibilité et l'équité (Côté 2006, p.387). »

La présupposée cohérence d'un système juridique est triviale lorsque l'on considère que l'objectif des lois est de guider l'action (Weinberger 2001, p.134). En effet, l'objectif d'un système juridique est de déterminer le cadre à l'intérieur duquel un agent peut ou doit agir. En légiférant de manière contradictoire, on obtient un cadre « vide », où aucune action ne peut être accomplie sans que l'on contrevienne à la loi. Ainsi, sans le présupposé de cohérence, un système juridique ne pourrait atteindre son objectif: si l'objectif de la loi est de guider l'action du citoyen, alors les lois doivent minimalement être cohérentes, sans quoi le citoyen ne pourra agir conformément aux actions qui lui sont prescrites.

La présomption de cohérence est non seulement fondamentale au discours juridique, mais est aussi un critère de rationalité. D'une part, il s'agit d'un principe fondamental puisqu'elle permet de justifier d'autres principes d'interprétation des lois. Par exemple, la présomption de cohérence se trouve au fondement de la règle *nemo intelligere possit antequam iterum perlegerit*: nul ne saurait comprendre avant d'avoir lu et relu le tout au complet.[3] Ce principe, qui dicte qu'une loi doit être comprise à la lumière de l'ensemble de la législation, sert à assurer que l'interprétation d'une loi n'engendre pas d'incohérence. D'autre part, la présomption de cohérence est un critère assurant la rationalité du discours juridique. En effet, la Cour suprême, citant Sullivan (1994, p.176), affirme que la présomption de cohérence vise l'atteinte d'un « cadre rationel ».[4] En ce sens, la présomption de cohérence des lois (après interprétation) vise à assurer et préserver la rationalité du discours juridique.

Suivant la Cour suprême[5], la présomption de cohérence est reconnue en Common law depuis le 16^{e} siècle[6] et a été reformulé par le Conseil privé[7]. Cette reformulation, qui exprime essentiellement qu'une loi doit être interprétée de manière à préserver la cohérence de la législation, a été réadoptée par la Cour suprême en 1952 dans l'affaire *The King c. Assessors of the Town of Sunny Brae*[8]. Plus tard, dans *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, la Cour suprême s'appuiera sur Sullivan (1994, p.176) afin de conclure que la présomption de cohérence vise non seulement l'obtention d'un

[3] *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, [1996] 3 RCS 919, paragraphe 207.

[4] *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, [1996] 3 RCS 919, paragraphe 209.

[5] *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, [1996] 3 RCS 919, paragraphe 207.

[6] Cf. *Chamberlain's Case* (1611), Lane 117, 145 E.R. 346.

[7] Voir *City of Victoria c. Bishop of Vancouver Island*, [1921] 2 A.C. 384 (C.P.), p. 388. Le Conseil privé est l'un des plus haut tribunaux du Royaume-Uni, qui était le plus haut tribunal au Canada avant que la Cour suprême ne soit instaurée. Voir aussi *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, [1996] 3 RCS 919, paragraphe 208.

[8] *The King c. Assessors of Sunny Brae (Town)*, [1952] 2 SCR 76, page 97.

« cadre rationel », mais qu'il s'agit surtout d'un principe « virtuellement irréfragable ».[9] En conséquence, la présomption de cohérence des lois (après interprétation) est un principe fondamental qui ne peut être réfuté sans contrevenir à la rationalité du discours juridique.

En droit, ce présupposé de cohérence se réalise de diverses façons. D'abord, la cohérence est présupposée de manière *horizontale* (Côté 2006, p.388). La cohérence horizontale s'applique à un même ensemble de normes, comme le Code civil, le Code criminel ou la Constitution canadienne. Ce type de cohérence présuppose que les normes à l'intérieur d'un même ensemble sont cohérentes les unes avec les autres. Par ailleurs, il y a aussi une présupposition de cohérence *verticale* (Côté 2006, p.388). La cohérence verticale s'applique entre les différents ensembles de normes: on présuppose que les différents ensembles de normes sont cohérents les uns avec les autres. Par exemple, le Code civil est présupposé cohérent avec la Constitution canadienne, tout comme le Code criminel est présupposé cohérent avec la Charte canadienne des droits et libertés. Ainsi, l'ensemble des lois est présupposé former un tout cohérent (Côté 2006, p.433). La cohérence du système juridique résulte du fait que l'on présuppose « la volonté d'un législateur logique qui, à l'intérieur de l'ensemble des lois sur une même matière, est censé procéder systématiquement, c'est-à-dire sans contradiction, et donner à des problèmes semblables des solutions semblables (Côté 2006, p.439). »[10]

La présomption de cohérence a notamment deux corollaires jouant un rôle important d'un point de vue fondationnel. Le premier est celui du principe de hiérarchisation des lois. En effet, la hiérarchie des lois permet de préserver la cohérence du système juridique, autant de manière verticale que horizontale. Alors que cette hiérarchie peut être explicitement mentionnée dans la loi, celle-ci peut aussi provenir de la construction des différents ensembles de normes (Côté 2006, pp.45,450). Au Canada, par exemple, aucune législation ne peut aller à l'encontre de la Constitution canadienne. En ce sens, la Constitution possède un statut supérieur aux autres législations, et en cas de conflit entre certaines normes celles de la Constitution prévaudront. En philosophie formelle, cette stratégie, qui consiste à introduire une (ou des) relation(s) de hiérarchie à l'intérieur d'un modèle sémantique afin de préserver la cohérence et de résoudre les conflits d'obligations a été utilisée notamment par Alchourrón et Makinson (1981).

Le second corollaire est que le discours juridique est régit par une relation de conséquence « logique ». En présupposant la cohérence du discours juridique, on présuppose par le fait même que les conclusions auxquelles on parvient par le biais d'inférences à partir de la loi sont cohérentes avec la législation. Ainsi, en acceptant le présupposé de cohérence, on accepte du même coup un principe de « conséquence logique », qui permet d'inférer adéquatement certaines choses à partir de la loi et des principes d'interprétation. De fait, en présupposant que le discours juridique est cohérent, on assume d'emblée que les inférences légales sont sujettes à des règles « logiques » gouvernant leur bon usage. En ce sens, la cohérence implique que l'on peut « déduire des normes expressément formulées

[9] *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, [1996] 3 RCS 919, paragraphe 209.

[10] La cohérence peut évidemment s'apercevoir sous d'autres angles. Par exemple, il doit y avoir cohérence entre la loi, les faits admis à un procès et le verdict d'un juge, tout comme il doit y avoir cohérence entre les différentes décisions prises aux tribunaux.

certaines règles implicites qui s'en dégagent logiquement (Côté 2006, p.422) ».

Cela dit, l'utilisation de l'expression « conséquence logique » au sein du discours juridique est au mieux métaphorique. En effet, cette notion est utilisée de manière intuitive, sans qu'il y ait de définition explicite de ce que signifie la conséquence logique au niveau du discours légal. Ainsi, on présuppose que le discours juridique est cohérent et qu'il existe des règles permettant de déterminer adéquatement les conclusions que l'on peut inférer, mais aucune définition légale de la « conséquence logique » n'est proposée.

Or, un des objectifs de cette thèse est précisément de faire un pas de plus en faveur d'une telle définition. Comme nous le verrons dans la prochaine partie, une des contributions théoriques de cette thèse est de fournir une définition formelle explicite de la notion de conséquence logique à l'œuvre au sein des inférences légales. Pour ce faire, nous distinguerons entre les inférences qui mettent en jeu des obligations inconditionnelles de celles qui se font sur la base d'obligations conditionnelles.

Du côté de la philosophie formelle, outre les relations de conséquences implicites aux divers systèmes qui ont été proposés afin de rendre compte des inférences normatives (et qui ont donc tenté de fournir une définition formelle explicite de la relation de conséquence logique à l'œuvre au sein des inférences légales), il est généralement admis que les inférences normatives sont sujettes à une forme ou une autre du *principe de conséquence déontique*. En ce qui nous concerne, nous adopterons l'interprétation proposée par Castañeda (1968, p.13), à savoir que si φ est obligatoire et que φ implique ψ, alors ψ est aussi obligatoire. Lorsqu'on le considère en termes d'interdiction, la validité (intuitive) de ce principe est triviale d'un point de vue légal: il serait absurde que φ soit interdit, que ψ entraîne φ mais que ψ soit permis. Autrement dit, il serait absurde de permettre une action ψ qui viole directement une interdiction $O\neg\varphi$.

Par ailleurs, une conséquence directe de la présomption de cohérence est l'adoption du schéma d'axiome (**D**) de la logique déontique standard. Cela n'implique cependant pas que la logique qui sera proposée aux chapitres 11 et 16 soit une logique modale ou encore une logique fortement normale (au sens de Åqvist 2002).[11]

Tout compte fait, il résulte de notre analyse que la cohérence est non seulement un principe fondamental au discours juridique, mais qu'il s'agit aussi d'un critère permettant d'en assurer le caractère rationnel. En outre, le principe de cohérence permet de dériver d'autres principes fondationnels pertinents, notamment ceux de la hiérarchie des lois afin de préserver la cohérence du système juridique et de la « conséquence logique » régissant les inférences légales correctes. Ces principes seront utilisés tout au long de la seconde partie, où notre tâche sera de définir une relation de conséquence logique adéquate aux inférences légales. La relation de hiérarchie sera plutôt utilisée lors de l'analyse de l'acceptabilité des prémisses.

Avant de poursuivre avec la contribution théorique de la thèse, voyons d'abord quelques arguments que l'on peut avancer contre les principaux courants en logique déontique. Cela nous permettra de mettre en évidence plusieurs points à prendre en considéra-

[11] Ni K ni KD n'est une extension de OL.

tion pour la modélisation du discours juridique, qui seront repris aux chapitre 10 et 17 afin de montrer comment la thèse permet de répondre à plusieurs lacunes que l'on retrouve au sein de la littérature.

Critique et arguments

Par souci de clarté et de concision, nous allons d'abord lister les arguments généraux pouvant être objectés aux principaux courants en logique déontique, en les justifiant d'un point de vue juridique. Soulignons que cette critique s'opère relativement à notre objectif de départ, à savoir l'utilisation de la logique aux fins de l'analyse des inférences légales. Cela fait, nous discuterons de chacun des types d'approches présentés jusqu'à présent en respectant l'ordre de présentation de la revue de la littérature. D'autres arguments seront insérés dans les sections appropriées lorsque cela sera pertinent.

D'emblée, une logique déontique adéquate à la modélisation du discours juridique doit être de type *Ought-to-do*. En effet, d'un point de vue légal, ce ne sont pas des propositions qui sont obligatoires mais bien des actions. Lorsqu'un acusé est inculpé d'un crime, on ne lui reproche pas d'avoir « rendu une certaine description du monde vraie »: on lui reproche d'avoir agit à l'encontre de certaines lois.

ARGUMENT 1: Une logique déontique adéquate à la modélisation du discours légal est de type *Ought-to-do*.

ARGUMENT 2: Ce sont les actions qui sont obligatoires, et non les propositions. Par conséquent, une logique déontique doit distinguer entre les actions et les propositions, et seules les actions doivent pouvoir être placées dans la portée des opérateurs déontiques.

Par ailleurs, la vérité des propositions légales ne dépend pas d'une forme naïve de correspondance avec une « réalité normative ». La vérité des propositions normatives, en droit, est le résultat d'un processus de construction. Comme le mentionne l'honorable Jean-Louis Baudoin, la vérité « se cache dans le discours du législateur qui d'autorité décrète ce qui est vrai et ce qui doit être cru et vu comme tel pour la bonne ordonnance du système juridique dans son ensemble (Baudoin 2010, p.5). »

ARGUMENT 3: Toute obligation dépend d'une norme, laquelle est établie par une certaine autorité (Alchourrón et Bulygin 1981, pp.97,102). La valeur de vérité d'une proposition de la forme $O\varphi$ ne dépend pas de la valeur de vérité de la proposition φ dans la portée de l'opérateur déontique. La valeur de vérité de $O\varphi$ dépend d'une norme, qui établit que l'action décrite par la proposition φ est obligatoire.

ARGUMENT 4: Il y a une dichotomie sémantique entre les propositions descriptives et les propositions normatives. La valeur de vérité d'une proposition descriptive ne dépend pas des mêmes conditions que de celle d'une proposition normative (cf. Hume 1740, Poincaré 1913 et Jørgensen 1937).

Argument 5: En vertu de l'argument 3 et suivant Chellas (1974, p.24), une logique déontique Δ ne doit donc pas admettre d'obligation absolue, c'est-à-dire de théorème de la forme $\vdash_\Delta O\top$, où $\top$ est une tautologie. Autement dit, il est possible d'avoir une situation où aucune obligation n'est en force.

Argument 6: Une obligation peut potentiellement être violée (Jones et Pörn 1985, p.279). De fait, une logique déontique Δ ne doit pas admettre de tautologie comme obligation dérivée, et donc elle doit satisfaire $\nvdash_\Delta O\varphi \supset O\top$.

Argument 7: Une conséquence de l'argument 2 est que l'itération d'un même opérateur déontique n'est pas acceptable, puisque seule une action peut se trouver dans la portée d'un opérateur déontique.

Argument 8: L'itération d'un même opérateur déontique n'est pas acceptable (Castañeda 1981, p.66). En effet, la signification d'un opérateur déontique change lorsque celui-ci est itéré (Jones et Pörn 1985, p.286).

Argument 9: Une logique déontique adéquate à la modélisation du discours juridique doit être fondée sur la présomption de cohérence normative.

Les corollaires de l'argument 9 sont:

Argument 10: Une logique déontique adéquate à la modélisation du discours juridique doit admettre le schéma d'axiome (**D**): une même action ne peut être à la fois obligatoire et interdite au même moment et selon la même autorité.

Argument 11: Dans un système qui admet plusieurs types d'obligations, et donc où il y a des conflits potentiels, une relation de hiérarchie doit pouvoir être utilisée afin de pouvoir résoudre les conflits d'obligations.

Argument 12: Une logique déontique adéquate à la modélisation du discours juridique doit inclure une relation de conséquence logique permettant de valider le principe de conséquence déontique.

Interprétation modale

L'interprétation modale de la logique déontique occupe une place importante dans la littérature. Même si, comme nous l'avons vu, plusieurs approches ne traitent pas l'obligation comme une modalité de type K, il n'en demeure pas moins que la logique déontique est souvent présentée comme une branche des logiques modales. De manière générale, les interprétations modales de la logique déontique, où l'obligation est interprétée comme une modalité O faisant changer la valeur de vérité d'une proposition φ dans sa portée, succombent aux arguments 1, 2, 3, 4, 7 et 8.

Cela dit, les systèmes normaux prêtent le flanc à encore plus d'objections. Une logique déontique normale, au sens de Åqvist (2002), est un système où l'obligation est axiomatisée comme une modalité de type K, et donc qui admet minimalement les théorèmes du système K. Un système fortement normal admet minimalement les théorèmes de KD. Sans compter les arguments découlant du paradoxe de Chisholm en faveur des logiques dyadiques ou des logiques non-monotones, mettant en évidence que les logiques déontiques normales ne traitent pas adéquatement des obligations conditionnelles et des conflits d'obligations, les systèmes normaux succombent aussi aux arguments 5 et 6. Lorsqu'un système est simplement normal, et non fortement normal, le système tombe aussi sous l'argument 10. Par ailleurs, un système normal ou fortement normal succombera aussi à l'argument 12, à moins d'admettre d'autres modalités ou certaines restrictions *ad hoc* sur le modèle sémantique (à ce sujet, voir Peterson et Marquis 2012).

Logique dyadique

Étant des interprétations modales, les arguments 1, 2, 3, 4, 7 et 8 peuvent s'objecter aux logiques dyadiques. De même, les logiques dyadiques normales ou fortements normales succombent aux arguments 5 et 6, et aussi à l'argument 12, à moins d'admettre d'autres modalités ou certaines restrictions *ad hoc* sur le modèle sémantique. On peut aussi objecter l'argument 10 à un système dyadique simplement normal.

Malgré que les logiques dyadiques permettent la résolution du paradoxe de Chisholm, il s'avère qu'elles ne permettent pas la modélisation adéquate des obligations conditionnelles, puisque soit elles admettent (sans restriction) le détachement factuel, soit elles l'excluent totalement, sans compter qu'elles ne sont pas en mesure de rendre compte des conflits d'obligations. Comme nous le verrons au chapitre 15, les raisonnements qui se font sur la base d'obligations conditionnelles sont beaucoup plus nuancés, et une logique adéquate à la représentation des raisonnements normatifs conditionnels doit admettre une forme restreinte de détachement.

Argument 13: Une logique déontique capable de représenter formellement les raisonnements légaux conditionnels doit être en mesure de permettre le détachement restreint.

Argument 14: Une logique déontique capable de représenter formellement les raisonnements légaux conditionnels doit bloquer l'augmentation.

Finalement, à l'instar des systèmes monadiques, la logique dyadique ne permet pas de gérer les conflits d'obligations.

Argument 15: Une logique déontique capable de représenter formellement les raisonnements légaux conditionnels doit être capable de représenter les différents types de défaisabilités (cf. van der Torre et Tan 1997).

Argument 16: Une logique déontique capable de représenter formellement les raisonnements légaux conditionnels doit éviter l'explosion.

Argument 17: Une logique déontique adéquate à la représentation du discours légal doit être en mesure de représenter de manière consistante les conflits (conditionnels) d'obligations.

Logique stit

Bien que la logique STIT tente de traiter de l'obligation par rapport à l'action, celle-ci succombe malgré tout aux arguments 1, 2, 3 et 4. De plus, si l'obligation est axiomatisée comme une modalité de type K, alors les mêmes arguments que pour les systèmes normaux et fortement normaux pourront être objectés.

Par ailleurs, la logique STIT n'offre pas une représentation formelle adéquate de l'action. En effet, la logique STIT réduit l'action à une description du monde, et cela n'est simplement pas représentatif de ce qu'est une action. Une description φ peut être le résultat désiré d'une action, tout comme elle peut être un résultat non désiré, ou même une conséquence (désirée ou non) de l'action. En ce sens, même si l'action engendre quelque chose dans le monde, l'action ne se réduit pas à la description qu'elle engendre: même si un agent i, en volant une banque (disons) fait advenir qu'un ancien crime est résolu (p. ex., i fait tomber une pile de vieux documents où un élément de preuve est trouvé), on ne dira pas que l'agent a accompli l'action de résoudre le crime même si dans les faits il est vrai que $[i : stit]\varphi$.

Logique dynamique

D'entrée de jeu, la logique dynamique réduit l'obligation à la réalisation d'une violation dans un état, et en ce sens elle tombe sous l'argument 3 puisqu'elle fait dépendre la valeur de vérité d'une proposition normative à la valeur de vérité d'une proposition descriptive. Par le fait même, elle est aussi sujette à l'argument 4. Cela dit, un argument important que l'on peut objecter à la logique dynamique est le suivant.

ARGUMENT 18: Une logique déontique appropriée doit reposer sur une algèbre d'action adéquate à la modélisation des actions humaines.

En effet, la logique dynamique est principalement une logique de la programmation. Or, comme nous le verrons au chapitre 14 (auquel le lecteur est référé pour une analyse détaillée du sujet), la logique dynamique n'est pas fondée sur une algèbre visant à représenter les action humaines. En ce sens, puisque la logique dynamique ne fournit pas une modélisation adéquate de l'action humaine, et que les inférences légales portent sur les actions humaines, il en résulte que la logique dynamique n'est pas suffisante à l'analyse des raisonnements juridiques.

Logique i/o

Les logiques I/O offrent un cadre de travail intéressant autre que celui de la logique modale K. Cela dit, d'un point de vue fondationnel, elles contreviennent à l'argument suivant.

ARGUMENT 19: Un système formel adéquat à la représentation du discours juridique doit laisser place a l'interprétation de la loi.

La pertinence d'une logique déontique quant à l'analyse des inférences légales est du côté de leur forme. En effet, la logique vise l'étude des structures, et, lorsqu'elle est utilisée comme outil en vue de l'analyse des raisonnements, elle permet l'analyse de la *structure* des inférences. Elle permet l'analyse de la structure des raisonnements, mais pas de leur contenu. En ce sens, l'évaluation des prémisses afin de déterminer si, en plus d'être valide, un argument est probant, ne relève pas de la logique (cf. Peterson 2013b). Bien que la logique puisse être utilisée afin d'analyser et d'aider les raisonnements des juristes, l'évaluation des prémisses et l'interprétation des lois sont de leur ressort.

La signification des lois n'est pas univoque, et les ensembles de normes ainsi que la jurisprudence sont des choses qui évoluent dans le temps. En ce sens, l'information pertinente à un jugement en 1996 n'est pas la même que celle qui serait pertinente si le même jugement avait plutôt été fait en 2011. Les lois et la jurisprudence ne sont pas des choses *fixes*, elles évoluent, et cette évolution ne peut pas être anticipée logiquement. D'où la nécessité d'un facteur humain dans l'évaluation des prémisses d'un argument légal: la vérité d'une prémisse légale ne peut pas être déterminée uniquement par un algorythme et requiert l'examen attentif d'un juriste. D'un point de vue méta-logique, un système formel doit donc laisser place à l'interprétation des lois afin de pouvoir modéliser adéquatement le discours juridique.

Or, les logiques I/O échouent à rendre compte de ce principe fondationnel. En considérant un système normatif comme un simple processus input-output, on présuppose que le système formel, les ensembles de normes ainsi que les ensembles de *count-as* contiennent d'emblée toute l'information pertinente et nécessaire à la prise de décision. Cette vision du discours légal est cependant réductrice et ne tient pas compte adéquatement du facteur humain prenant place lors des raisonnements juridiques. Dans les logiques I/O, on présusuppose que l'ensemble des *count-as* est bien défini. Au niveau du discours juridique, cela est toutefois trivialement faux: la question de savoir si une action particulière *compte pour* une action qui est interdite, ou encore si une action particulière tombe dans la portée d'une loi ou d'un jugement sont souvent l'objet de litige.

En présupposant qu'un système normatif est bien défini et possède toute l'information pertinente et nécessaire à la prise de décision, les logiques I/O ne parviennent pas à rendre compte adéquatement du principe de conséquence déontique. Brièvement, la conséquence déontique permet de conclure des obligations dérivées à partir d'obligations fixes. Le principe est tel que si φ est obligatoire et que φ implique ψ, alors ψ est aussi obligatoire. Au sein des logiques I/O, la conséquence déontique se formalise par un ensemble générateur $G = \{(\top, x), (\top, x \to y)\}$, où l'on obtient $y \in out_i(G, x)$ par (AND) et (WO).

Cependant, cela pose problème lors de l'interprétation normative de G. En effet, si G représente un code normatif, c'est-à-dire un ensemble de normes, il est peu plausible que celui-ci contienne d'emblée $(\top, x \to y)$. Le propre de la conséquence déontique est de prendre un ensemble fini de normes afin d'en dériver les conséquences (dénombrables). Considérant que le propre d'une obligation dérivée est, par définition, de ne pas être une obligation fixe, il s'ensuit que les obligations dérivées pour un système normatif donné ne peuvent pas simplement être déterminées à l'aide d'un ensemble générateur fixé au départ.

Lorsque l'on cherche à déterminer si $O\psi$ est la conséquence déontique de $O\varphi$, il faut évaluer premièrement s'il est vrai que $O\varphi$, et deuxièmement s'il est vrai que $\varphi \supset \psi$. En ce sens, l'évaluation d'une conséquence déontique se fait à partir d'un ensemble générateur de normes *et* d'un ensemble de propositions qui sont postulées vraies. Même si la question de savoir si ψ *compte pour* φ peut être d'entrée de jeu fixée au sein de G, il n'en demeure pas moins qu'il est possible que cela soit l'objet d'un litige. En ce sens, un raisonnement légal ne se réduit pas simplement à un processus intput-output qui contient toute l'information pertinente et nécessaire à la prise de décision.

Évidemment, on pourrait répondre à cette objection qu'un système normatif opère selon un processus input-output, et que les résultats obtenus (les outputs) dépendent de ce que l'on met dedans (le input) au départ. En ce sens, le système normatif, conçu comme un procédé input-output, n'est qu'un outil prêt à être utilisé par les juristes. Il y a la boîte, et il y a ce que l'on met dans la boîte.

À cette objection, on pourrait alors répondre que les logiques I/O ne rendent pas compte adéquatement du discours juridique dans la mesure où elles succombent aux arguments 5 et 6.

Approches algébriques

La critique que l'on peut faire à l'égard des approches algébriques s'insèrent dans la lignée de l'argument 18. Dans le cas des approches qui utilisent des algèbres d'action, il n'est pas évident que la structure proposée soit représentative des actions humaines. En fait, mis à part les logiques dynamiques ainsi que l'approche de Lucas, les systèmes fondés sur des algèbres d'action utilisent des algèbres de Boole. Cependant, comme nous le verrons au chapitre 14, la logique de l'action humaine, si elle existe, n'a certainement pas la structure d'une algèbre de Boole (le lecteur est référé au chapitre susmentionné pour une analyse détaillée).

Logique non-monotone

De manière usuelle, la logique non-monotone est présentée comme une alternative aux logiques monadiques et dyadiques, permettant de répondre à la fois aux problèmes du détachement, de l'augmentation et aux conflits d'obligations. Cela dit, les logiques déontiques non-monotones sont souvent développées en vue de l'intelligence artificielle, où l'objectif est de développer un cadre de travail permettant de modéliser adéquatement les inférences normatives quotidiennes.

Or, il s'avère que la non-monotonie de l'inférence normative quotidienne vient principalement du fait que les agents ne possèdent pas toute l'information pertinente à leur raisonnement. Le problème de l'augmentation, par exemple, vient du fait qu'il est possible que l'on ajoute de l'information pertinente qui viendra changer la conclusion à laquelle on parvient. Cela dit, dès que toute l'information pertinente à un raisonnement normatif est présente, la non-monotonie de l'inférence normative et le problème de l'augmentation disparaissent. En droit, par exemple, un juge prendra en considération la législation, la jurisprudence ainsi que la doctrine pertinente et prendra sa décision à la lumière des faits admissibles. Si l'on fait appel à cette décision en Cour d'appel et que l'on cherche à défaire l'inférence du juge de première instance, alors on tentera d'établir qu'il y a eu une erreur de droit et que ce n'est pas toute la législation et la jurisprudence pertinente qui ont été prises en compte (sans compter les questions d'interprétation). En ce sens, une logique déontique adéquate à la représentation du discours juridique doit être en mesure de représenter que certaines inférences sont défaisables alors que d'autres ne le sont pas.

En logique non-monotone, on « complexifie » la syntaxe et la sémantique afin de répondre aux objections contre les logiques standards. Comme nous le verrons au chapitre 15, il est cependant possible de répondre à ces objections en « simplifiant » le système utilisé pour raisonner à l'aide d'obligations conditionnelles.

Notons qu'en plus de modéliser les conflits d'obligations et leur résolution, les logiques déontiques non-monotones visent aussi à représenter les conflits irréductibles d'obligations, c'est-à-dire les conflits qui ne peuvent être résolus. En ce qui nous concerne, nous sommes d'avis qu'une théorie qui admet des conflits irréductibles d'obligations est irrationnelle. L'objectif d'une théorie normative est de déterminer l'action d'un individu, et lorsqu'il y a conflit irréductible, la théorie ne permet pas de dire comment l'agent devrait agir. En ce sens, une théorie qui admet des conflits irréductibles d'obligations est une théorie qui d'emblée ne peut atteindre son objectif (qui est de guider l'action). D'ailleurs, en vertu de la présomption de cohérence, cela est impossible en droit: tout conflit sera résolu, localement (au cas par cas) ou globalement, soit par une nouvelle législation ou par un jugement. Rappelons-nous que la cohérence est un critère assurant la rationalité du discours juridique.

Par ailleurs, soulignons que l'objectif principal de cette thèse n'est pas de modéliser les inférences normatives quotidiennes, mais est plutôt de développer un cadre de travail permettant d'*évaluer* les inférences normatives. Notre tâche n'est pas descriptive mais prescriptive. L'objectif n'est pas de déterminer comment les inférences normatives se *font*, mais bien de déterminer comment elles *devraient se faire*.

* * *

Cela conclut la première partie de cette thèse. Passons maintenant à la contribution théorique.

Partie II

Articles

Chapitre 10

Introduction

Objectifs théoriques

D'un point de vue théorique, l'objectif de la présente thèse est de montrer qu'il existe d'autres avenues satisfaisantes, qui n'ont pas encore été explorée dans la littérature, afin de modéliser les inférences normatives conditionnelles et inconditionnelles. Ainsi, l'une des principales motivations de cette thèse est le rejet de trois dogmes: l'interprétation modale de la logique déontique, l'utilisation des algèbres de Boole pour modéliser l'action humaine et la nécessité des logiques non-monotones ou adaptatives afin de modéliser les conflits d'obligations et les obligations conditionnelles. On voit donc se tracer les trois principaux axes sur lesquels sont fondés les articles de cette deuxième partie:

1. développement d'un cadre conceptuel permettant l'analyse de la validité des inférences légales inconditionnelles;
2. développement d'une logique de l'action adéquate à la représentation des actions humaines;
3. modélisation des raisonnements normatifs conditionnels et des conflits d'obligations.

En premier lieu, cette thèse vise le développement d'un cadre de travail différent de la logique modale K afin de modéliser adéquatement les inférences normatives inconditionnelles et non-conflictuelles. Cela fait l'objet principalement des chapitres 11 et 12, où la logique proposée vise la représentation de la transmission de la propriété « obligation » entre les actions.

Actuellement, on ne trouve pas de cadre de travail permettant l'analyse des inférences normatives inconditionnelles et visant à être enseigné dans les cours de pensée critique et d'introduction à la logique. En effet, seules la logique propositionnelle classique et le calcul de premier ordre sont usuellement présentés dans les cours d'introduction, et il faut attendre un cours de logique avancée (lorsqu'il y en a) où l'on traite de la logique modale avant de pouvoir même penser à introduire l'étudiant à la logique déontique et à l'analyse des inférences normatives.

Historiquement, la logique est apparue afin d'offrir un cadre de travail permettant l'analyse des raisonnements mathématiques et scientifiques. En quelques mots, l'objectif était d'obtenir une méthode capable de garantir qu'une conclusion peut effectivement être dérivée de certaines hypothèses. Cela dit, au fils du temps, la logique s'est vue être appliquée à l'analyse des inférences en général. Ainsi, ayant perdu de vue son objectif originel, la logique est utilisée afin d'analyser des raisonnements qui ne sont pas nécessairement mathématiques ou scientifiques. En philosophie, par exemple, il est pratique courante d'initier les étudiants à la logique propositionnelle et au calcul de premier ordre de façon a les préparer à bien structurer leur pensée et à pouvoir analyser la forme de différents arguments.

Or, il s'avère que les inférences *normatives* ont une place importante au sein des débats philosophiques. À ce titre, on peut juger que les cours de pensée critique et d'introduction à la logique échouent à préparer l'étudiant à l'analyse des raisonnements. En effet, nonobstant le dilemme de Jørgensen, ces logiques sont insuffisantes à la modélisation des inférences normatives, qui requièrent un degré d'analyse plus fin. Par surcroît, même si l'on restreint l'utilisation de la logique déontique standard à l'analyse des obligations inconditionnelles et non-conflictuelles, il n'en demeure pas moins que celle-ci possède des lacunes philosophiques considérables, comme nous l'avons exposé au chapitre 9. En conséquence, un des objectifs de cette thèse est de répondre à ces deux insuffisances en fournissant un système de base alternatif à la logique modale K pour traiter des obligations inconditionnelles et qui puisse être adapté à des cours de pensée critique afin d'analyser les inférences normatives. L'alternative à la logique modale K est présentée au chapitre 11 et le chapitre 12 en fait l'adaptation pour l'analyse des raisonnements inconditionnels.

Dans un second temps, cette thèse propose une analyse des différentes logiques de l'action que l'on retrouve au sein de la littérature et propose une nouvelle approche qui, selon nous, est plus appropriée à la modélisation des actions humaines. Alors que la majorité des logiques de l'action vise une application en informatique, ce sont usuellement les algèbres de Boole qui sont utilisées afin de représenter l'action humaine. Nous tâchons d'expliquer en quoi ces logiques échouent à modéliser adéquatement l'action humaine en insistant sur les propriétés des connecteurs logiques utilisés. Cet objectif est atteint au chapitre 14, où une distinction entre une logique de l'action $\mathcal{AL}$ et une logique propositionnelle d'action $\mathcal{PAL}$ est faite.

Troisièmement, les problèmes reliés aux inférences normatives conditionnelles et aux conflits d'obligations sont analysés au chapitre 15. Les principaux arguments en faveur de fondations non-monotones pour la logique déontique sont exposés et les problèmes de l'augmentation et du détachement ainsi que les conflits d'obligations sont analysés à la lumière de la théorie des catégorie. D'un point de vue théorique, la contribution de ce chapitre est de montrer que plusieurs problèmes fondamentaux à la modélisation des inférences normatives conditionnelles peuvent être corrélés aux propriétés structurelles des logiques utilisées lorsque celles-ci sont analysée selon la perspective de la théorie des catégories.

Finalement, le chapitre 16 réunit les trois objectifs de la thèse et propose un cadre de travail unificateur qui incorpore une logique des obligations inconditionnelles et une logique pour les inférences normatives conditionnelles, les deux étant fondées sur une logique de l'action pertinente à la modélisation des actions humaines. Ce chapitre fait la synthèse des articles précédents et propose une logique typée (cf. chapitre 13) qui réunit à la fois une logique de l'action, une logique propositionnelle d'action (cf. chapitre 14), une logique pour les obligations inconditionnelles (cf. chapitre 11) et une logique pour les inférences normatives conditionnelles (cf. chapitre 15). Cette logique typée est basée principalement sur les résultats présentés au chapitre 13, qui fait le pont entre les travaux initiaux de Lambek en linguistique, la théorie des catégories et la logique déontique en montrant la pertinence de l'utilisation d'une logique typée afin d'éviter (ou de résoudre) les paradoxes. À l'instar de la logique présentée au chapitre 11, le système présenté au chapitre 16 vise explicitement la modélisation des inférences légales et est fondé sur des principes admis en droit.

Les relations entre les chapitres sont représentées à la figure 10.1. Notons que chaque chapitre a fait l'objet d'un article et a été rédigé en vue d'être autonome. Ainsi, malgré cette interdépendance, chaque article possède ses objectifs propres et contient tout le matériel nécessaire à sa compréhension.

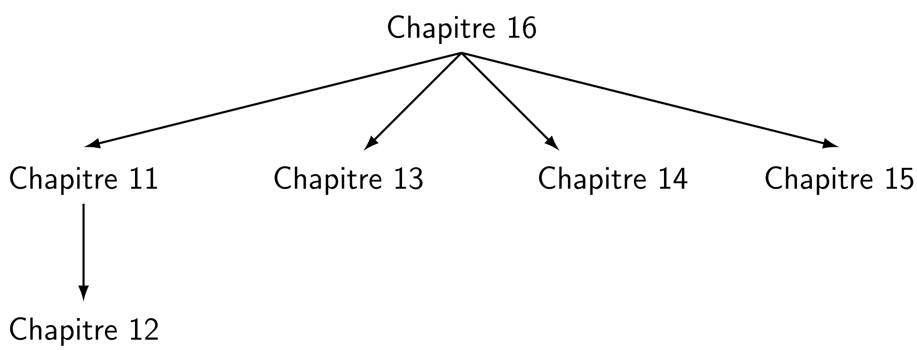


Figure 10.1: Structure des chapitres

Cadre théorique

Avant d'entamer la contribution théorique de cette thèse, il convient d'exposer brièvement les fondements de notre approche, dont les détails techniques seront exposés aux chapitres 13, 14, 15 et 16. Dans les chapitres qui suivent, la logique est abordée d'un point de vue syntaxique sous l'angle de la théorie des catégories. Dans une perspective épistémique, la théorie des catégories offre un cadre de travail riche et unificateur permettant de réunir divers domaines sur les mêmes fondations.[1] En ce qui nous concerne, la théorie des caté-

[1] Le lecteur est référé à Marquis (2009) pour des arguments en faveur des vertus épistémiques de la théorie des catégories. Voir aussi Baez et Stay (2011) pour une comparaison entre la physique, la logique et l'informatique théorique selon les catégories.

gories est utilisée comme cadre de travail en vue de l'analyse de la théorie de la preuve de diverses logiques.

Cette compréhension de la logique s'insère dans la tradition engendrée par Joachim Lambek, qui voyait un parallèle net entre la notion de catégorie et celle de système déductif. Sans entrer dans les détails techniques, l'idée est de concevoir une logique comme un *système déductif*, c'est-à-dire une catégorie où les objets sont des propositions et les flèches sont des preuves. L'intérêt d'un tel regard est que cela permet de déterminer l'ordre précis dans lequel les connecteurs logiques sont introduits et que cela permet d'expliciter les relations qui se trouvent entre les propriétés des connecteurs et la structure catégorique de la logique. Ainsi, même si d'emblée certains systèmes semblent incommensurables, la théorie des catégories permet de comparer diverses logiques selon les propriétés de leurs connecteurs et offre un moyen de les classifier selon leur structure catégorique (i.e., comme différentes catégories monoïdales).

Or, il s'avère que les problèmes relatifs à la modélisation des inférences normatives conditionnelles et des conflits d'obligations sont reliés à la structure cartésienne des logiques utilisées. Similairement, on peut montrer que les différentes logiques d'actions échouent à modéliser adéquatement l'action humaine de par le fait que celles-ci n'ont pas la structure monoïdale appropriée.

Cette thèse doit donc être vue comme une contribution à la philosophie formelle, où la théorie des catégories est appliquée à la logique afin de répondre aux problèmes philosophiques qui résultent de la modélisation des inférences normatives.

Résumés des articles

Formal philosophy and legal reasoning: The validity of legal inferences

Cet article propose un cadre de travail alternatif à la logique modale K pour l'analyse des inférences légales. Le système proposé est fondé sur certains principes inhérents au droit, notamment la présomption de cohérence du discours légal, le fait que la hiérarchie des normes sert à préserver cette cohérence et le fait que les inférences légales soient sujettent à un principe de conséquence déontique. La logique OL, représentant la logique de l'obligation, est fondée sur ces principes et les résultats d'adéquation et de complétude du système sont fournis.

Normative inferences and validity: A heuristic

Ce texte reprend les résultats obtenus au chapitre 11 et les adapte en vue d'introduire les lecteurs non-initiés à la logique à l'analyse des inférences normatives. La nécessité de définir un cadre de travail alternatif aux logiques propositionnelle et de premier ordre est expliquée. La méthode des arbres sémantiques est appliquée à l'analyse de la validité des arguments, puis l'acceptabilité des prémisses normatives est étudiée.

From linguistics to deontic logic via category theory

Le paradoxe de Forrester est analysé à la lumière des travaux de Lambek (1958) en grammaire catégorielle. Son calcul syntaxique est présenté dans le cadre de la théorie des catégories et le paradoxe de Forrester est analysé selon ce cadre de travail. Nous montrons comment une syntaxe typée permet de résoudre ou d'éviter certains problèmes auxquels la logique déontique fait face.

A categorical analysis of action logics

Cet article propose une analyse des différentes logiques de l'action selon la perspective de la théorie des catégories. Les logiques de l'action sont comparées en fonction des propriétés de leurs connecteurs et sont classifiées en tant que différentes catégories monoïdales. Nous proposons une interprétation philosophique des connecteurs logiques d'action et distinguons entre une logique d'action et une logique propositionnelle d'action, conformément à la distinction que l'on retrouve au sein de la littérature entre les *actions* et les *propositions*.

Contrary-to-duty reasoning: A categorical approach

Ce texte met en évidence que plusieurs problèmes relatifs à la modélisation des inférences normatives conditionnelles sont reliés aux propriétés structurelles des logiques utilisées. Nous montrons qu'il est possible de répondre aux arguments usuels en faveur des logiques déontiques dyadiques et non-monotones en utilisant une logique qui possède la structure d'une catégorie $*$-autonome. Cela permet de modéliser les inférences normatives conditionnelles à l'aide d'une logique monadique sans ajouter d'opérateur additionnel.

The categorical imperative: Category theory as a foundation for deontic logic

Cet article introduit une logique visant à modéliser le discours légal canadien. Un système déductif déontique est défini par deux fibrations, utilisant une logique visant à modéliser les obligations inconditionnelles et une autre pour la représentation des inférences normatives conditionnelles et des conflits d'obligations. Une syntaxe typée est utilisée et il est montré comment le système déductif déontique permet l'analyse des inférences légales.

Chapitre 11

Formal philosophy and legal reasoning: The validity of legal inferences

Abstract

The aim of the present paper is to introduce a method to test the validity of legal inferences. We begin by presenting the rationale of our method and then we expose the philosophical foundations of our analysis. If formal philosophy is to be of help to legal discourse, then it must first reflect upon the law's fundamental characteristics that should be taken into account. Our analysis shows that (Canadian) legal discourse possesses three fundamental characteristics which ought to be considered if one wants to represent the formal structure of legal arguments. These characteristics are the presupposed consistency of legal discourse, the fact that there is a hierarchy between norms and obligations to preserve this consistency and the fact that legal inferences are subjected to the principle of deontic consequences. We present a formal deontic logic which is built according to these characteristics and provide the completeness results. Finally, we present a semi-formal method (based on the proposed deontic logic) to test the validity of legal inferences. This paper contributes to the literature insofar as it provides a method that covers a portion of the intuitive validity of legal inferences which is not covered by other frameworks.

Keywords: Legal obligations, Legal discourse, Normative consistency, Deontic consequence, Non Kripkean semantics, Deontic logic, Contingency

Introduction

Deontic logic began with the work of von Wright (1951) and has since been interpreted in many different ways.[1] Although the approaches that we find nowadays within the literature vary significantly, most share a common feature: the use of possible world semantics

[1] There was also a previous proposal made by Mally (1926), but it did not have as much impact as von Wright's at the time.

and the interpretation of deontic logic as a modal logic.[2] Among these approaches the reader will find monadic, dyadic, temporal, non-monotonic, first-order, dynamic and *stit* deontic logics (for an overview, see Peterson 2011).[3] The modal interpretations are usually characterized by the fact that the truth value of a normative proposition depends upon the truth value of the descriptive proposition in the scope of the deontic operator. Modal interpretations are usually of the *Ougth-to-Be* type, as opposed to the *Ought-to-Do* interpretation of deontic modalities (for the distinction, see von Wright 1999b). In the modal interpretations, a deontic proposition $O\varphi$ usually expresses that a specific description φ of the world *ought to be* the case, or that the world *ought to be* in a specific state. However, one can nonetheless find some modal interpretations of the *Ought-to-Do* type, which are characterized by the fact that the proposition in the scope of a deontic operator is instead interpreted as the name of an action. This kind of interpretation can be found notably within the dynamic approach to deontic logic[4], where an action is forbidden when its performance implies that the world is in a state within which there is a violation V.[5] It can also be found in deontic logics based upon a Boolean algebra.[6]

The present paper does not aim at criticizing these approaches but rather aims at providing a new one, based on different philosophical foundations. We propose an interpretation of deontic logic based upon the analysis of (Canadian) legal norms. Our goal is not to provide an analysis of how we use 'ordinary deontic language' or how 'ordinary deontic reasonings' work, as Castañeda (1981, p.38) would say, but is rather to show how a normative reasoning *should* work. The objective is to define a proper consequence relation according to some basic properties that govern a 'correct' use of a (legal) normative inference. We begin by exposing the rationale behind our framework and present the philosophical foundations of our system. The formal deontic logic and the completeness results are then presented. On these grounds, a semi-formal method that can be used to analyze the validity of legal inferences is introduced. The method consists in testing the validity of an argument through its graphical representation. Some insights regarding the analysis of the premises of a legal argument are given and the paradoxes of deontic logic are discussed. We conclude in the last section by summarizing the limitations of our

[2] Some approaches do not use possible world semantics, see for instance Makinson and van der Torre (2000).

[3] See for example Castañeda (1981), Schotch and Jennings (1981), Jones and Pörn (1985; 1986), Hansson (1990) or Jones (1991) for monadic deontic logic; Al-Hibri (1978), Chellas (1974), Mott (1973) and van Fraassen (1972) for dyadic deontic logic; Chellas (1969), McKinney (1977), Åqvist and Hoepelman (1981), Thomason (1981), van Eck (1982a; 1982b) or Bailhache (1991; 1993) for temporal deontic logic and Nute (1997) for non monotonic deontic logic. See also Broersen (2011a), Carmo and Pacheco (2001), Carmo and Jones (2002), Horty (2001) for multi-modal deontic logics, and Belnap and Perloff (1988), Broersen (2008; 2009; 2011a; 2011b), Carmo and Pacheco (2000; 2001), Horty and Belnap (1995), Horty (2001), Pacheco and Carmo (2003), Pacheco and Santos (2004) and Xu (1995) for *stit* logics. The reader may consult Wolenski (1990) for a sketch of possible world semantics in deontic logic.

[4] See Meyer (1987, 1988) for the introduction of dynamic deontic logic and, among others, see Royakkers (1998), Broersen (2004), Hughes and Royakkers (2008), Segerberg (2009), Anglberger (2009), Demolombe (2014), Segerberg (2012) or Prisacariu and Schneider (2012).

[5] This type of reduction is inspired by the work of Anderson (1958a).

[6] See for instance Segerberg (1982b) and Trypuz and Kulicki (2009; 2010).

approach and we present avenues for future research.

Philosophical assumptions

Contra the modal interpretation of deontic logic, we do not consider deontic propositions as being of the *Ought-to-Be* type. Following Solt (1984, p.350), the truth value of a normative proposition $O\varphi$ for the actual world does not depend upon the truth value of φ at any 'deontic alternative'. The truth value of a deontic proposition $O\varphi$ does not rely on the performance value of φ in every accessible 'deontically perfect world'. The fact that φ is obligatory depends upon the existence of a norm, which is established by some authority (Alchourrón and Bulygin 1981, pp.97,102) and aims to guide one's actions (Weinberger 2001, p.134). This is consistent with the legal adage *nullum crimen sine lege*: there is no crime without law, nor any obligation without a norm. Therefore, the truth value of a normative proposition $O\varphi$ does not depend upon the truth value of the descriptive proposition φ in the scope of the deontic operator but depends upon the fact that there is a norm which makes φ obligatory.[7] Following the same reasoning, the truth value of a deontic proposition does not depend upon the fact that there is a 'violation' in every state where the action is performed.

We do not wish to analyze legal obligations within the framework of modal logics since the operator O_i does not behave as a modality of the type $\Box_i$. Although O_i marks the property of an action, this operator is not interpreted as a predicate since it can be applied to combinations of actions, which are expressed by molecular compounds of descriptive formulas. The operator O_i is not interpreted as a modality of type $\Box_i$ since we do not wish to obtain formulas such as $O_i(A \vee \neg A)$ or $O_iA \supset O_i(B \vee \neg B)$, which are dubious from a legal point of view.

It is noteworthy that we are not considering legal discourse as a *normative system*, that is a "set of agents (human or artificial) whose interactions can fruitfully be regarded as norm-governed (Carmo and Jones 2002, p.265)". This point is important: we are not trying to *describe* how agents interact within a normative system. Rather, we wish to define a semantical consequence relation for the validity of legal inferences. Our approach is *normative* rather than *descriptive*. It concerns the analysis of legal reasoning from a critical point of view. This is mainly why we will not use *stit* or dynamic logic: we are not focusing on the notion of *agency*. Rather, we are analyzing the semantical consequence relation within a legal argument.

According to the semantical dichotomy between facts and norms (cf. Jørgensen 1937), descriptive and normative propositions are not true in the same conditions. While a descriptive proposition is true or false with regard to the world it describes, the truth of a normative proposition (which says how the world should be or how people should act, rather than how the world is) depends on the existence of a norm, which is established by some authority. We distinguish between a norm and a normative proposition: the former is

[7] This position is also consistent with the legal literature. See Baudoin (2010).

neither true nor false while the latter can be true or false (the truth value of the normative proposition depending upon the existence of a norm). A normative proposition expresses that an action (either a specific action or a class of specific actions) possesses a deontic property. For example, from the Canadian Criminal Code we can conclude that the action 'stealing a red bicycle' possesses the propriety 'forbidden' or, equivalently, that the action negation 'not stealing a red bicycle' possesses the property 'obligatory'. Assuming that an action possesses a deontic property, we want to develop a basic logic that can represent how this property can be transmitted from one action to another.

One property of legal discourse is that there are no obligations without norms. Following Chellas (1974, p.24), we want to be able to represent situations where there are no obligations, and thus our system must not include 'absolute' or 'unconditional' obligations. Likewise, following Jones and Pörn (1985, p.279), it is always possible for someone to act against one's obligations. Hence, tautologies are not obligations unless there is a norm that makes it so. This is the principle of normative contingency, first introduced by von Wright (1951). Formally, we do not want our system to validate theorems of the form $\vdash O_i\top$ or $\vdash A \supset O_i\top$ (with $\top$ a tautology). Also, since the meaning of a deontic operator changes when it is iterated (Jones and Pörn 1985, p.286), it follows that, according to Castañeda (1981, p.66), it must not be possible to iterate the same deontic operator. For now, our system will concentrate on propositions within which there is only one type of deontic operator (one authority) that cannot be iterated. Finally, we will not be considering mixed propositions (i.e., propositions composed with both descriptive and normative atoms). Indeed, it is unclear whether or not a material conditional is sufficient to represent the transmission of truth value between descriptive and normative propositions since the truth-value assignment for a descriptive proposition differs from the truth-value assignment of a normative one. The connective '$\supset$' does not preserve truth from a descriptive proposition to a normative one. When one tries to model deontic conditionals and contrary-to-duty reasoning, one faces the problems of augmentation and detachment (cf. Jones 1991). For example, it is possible to have a situation where $p \supset P_iq$ is true but $(p \wedge p') \supset P_iq$ is false.[8] Hence, the descriptive antecedent of a deontic conditional cannot be augmented, as the normative consequent cannot always be detached.[9] These are also arguments in favour of non-monotonic foundations for deontic logic.[10] This, however, will be the topic of another paper.[11] For now, we will not be considering mixed formulas and will only concentrate on normative ones.

The deontic logic we propose is based upon an analysis of Canadian legal discourse. Since legal norms are meant to guide one's actions, it follows that norms must be consistent since it would be impossible to act accordingly with an inconsistent set of norms. Thus, we assume the criterion of *normative consistency*: obligations are supposed to be consistent

[8] If Paul has a driver's license, then he is permitted to drive, but if in addition he is drunk, then he is not.

[9] This is why dyadic deontic logics were introduced. See Loewer and Belzer (1983) and Vorobej (1986) for a discussion.

[10] As a result, we will not be considering defeasible reasoning (see for example Horty (1994), Sartor (1994), van der Torre and Tan (1997) or Governatori and Rotolo (2004)).

[11] See Peterson (2014b; 2014d).

(at least after interpretation) since the legislator is presupposed to be rational, meaning that it is assumed that he thinks rationally and logically (Côté 2006, p.387). The set of norms created by the legislator is therefore presupposed to be consistent, and thus legal obligations must not be interpreted as contradictory. Consistency is a rational criterion that ensures the accessibility, the authority and the equity of the law (Côté 2006, p.387). Canadian legal discourse is considered as forming a 'logical system' which is *horizontally* and *vertically* consistent (after interpretation). As such, obligations from the same set of norms must be consistent with each other, as are the different sets of norms (Côté 2006, p.388). In other words, the set of all legal laws is presupposed to form a consistent whole (Côté 2006, p.433). Consistency is a rational criterion that enables one to judge the value of a set of norms, which can be examined through the consistency of the set of obligations that it generates.[12]

The relation of hierarchy between norms is meant to preserve a legal system's consistency. It is necessary to insure *vertical* and *horizontal* consistency. Since there can be (*a priori*) contradictions between different sets of laws or different laws within a set of norms, hierarchy enables us to preserve the consistency of the whole system by resolving potential conflicts of obligations. Moreover, laws are constructed in an hierarchical manner. For example, in Canada, no legislation can go against the Canadian Constitution. The fact is that there are sets of laws that are superior to others by construction, although these relations of hierarchy can also be explicitly mentioned in the law itself (Côté 2006, p.45 and p.450).[13]

Legal reasoning is characterized by the fact that it is hypothetical: a legal conclusion cannot be drawn without a legal premise. This is consistent with semantical dichotomy. A valid legal inference implies that there is a set of hypothetical norms (or obligations) from which legal conclusions are drawn. Thus, an obligation is conditional to a set of norms (or principles), which is established by some authority. The fact that obligations are derived from norms implies that normative inferences are governed by a specific criterion of validity. Since there is a finite number of norms, it would seem that there is also a finite number of obligations that will derive directly from these norms. For example, we can derive from the law that it is forbidden to steal. However, norms are often formulated in order to be applied to a class of specific actions. The obligation that derives directly from the law is to *not steal*, but clearly it is meant to be applied to a class of actions: not stealing a car, not stealing money, not stealing Paul's money, not stealing Peter's car, not stealing Paul's car and Peter's bicycle, etc. In other words, there is a much greater number of obligations that can be derived from a specific norm. But how do we legitimately conclude them since they are not *explicitly* mentioned by the norms?

To answer this question, we follow Alchourrón and Bulygin (1981, p.102) and distinguish between *fixed* and *derived* obligations. A fixed obligation stems from a norm and can

[12] The fact that many deontic logics that satisfy the axiom schema (**D**) cannot represent conflicting obligations is often seen as an argument in favour of non-monotonic deontic logic (see for example Horty 1997). Since we only concentrate on one authority, normative consistency can be assumed, otherwise the legal authority would force us to act within an impossible frame.

[13] The use of hierarchy to solve normative conflicts was advocated by Alchourrón and Makinson (1981).

be general in order to be applied to a class of specific actions. A finite number of norms will hence entail a finite number of fixed obligations. A derived obligation can be inferred from a fixed one. From *it is forbidden to steal* we can derive *Paul has the obligation to not steal Peter's money.* The rule which governs this type of inference is the *principle of deontic consequences* (Castañeda 1968, p.13)[14]: if A is an obligation and A implies B, then B is also an obligation. The mere formulation of the law implies that there are derived obligations. In *French Civil Law* it is obvious since the law is formulated in a general way and applies implicitly to each particular case that falls within its scope. Even though it seems to be the contrary in *Common Law*, where the law is constructed upon each particular judgment, the same principle applies nonetheless since a judgment applies to a class of particular actions.[15] Even if a judgment comes from a particular action, the case law applies to other similar cases. Put differently, there are actions that are obligatory (or forbidden) even though they are not mentioned explicitly by the law (or the case law). The law is not an enumeration of all possible cases. The law says that it is forbidden to steal. What it means is that it is forbidden to do any action that implies (or *count-as*) stealing.[16] Furthermore, normative consistency implies that legal obligations are governed by the principle of deontic consequences. Indeed, assuming that the legislator is rational, it is possible to deduce from the law certain things that logically follow (Côté 2006, p.422). The aim of the present paper is to define this consequence relation.

To sum up, the analysis of Canadian legal norms brought to light three fundamental characteristics of legal discourse, namely its presupposed consistency, the fact that hierarchy is meant to preserve that consistency and the validity of the principle of deontic consequences, which enables one to infer derived obligations from fixed ones. If one is to apply formal philosophy to law, then one must take these characteristics into account.

The logic of legal obligations

According to these characteristics, the logic we propose relies upon a consequence relation which is defined as a function of normative consistency and deontic consequences. For now, we do not need to include hierarchy within the formal definition since we are only considering one type of deontic operator which cannot be iterated. However, as we will see, hierarchy will play a role for the analysis of the soundness of an argument. The basic idea of our approach is to define a consequence relation which can represent how the property 'obligatory' can be transmitted from an action (or a combination of actions) to another. We follow Castañeda (1981, p.46) and assume a distinction between different types of obligations, where the type of an obligation depends upon the set of norms from which it can be derived (Alchourrón and Bulygin 1981, p.120).

[14] See also van Fraassen (1972, p.421).

[15] Note that Quebec's legislation (in Canada) is composed of both French Civil Law and Common Law.

[16] We are not pretending that logic can resolve the problem of interpreting the law. We only say that laws are formulated in order to include different kinds of actions, which has nothing to do with deciding whether the action falls within the scope of the law or not.

According to the semantical dichotomy between facts and norms, the idea is to develop a framework which distinguishes between factual and normative truth. As such, the language $\mathcal{L}$ will be divided in two sub-languages $\mathcal{L}_{PL}$ and $\mathcal{L}_{OL}$, respectively representing the language of descriptive propositions expressed by propositional logic (PL) and the language of normative propositions expressed by the logic of (unconditional) obligations OL.[17] Similarly, the semantical model $\mathfrak{M}$ will be divided in two sub-models $\mathcal{M}$ and $\mathcal{N}$, the first being descriptive and the other normative. We assume that the model $\mathcal{M}$ of $\mathcal{L}_{PL}$ is a standard model of propositional logic. A proposition is said to be derivable in $\mathcal{L}$ if and only if either it is a consequence of PL or it is a consequence of OL. Similarly, a proposition A is said to be *valid* for $\mathfrak{M}$ if and only if either it is *descriptively* or *normatively* valid. For the rest of this paper, we concentrate only upon the normative fragments $\mathcal{L}_{OL}$ and $\mathcal{N}$.

Syntax

The formulas of OL have the form O_iA, where O_i is a single type of obligation and A is a proposition that refers to an action or a combination of actions. We assume a language

$$\mathcal{L} = \{(,), Prop, \neg, \supset, O_i\}$$

where $Prop = \{p_1, \dots, p_n, \dots\}$ is a denumerable set of propositional descriptive atoms and descriptive propositions are understood as descriptions of actions (not any descriptions). The other logical connectives $\wedge$, $\vee$, $\equiv$, F_i and P_i are defined as usual, with:

$$F_iA =_{def} O_i\neg A \qquad \text{(Interdiction)}$$
$$P_iA =_{def} \neg O_i\neg A \qquad \text{(Weak permission)}$$

We use A, B and C as meta-variables. The set $WFF_{\mathcal{L}}$ of well-formed formulas of $\mathcal{L}$ is defined recursively by:

$$p_i \in WFF_{\mathcal{L}_{PL}} \text{ for all } p_i \in Prop \tag{11.1}$$
$$\text{if } A, B \in WFF_{\mathcal{L}_{PL}}, \text{ then } \neg A,\ A \supset B \in WFF_{\mathcal{L}_{PL}} \tag{11.2}$$
$$\text{if } A \in WFF_{\mathcal{L}_{PL}}, \text{ then } O_iA \in WFF_{\mathcal{L}_{OL}} \tag{11.3}$$
$$\text{if } A, B \in WFF_{\mathcal{L}_{OL}}, \text{ then } \neg A,\ A \supset B \in WFF_{\mathcal{L}_{OL}} \tag{11.4}$$
$$\text{if } A \in WFF_{\mathcal{L}_{PL}} \text{ or } A \in WFF_{\mathcal{L}_{OL}}, \text{ then } A \in WFF_{\mathcal{L}} \tag{11.5}$$

Thus defined, $WFF_{\mathcal{L}}$ does not contain any mixed formulas or any formula where there is an iteration of a deontic operator. The axiom schema for normative consistency

[17] Since we assumed that every obligation derives from a norm (established by some authority), it follows that every obligation is 'conditional' in nature (i.e., obligations are conditional to the existence of norms). This type of conditionality, however, is different from the distinction between *conditional* and *unconditional* obligations. A conditional obligation is formulated in such a way that it specifies in which context the obligation holds. An unconditional obligation is understood as an obligation that always holds, unless specified by another norm in some specific context. Note that in such a context, OL would not apply anymore and we would need a logic for conditional normative reasoning.

is represented by (A1), which is propositionally equivalent to the axiom (**D**) of standard deontic logic.

$$\neg(O_i A \wedge O_i \neg A) \tag{A1}$$

The principle of deontic consequences is represented by the rule of inference (R1) and its use is restricted by two conditions.

$$\begin{array}{l|ll} 1 & \vdots & \\ 2 & O_i A_1 \wedge \cdots \wedge O_i A_n & (n \geq 1) \\ 3 & (A_1 \wedge \cdots \wedge A_n) \supset B & \\ 4 & O_i B & \text{R1 l.2,3} \end{array} \tag{R1}$$

First, $\{A_1, \ldots, A_n\}$ cannot be the empty set $(n \geq 1)$, otherwise B would be a theorem of PL and there would be a set of absolute obligations. Since every obligation is conditional to a norm, it follows that B must be a consequence that depends upon $A_1 \wedge \cdots \wedge A_n$. This condition is meant to preserve the principle of normative contingency. Put differently, (R1) cannot be used if there is no normative hypothesis. Second, the conditional

$$(A_1 \wedge \cdots \wedge A_n) \supset B$$

must be either a theorem of PL or a hypothesis. As a result, (R1) cannot be applied to *any* material conditional but can only be applied to a conditional which is either a theorem of PL or a hypothesis. For instance, the following use would be an *incorrect* application of (R1).

$$\begin{array}{l|ll} 1 & p & \text{H} \\ 2 & O_i q & \text{H} \\ \hline 3 & p \supset (q \supset p) & PL \\ 4 & q \supset p & \text{MP 1,3} \\ 5 & O_i p & \text{INCORRECT USE OF R1} \end{array}$$

Let us note that since (R1) can be applied to an hypothesis, it follows that this rule is invalid in a standard system such as KD. We introduce OL in a natural deduction system. We say that A is a theorem of OL, written $\vdash_{OL} A$, when there is a proof of A without the use of any hypothesis. In addition to (R1), we assume the rules of PL (and hence also the derived rules of PL) of Garson (2006, p.35), that is:

1. hypothesis (H);
2. reiteration (Reit);
3. detachment ($\supset$ *out*);

4. conditional proof ($\supset in$);

5. double negation (DN).

Since negation is taken as a primitive, we also need a rule to govern its introduction:

$$
\begin{array}{l|l l}
1 & A & \text{H} \\
\cline{2-2}
2 & \vdots & \\
3 & B & \\
4 & \neg B & \\
\multicolumn{1}{l}{5} & \multicolumn{1}{l}{\neg A} & (\neg in)
\end{array}
$$

Although OL contains only normative propositions, the proofs of the theorems are done in $\mathcal{L}$. Thus constructed, OL (and moreover $\mathcal{L}$) behaves at the propositional level according to the rules of PL. We write 'PL' as a justification in a proof when the formula is a theorem of PL.[18]

Semantics

Let $Act = \{\Gamma_A : \Gamma_A \text{ is a positive action}\}$ be a denumerable set of *positive* actions (e.g., walking, talking, stealing, etc.), where Γ_A stands for the action described by the proposition A, and $\overline{Act} = \{\overline{\Gamma}_A : \overline{\Gamma}_A \text{ is a negative action}\}$ a denumerable set of *negative* actions (e.g., not walking, not talking, not stealing, etc.). The set $\mathbb{A} = Act \cup \overline{Act}$ is thus the denumerable set of all possible actions (actions in Act do not need to be atomic).[19] Each action in $\mathbb{A}$ can be described by a proposition of WFF_{PL}.

Let the sequence $s^\alpha =< s_1^\alpha, ..., s_n^\alpha, ... >$ be an arbitrary enumeration of Act and the sequence $\overline{s}^\alpha =< \overline{s}_1^\alpha, ..., \overline{s}_n^\alpha, ... >$ an enumeration of $\overline{Act}$, where $\overline{s}_i^\alpha$ refers to the negation of the action s_i^α. A propositional variable p_i (or a molecular compound A_i) refers to an action s_i^α member of an arbitrary sequence s^α, while $\neg p_i$ (or $\neg A_i$) refers to $\overline{s}_i^\alpha$. If A_i refers to s_i^α, then s_i^α is the action Γ_{A_i} (described by A_i), that is the i^{th} member of s^α. An arbitrary sequence s^α (and its counterpart $\overline{s}^\alpha$) is an interpretation of the language $\mathcal{L}$. It assigns a descriptive proposition of WFF_{PL} to each object of the domain (i.e., each action).[20]

[18] One the one hand, it should be noted that our syntactical system is similar to that of von Wright (1951) (see the axiomatization in von Wright (1967)). However, the two are not equivalent, since $\nvdash_{OL} O_i(A \supset A) \supset O_i(A \vee \neg A)$. On the other hand, even if (R1) may look like the Rule 0 in Alchourrón (1990), they differ from the important fact that $\{A_1, ..., A_n\}$ cannot be the empty set. Also, his syntactical system is equivalent to KD, which is not equivalent to OL since (R1) is invalid in the standard system.

[19] Note that our understanding of 'action' includes both action types and action tokens.

[20] To be precise, it assigns an equivalence class of propositions to an equivalence class of actions.

In order to formalize semantically the principle of deontic consequences, we assume that the set $\mathbb{A}$ is pre-ordered by '$\sqsubseteq$'. For example, if we suppose that

$$\begin{aligned} p_1 &= \text{Peter steals a red bicycle} \\ p_2 &= \text{Peter steals a bicycle} \\ p_3 &= \text{Peter steals} \end{aligned}$$

then we have $\Gamma_{p1} \sqsubseteq \Gamma_{p2} \sqsubseteq \Gamma_{p3}$. This allows us to give a semantical account of the relation of implication between actions. If an action Γ_A implies another action Γ_B, then $\Gamma_A \sqsubseteq \Gamma_B$. Thus, if we assume that 'stealing implies not violating the law' is true in some (descriptive) interpretation, then we assume that 'stealing' $\sqsubseteq$ 'not violating the law' in some normative model. Let $\mathcal{M}$ be a standard model of PL. In other words, if $\models_{\mathcal{M}} p_1 \supset p_2$ (i.e., $p_1 \supset p_2$ is assumed to be true in $\mathcal{M}$), then $\Gamma_{p_1} \sqsubseteq \Gamma_{p_2}$ holds in the normative interpretation. We do not assume the converse because it would lead us to ideality. (It is not because an entailment between actions is assumed in a normative interpretation that it is necessarily true in the descriptive one. For instance, one can assume that 'it is obligatory that *if one is in a public place, then one is not naked*' holds in a normative model while the conditional in the scope of the operator is false in a descriptive one.)

Let $\mathcal{N} = \langle W, \mathbb{A}, \sqsubseteq, a\rangle$ be a normative model, where $W \neq \varnothing$ is the universe of discourse and contains normative propositions (which are members of $WFF_{\mathcal{L}_{OL}}$), $\mathbb{A}$ is a denumerable set of actions pre-ordered by '$\sqsubseteq$' and $a : W \longrightarrow \{\top, \bot\}$ a function which assigns truth values to propositions in W. Let $\mathcal{O}$ be a proper subset of $\mathbb{A}$ ($\mathcal{O} \subsetneq \mathbb{A}$). Informally, $\mathcal{O}$ is a set of actions which have the property 'obligatory', meaning that $\mathcal{O}$ is the extension of the concept 'obligatory' within a normative interpretation.[21] The set has to satisfy three conditions:

$$\text{If } \Gamma_A \in \mathcal{O}, \text{ then } \overline{\Gamma}_A \notin \mathcal{O} \tag{C1}$$

$$\text{If } \Gamma_A \in \mathcal{O}, \text{ then } \Gamma_B \in \mathcal{O} \text{ for any } \Gamma_A \sqsubseteq \Gamma_B \text{ such that } \not\models_{PL} B \tag{C2}$$

$$\text{If } \Gamma_A \in \mathcal{O} \text{ and } \Gamma_A \sqsubseteq \Gamma_{B\supset C}, \text{ then } \Gamma_B \sqsubseteq \Gamma_C \tag{C3}$$

The first condition represents normative consistency and the second implies that $\mathcal{O}$ is closed 'upwards' when B is not a tautology of the (classical) propositional calculus. The second part of C2 insures that normative contingency is respected. The third condition says that if the action described by A is semantically linked to the action described by $B \supset C$, then the action described by B is semantically linked to the action described by C insofar as the action described by A is obligatory. C2 and C3 allow us to represent the principle of deontic consequences. For a normative interpretation $\mathcal{N}$, $a_{\mathcal{N}}(A) = \top$ if and only if there is a sequence $s^\alpha =< s_1^\alpha, ..., s_n^\alpha, ... >$ which satisfies A. We now define satisfaction recursively.

Definition 11.1. *For any A, s^α* satisfies *A if and only if*

[21] If there are more than one authority, $\mathcal{O}$ can be indexed by $_i$.

1. *If A is O_iB, then*

   (a) *If B is p_i, then s^α satisfies A iff s^α satisfies p_i iff $s^\alpha_i \in \mathcal{O}$ (i.e. the i^{th} member of s^α is a member of $\mathcal{O}$).*

   (b) *If B is $\neg C$, then s^α satisfies A iff $\overline{s}^\alpha$ satisfies C.*

      i. *$\overline{s}^\alpha$ satisfies p_i iff $\overline{s}^\alpha_i \in \mathcal{O}$*

      ii. *$\overline{s}^\alpha$ satisfies $\neg D$ iff $\overline{\overline{s}}^\alpha = s^\alpha$ satisfies D*

      iii. *$\overline{s}^\alpha$ satisfies $D \supset E$ iff s^α satisfies D and $\overline{s}^\alpha$ satisfies E*

   (c) *If B is $C \supset D$, then s^α satisfies A iff either*

      i. *there is $s^\alpha_m = \Gamma_{C \supset D}$ such that $s^\alpha_m \in \mathcal{O}$ or*

      ii. *s^α satisfies D (provided that Γ_C is not Γ_D).*[22]

2. *If A is $\neg B$, then s^α satisfies A iff s^α does not satisfy B.*

3. *If A is $B \supset C$, then s^α satisfies A iff s^α does not satisfy B or s^α satisfies C.*[23]

We say that a normative proposition A is *normatively valid*, written $\models A$, when it is true for any normative interpretation $\mathcal{N}$.[24]

Completeness

The following lemmas will be useful. As a notational convention, we will write s^α_A to refer to the action Γ_A (i.e., the action described by the descriptive proposition A), and which is the m^{th} member of s^α. We also write $s^\alpha_{\neg A}$ instead of $\overline{s}^\alpha_A$.

Lemma 11.1. *If s^α satisfies O_iA, then $s^\alpha_A \in \mathcal{O}$.*

Proof. We proceed inductively on the length of the formula. Suppose that s^α satisfies O_iA but that $s^\alpha_A \notin O$. (The inductive step (HI) is that if the property holds for $l = n$, then it also holds for $l = n + 1$.)

1. A is p_i, thus s^α satisfies p_i and $s^\alpha_{p_i} \in \mathcal{O}$.

2. A is $\neg B$, thus $\overline{s}^\alpha$ satisfies B.

   (a) B is p_i, thus $\overline{s}^\alpha$ satisfies p_i and $s^\alpha_{\neg p_i} \in \mathcal{O}$.

[22] Informally, the first condition means that the conditional represents a fixed obligation while the second represents a derived obligation. The condition that Γ_C is not Γ_D is meant to preserve normative contingency.

[23] Let us note that it would be incorrect to infer from the fact that $\overline{s}^\alpha$ satisfies p_i that it also satisfies $A \supset p_i$ from condition 1c). Rather, if $\overline{s}^\alpha$ satisfies p_i, then s^α satisfies $\neg p_i$, and thus s^α satisfies $A \supset \neg p_i$.

[24] There is an equivalence of notation between $\models_{\mathcal{N}} A$ and $a_{\mathcal{N}}(A) = \top$. Also, from this definition it follows that $a_{\mathcal{N}}(A) = \top \Leftrightarrow a_{\mathcal{N}}(\neg A) = \bot$.

(b) B is $\neg C$, thus $\overline{s}^\alpha$ satisfies $\neg C$, meaning that s^α satisfies C and by (HI) $s^\alpha_C \in \mathcal{O}$, and thus $s^\alpha_{\neg\neg C} \in \mathcal{O}$ by C2 since $\models_{PL} C \supset \neg\neg C$ and $\Gamma_C \sqsubseteq \Gamma_{\neg\neg C}$.

(c) B is $C \supset D$, thus $\overline{s}^\alpha$ satisfies $C \supset D$, meaning that s^α satisfies C and $\overline{s}^\alpha$ satisfies D. But by (HI) $s^\alpha_C \in \mathcal{O}$ and $s^\alpha_{\neg D} \in \mathcal{O}$, and since $\models_{PL} C \supset (\neg D \supset \neg(C \supset D))$, we have $s^\alpha_C \sqsubseteq s^\alpha_{\neg D \supset \neg(C \supset D)}$, thus by C3 we have $s^\alpha_{\neg D} \sqsubseteq s^\alpha_{\neg(C \supset D)}$, and by C2 we obtain $s^\alpha_{\neg(C \supset D)} \in \mathcal{O}$ since $s^\alpha_{\neg D} \in \mathcal{O}$.

3. A is $B \supset C$, thus either

(a) s^α satisfies $\Gamma_m = B \supset C$ and thus $s^\alpha_{B \supset C} \in \mathcal{O}$

(b) s^α satisfies C ($\Gamma_B \neq \Gamma_C$) and by (HI) $s^\alpha_C \in \mathcal{O}$ and by C2 $s^\alpha_{B \supset C} \in \mathcal{O}$ since $\models_{PL} C \supset (B \supset C)$ and thus $\Gamma_C \sqsubseteq \Gamma_{B \supset C}$.

□

Lemma 11.2. *If $s^\alpha_A \in \mathcal{O}$, then s^α satisfies O_iA.*

Proof. We proceed inductively on the length of the formula. Suppose that $s^\alpha_A \in \mathcal{O}$ but that s^α does not satisfy O_iA.

1. A is p_i, thus $s^\alpha_{p_i} \in \mathcal{O}$ and so s^α satisfies p_i.

2. A is $\neg B$, thus $\overline{s}^\alpha$ does not satisfy B.

(a) B is p_i, thus $s^\alpha_{\neg p_i} \in \mathcal{O}$ and so $\overline{s}^\alpha$ satisfies p_i

(b) B is $\neg C$, thus $\overline{s}^\alpha$ does not satisfy $\neg C$, meaning that s^α does not satisfy C. Since $\models_{PL} \neg\neg C \supset C$, we obtain $s^\alpha_{\neg\neg C} \sqsubseteq s^\alpha_C$ and by C2 we have $s^\alpha_C \in \mathcal{O}$, and by (HI) s^α satisfies C.

(c) B is $C \supset D$, thus $\overline{s}^\alpha$ does not satisfy $C \supset D$, meaning that either s^α does not satisfy C or $\overline{s}^\alpha$ does not satisfy D. By hypothesis, we have $s^\alpha_{\neg(C \supset D)} \in \mathcal{O}$, and by PL we obtain $s^\alpha_{\neg(C \supset D)} \sqsubseteq s^\alpha_C$ and $s^\alpha_{\neg(C \supset D)} \sqsubseteq s^\alpha_{\neg D}$ since $\models_{PL} \neg(C \supset D) \supset C$ and $\models_{PL} \neg(C \supset D) \supset \neg D$. Therefore, by C2 we obtain $s^\alpha_C \in \mathcal{O}$ and $s^\alpha_{\neg D} \in \mathcal{O}$, which by (HI) implies that s^α satisfies C and $\overline{s}^\alpha$ satisfies D.

3. A is $B \supset C$, thus s^α does not satisfy $\Gamma_m = B \supset C$ and either s^α does not satisfy C or $s^\alpha_B = s^\alpha_C$. However, by definition if s^α does not satisfy Γ_m, it implies that $s^\alpha_m \notin O$, that is $s^\alpha_{B \supset C} \notin O$.

□

Lemma 11.3. *If $\Gamma_A \sqsubseteq \Gamma_B$, $\not\models_{PL} B$ and s^α satisfies O_iA, then s^α satisfies O_iB.*

Proof. Assume $\Gamma_A \sqsubseteq \Gamma_B$, $\not\models_{PL} B$ and s^α satisfies O_iA but s^α does not satisfy O_iB. By lemma 11.1 we have $s^\alpha_A \in \mathcal{O}$, thus by C2 $s^\alpha_B \in \mathcal{O}$ and by lemma 11.2 s^α satisfies O_iB. □

Lemma 11.4. *(A1) preserves validity.*

Proof. Suppose that (A1) does not preserve validity. Thus, we have $\vdash_{OL} \neg(O_i A \wedge O_i \neg A)$ and $\not\models_{\mathcal{N}} \neg(O_i A \wedge O_i \neg A)$. However, if $\not\models_{\mathcal{N}} \neg(O_i A \wedge O_i \neg A)$, then $\models_{\mathcal{N}} O_i A \wedge O_i \neg A$. It follows that s^α satisfies $O_i A$ and $O_i \neg A$, meaning that s^α satisfies A and it satisfies $\neg A$. By lemma 11.1, it implies that both $s^\alpha_A \in \mathcal{O}$ and $s^\alpha_{\neg A} \in \mathcal{O}$, which contradicts C1. $\square$

Lemma 11.5. *(R1) preserves validity.*

Proof. Assume that (R1) does not preserve validity. This means that we have a situation where $O_i A \vdash_{OL} O_i B$ is obtained by the use of (R1) but $O_i A \not\models_{\mathcal{N}} O_i B$, that is $\models_{\mathcal{N}} O_i A$ and $\not\models_{\mathcal{N}} O_i B$, and so s^α satisfies $O_i A$ but does not satisfy $O_i B$. By the use of (R1), we know that $A \supset B$ is true by hypothesis and that $\not\models_{PL} B$, and moreover that $\Gamma_A \sqsubseteq \Gamma_B$. Therefore, by lemma 11.3 s^α satisfies $O_i B$. $\square$

Theorem 11.1 (Adequacy)**.** *If $\vdash_{OL} A$, then $\models A$.*

Proof. Since OL is based upon PL, it suffices to show that (A1) and (R1) preserve validity, which follows from lemmas 11.4 and 11.5. $\square$

Lemma 11.6 (Lindenbaum's lemma)**.** *OL has a maximally consistent extension.*

Suppose $K = \bigcup_0^\infty K_i$, with $K_0 = OL$ the smallest set of wffs of $\mathcal{L}$ closed under the rules of PL, (A1) and (R1). Let $A_1, \ldots, A_n, \ldots$ be an arbitrary enumeration of OL's wffs. If $\vdash_{K_{n-1}} \neg A_n$, then $K_n = K_{n-1}$, else $K_n = K_{n-1} \cup \{A_n\}$. This way, we have K_i extension of OL for all $i \geq 0$. It is obvious that K is maximal since by construction either $A_i \in K$ or $\neg A_i \in K$ for all i. We now show that K is consistent.

Proof. Assume $K \vdash \bot$. Thus, there is a finite proof of $\bot$ from a finite subset of K, meaning that there is K_n such that $K_n \subset K$ and $K_n \vdash \bot$. If K_n is inconsistent, it implies that by construction the proposition A_n added to K_{n-1} broke K_n's consistency. However, such a situation is impossible. Indeed, this means that $K_n \vdash A_n \wedge \neg A_n$, with $K_n = K_{n-1} \cup \{A_n\}$ and $K_{n-1} \vdash \neg A_n$. But if $K_{n-1} \vdash \neg A_n$, then $K_n = K_{n-1}$, thus $K_n \nvdash \bot$. And if $K_{n-1} \nvdash \neg A_n$, then $K_n = K_{n-1} \cup \{A_n\}$, thus $K_n \nvdash \bot$ since $K_n \nvdash \neg A_n$. Therefore, K is consistent.

$\square$

Lemma 11.7. *$\mathcal{N}$ is a model of K.*

Proof. Assume that $\mathcal{N}$ is not a model of some maximally consistent extension K. It follows that there is a proposition A such that $\vdash_K A$ and $K \not\models_{\mathcal{N}} A$. But if $\vdash_K A$, it means that there is a finite proof of A from a finite subset of K, thus $\vdash_{K_n} A$ and $K_n \not\models_{\mathcal{N}} A$. However, since every subset of K is an extension of OL, it implies that $K_n \vdash_{OL} A$. By theorem 11.1, if $K_n \vdash_{OL} A$, then $K_n \models_{\mathcal{N}} A$, which contradicts our first hypothesis. Therefore, $\mathcal{N}$ is a model of K.

$\square$

Theorem 11.2 (Completeness)**.** *If $H \models C$, then $H \vdash_{OL} C$.*

Proof. Assume that there is a maximally consistent extension of OL where $H \models_{\mathcal{N}} C$ and $H \nvdash_K C$. Since K is maximally consistent, it follows that $H \vdash_K \neg C$, and since $\mathcal{N}$ is a model of K it implies that $H \models_{\mathcal{N}} \neg C$. However, if $H \models_{\mathcal{N}} C$, then $H \not\models_{\mathcal{N}} \neg C$, which contradicts our first hypothesis. Therefore, if $H \models_{\mathcal{N}} C$, then $H \vdash_K C$ for any $\mathcal{N}$.

□

The reader may note that the deduction theorem does not need to be proven in a natural deduction system since it follows immediately from the conditional proof rule.

Validity of legal reasoning

Graphical representation

The idea that lies behind this approach is that, according to the semantical dichotomy, a scenario w is divided in two parts : one descriptive ($\mathfrak{D}$) and the other normative ($\mathfrak{N}$), which contains a set of obligations $\mathcal{O}$.[25] A scenario w is inconsistent if either $\mathfrak{D}$ or $\mathfrak{N}$ is. The scenario can be seen as a model $\mathfrak{M}$ where the descriptive part represents $\mathcal{M}$ and the normative one $\mathcal{N}$. The validity of an argument is tested via the notion of a *counterexample*, that is, a scenario w in which the premises are assumed to be true but the conclusion is false. Since a valid argument does not possess any counterexample, it follows that if the scenario is consistent, then the argument is invalid. As such, if the scenario is consistent, there is a truth-value assignment which makes the premises true but the conclusion false. Otherwise, if the scenario is inconsistent (i.e., it is impossible for the premises to be true while the conclusion is false), then the argument is valid. The propositional rules (figure 11.1) representing schematically truth conditions for complex propositions are quite straightforward (cf. Garson 2006, p.91).

$$
\begin{array}{c} A \supset B \\ \diagup \quad \diagdown \\ \neg A \quad\; B \end{array}
\qquad
\begin{array}{c} \neg(A \supset B) \\ | \\ A \\ \neg B \end{array}
\qquad
\begin{array}{c} A \wedge B \\ | \\ A \\ B \end{array}
\qquad
\begin{array}{c} \neg(A \wedge B) \\ \diagup \quad \diagdown \\ \neg A \quad \neg B \end{array}
\qquad
\begin{array}{c} A \vee B \\ \diagup \quad \diagdown \\ A \quad\; B \end{array}
\qquad
\begin{array}{c} \neg(A \vee B) \\ | \\ \neg A \\ \neg B \end{array}
\qquad
\begin{array}{c} \neg\neg A \\ | \\ A \end{array}
$$

Figure 11.1: Propositional rules

The semantical dichotomy implies that normative and descriptive atoms are not true in the same conditions. Otherwise, complex normative formulas behave schematically as complex descriptive formulas since both are composed of the same logical connectives (figure 11.1). The difference between the truth value of normative and descriptive formulas can be seen at the atomic level. While the truth of a descriptive proposition can be

[25] This section summarizes the results presented in Peterson (2013a).

represented by the fact that it belongs to the descriptive part of a scenario, the truth of a normative proposition depends upon a norm established by some authority. A normative proposition A is obligatory if and only if there is a norm which makes the action described by A (i.e., Γ_A) an obligation. As such, if a normative proposition O_iA is true for a scenario w, then there is a norm (established by authority i) which makes Γ_A obligatory and the action described by A pertains to a set of obligations. Considering that the logical connectives can (classically) be reduced to compositions of $\neg$ and $\supset$, we only represent schematically the truth conditions for O_ip, $O_i\neg A$ and $O_i(A \supset B)$ (figure 11.2). Note, however, that derived rules could easily be constructed from the lemmas that the reader will find at the end of this section. These rules, combined with conditions C1, C2 and C3, allow us to test the validity of a normative inference. Their construction is possible according to lemmas 11.1 and 11.2.

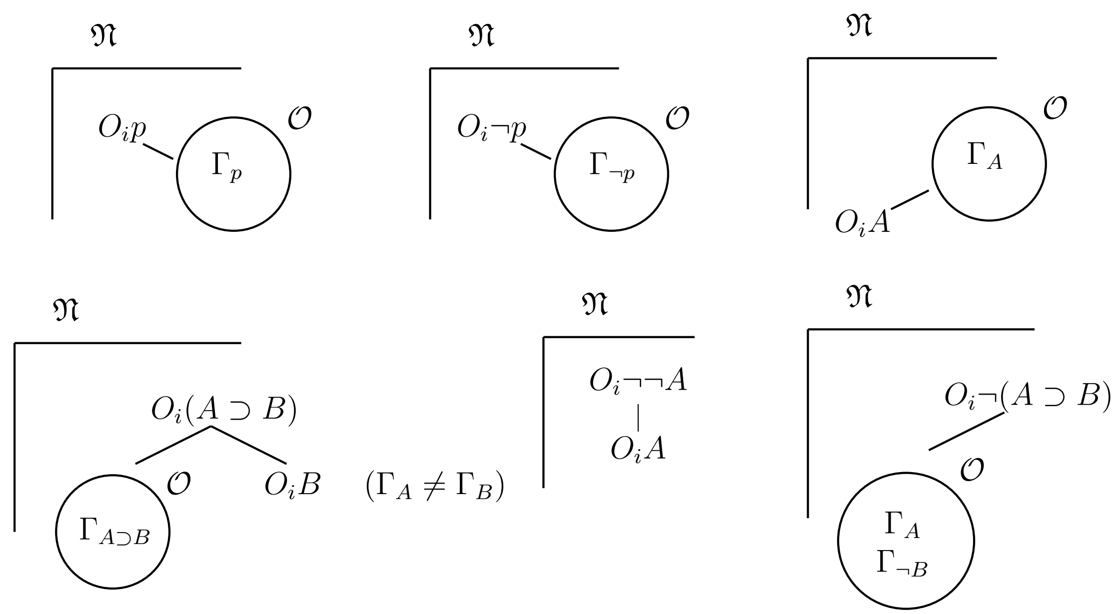


Figure 11.2: Truth conditions of normative atoms

As an example, let us test (R1)'s validity with $n = 1$ (figure 11.3). Assume that w is a counter-example for (R1). Then, place the assumptions in their respective descriptive or normative part of w. According to the rules for normative atoms, Γ_p is in $\mathcal{O}$. Since $p \supset q$ is assumed to be true in the descriptive part of w, it follows that $\Gamma_p \sqsubseteq \Gamma_q$ holds in its normative part. By C2 we obtain Γ_q in O, hence Oq can be concluded in the normative part of w. The only branch in the normative part closes (schematically represented by $\times$) since it contains a contradiction (i.e., O_iq and $\neg O_iq$). Therefore, there is no possible scenario in which the premises of (R1) are true while its conclusion is false, hence the proof of (R1)'s validity.

We now list some semantical properties of the model. These properties are useful for both the formal semantical proofs and the graphical representation. The following lemmas can be used to construct derived rules.

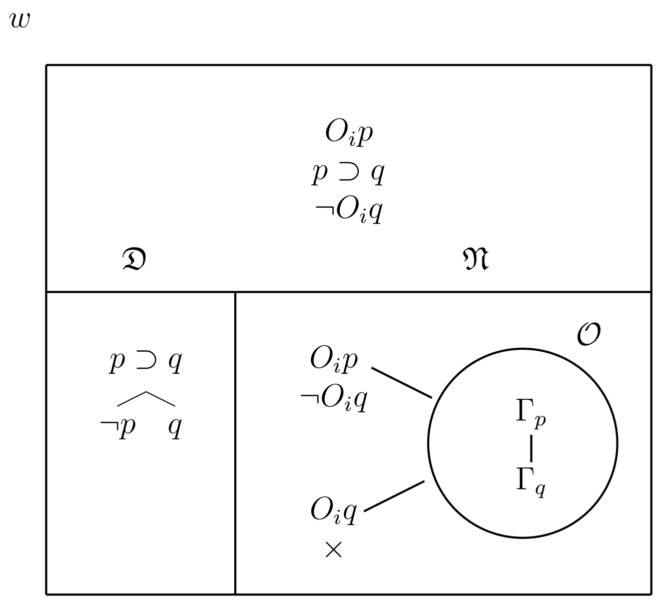


Figure 11.3: Test of (R1)'s validity

Lemma 11.8. *If $\Gamma_A \in \mathcal{O}$, then $\Gamma_{B \supset A} \in \mathcal{O}$ for any B.*

Proof. Assume that $\Gamma_A \in \mathcal{O}$. By PL we know that $A \supset (B \supset A)$ is true in $\mathfrak{D}$, thus $\Gamma_A \sqsubseteq \Gamma_{B \supset A}$ holds in $\mathfrak{N}$. Hence, by C2 $\Gamma_{B \supset A} \in \mathcal{O}$. □

Lemma 11.9. *If $\Gamma_A \in \mathcal{O}$ and $\Gamma_B \in \mathcal{O}$, then $\Gamma_{A \wedge B} \in \mathcal{O}$.*

Proof. Assume $\Gamma_A \in \mathcal{O}$ and $\Gamma_B \in \mathcal{O}$. By PL we have $A \supset (B \supset (A \wedge B))$ true in $\mathfrak{D}$, thus $\Gamma_A \sqsubseteq \Gamma_{B \supset (A \wedge B)}$ holds in $\mathfrak{N}$. By C2, we obtain $\Gamma_{B \supset (A \wedge B)} \in \mathcal{O}$, and by C3 $\Gamma_B \sqsubseteq \Gamma_{A \wedge B}$. Therefore, by C2 we have $\Gamma_{A \wedge B} \in \mathcal{O}$. □

Lemma 11.10. *If $\Gamma_{A \wedge B} \in \mathcal{O}$, then $\Gamma_A \in \mathcal{O}$.*

Proof. Assume $\Gamma_{A \wedge B} \in \mathcal{O}$. By PL we have $(A \wedge B) \supset A$ true in $\mathfrak{D}$, thus $\Gamma_{A \wedge B} \sqsubseteq \Gamma_A$ holds in $\mathfrak{N}$, and by C2 $\Gamma_A \in \mathcal{O}$. □

Lemma 11.11 (Deontic detachment)**.** *If $\Gamma_A \in \mathcal{O}$ and $\Gamma_{A \supset B} \in \mathcal{O}$, then $\Gamma_B \in \mathcal{O}$.*

Proof. Assume that $\Gamma_A \in \mathcal{O}$ and $\Gamma_{A \supset B} \in \mathcal{O}$. Then by lemma 11.3 $\Gamma_{A \wedge (A \supset B)} \in \mathcal{O}$. Since by PL $(A \wedge (A \supset B)) \supset B$ is true in $\mathfrak{D}$, we have $\Gamma_{A \wedge (A \supset B)} \sqsubseteq \Gamma_B$ in $\mathfrak{N}$, and by C2 $\Gamma_B \in \mathcal{O}$. Equivalently, assume that $\Gamma_A \in \mathcal{O}$ and $\Gamma_{A \supset B} \in \mathcal{O}$. By C3 $\Gamma_A \sqsubseteq \Gamma_B$ holds in $\mathfrak{N}$ and by C2 $\Gamma_B \in \mathcal{O}$. □

Lemma 11.12 (De Morgan)**.** $\Gamma_{A\wedge B} \in \mathcal{O}$ *if and only if* $\Gamma_{\neg(\neg A\vee\neg B)} \in \mathcal{O}$.

Proof. ($\Rightarrow$) Assume $\Gamma_{A\wedge B} \in \mathcal{O}$. Since by PL we have $(A \wedge B) \supset \neg(\neg A \vee \neg B)$ true in $\mathfrak{D}$, we have $\Gamma_{A\wedge B} \sqsubseteq \Gamma_{\neg(\neg A\vee\neg B)}$ in $\mathfrak{N}$, thus $\Gamma_{\neg(\neg A\vee\neg B)} \in \mathcal{O}$ by C2.

($\Leftarrow$) Assume $\Gamma_{\neg(\neg A\vee\neg B)} \in \mathcal{O}$. Since by PL we have $\neg(\neg A \vee \neg B) \supset (A \wedge B)$ true in $\mathfrak{D}$, we have $\Gamma_{\neg(\neg A\vee\neg B)} \sqsubseteq \Gamma_{A\wedge B}$ in $\mathfrak{N}$, thus $\Gamma_{A\wedge B} \in \mathcal{O}$ by C2. □

Lemma 11.13. *The next lemmas are consequences of (De Morgan) or lemma 11.8.*

$$\textit{If } \Gamma_A \in \mathcal{O}, \textit{ then } \Gamma_{A\vee B} \in \mathcal{O} \tag{11.6}$$

$$\textit{If } \Gamma_A \in \mathcal{O}, \textit{ then } \Gamma_{\neg(\neg A\wedge\neg B)} \in \mathcal{O} \tag{11.7}$$

$$\textit{If } \Gamma_{\neg A} \in \mathcal{O}, \textit{ then } \Gamma_{\neg A\vee\neg B} \in \mathcal{O} \tag{11.8}$$

$$\textit{If } \Gamma_{\neg A} \in \mathcal{O}, \textit{ then } \Gamma_{\neg(A\wedge B)} \in \mathcal{O} \tag{11.9}$$

$$\textit{If } \Gamma_{\neg(A\vee B)} \in \mathcal{O}, \textit{ then } \Gamma_{\neg A} \in \mathcal{O} \tag{11.10}$$

$$\textit{If } \Gamma_{\neg(\neg A\vee\neg B)} \in \mathcal{O}, \textit{ then } \Gamma_{A} \in \mathcal{O} \tag{11.11}$$

Soundness of legal inference

Formal logic is relevant to legal discourse and critical thinking insofar as it enables one to have a well-defined concept of validity, and hereby to have a method to prove that inferences are valid or not. However, each formal framework that tries to model the validity of our natural language possesses its limits, the natural language being vast and rich and, as we will see below, ours possesses its limitations too. Hence, it might be useful to present some informal considerations one must take into account when analyzing normative inferences. A *sound* inference is valid and has true (or acceptable) premises (cf. Peterson 2013b). As such, not all valid inferences are sound. Without pretending to be exhaustive, we provide the reader with some philosophical insights regarding the acceptability of the premises within a legal argument (see also Peterson 2013a).

A sound inference must be valid. Note, though, that our framework possesses its limits (e.g., it cannot model deontic conditionals), and therefore it is not because an argument is invalid within OL that it must be rejected. When the logical form of the argument falls within the scope of OL, the second step after verifying its validity is to determine whether or not its premises are acceptable. Since the truth of a normative proposition depends upon a norm, which is established by some authority, the first thing to do is to contextualize the normative propositions to a legal authority: is it true according to the Constitution? the Civil Code? the Criminal Code? the Canadian Charter of Rights and Freedom? It is noteworthy that some inferences can be fallacious when the meaning of 'ought' is not the same throughout the argument.

In addition to the normative propositions, the descriptive propositions within the argument should also be contextualized to some authority. For instance, the meaning of 'Paul respects his neighbour' will vary depending whether the legal authority is the Civil Code or the Criminal Code or the Canadian Charter of Rights and Freedoms. Then, one must determine what legally counts as a lack of respect from the Civil Code's point of view. In other words, one must answer the question 'what action counts as A from the authority's standpoint?'.[26]

The context of the argument must also be considered to establish the truth of the normative premises. For example, if 'ought' is understood legally, the proposition 'Paul ought to tell the truth' is not true in general but is true if Paul is providing testimony in a court of justice. Furthermore, the descriptive propositions used together with (R1) must be analyzed in terms of necessity and sufficiency. One must determine whether or not the conditional represents a semantical entailment between actions and see if the consequent is necessarily entailed by the action described in the antecedent.

Also, since our framework is not able to deal properly with contrary-to-duty reasoning and deontic conditionals, it might happen that conflicting obligations arise within a given situation. For example, consider the following argument (and assume a context in which premises 2 and 3 hold).

1. Paul ought to rescue Peter from drowning.
2. If Paul rescues Peter, then Paul calls 911.
3. If Paul calls 911, then Paul breaks into Sam's house.

$\therefore$ Paul ought to break into Sam's house.

Assuming that breaking and entering is legally forbidden, this yields a conflict of obligations when we add the following implicit premise.

4. Paul ought to not break into Sam's house.

Alchourrón and Makinson (1981) suggested that conflicts of obligations can be resolved by introducing a relation of hierarchy (or priority) between obligations. This idea was used by many logicians in non-monotonic deontic logic. Among others, van der Torre and Tan (1997) provided an insightful analysis of how conflicts of obligations can emerge in various situations. According to their analysis, the aforementioned example is a case of *weak overridden defeasibility*, where the contradiction appears because there is a conflict between the *prima facie* obligation to not break into Sam's house and the obligation to rescue Peter under the circumstances that he is drowning. Even though Paul has in general the obligation to not break into Sam's house, there might be specific situations in which it can be excused (but not allowed nor justified), namely if it is necessary to save Peter's life.

The assumption that premise 3 holds implies that Paul could argue that he should not be punished for breaking and entering since it was necessary to break into Sam's house to save Peter's life. Indeed, under specific circumstances, the defense of necessity

[26] For an analysis of 'count as', see Jones and Sergot (1996) and Boella and van der Torre (2006c).

can be invoked to excuse (but not justify) the violation of the law.[27] In the aforementioned example, Paul's actions are not morally voluntary given that he has no realistic choice: all things considered, it is reasonable to expect that Paul will break into Sam's house to call 911 if it is indeed the only possible course of action he has to save Peter's life. In such a context, the obligation to not break into Sam's house is *overshadowed* (as opposed to *cancelled*) by the obligation to break into Sam's house so that Paul can save Peter's life. Hence, in this context, Paul can be excused to break into Sam's house. Even though the obligation to not break into Sam's house is still in force (i.e., it does not simply disappear when a conflict arises), it is not *actual* in the sense that it should not be guiding Paul's actions in that context. Having this information at hand, one can thus reject premise 4 and the argument can be formalized within our framework while avoiding the conflict.

At last, the premises of a normative inference can be analyzed in the light of the ought-implies-can principle. Instead of the aforementioned example, consider the following argument.

1. Paul ought to rescue Peter from drowning.
2. If Paul rescues Peter from drowning, then Paul jumps in the water to retrieve Peter.

∴ Paul ought to jump in the water to retrieve Peter.

If for some reasons Paul cannot swim (e.g., if Paul has two broken arms), then premise 2 is a violation of the ought-implies-can principle. Thus, it must be rejected since in this case 'rescuing Peter from drowning' does not entail that Paul jumps in the water to retrieve him.

The paradoxes of deontic logic

Having presented some insights regarding the analysis of the soundness of legal inferences, let us now examine how our framework deals with the notorious *paradoxes* of deontic logic. Following Åqvist's (2002) presentation, a formula A is a paradox for a deontic logic Δ either if it is derivable in Δ but its translation does not seem derivable within the natural normative language, or it is *not* derivable in Δ but its translation seems derivable within the natural normative language. In other words, a formula is a paradox either when it can be derived but it should not, or when it cannot be derived but it should.

Ross's paradox

Ross's Ross (1941) disjunctive paradox appeared in reaction to Jørgensen's (1937) dilemma and aimed to show that a logic for imperatives does not behave similarly to propositional logic. For instance, the satisfaction of the imperative *Mail the letter!* does not imply that *Mail the letter or burn it!* is also satisfied.

[27] This has been accepted by the Supreme Court. See *Perka c. La Reine, [1984] 2 R.C.S. 232*.

Theorem 11.3. $\vdash_{OL} O_ip \supset O_i(p \vee q)$

Proof. Assume $\models_{\mathcal{N}} O_ip$ but $\models_{\mathcal{N}} \neg O_i(p \vee q)$. Thus, s^α satisfies O_ip but does not satisfy $O_i(p \vee q)$, which by definition is $O_i(\neg p \supset q)$. Thus, s^α satisfies p and $s_p^\alpha \in O$ but s^α does not satisfy $\neg p \supset q$, and so s^α does not satisfy $\Gamma_m = \neg p \supset q$. By PL, we have $\Gamma_p \sqsubseteq \Gamma_m$, and so by C2 s^α satisfies Γ_m.

$$\begin{array}{lll}
1 & \quad O_ip & \text{H} \\
2 & \quad p \supset (p \vee q) & PL \\
3 & \quad O_i(p \vee q) & \text{(R1) 1,2} \\
4 & \vdash_{OL} O_ip \supset O_i(p \vee q) & \supset \textit{in } \text{1-3}
\end{array}$$

□

Prior's paradox

Prior's (1954) paradox of derived obligations was an objection against von Wright's (1951) initial approach to deontic logic. It aimed to show that von Wright's notion of *commitment*, represented by $O(p \supset q)$, fails in contexts of derived obligations: it is not because p is forbidden that doing p *commits* one to do q.

Theorem 11.4. $\vdash_{OL} F_ip \supset O_i(p \supset q)$

Proof. Assume $\models_{\mathcal{N}} F_ip$ but $\models_{\mathcal{N}} \neg O_i(p \supset q)$. By definition it means that s^α satisfies $O_i\neg p$, thus s^α satisfies $\neg p$, meaning that $\overline{s}^\alpha$ satisfies p, and s^α does not satisfy $O_i(p \supset q)$, thus does not satisfy $p \supset q$. Therefore, s^α does not satisfy $\Gamma_m = p \supset q$ and $s_m^\alpha \notin O$. However, we have $\overline{s}_p^\alpha \sqsubseteq \Gamma_m$ since $\models_{PL} \neg p \supset (p \supset q)$, and so by C2 $s_m^\alpha \in O$.

$$\begin{array}{lll}
1 & \quad F_ip & \text{H} \\
2 & \quad O_i\neg p & \text{def F 1} \\
3 & \quad \neg p \supset (p \supset q) & PL \\
4 & \quad O_i(p \supset q) & \text{(R1) 2,3} \\
5 & \vdash_{OL} F_ip \supset O_i(p \supset q) & \supset \textit{in } \text{1-4}
\end{array}$$

□

Theorem 11.5. $\vdash_{OL} O_ip \supset O_i(q \supset p)$

Proof. Assume $\models_{\mathcal{N}} O_ip$ but $\models_{\mathcal{N}} \neg O_i(q \supset p)$. Thus, s^α satisfies O_ip but does not satisfy $O_i(q \supset p)$, meaning that s^α satisfies p, and so $s_p^\alpha \in O$, but s^α does not satisfy $q \supset p$.

Therefore, s^α does not satisfy $\Gamma_m = (q \supset p)$ and $s^\alpha_m \notin O$. However, $s^\alpha_p \sqsubseteq s^\alpha_m$ since $\models_{PL} p \supset (q \supset p)$, and thus by C2 $s^\alpha_m \in O$.

1	$O_i p$	H
2	$p \supset (q \supset p)$	PL
3	$O_i(q \supset p)$	(R1) 1,2
4	$\vdash_{OL} O_i p \supset O_i(q \supset p)$	$\supset$ *in* 1-3

□

Although it might surprise the reader, we think that Ross's and Prior's paradoxes are both desirable consequences of OL. Both paradoxes can be compared with theorem 11.6 since there is an equivalence between $O_i(p \vee q)$ and $O_i\neg(\neg p \wedge \neg q)$, and between $O_i(p \supset q)$ and $O_i\neg(p \wedge \neg q)$.

Theorem 11.6. $\vdash_{OL} F_i p \supset F_i(p \wedge q)$

Proof. Assume $\models_{\mathcal{N}} F_i p$ but $\models_{\mathcal{N}} \neg F_i(p \wedge q)$. By definition it means that s^α satisfies $O_i\neg p$ but does not satisfy $O_i\neg(p \wedge q)$. Thus, s^α satisfies $\neg p$, meaning that $\overline{s}^\alpha$ satisfies p and so $\overline{s}^\alpha_p \in O$, but s^α does not satisfy $\neg(p \wedge q)$, which by definition is $p \supset \neg q$. Therefore, s^α does not satisfy $\Gamma_m = p \supset \neg q$, and so $s^\alpha_m \notin O$. However, $\overline{s}^\alpha_p \sqsubseteq \Gamma_m$ since $\models_{PL} \neg p \supset (p \supset \neg q)$, and so by C2 $s^\alpha_m \in O$.

1	$F_i p$	H
2	$O_i\neg p$	def F 1
3	$\neg p \supset \neg(p \wedge q)$	PL
4	$O_i\neg(p \wedge q)$	(R1) 2,3
5	$F_i(p \wedge q)$	def F 4
6	$\vdash_{OL} F_i p \supset F_i(p \wedge q)$	$\supset$ *in* 1-5

□

Recall that OL aims to model how the property *obligatory* is transmitted from an action to another. Thus, while Ross's paradox says that if the action described by p is obligatory, then any conjunction of actions which includes the action described by $\neg p$ is forbidden, Prior's paradox says that if the action described by p is forbidden, then any conjunction of actions which includes the action described by p is also forbidden. Hence the comparison with theorem 11.6: if the action described by p is forbidden, then so is the action described by $p \wedge q$ for any q. When considered from a legal point of view, these theorems are desirable. If an action is legally forbidden, then any conjunction of action which includes the forbidden action is also forbidden, notwithstanding the nature of the other action. If it is forbidden to steal a car, then it is also forbidden to steal a car while eating ice cream.

Chisholm's paradox

Chisholm's (1963) contrary-to-duty paradox is the most damaging for monadic deontic logic. It shows that monadic frameworks are not able to properly represent conditional obligations. Although Prior's paradox motivated von Wright (1956; 1967) to introduce dyadic deontic logic (see Åqvist (2002) for an overview of dyadic deontic logic and Tomberlin (1981) for a discussion of conditional obligations), the need for a dydadic deontic logic arises when one considers Chisholm's puzzle. While the following sentences are perfectly consistent within the natural normative language, the conjunction of their translation is inconsistent in standard monadic deontic logic, i.e., the modal logic KD (cf. Åqvist 2002, p.155).[28]

1. Paul ought to not steal.
2. It ought to be that if Paul does not steal, then he does not give back what was stolen.
3. If Paul steals, then he ought to give back what was stolen.
4. Paul steals.

These are translated in KD by:

$$O\neg p \tag{11.12}$$
$$O(\neg p \supset \neg q) \tag{11.13}$$
$$p \supset Oq \tag{11.14}$$
$$p \tag{11.15}$$

Propositions 11.13 and 11.14 must be translated differently since they are independent within the natural normative language (cf. Åqvist 2002). Translating 11.13 by

$$\neg p \supset O\neg q \tag{11.16}$$

would make 11.16 a consequence of 11.15, as translating 11.14 by

$$O(p \supset q) \tag{11.17}$$

would make 11.17 a consequence of 11.12.

We agree with the literature that material conditionals cannot adequately model deontic conditionals. As such, we do not think that $O(p \supset q)$ is the appropriate formulation to model an obligation Oq conditional to a context p. Since mixed formulas are not

[28] See Åqvist (1967), Decew (1981) or Tomberlin (1981, 1983) for a discussion of the paradox.

available within our framework, it follows that Chisholm's paradox cannot be formulated appropriately. The available translations are:

$$O_i \neg p \tag{11.18}$$
$$O_i(\neg p \supset \neg q) \tag{11.19}$$
$$O_i(p \supset q) \tag{11.20}$$
$$p \tag{11.21}$$

$$O_i \neg p \tag{11.22}$$
$$O_i \neg p \supset O_i \neg q \tag{11.23}$$
$$O_i(p \supset q) \tag{11.24}$$
$$p \tag{11.25}$$

$$O_i \neg p \tag{11.26}$$
$$O_i(\neg p \supset \neg q) \tag{11.27}$$
$$O_i p \supset O_i q \tag{11.28}$$
$$p \tag{11.29}$$

$$O_i \neg p \tag{11.30}$$
$$O_i \neg p \supset O_i \neg q \tag{11.31}$$
$$O_i p \supset O_i q \tag{11.32}$$
$$p \tag{11.33}$$

It is noteworthy that each of these translations is consistent within our language. All imply $O_i \neg q$ but none implies $O_i q$ since $\Gamma_p \notin \mathcal{O}$. However, they fail to model Chisholm's paradox insofar as the propositions within these translations are not independent from each other. Indeed, 11.20 (=11.24) is a consequence of 11.18 (=11.22) in virtue of theorem 11.4, and 11.28 (=11.32) is a consequence of 11.26 (=11.30) and theorem 11.7.

Theorem 11.7. $\vdash_{OL} O_i\neg p \supset (O_i p \supset O_i q)$

Proof.

1	$O_i\neg p$	H
2	$\neg(O_i\neg p \wedge O_i\neg\neg p)$	(A1)
3	$O_i p\neg \supset \neg O_i\neg\neg p$	def $\wedge$ 2
4	$\neg O_i\neg\neg p$	$\supset$ *out* 1,3
5	$\neg O_i\neg\neg p \supset (O_i\neg\neg p \supset O_i q)$	PL
6	$O_i\neg\neg p \supset O_i q$	$\supset$ *out* 4,5
7	$\quad O_i p$	H
8	$\quad p \supset \neg\neg p$	PL
9	$\quad O_i\neg\neg p$	(R1) 5,6
10	$\quad O_i\neg\neg p \supset O_i q$	(Reit) 6
11	$\quad O_i q$	$\supset$ *out* 4,5
12	$O_i p \supset O_i q$	$\supset$ *in* 7-11

□

Chisholm's paradox is thus a limitation of our framework. However, the aim of this paper was not to model deontic conditionals but was rather to define a proper consequence relation for inferences within which there are no mixed formulas. The aim was to model the transmission of the property *obligatory* between *normative* propositions. The discussion of contrary-to-duty reasoning and deontic conditionals goes far beyond this paper and is an avenue for future research (see Peterson 2014d).

Detachment and Augmentation

Chisholm's paradox is related to the problems of *detachment* and *augmentation* of deontic conditionals (cf. Jones 1991).[29] The problem of (factual) detachment can be summarized as follows: although the detachment of a conditional obligation from its context seems suitable, as in Chisholm's paradox where we want to derive that Paul ought to give back what was stolen from the contrary-to-duty $p \supset Oq$ and the fact that Paul stole, there are situations where some extra conditions can make detachment undesirable. For instance, even though Paul has the conditional obligation to drive his wife to the hospital if she is near delivering, he does not have this obligation under the conditions that she is near delivering *and* (for some reason) he is drunk. As such, only a form of *restricted* detachment seems desirable, namely when we can insure that no other information will thwart

[29] There is a distinction between *deontic* detachment, which is expressed by lemma 11.11, and *factual* detachment, which is the detachment problem we are discussing within this section. See Loewer and Belzer (1983) for the distinction.

detachment.[30]

The problem of (unrestricted) detachment is related to the problem of augmentation. In a nutshell, the problem of augmentation comes from the fact that one cannot strengthen the antecedent of a deontic conditional. Hence,

$$\frac{p \supset Oq}{(p \wedge r) \supset Oq}$$

is not a valid inference pattern for deontic conditionals since the other conditions r might block the detachment of Oq. If one wants to model Chisholm's paradox and contrary-to-duty reasoning, then one will need to address these problems. This, which is also why we only concentrated upon the consequence relation of normative inferences in which there are no mixed formulas, will be the subject of another paper.[31]

Forrester's paradox

Forrester's (1984) paradox aims to show that (R1) is not an acceptable rule for normative reasoning. The paradox is similar to Prior's (1958) initial good Samaritan paradox and Nozick's and Routley's (1962) robbery paradox.[32] Assume a legal system in which a murder must be gentle. The original formulation of the paradox uses a conditional obligation (i.e., if Jones murders Smith, then he ought to do it gently)[33], but since in this respect it does not tell us more than what was already pointed out by Chisholm, we will only present the argument against (R1). The paradox follows from:

1. Jones ought to murder Smith gently.
2. If Jones murders Smith gently, then Jones murders Smith.
3. Jones ought to murder Smith.

According to Forrester, (R1) must be rejected since it allows one to conclude that Jones ought to murder Smith in a context where gentle murders are preferable to violent ones. The validity of this inference in our framework follows from lemma 11.14.

Lemma 11.14. *If* $\models_{\mathcal{N}} O_i p$ *and* $\models_{\mathcal{M}} p \supset q$*, then* $\models_{\mathcal{N}} O_i q$*.*

This represents the principle of deontic consequences. Notice that both $\models_{\mathcal{N}} O_i p$ *and* $\models_{\mathcal{M}} p \supset q$ *say that the propositions are true (respectively in a normative model* $\mathcal{N}$ *and a*

[30] On this point, see Decew (1981), Loewer and Belzer (1983), Vorobej (1986), Jones (1991), Alchourrón (1996) and Bonevac (1998).

[31] See Peterson (2014d).

[32] The good Samaritan paradox was reformulated by Åqvist (1967).

[33] The original paradox also insists on the contradiction between the assumption that *Jones ought to not murder Smith* and the conclusion from (R1) that *Jones ought to murder Smith.*

descriptive model $\mathcal{M}$), which does not mean that the propositions are valid, that is, true in every models.

Proof. Assume $\models_{\mathcal{N}} O_i p$ and $\models_{\mathcal{M}} p \supset q$ but $\models_{\mathcal{N}} \neg O_i q$. Therefore, s^α satisfies $O_i p$ but does not satisfy $O_i q$, meaning that it satisfies p but not q, thus $s^\alpha_p \in O$ but $s^\alpha_q \notin O$. However, by hypothesis $\models_{\mathcal{M}} p \supset q$, so $\Gamma_p \sqsubseteq \Gamma_q$, and therefore s^α satisfies q by C2. □

However, despite its validity, this argument is not sound. As it was argued by Jacquette (1986, p.762), the first premise is to be rejected. Of course, if one accepts a legal system where Jones ought to murder Smith gently, then one accepts a system where Jones ought to murder Smith! But the obligation of 'murdering gently' must be rejected in favour of the obligation of 'not murdering violently', and this does not lead to paradoxical results.[34]

Lemmon's paradox

Lemmon's (1962) paradox insists on the possibility of moral dilemmas and conflicting obligations. Although it is possible to face conflicting obligations in real life situations, a deontic logic which accepts (A1) as an axiom for normative consistency cannot model such situations. Moreover, when such a conflict happens, a deontic logic which satisfies (A1) will dictate that anything is obligatory insofar as any extension of classical propositional logic validates $\vdash \perp \supset A$ for any A (cf. Horty 1997).[35] Thus, from Paul's obligation to pick up his kids at the kindergarten by 5 p.m. and from his obligation to help the victim of a car accident that just happened, Paul can conclude that he ought to commit adultery, assuming that he picks up his kids by 5 p.m. if and only if he does not help the victim.

This, however, is not a paradox for OL. Indeed, axiom (A1) says that an authority i cannot state that both A and $\neg A$ are obligatory at the same time. This is acceptable insofar as norms are meant to guide one's actions, and it would be impossible to act within the frame dictated by inconsistent norms. That being said, the obligation to pick up his kids in the aforementioned example is not of the same type as his obligation to help the victim of the car accident. Hence, the conflict can be translated by $O_i p \wedge O_j \neg p$, which is not inconsistent in OL. In the eventuality that the conflict arises from contrary-do-duty reasoning or conditional obligations, it can be solved through the relation of hierarchy as we saw previously.

Aggregation

Aggregation of obligations is sometimes contested insofar as it can lead to conflicting obligations (see for instance Schotch and Jennings (1981) and Horty (1997)). These conflicts, however, often arise from different types of obligations. Since OL concerns only one type

[34] For an analysis of Forrester's paradox, see Peterson and Marquis (2012).
[35] See also theorem 11.7 above.

of obligation, it is acceptable to accept that if an authority makes A and B obligatory, then it makes both obligatory together. This is represented by theorem 11.8.[36]

Theorem 11.8. $\vdash_{OL} O_i(p \wedge q) \equiv (O_i p \wedge O_i q)$

Proof. It follows from theorems 11.9 and 11.10. □

Theorem 11.9. $\vdash_{OL} O_i(p \wedge q) \supset (O_i p \wedge O_i q)$

Proof. Assume $\models_{\mathcal{N}} O_i(p \wedge q)$ but $\models_{\mathcal{N}} \neg(O_i p \wedge O_i q)$. Thus, s^α satisfies $O_i(p \wedge q)$, which by definition means that s^α satisfies $\neg(p \supset \neg q)$, and so $\overline{s}^\alpha$ satisfies $p \supset \neg q$. Thus, s^α satisfies p and $\overline{s}^\alpha$ satisfies $\neg q$, meaning that s^α satisfies q. However, s^α does not satisfy $(O_i p \wedge O_i q)$, meaning that either (a) s^α does not satisfy $O_i p$ or (b) it does not satisfy $O_i q$.

(a) If s^α does not satisfy $O_i p$, it implies that it does not satisfy p, which is inconsistent with our hypothesis

(b) If s^α does not satisfy $O_i q$, it implies that it does not satisfy q which is also inconsistent with our hypothesis

1	$O_i(p \wedge q)$	H
2	$(p \wedge q) \supset p$	PL
3	$O_i p$	(R1) 1,2
4	$(p \wedge q) \supset q$	PL
5	$O_i q$	(R1) 1,4
6	$O_i p \wedge O_i q$	$\wedge in$ 3,5
7	$\vdash_{OL} O_i(p \wedge q) \supset (O_i p \wedge O_i q)$	$\supset in$ 1-6

□

Theorem 11.10. $\vdash_{OL} (O_i p \wedge O_i q) \supset O_i(p \wedge q)$

Proof. Assume $\models_{\mathcal{N}} O_i p \wedge O_i q$ but $\models_{\mathcal{N}} \neg O_i(p \wedge q)$. Thus, s^α satisfies $O_i p$ and satisfies $O_i q$, meaning that it satisfies p and it satisfies q, and so $s^\alpha_p, s^\alpha_q \in O$, but it does not satisfy $O_i(p \wedge q)$. By definition, it implies that s^α does not satisfy $\neg(p \supset \neg q)$, thus $\overline{s}^\alpha$ does not satisfy $p \supset \neg q$. Thus, either s^α does not satisfy p, which is inconsistent with our

[36] This principle can be contested if we consider contrary-to-duty reasoning and conditional obligations. As the reader will see, we will change our mind in Peterson (2014b) and reject the aggregation principle on the grounds that $\wedge$ does not adequately represent action conjunction.

hypothesis, or $\overline{s}^\alpha$ does not satisfy $\neg q$, meaning that s^α does not satisfy q, which is also inconsistent with our hypothesis.

$$
\begin{array}{lll}
1 & \quad O_i p \wedge O_i q & \text{H} \\
2 & \quad (p \wedge q) \supset (p \wedge q) & PL \\
3 & \quad O_i(p \wedge q) & \text{(R1) 1,2} \\
4 & \vdash_{OL} (O_i p \wedge O_i q) \supset O_i(p \wedge q) & \supset in \text{ 1-3}
\end{array}
$$

□

Contingency

The principle of normative contingency was first advocated by von Wright (1951). It has since then been defended by Chellas (1974) and Jones and Pörn (1985). This principle implies that tautologies are neither unconditional nor derived obligations. As it is shown by the following lemmas, our approach respects contingency.

Lemma 11.15. $\nvdash_{OL} O_i p \supset O_i(\neg p \supset \neg p)$

Proof. We show that there is a model $\mathcal{N}$ such that $\models_\mathcal{N} \neg(O_i p \supset O_i(\neg p \supset \neg p))$. Suppose $\models_\mathcal{N} O_i p$ and $\models_\mathcal{N} \neg O_i(\neg p \supset \neg p)$. So s^α satisfies $O_i p$ and thus $s^\alpha_p \in O$ but s^α does not satisfy $O_i(\neg p \supset \neg p)$, meaning that s^α does not satisfy $\neg p \supset \neg p$ and either s^α does not satisfy $\neg p$, thus $\overline{s}^\alpha$ does not satisfy p, or $\Gamma_{\neg p} = \Gamma_{\neg p}$. Both cases are consistent with C1–C3 and the definition of satisfaction. □

Lemma 11.16. $\nvdash_{OL} O_i p \supset O_i(p \vee \neg p)$

Proof. We show that there is a model $\mathcal{N}$ such that $\models_\mathcal{N} \neg(O_i p \supset O_i(p \vee \neg p))$. Assume $\models_\mathcal{N} O_i p$ and $\models_\mathcal{N} \neg O_i(p \vee \neg p)$. Thus, s^α satisfies $O_i p$ but does not satisfy $O_i(p \vee \neg p)$. By definition, it means that s^α satisfies p but does not satisfy $\neg p \supset \neg p$ (theorem 11.15). □

Lemma 11.17. $\nvdash_{OL} O_i p \supset O_i(p \supset p)$

Proof. We show that there is a model $\mathcal{N}$ such that $\models_\mathcal{N} \neg(O_i p \supset O_i(p \supset p))$. Assume $\models_\mathcal{N} O_i p$ and $\models_\mathcal{N} \neg O_i(p \supset p)$. Thus, s^α satisfies $O_i p$ but does not satisfy $O_i(p \supset p)$, meaning that s^α satisfies p and $s^\alpha_p \in O$, but s^α does not satisfy $p \supset p$. Therefore, s^α does not satisfy $\Gamma_m = p \supset p$, which is consistent with C2 since $\models_{PL} p \supset p$, and either s^α does not satisfy p, which is inconsistent with our hypothesis, or $\Gamma_p = \Gamma_p$, which is true. Therefore, it is possible that s^α satisfies $O_i p$ but does not satisfy $O_i(p \supset p)$. □

Lemma 11.18. $\nvdash_{OL} O_i(p \vee \neg p)$

Proof. It follows from theorem 11.16. $\square$

Lemma 11.19. $\nvdash_{OL} O_i(\neg p \supset \neg p)$

Proof. It follows from theorem 11.15. $\square$

Also, it must be possible to represent that some actions are neither obligatory nor forbidden, which follows from the following lemma.

Lemma 11.20. $\nvdash_{OL} O_i p \vee O_i \neg p$

Proof. Assume $\models_{\mathcal{N}} \neg O_i p \wedge \neg O_i \neg p$. Thus, s^α does not satisfy $O_i p$ and does not satisfy $O_i \neg p$, meaning that s^α does not satisfy p and does not satisfy $\neg p$. Therefore, $\overline{s}^\alpha$ does not satisfy p and $s^\alpha_p \notin O$ and $\overline{s}^\alpha_p \notin O$, which is not inconsistent. $\square$

This would have been violated if we had assumed an 'if and only if' instead of a simple 'if then' in condition C1.

Disjunctive permission

A consequence of OL is that if an action is permitted, then a disjunctive action which includes said action is also permitted. Hence, if it is permitted to walk in the park, then it is permitted to either walk in the park or rob a bank. This reasoning, which follows from theorem 11.11, is also derivable in standard deontic logic (it is actually derivable in a normal deontic logic, in the sense of Åqvist (2002, p.155), i.e., an extension of the modal logic K). But this paradox is no more puzzling than the fact that from p's truth one can infer the truth of $p \vee q$ in PL. Moreover, theorem 11.11 appears as a desirable consequence when one considers theorem 11.12.

Theorem 11.11. $\vdash_{OL} P_i p \supset P_i(p \vee q)$

Proof. It follows from theorem 11.12 by contraposition. $\square$

Theorem 11.12. $\vdash_{OL} F_i(p \vee q) \supset F_i p$

Proof. Assume $\models_{\mathcal{N}} F_i(p \vee q)$ but $\models_{\mathcal{N}} \neg F_i p$. By definition it means that s^α satisfies $O_i \neg(p \vee q)$, thus s^α satisfies $\neg(p \vee q)$, which by definition is $\neg(\neg p \supset q)$, and s^α does not satisfy $\neg p$, thus $\overline{s}^\alpha$ does not satisfy p and $\overline{s}^\alpha_p \notin O$. Since s^α satisfies $\neg(\neg p \supset q)$, $\overline{s}^\alpha$ satisfies

$\neg p \supset q$, thus s^{α} satisfies $\neg p$, and so $\overline{s}^{\alpha}$ satisfies p.

1	$F_i(p \vee q)$	H
2	$O_i\neg(p \vee q)$	def F 1
3	$\neg(p \vee q) \supset \neg p$	PL
4	$O_i\neg p$	(R1) 2,3
5	$F_i p$	def F 4
6	$\vdash_{OL} F_i(p \vee q) \supset F_i p$	$\supset in$ 1-5

□

Theorem 11.12 says that if a disjunction of actions is forbidden, then each member of the disjunction is forbidden too. This is a desirable consequence. For instance, if Paul tells his son Peter that he is forbidden to either take the car or the motorcycle, we do not expect Peter to answer 'Fine! I'll take the motorcycle then!'.

Conclusion

To sum up, we introduced a method to test the formal validity of legal inferences. Although the main contribution of this paper was to develop a formal method to test validity, we also provided the reader with a brief analysis of some important aspects of sound normative inferences. It is our view that if formal philosophy is to be of help to legal discourse, then it must first reflect upon the law's fundamental characteristics that should be taken into account. We provided the reader with a brief analysis of Canadian legal discourse and we exposed three fundamental characteristics which ought to be considered if one wants to represent the formal structure of legal arguments. These characteristics are the presupposed consistency of legal discourse, the fact that there is a hierarchy between norms and obligations to preserve this consistency and the fact that legal inferences are subjected to the principle of deontic consequences. The formal logic that was built according to these characteristics is restricted to normative inferences in which there is only one type of obligation, no iteration of deontic operator and no mixed formulas (i.e., no conditional obligations). This is both a strength and a weakness of our approach. It is a strength insofar as the paradoxes which use these properties cannot be formulated in our framework. But it is a weakness since we cannot formulate contrary-to-duty reasoning and conditional obligations, which are of central importance in legal reasoning.

The main contribution of this paper is that our method covers a portion of the intuitive validity of legal inferences that was not covered by other monadic frameworks in the literature. For instance, the principle of deontic consequences is not valid in standard deontic logic. Although the principle of deontic consequences can be formulated (and validated) in a multi-modal deontic logic which includes a minimal necessity operator $\Box$, our aim was to develop a system which can easily be applied to test the validity of some basic legal inferences and where the principles of contingency and deontic consequences are handled with very little formalism. Our intention was to introduce an alternative to the modal logic K, which is usually used as a building block for the construction of deontic logics. This framework also provides a method to graphically represent legal arguments and test their validity. It is our view that our framework is better suited than K to formalize the transmission of the property 'legally obligatory' between actions.

For future research, we intend to extend this method to arguments in which there are different types of obligations which can be iterated and are linked by a relation of hierarchy to prevent conflicts of obligations. We will also work on the representation of mixed formulas, conditional obligations and contrary-to-duty reasoning. It would be interesting to see if (and how) this system can deal with agency and incorporate (or be incorporated in) the frameworks of dynamic and *stit* logics. Another avenue will be to redefine the propositions within the scope of the deontic operator O_i. Instead of using descriptive formulas that refer to actions, the satisfaction of normative atoms could be redefined in terms of an action algebra.[37] However, this would require to study carefully action logics and determine the proper formal framework to represent the structure of actions.[38] Finally, another avenue for future research will be to see how this approach can be reformulated within the framework of categorical logic.

* * *

Cet article répond en partie au premier objectif de cette thèse, à savoir proposer un cadre d'analyse différent de la logique modale K et de la logique déontique standard pour la modélisation des inférences normatives inconditionnelles et non-conflictuelles. Le prochain chapitre adapte le présent système en vue d'une introduction à l'analyse des inférences normatives et s'adresse à un public qui ne maîtrise pas nécessairement la logique. Les idées aux fondements de OL seront reprises au chapitre 16 quant à la modélisation des inférences inconditionnelles.

[37] See Peterson (2014b).
[38] See Peterson (2014a).

Chapitre 12

Normative inferences and validity: A heuristic

Abstract

Although there are various formal methods to analyze the validity of arguments, undergraduate students are often only introduced to classical propositional logic and first-order logic as tools to analyze reasoning. Classical and predicate logics are tools that can be used to study declarative statements (i.e., statements which have the potential to be true or false). The notion of truth is often understood using a correspondence theory of truth, where a sentence φ is said to be true when it is an adequate description of the world. But what about normative propositions? Are they declarative? To answer this question is to enter into a widely controversial debate in metaethics. But one can safely assume that if they are, then it seems that they are not potentially true or false in the same way that descriptive propositions are. Thus, it is unclear whether classical logic and predicate logic are tools that can help in the study of normative reasoning, notwithstanding that many valid normative inferences are invalid in these frameworks. But the fact is that normative reasoning is at the core of philosophical reflections. So what are the tools one can use to analyze normative inferences? In this paper, we present an informal method based on the formal deontic logic OL, which provides an easy and efficient tool to analyze the validity of normative inferences.[1]

Keywords: Deontic logic, Normative consistency, Deontic consequence, Argumentation, Reasoning, Critical thinking

Introduction

One of the main purposes of argumentation theory is to provide rational criteria that enable one to judge the value of reasons. Although the aim of argumentation is to *convince* of a thesis, rational thinking, at least as it was conceived in Greek philosophy, is not only

[1] This chapter has been published in Peterson (2013a).

concerned about the aim (convincing) but also about the manner employed to persuade. One must not simply *convince* but must convince *rightfully*. But how, from a rational point of view, does one *rightfully* convince someone of a thesis? The obvious answer would be to say that one rightfully convinces someone of a thesis by means of rational arguments, but this, if not circular, would simply beg another question: what is a rational argument? In a wider sense, an *argument* can be thought of as a form of persuasion, where one provides reasons or data to justify a standpoint. In a narrower sense, an *argument* can be thought of as a sequence of sentences, where there is at least one premise and a conclusion. The latter is the definition we will adopt throughout this paper. Understood this way, an argument is similar to a reasoning or an inference. This distinction is similar to Walton's (1997) distinction between an argument, which uses reasoning to convince and correspond to our conception of an argument in a wider sense, and a reasoning, that is, a sequence of inferences, which corresponds to our conception of an argument in a narrower sense. An argument can be *complex* and contain intermediary conclusions. In this case, a complex argument can be thought of as an embodiment of simple arguments. But what distinguishes an argument that is *rational* from one that is not? Although questions regarding the definition of rationality are of philosophical significance, to try to answer them would lead us astray from our main objective. Instead of trying to provide a definition of rationality, our analysis will simply assume the criterion of consistency. Assuming consistency as a minimal criterion of rationality is not too demanding.[2] The least we can expect from an interlocutor is to be not contradictory. A rational argument is thus understood as an argument which at least respects the criterion of *consistency*.[3]

But what does it mean for an argument to respect the criterion of consistency? Obviously, the premises of an argument must be consistent with each other, otherwise *ex falso sequitur quodlibet*. But consistency can be more subtle than that. Consider the two following sentences:

(1) Paul ought to respect the law.

(2) It is not the case that Paul ought to respect the law.

Clearly, these two propositions are in contradiction with each other: the first says φ while the other says *not-*φ. What about (1) and (3)?

(3) Paul ought to not respect the law.

Are they in contradiction? The sentences (2) and (3) do not have the same meaning, and so (3) does not have the form *not-*φ. While in (2) the negation applies to the whole sentence, it is inside the scope of the 'ought to' in (3). However, although (1) and (3)

[2] Of course, advocates of paraconsistency might reject that assumption.

[3] It would be accurate to distinguish between the syntactical consistency of a formal language and the intuitive coherence of our natural language. While the former is well defined, the latter is not. We will use the term 'consistency' throughout this paper despite this distinction.

are not inconsistent from a propositional point of view, they seem inconsistent from a *normative* one. Indeed, both cannot be true at the same time (even though they could be false together).[4] This brings to light an interesting point: if we accept that consistency is a minimal criterion of rationality, then we must assume both propositional and normative consistency as foundational principles. We defined an argument as being a sequence of sentences where a conclusion is inferred from some premises. This relation of inference between the premises and the conclusion is a *consequence relation*. Formally, consistency relies upon the definition of the consequence relation, and thus an argument's rationality can be examined through it. But what are the properties of a proper consequence relation for normative inferences?

The aim of the present paper is to answer this question with the help of formal deontic logic, but without the formal framework. We wish to provide the reader with a tool to analyze the validity of normative inferences without requiring any background expertise in formal logic. First, we will explain why there is a need for a framework other than the classical ones, that is, classical propositional logic (**CL**) and first-order logic (**FOL**), with which most philosophers are familiar. That being done, we will expose how our approach should be understood from the point of view of critical thinking and informal logic. Then, we will present an informal method to analyze a normative inference's validity and provide some philosophical insights regarding its soundness.

Why another framework?

Logic is about abstract structures, and hence it is a useful and powerful tool for argumentation theory. To see this, consider the following definition.

Definition 1 (Validity)**.** An argument is *valid* if and only if it is impossible for its conclusion to be false while its premises are true.

In other words, an argument is valid if and only if the truth of the premises necessarily entails the truth of the conclusion. Now, consider the following well-known first-order reasoning.

P_1	Socrates is a man.
P_2	All men are mortal.
C	Socrates is mortal.

Intuitively, this reasoning seems valid (and formally it is). But how can we *prove* that it is valid? To prove it, one needs a well-defined method to prove that it is actually impossible for the premises to be true while the conclusion is false. The aforementioned

[4] In this respect, $O\varphi$ and $O\neg\varphi$ can be seen as contraries rather than contradictory propositions. That being said, assuming the axiom for normative consistency, the conjunction $O\varphi \wedge O\neg\varphi$ can still be seen as a contradiction given that it is logically impossible for $O\varphi \wedge O\neg\varphi$ to be true.

definition of validity relies upon a central concept, namely the concept of *possibility*. If one wants a well-defined notion of *validity*, then one will need a well-defined notion of *possibility*, otherwise the definition is vague, imprecise and relies only upon an intuitive understanding of this notion. This is precisely why logic is useful for critical thinking: while the intuitive conception of validity is not well defined, the formal one is. The formal conception of validity relies upon the notion of *logical* possibility, which in turn depends on the (well defined) consequence relation of the logic.

Informally, the notion of intuitive validity for our natural language can be represented as a set of *intuitively valid* inferences (NL), as in figure 12.1.[5] A logic can be thought of as a set of *formally valid* inferences (i.e., a set of theorems), which cover a portion of the intuitively valid inferences. Formal logic is about the representation of structures. This representation is an abstraction, and there are different levels of abstraction. Consider for example Aristotle and Plato. Clearly, these two individuals were different. However, if one shifts perspective and looks at them from another standpoint, then one can see many similarities. For instance, assuming that there is a proper definition of *human*, then this definition can be seen as a criterion of identity. In this respect, one could say that Aristotle and Plato are 'identical'. One can have an even more general point of view and consider them as *mammals*. As *humans*, Plato and Aristotle are the same, and from that perspective they are different from Bucephalus. However, as *mammals*, these three individuals are not different.

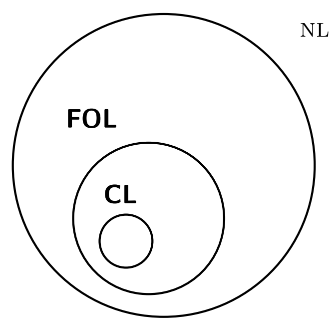


Figure 12.1: Validity and natural language

It is the same kind of abstraction that is done when logic is applied to the natural language. Consider the following sentences.

(1) Socrates is a man.

(2) Felix is a cat.

(3) John is the brother of Sam.

[5] **CL** is not actually included in **FOL** since the formulas of **CL** are not well-formed in **FOL** (and vice versa). There is however a common denominator since **FOL** behaves at the propositional level according to the rules of **CL**. On this point, see Peterson (2013b, chapter 7).

The *propositional* point of view of **CL** is the most general one. It identifies the logical connectives (i.e., negation $\neg$, conjunction $\wedge$, disjunction $\vee$, entailment $\supset$) in the natural language and then considers *atomic propositions*, from which one obtains *complex propositions* with the combination of atoms and logical connectives. From this standpoint, these three sentences are atomic and can be represented by propositional variables. But one can refine that perspective with **FOL** and see similarities between these atomic sentences. For instance, (1) and (2) are both of the form 'x is P', with x an object and P some property. From this perspective, both propositions are different from (3), which is of the form 'x is in relation R with y'. In short, depending on the level of abstraction your looking from, formal logic helps to represent the structure of a portion of our natural language.

When applied to natural language, logic is often viewed as the study of the transmission of truth value between propositions. In other words, logic is *truth functional*, meaning that the truth value of a complex proposition depends upon the truth value of its logical connectives and its atoms. Assuming that logic studies the transmission of truth values between propositions, it follows that atomic sentences should (at least in principle) have the potential to be true or false. This is why logic is often understood as the study of *declarative* statements. A statement is *declarative* if it is (in principle) subject to truth-value assignment. Returning to figure 12.1, **CL** allows one to determine which inferences of our natural language are valid propositionally, while **FOL** gives us a richer insight and determines which are valid from a first-order standpoint (i.e., the transmission of properties between objects). In both cases, the portion of the natural language which is studied contains only declarative statements.

Declarative statements are usually understood as sentences that have the potential to be true or false in regards to the *correspondence theory of truth*: a sentence is true if and only if it is an accurate description of the reality (e.g., 'snow is white' is true if and only if snow *is* white).[6] In short, a sentence has the potential to be true or false if it says something about how the world is. Thus, when truth is understood as a correspondence with the world, declarative statements are understood as *descriptive* propositions. What about proposition (4)?

(4) Paul should tell the truth.

Is it declarative? Clearly, (4) does not say how the world *is* but rather says how the world *should be*, and thus it is not descriptive. In his *Treatise of human nature*, David Hume (1993) introduced the semantical dichotomy between *facts* and *norms*, also known as the *is-ought thesis*.[7] There is a semantical dichotomy insofar as one cannot infer a normative proposition from a set of purely descriptive sentences, and vice versa. In other words, inferring an *is* from a *should* always results in invalid reasoning, and vice versa. According to this dichotomy, normative propositions (i.e., propositions that says how the world should be or how people should act) are semantically different from descriptive ones.

[6] See Tarski (1944) for a formal definition.
[7] See also Poincaré (1913) and Jørgensen (1937).

Therefore, if normative propositions have the potential to be true or false, then it is not in the same way that descriptive propositions do.

There are two options: either normative propositions are declarative and thus we must adopt a theory other than the (realistic) correspondence theory of truth or we keep this theory and reject normative propositions as declarative statements.[8] This is, in a nutshell, the underlying issue behind the debate between cognitivism and non-cognitivism in ethical philosophy (cf. van Roojen 2011). In the present paper, we do not pretend to offer a definitive solution to this debate. However, for practical reasons, we will nonetheless assume that normative propositions are declarative and thus reject the realistic interpretation of the correspondence theory of truth.[9] This choice is motivated by the fact that if one accepts that normative propositions are not declarative, then it seems that logic cannot be applied to normative inferences since logic concerns the transmission of truth values between propositions.[10] However, normative inferences are an integral part of our day to day life, if not from an ethical point of view, then at least from a legal one. Moreover, there are some normative inferences that intuitively seem valid while others do not. Hence, we assume that normative propositions are declarative so that we can use (deontic) logic to formalize normative validity.

But if normative statements are declarative, then normative inferences can be analyzed within the framework of **CL** or **FOL**. So why would we need another framework? Without entering into the formal details, **CL** and **FOL** are not powerful enough to represent the formal validity of normative inferences. Consider the following example.

P_1	It is forbidden to steal
C	Paul ought to not steal

Assuming a predicate F for interdiction and a binary relation O for obligation, this inference is clearly invalid in **CL** and **FOL** since it would be (respectively) translated by a contingent conditional of the form $\varphi \supset \psi$ and $Fx \supset \neg pOx$. Moreover, the translation of the conclusion by $\neg pOx$ would not be accurate since the negation in C is applied inside the 'ought to' while it is applied to the whole sentence in 'Paul does not ought to steal ($\neg pOx$)'.

The usual formal framework to analyze normative inferences is *deontic logic*, which is (roughly) the logic of obligations, permissions and interdictions.[11] Although there are various approaches in deontic logic[12], these are formulated in reaction to the basic framework

[8] See for example Ayer (1946) for the latter.

[9] It should be noted that 'correspondence theory' is not understood here as a class including every type and variation of the correspondence theory of truth. Rather, we refer to the (naive) realistic conception which says that a proposition φ is *true* in as much as it corresponds to what *is* or to (some form of normative) *reality*.

[10] This is Jorgensen's (1937) dilemma. But this problem is not unsolvable. Although truth is a useful tool in logic, it is dispensable (e.g., think of proof theory).

[11] See McNamara (2010) for an introduction.

[12] See Peterson (2011) for an overview.

of *standard deontic logic* (i.e., the modal logic KD).[13] The following argument is invalid in the standard system, and thereby the reader may see why we will use a framework other than KD in the next section.[14]

P_1	If Paul respects the law, then Paul does not steal
P_2	Paul must respect the law
C	Paul must not steal

As we will see, this argument is an instance of the *principle of deontic consequences*.

Informal logic and argumentation theory

Despite the semantical dichotomy, approaches in the literature do not put much emphasis on normative inferences. Although the reader can find conceptual clarifications regarding normative statements within the deontic logic literature[15], these considerations are not taken into account when logic is used to analyze inferences[16]. Even informal logic does not distinguish between normative and descriptive inferences.[17] Moreover, informal logic tends to treat normative statements as if they were descriptive (predicative) statements[18], which is a conceptual mistake according to the dichotomy.

This is not to say that formal logic is superior to informal logic. It is rather to say that formal logic can help informal logic in clarifying some conceptual issues regarding normative inferences. Most approaches in informal logic and critical thinking reject formal methods as tools to analyze arguments (Walton and Godden 2007, p.3), mainly because formal logic studies *forms* of arguments and "the form of the argument is not sufficient to enable one to arrive at an evaluation of the argument as weak or strong, reasonable or fallacious (Walton 1995, p.45)." But this conception of formal logic as a method that enables one to determine the strength of an argument is misleading: formal logic is a tool which enables one to determine if an argument is valid or not. When one evaluates the strength of an argument, one is studying its *soundness* rather than its validity.

Definition 2 (Soundness)**.** An argument is *sound* if and only if:

1. it is valid;

2. it has true premises.

[13] See Chellas (1980) or Garson (2006) for an introduction to modal logics.
[14] Note that some multi-modal approaches validate this reasoning. See Peterson and Marquis (2012).
[15] See for instance Horty (2001).
[16] See for example Gabbay et al. (2002).
[17] See for example Walton (1997), Walton and Gordon (2005), Walton et al. (2010) or Macagno and Walton (2010; 2012).
[18] See for example Walton (2010), Macagno and Walton (2009) and Gordon and Walton (2009). Pincione (2009) merely makes the distinction but uses Kelsen's notion of validity instead of truth in order to be able to use the usual frameworks.

Soundness cannot be determined by logic alone. Formal logic enables one to prove that some arguments in the natural language are valid. When this is done, one must evaluate the acceptability of the premises and this is not, unfortunately, a well defined process. Following Tomassi (1999), formal and informal logic are to be seen as *complementary* rather than incompatible.

Normative inferences: a heuristic

Prerequisite

The present section is based upon the formal framework OL, which is a (deontic) logic of obligation. Prior knowledge of this work is not necessary but the reader interested in the formal details and the philosophical justifications of the system can consult Peterson (2014e). A *normative inference* is understood as an argument within which the conclusion is a normative proposition. It often has at least one normative premise but can also have none, as we will see with the formulation of the principle of deontic consequences. Let $Prop = \{p, q, r, ...\}$ be a denumerable set of atomic descriptive propositions. Here, descriptions are not understood as in **CL**. Rather than being *any* kind of description, propositional atoms are understood as descriptions of *actions* (but not descriptions of *facts*). A *descriptive proposition* is either a descriptive atom or a complex formula composed of descriptive atoms and logical connectives. Let $Act = \{\Gamma_1, ..., \Gamma_n, ...\}$ be a set of actions (not considered as declarative statements). These actions do not need to be atomic. For our purpose, it suffices to understand that each action can be descripbed by a descriptive proposition. Thus, every action in *Act* can be described by a descriptive proposition (either atomic or complex). We will write Γ_A for the action described by the descriptive formula A. We assume that *Act* is pre-ordered by $\sqsubseteq$ in such a way that if $A \supset B$ is taken to be true in the descriptive part of a scenario, then $\Gamma_A \sqsubseteq \Gamma_B$ holds in its corresponding normative part. Informally, $\sqsubseteq$ represents the relation of entailment between actions. (We do not assume the converse because it would lead us to ideality.)

An atomic normative proposition is of the form O_iA ('it is obligatory that A'), where O_i is a deontic operator for obligation and A is a descriptive formula (either atomic or complex). The subscript $_i$ stands for a specific type of obligation (e.g., legal, moral, contractual, etc.). From O_i we can define the notion of (weak) permission $P_iA =_{def} \neg O_i \neg A$ and interdiction $F_iA =_{def} O_i \neg A$.[19] In short, A is permitted when $\neg A$ is not obligatory and A is forbidden when it is not permitted. A *normative proposition* is either a normative atom or a complex formula composed of normative atoms and logical connectives. For technical reasons and because the meaning of a deontic operator changes when it is iterated[20], we will concentrate only on normative propositions in which there is only one type of deontic operator and where there is no iteration of operators. For example, a sentence

[19] These definitions are sometimes contested, mainly because this does not represent the concept of strong permission (cf. Makinson and van der Torre 2003a).

[20] See Castañeda (1981) and Jones and Pörn (1985).

such as ‘it ought to be that stealing is not permitted’ is not studied by our method.

As we saw, normative validity is subtler than propositional validity. In addition to propositional validity, there seems to be a finer criterion at work in the transmission of the property ‘obligation’ between actions. Even normative consistency is finer than mere propositional consistency (cf. Royakkers 1998). Considering that norms aim at guiding one’s actions (Weinberger 2001) and that it would be impossible to act accordingly to an inconsistent set of norms, it follows that norms must (at least after interpretation) prescribe consistent actions. Normative consistency can be represented by the axiom schema (**D**) of standard deontic logic, which states that it is false that both A and $\neg A$ are obligatory (according to the same authority at the same time).[21]

$$\neg(O_i A \wedge O_i \neg A) \tag{\textbf{D}}$$

This also means that if A is obligatory, then A is permitted. This axiom is taken to be true in all interpretations, hence valid. (**D**) allows us to reduce normative consistency to propositional consistency. In addition to (**D**), the validity of normative inferences is ruled by the *principle of deontic consequences*. Intuitively, this principle means that:

> If an act A entails an act B, then (1) the obligatoriness of A entails the obligatoriness of B, and (2) the forbiddenness of B entails the forbiddenness of A (Castañeda 1968, p.13).

This principle allows one to infer a normative conditional from a descriptive one. At first glance, this may seem in contradiction with the semantical dichotomy. However, this does not mean that we infer a norm from a fact, but rather that we infer a normative entailment from a descriptive entailment. This is not to say that from ‘Paul steals’ one can infer ‘it is obligatory that Paul steals’: it is rather to say that from the descriptive entailment ‘if Paul takes Peter’s money without his consent, then Paul steals’ it is possible to infer the normative entailment ‘if it is forbidden that Paul steals, then it is forbidden that Paul takes Peter’s money without his consent’.

This principle can be reformulated: if A is obligatory and A implies B, then B is also obligatory. Formally, this can be represented by the rule (Δ),

$$\begin{array}{r|ll}
1 & \vdots & \\
2 & O_i A_1 \wedge \cdots \wedge O_i A_n & (n \geq 1) \\
3 & (A_1 \wedge \cdots \wedge A_n) \supset B & \\
4 & O_i B & \Delta\ 2{,}3
\end{array} \tag{Δ}$$

The rule can be applied either when the conditional is a hypothesis or when it is a logical consequence of propositional logic, provided it respects two conditions. First,

[21] Note that the usual formulation of (**D**) is rather $O_i A \supset \neg O_i \neg A$. The two formulas are propositionally equivalent.

following Chellas (1974, p.24), it is always possible to have a situation in which there is no norm, and thus it must be possible to represent a situation where there is no obligation at all. In short, there must be no 'absolute' or 'unconditional' obligations. This is why (Δ) is restricted by the condition that $n \geq 1$. Second, following Jones and Pörn (1985, p.279), it is always possible for someone to act against one's obligations, and thus a tautology is not an obligation unless there is a norm that makes it so. Hence, B must not be a tautology.[22]

Validity

From a semantical point of view, (**D**) and (Δ) are represented by the following conditions. Assume that $\mathcal{O}$ is a set of obligations, then $\mathcal{O}$ has to satisfy:

C1 if $\Gamma_A \in \mathcal{O}$, then $\Gamma_{\neg A} \notin \mathcal{O}$;

C2 if $\Gamma_A \in \mathcal{O}$, then $\Gamma_B \in \mathcal{O}$ for any $\Gamma_A \sqsubseteq \Gamma_B$ such that $\not\models_{\mathsf{CL}} B$;

C3 If $\Gamma_A \in \mathcal{O}$ and $\Gamma_A \sqsubseteq \Gamma_{B \supset C}$, then $\Gamma_B \sqsubseteq \Gamma_C$

An argument's validity can be tested through a counterexample: assume a scenario w in which the premises are true but the conclusion is false and see if it is inconsistent. If it is, then the argument is valid (and if it is not inconsistent (i.e., consistent), then it is invalid). A scenario w is divided in two parts, one descriptive ($\mathfrak{D}$) and the other normative ($\mathfrak{N}$).[23] The normative part contains a set of obligations $\mathcal{O}$. As an example, let us test (Δ)'s validity with $n = 1$ (figure 12.4).

Assume that w is a counter-example for (Δ). Then, place the assumptions in their respective descriptive or normative part of w. The propositional rules (figure 12.2) representing schematically truth conditions for complex propositions are quite straightforward (cf. Garson 2006, p.91). The semantical dichotomy implies that normative and descriptive

$$
\begin{array}{c} \neg(A \supset B) \\ | \\ A \\ \neg B \end{array}
\qquad
\begin{array}{c} A \supset B \\ \diagup \quad \diagdown \\ \neg A \quad\; B \end{array}
\qquad
\begin{array}{c} A \wedge B \\ | \\ A \\ B \end{array}
\qquad
\begin{array}{c} \neg(A \wedge B) \\ \diagup \quad \diagdown \\ \neg A \quad\; \neg B \end{array}
\qquad
\begin{array}{c} A \vee B \\ \diagup \quad \diagdown \\ A \quad\; B \end{array}
\qquad
\begin{array}{c} \neg(A \vee B) \\ | \\ \neg A \\ \neg B \end{array}
\qquad
\begin{array}{c} \neg\neg A \\ | \\ A \end{array}
$$

Figure 12.2: Propositional rules

propositions are not true in the same conditions. While the truth of a descriptive proposition can be represented by the fact that it belongs to the descriptive part of a scenario, the truth of a normative proposition depends upon a norm established by some authority

[22] For technical reasons, we do not consider mixed propositions (which contain both normative and descriptive atoms), mainly because it is unclear that $\supset$ preserve truth between descriptive and normative formulas (e.g., it is possible to have $p \supset P_i q$ true but $(p \wedge r) \supset P_i q$ false). This is related to the problem of detachment for deontic conditional (see Vorobej 1986).

[23] w is inconsistent if either $\mathfrak{D}$ or $\mathfrak{N}$ is.

(Alchourrón and Bulygin 1981, pp.97,102): A is obligatory if and only if there is a norm which makes the action described by A (i.e., Γ_A) an obligation. From a propositional standpoint, complex normative formulas behave schematically as complex descriptive formulas since both are composed of the same logical connectives (figure 12.2). The difference of truth conditions between normative and descriptive propositions can be seen through the representation of normative atoms: if O_iA is true for w, then

1. there is a norm (established by authority i) which makes A obligatory;
2. the action described by A pertains to a set of obligations.

Since logical connectives can (classically) be reduced to some composition of $\neg$ and $\supset$, we will only represent schematically truth conditions for O_ip, $O_i\neg A$ and $O_i(A \supset B)$ (figure 12.3).[24]

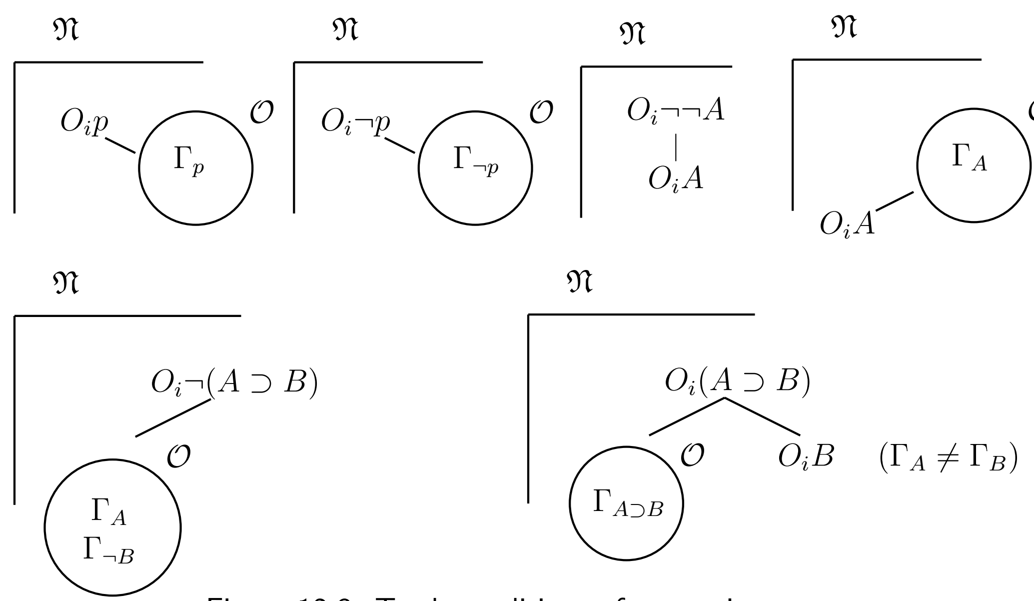


Figure 12.3: Truth conditions of normative atoms

These rules, combined with conditions C1, C2 and C3, allow us to verify the validity of a normative inference. Returning to figure 12.4, we can put the action Γ_p in the set $\mathcal{O}$ by the rules for normative atoms. Since $p \supset q$ is assumed to be true in the descriptive part of w, it follows that $\Gamma_p \sqsubseteq \Gamma_q$ holds in its normative part. By C2 we obtain Γ_q in $\mathcal{O}$, meaning that we can conclude Oq in the normative part of w. The only branch in the normative part closes (schematically represented by $\times$) since it contains a contradiction (i.e., O_iq and $\neg O_iq$). Therefore, there is no possible scenario in which the premises of (Δ) are true while its conclusion is false, hence the proof of (Δ)'s validity.

This proof can be rephrased as following. Assume that p is obligatory, that p implies q but that q is not obligatory. Since p is obligatory, it follows that the action described by p pertains to a set of obligation. Considering that p implies q is taken to be true in

[24] The definitions are $A \wedge B =_{def} \neg(A \supset \neg B)$, $A \vee B =_{def} \neg A \supset B$ and $A \equiv B =_{def} (A \supset B) \wedge (B \supset A)$.

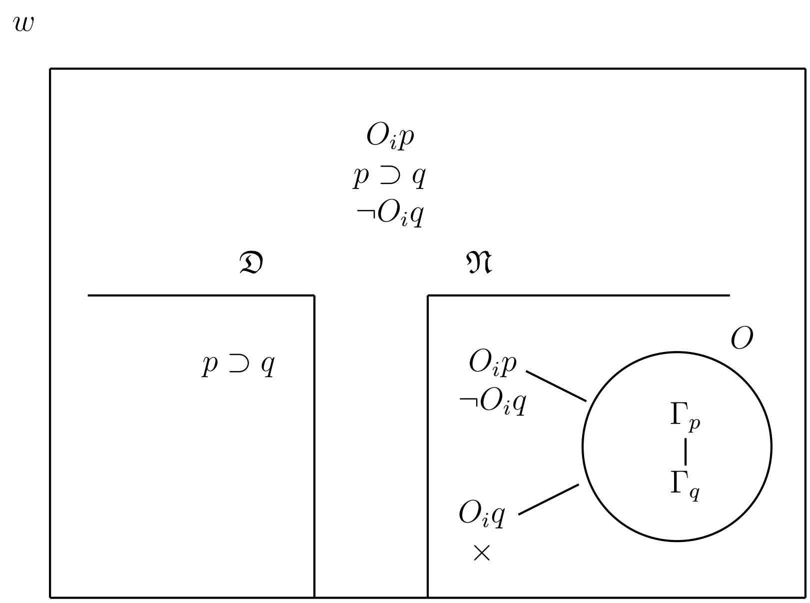


Figure 12.4: Test of validity

the descriptive part of w, it follows that there is a semantical relation between the action described by p and the action described by q. From this relation and the condition C2, we can conclude that the action described by q also pertains to the set of obligations, and thus q is obligatory, which contradicts our hypothesis.

Soundness

In addition to the formal method developed so far, one might want some hints to test the soundness of a normative inference. Without pretending to be exhaustive, we only give some philosophical insights regarding relevant questions one might try to answer. A sound inference is valid and has true premises. The first question is thus: are the normative premises true? The truth of a normative proposition depends upon the fact that there is an authority which establishes that norm. Hence, one must first contextualize the normative proposition to an authority: is it a legal ought? a moral ought? If it is a moral ought, according to which ethical theory is it understood? Utilitarianism? Kant's deontology? When the authority is specified, it will be easier to tell if the normative statement is actually true. It is noteworthy that some inferences can be fallacious when the meaning of 'ought' is not the same throughout the argument.

The context of the argument is also relevant. Consider the premise 'Paul ought to tell the truth'. Assume that 'ought' is understood legally. While one is not legally obliged to tell the truth in one's day to day life, one *is* however legally obliged to tell the truth in a court of justice.

In addition to the truth of the normative premises, it might be useful to analyze the descriptive propositions used in conjunction with (Δ). The conditional must be analyzed in terms of necessity and sufficiency. Is the conditional representing a semantical relation of entailment between two actions? Consider for example the following premise.

P_1	If Paul does not steal, then Paul does not break Peter's window.

Is this a necessary entailment between two actions? Is the action described by 'Paul steals' necessarily entailed by 'Paul breaks Peter's window'? Clearly not: it is perfectly plausible that Paul breaks Peter's window without stealing (e.g., while playing hockey).

Moreover, one must contextualize the descriptive proposition in regards to the normative authority. Consider the following argument.

P_1	Paul ought to respect his neighbour.
P_2	If Paul respects his neighbour, then Paul does not insult his neighbour.
C	Paul ought to not insult his neighbour.

If one understands 'ought' in terms of common sense morality, then it is acceptable to see an insult as a lack of respect. But if one understands 'ought' legally, then it is not clear that P_2 is an acceptable premise. Assuming that the first premise is legally true, the question would be: what legally counts as a lack of respect? Would Paul insulting his neighbour count as violating his legal obligation to respect his neighbour? It seems not, since 'respecting his neighbour' would in all likelihood legally mean to respect his neighbour's rights, property, belongings, etc. In short, one must answer the question 'what action counts as *A* from the authority's standpoint?'.[25]

Now, consider the following argument and assume a context in which premises P_2 and P_3 are acceptable.

P_1	Paul ought to rescue Peter from drowning.
P_2	If Paul rescues Peter from drowning, then Paul calls 911.
P_3	If Paul calls 911, then Paul breaks into Sam's house.
C	Paul ought to break into Sam's house.

Assuming that breaking and entering is legally forbidden, this argument leads to a conflict of obligations, hence to a contradiction into our current framework. The contradiction can be obtained with the addition of P_0.

P_0	Paul ought to not break into Sam's house.

[25] For an analysis of 'count as', see Jones and Sergot (1996) and Boella and van der Torre (2006c).

Despite its validity, this argument is however not sound. Indeed, either P_0 or P_1 must be rejected. In order to see this, let us borrow from legal discourse the concept of *hierarchy*. Even though there are often (*a priori*) inconsistencies in the law, a legal system preserves its consistency *after interpretation* by means of hierarchy (Côté 2006). The conflict of obligations can be overruled with the relation of hierarchy since it allows one to prioritize P_1 over P_0. Indeed, it is likely that preserving the integrity of one's life would overrule preserving the integrity of one's property: if the only possible course of action for Paul to save Peter's life is to break into Sam's house, then the obligation to save Peter will have priority. Hence, premise P_0 can be rejected and the conflict of obligations is solved.

A final clue is to consider the principle *ought-implies-can*. Consider the following argument.

P_1	Paul ought to rescue Peter from drowning.
P_2	If Paul rescues Peter from drowning, then Paul jumps in the water to retrieve Peter.
C	Paul ought to jump in the water to retrieve Peter.

If Paul cannot swim, then the action described by 'rescuing Peter from drowning' does not entail the action described by 'jumping in the water to retrieve Peter'. Thus, in this context, one could reject premise P_2 on the ground that it is a violation of the ought imply can principle.

Conclusion

To sum up, we introduced a method to test the formal validity of normative inferences. This method is restricted to normative inferences in which there is only one type of obligation and where there is no iteration of deontic operator. For technical reasons, we also did not study 'mixed' formulas, and as such the method does not apply to conditional normative reasoning. A formal method for testing the validity of normative inferences is a useful tool in critical thinking since it allows one to obtain a definitive proof of an argument's validity. This follows from the fact that formal validity is well defined while intuitive validity is not. Our method covers a portion of the intuitive validity of normative inferences which was not covered by other frameworks in the literature. In addition to the fact that the usual frameworks (**CL** and **FOL**) used in critical thinking are not rich enough to represent the formal validity of normative inferences, these framework are not able to distinguish between truth conditions of normative and descriptive propositions, and thus are not able to represent the semantical dichotomy between facts and norms. Although the main contribution of this paper was to develop a formal method to test validity for normative inferences, we also provided the reader with a brief analysis of some important aspects of sound normative inferences.

All things considered, we introduced a method which can help to determine the 'validity part' of a sound normative inference. As a contribution to the field of critical thinking, it is our view that this method would be best used in conjunction with informal approaches to verify the acceptability of a normative inference's premises.

* * *

Le présent chapitre répondait à la seconde partie du premier objectif de cette thèse, soit de fournir un cadre de travail en vue de l'analyse des inférences normatives inconditionnelles s'adressant à un public non ou peu initié à la logique philosophique. Il se base sur les résultat du chapitre précédent et vise à expliciter la nécessité et la pertinence d'une analyse plus fine quant à la modélisation des inférences normatives. Le prochain article ne répond pas directement à l'un des objectifs de la thèse mais introduit néanmoins la théorie des types pertinente à ce qui sera présenté plus tard au chapitre 16. Nous avons préféré introduire ce chapitre dès maintenant considérant que celui-ci explicite davantage, à l'instar du chapitre 14, les notions relatives à la théorie des catégories.

Chapitre 13

From linguistics to deontic logic via category theory

Abstract

The present paper aims to bridge the gap between deontic logic, categorial grammar and category theory. We propose to analyze Forrester's (1984) paradox through the framework of Lambek's (1958) syntactic calculus. We first recall the definition of the syntactic calculus and then explain how Lambek (1988) defines it within the framework of category theory. Then, we briefly present Forrester's paradox in conjunction with standard deontic logic, showing that this paradox contains some features that reflect many problem within the literature. Finally, we analyze Forrester's paradox within the framework of the syntactic calculus and we show how a typed syntax can provide conceptual insight regarding some of the problems that deontic logic faces.

Keywords: Categorial grammar, Joachim Lambek, Forrester's paradox, Modal logic, Deductive systems, Type theory

Introduction

Categorial grammar, understood as a (well defined) formal syntax, was introduced in the work of Ajdukiewicz (1935) and Bar-Hillel (1953), although the expression 'categorial grammar' appeared later (see Bach 1988). In contrast with approaches that concentrate on the phrase's structure, categorial grammar analyzes sentences in terms of *syntactical categories* (Bach 1988, p.1). In 1958, Lambek introduced a syntactical calculus based on Gentzen's (1934) sequent calculus as a tool to analyze these *syntactical categories*. Later, Lambek (1968, 1969) realized that his syntactic calculus behaved as a *closed category*, and so Lambek (1988) revisited categorial grammar within the framework of category theory. Although the term 'category' is used to refer to *syntactical categories* in categorial grammar, 'category' in the context of category theory refers to the concept introduced in the work of Eilenberg and Mac Lane (1945). Following Lambek (1988, p.298), we shall restrict the use of the term 'category' to category theory only. We will speak of *syntactic types* rather than *syntactic categories* to avoid confusion.

Also introduced in the fifties but utterly unrelated to categorial grammar and category theory, von Wright (1951) proposed to construct a (monadic) deontic logic in analogy with alethic modal logic to formalize the behaviour of terms such as *obligatory*, *permitted* or *forbidden*. It did not take much time for his work to become the target of multiple objections, now known as the *paradoxes* of deontic logic. The first objection was formulated by Prior (1954) and aimed to show that von Wright's notion of *commitment* fails in context of derived obligations. Von Wright (1956) took this paradox very seriously and answered it by introducing the building blocks of what is nowadays known as *dyadic deontic logic*. Even though dyadic deontic logic attracted the interest of some philosophers (for instance Rescher 1958), the initial system of von Wright stayed a choice target for the paradoxes. In 1958, Prior presented the paradox of the good Samaritan, precursor of Nozick's and Routley's (1962) robbery paradox, which aimed to show that the relation of deontic consequence in von Wright's system leads to paradoxical results. But the *coup de grâce* against his approach was given by Chisholm (1963), who showed that monadic deontic logic is simply not powerful enough to represent conditional obligations, a central notion for normative reasoning. Dyadic deontic logic was further developed by von Wright (1967) and has since then become a distinct field of study in philosophical logic (for an overview, see Tomberlin 1981). Chisholm's objection was however not literally a mortal blow to monadic deontic logic. After the introduction of possible world semantics by Kripke (1963), von Wright's system was analyzed in the framework of modal logics, which gave rise to the well-known *standard system of deontic logic* (SDL for short).[1] Despite many objections against SDL as a logic that represents deontic entailment, it remains a central point in the literature insofar as philosophers either build their systems as extensions of SDL or position themselves according to it, pointing out which principles they reject and why. Some times after the development of SDL, Forrester (1984) presented an objection against the standard system to show that SDL's notion of consequence simply does not represent the relation of deontic entailment within the natural language.

Although deontic logic was introduced roughly at the same time as categorial grammar and category theory, it has been developed and utilized independently from both these fields.[2] The present paper aims at making the bridge between deontic logic, on the one hand, and category theory and categorial grammar, on the other. This will be done by analyzing Forrester's paradox through the categorical framework of Lambek's syntactic calculus. We will first present the syntactical calculus in the context of category theory, and then we will briefly present SDL and Forrester's paradox, explaining why we concentrate on this paradox rather than others. We will analyze the paradox through the framework developed so far and conclude in the last section with remarks for future research.

[1] Note that von Wright's initial system is not equivalent to the standard system.

[2] To the best of our knowledge, there has been no approach joining categorial grammar with deontic logic and scant few regarding deontic logic and category theory. Johanson (1996) used category theory to model normative systems but there was no follow up, and Lucas (2006, 2007, 2008) used category theory to model actions in deontic contexts.

Lambek's syntactic calculus

The general idea of categorial grammar is to analyze sentences in terms of syntactical types rather than to concentrate on the phrase's structure (Bach 1988, p.1). Assuming a finite set of primitive types $\mathcal{T} = \{\tau_1, \ldots, \tau_n\}$, the set of complex types is defined recursively from the language $\mathcal{L} = \{(,), \otimes, /, \backslash, \mathcal{T}\}$ by:

$$\varphi := \tau_i \mid \varphi \otimes \psi \mid \varphi/\psi \mid \varphi\backslash\psi$$

The operator $\otimes$ expresses concatenation of types and $\varphi \otimes \psi$ is read 'φ times ψ'. The types φ/ψ and $\varphi\backslash\psi$ are respectively read 'φ over ψ' and 'φ under ψ'. Here, a syntactical type is understood as a linguistic functor (not to be confused with a functor between two categories), which takes a type and transforms it into another. For instance, if we assume a set $\mathcal{T} = \{s, n\}$ of two primitive types s (for declarative sentences) and n (for nouns), then the type $n\backslash s$ is understood as a linguistic functor which takes a noun and transforms it into a declarative sentence *from the right*. Similarly, the type s/n would be understood as a linguistic functor which takes a type n and transforms it into a declarative sentence *from the left*. Let us borrow two examples from Lambek (1958, p.156) to see how it works.

$$\begin{array}{cc} \text{Paul} & \text{steals} \\ n & n\backslash s \end{array}$$

While 'Paul' is of type n, 'steals' is of the type that takes a noun and transforms it into something of type s from the right. Similarly, the word 'poor' in the following sentence is of the type that takes a noun and transforms it into a noun from the left, thus n/n.

$$\begin{array}{ccc} \text{poor} & \text{Paul} & \text{steals} \\ n/n & n & n\backslash s \end{array}$$

A word must be of *at least* one type, but it can have different types depending on the context. For instance, 'steals' in the latter example could also be understood as of the type $((n/n) \otimes n)\backslash s$, that is, of the type which takes something of type $(n/n) \otimes n$ and transforms it into a declarative sentence from the right. By concatenation, the sentence 'poor Paul steals' is of the following type.

$$((n/n) \otimes n) \otimes (n\backslash s)$$

Having this in mind, an obvious question would be to ask: how can one prove that something which is of type $((n/n) \otimes n) \otimes (n\backslash s)$ is of type s? Put differently, how can we prove that 'poor Paul steals' is a sentence? This is where Lambek's syntactic calculs enters the picture.

The syntactic calculus is defined on the grounds of a *deductive system*, which in turn is defined within the framework of category theory. Let us consider first the axiomatic definition of a category (cf. Mac Lane 1971, pp.7-8).

Definition 13.1. *A category $\mathcal{C}$ is composed of:*

1. *$\mathcal{C}$-objects;*
2. *$\mathcal{C}$-arrows;*
3. *an operation which assigns to each $\mathcal{C}$-arrow two $\mathcal{C}$-objects (the domain and the codomain);*
4. *an operation which assigns to each pairs of $\mathcal{C}$-arrows f and g for which g's domain is f's codomain the* composite arrow *gf and which respects the associative law, meaning that the diagram*

$$\begin{tikzcd}
a \arrow[r, "f"] \arrow[d, "h(gf)=(hg)f"'] \arrow[dr, "gf"] & b \arrow[d, "g"] \arrow[dl, "hg"'] \\
d & c \arrow[l, "h"]
\end{tikzcd}$$

   commutes;

5. *an operation which assigns to each $\mathcal{C}$-object an identity arrow $1_a : a \longrightarrow a$ which respects the* identity law, *meaning that the diagram*

$$\begin{tikzcd}
a \arrow[r, "f"] \arrow[dr, "f"'] & b \arrow[d, "1_b"] \arrow[dr, "g"] & \\
& b \arrow[r, "g"'] & c
\end{tikzcd}$$

   commutes.

Now, consider Lambek's (1988, pp.302,307) definition of a deductive system.

Definition 13.2. *A* deductive system *is composed of:*

1. *types, considered as objects;*
2. *proofs, considered as arrows between types;*
3. *the identity arrow $1_A : A \longrightarrow A$ for any type A respecting the identity law;*
4. *composition $gf : A \longrightarrow C$ of arrows $f : A \longrightarrow B$ and $g : B \longrightarrow C$ respecting the associative law.*

$$f1_A = f = 1_B f \tag{Identity law}$$
$$h(gf) = (hg)f \tag{Associative law}$$

As such, a deductive system can be defined as a category where the arrows are proofs (deductions) and the objects are formulas (types).[3] Now, consider the definition of a *monoidal category* (cf. Mac Lane 1971, p.161).[4]

Definition 13.3. *A* monoidal category *is a category $\mathcal{C}$ composed of:*

1. *a tensor product $\otimes$ (i.e., a functor $\otimes : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$);*
2. *a unit object 1 such that there are arrows (natural isomorphisms) f_x and g_x;*

$$f_x : 1 \otimes x \longrightarrow x$$
$$g_x : x \otimes 1 \longrightarrow x$$

3. *an arrow $a_{x,y,z}$ for associativity (which is a natural isomorphism);*

$$a_{x,y,z} : (x \otimes y) \otimes z \longrightarrow x \otimes (y \otimes z)$$

4. *the following commutative diagrams, representing respectively the triangle and pentagon identities.*

$$\begin{tikzcd}
(x \otimes 1) \otimes y \arrow[rr, "a_{x,1,y}"] \arrow[dr] & & x \otimes (1 \otimes y) \arrow[dl] \\
& x \otimes y &
\end{tikzcd}$$

$$\begin{tikzcd}
& (w \otimes x) \otimes (y \otimes z) \arrow[r] & w \otimes (x \otimes (y \otimes z)) \\
((w \otimes x) \otimes y) \otimes z \arrow[ur] \arrow[dr] & & \\
& (w \otimes (x \otimes y)) \otimes z \arrow[r] & w \otimes ((x \otimes y) \otimes z) \arrow[uu]
\end{tikzcd}$$

In a nutshell, a monoidal category is a category $\mathcal{C}$ together with an associative tensor product and a unit. From the definition of a deductive system, Lambek (1988, p.302) proposes the following definition for the syntactic calculus. As we will see, a deductive system is defined as a monoidal category given that concatenation of types is an associative tensor product that comes with a unit I.

[3] It should be noted that Lambek (1988, 1999) proceeds backwards and defines a category on the grounds of a deductive system instead of defining a deductive system as a category, although he himself points out that this definition of a category is unorthodox (cf. Lambek 1988, p.298). Since Lambek's work, it has been shown that there is a strong connection between logical systems and different kinds of categories. As such, from a conceptual point of view, it is better to define a deductive system as a category that respects specific criteria, and this is what will be done within this paper.

[4] See also the $n$lab entry on monoidal category: `http://ncatlab.org/nlab/show/monoidal+category`.

Definition 13.4. *The* syntactic calculus *is a deductive system with a special type I and which is closed under three binary operations $\otimes$, $/$ and $\backslash$ that satisfy the axiom schemata (ρ_A)–(α_A^{-1}), which are natural transformations, and the rules of inference (f^*)–(^+g).*

$$\rho_A : A \otimes I \longrightarrow A \tag{ρ_A}$$
$$\rho_A^{-1} : A \longrightarrow A \otimes I \tag{ρ_A^{-1}}$$
$$\lambda_A : I \otimes A \longrightarrow A \tag{λ_A}$$
$$\lambda_A^{-1} : A \longrightarrow I \otimes A \tag{λ_A^{-1}}$$
$$\alpha_{A,B,C} : (A \otimes B) \otimes C \longrightarrow A \otimes (B \otimes C) \tag{$\alpha_{A,B,C}$}$$
$$\alpha_{A,B,C}^{-1} : A \otimes (B \otimes C) \longrightarrow (A \otimes B) \otimes C \tag{$\alpha_{A,B,C}^{-1}$}$$

$$\frac{g : A \longrightarrow C/B}{g^+ : A \otimes B \longrightarrow C}\,(g^+) \qquad \frac{g : B \longrightarrow A\backslash C}{{}^+g : A \otimes B \longrightarrow C}\,({}^+g)$$

$$\frac{f : A \otimes B \longrightarrow C}{f^* : A \longrightarrow C/B}\,(f^*) \qquad \frac{f : A \otimes B \longrightarrow C}{{}^*f : B \longrightarrow A\backslash C}\,({}^*f)$$

Taken together, the axiom schemata (ρ), (ρ^{-1}), (λ) and (λ^{-1}) make I into a nullary type (we can think of I as a space), and the axiom schemata (α) and (α^{-1}) state that $\otimes$ respects associativity. Note that Lambek (1988, p.309) assumes the following proofs, and as such the syntactic calculus satisfies the triangle and pentagon identities.

$$m_1 : A \otimes B \longrightarrow (A \otimes I) \otimes B$$
$$m_2 : (A \otimes I) \otimes B \longrightarrow A \otimes (I \otimes B)$$
$$m_3 : A \otimes (I \otimes B) \longrightarrow A \otimes B$$
$$n_1 : ((A \otimes B) \otimes C) \otimes D \longrightarrow (A \otimes B) \otimes (C \otimes D)$$
$$n_2 : (A \otimes B) \otimes (C \otimes D \longrightarrow A \otimes (B \otimes (C \otimes D))$$
$$n_3 : A \otimes (B \otimes (C \otimes D)) \longrightarrow A \otimes ((B \otimes C) \otimes D)$$
$$n_4 : A \otimes ((B \otimes C) \otimes D) \longrightarrow (A \otimes (B \otimes C)) \otimes D$$
$$n_5 : (A \otimes (B \otimes C)) \otimes D \longrightarrow ((A \otimes B) \otimes C) \otimes D$$

To understand properly the categorical structure induced on the syntactic calculus via the operations $/$ and $\backslash$, we must introduce the notion of a *closed category*. Following the nLab entry on closed monoidal categories[5], a monoidal category is closed insofar as there is an isomorphism between the class of morphisms (of $\mathcal{C}$) from $A \otimes B$ to C and the class of morphisms from A to $[B, C]$, with $[B, C]$ a right adjoint functor of $\otimes$. In other words, we have:

$$\mathrm{Hom}_{\mathcal{C}}(A \otimes B, C) \cong \mathrm{Hom}_{\mathcal{C}}(A, [B, C])$$

[5] `http://ncatlab.org/nlab/show/closed+monoidal+category`

In Lambek's terminology, the syntactic calculus is defined as a *biclosed* monoidal category, that is, a monoidal category which is right closed and left closed.[6] In the syntactic calculus, $/ : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$ is a right adjoint of $\otimes$ inasmuch as the isomorphism can be proven from (f^*) and (g^+). Hence, the syntactic calculus, when considered as a monoidal category, is *left closed*. But it is also *right closed* seeing that $\otimes$ also possesses another right adjoint $[A, C]$, namely $\backslash : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$, such that:

$$\mathrm{Hom}_{\mathcal{C}}(A \otimes B, C) \cong \mathrm{Hom}_{\mathcal{C}}(B, [A, C])$$

This isomorphism can be proven from $({}^*f)$ and $({}^+g)$. Lambek defines the syntactical calculus as a *biclosed* monoidal category given that it is a monoidal category where the tensor product $\otimes$ possesses two right adjoints, namely $\backslash$ and $/$, and as such it is both left and right closed.

It is noteworthy that it is not assumed that there are proofs from $A \otimes B$ to $B \otimes A$, nor from $B \otimes A$ to $A \otimes B$. Hence, the syntactic calculus is not a *symmetric* biclosed monoidal category, nor a *braided* biclosed monoidal category. Actually, the fact that the syntactic calculus is *not* a symmetric monoidal closed category implies that the two right adjoints are different.

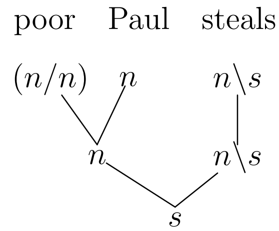
Returning to the aforementioned example, we can now prove within the syntactic calculus that a sentence of type $((n/n) \otimes n) \otimes (n\backslash s)$ is of type s. To do so, it suffices to show that there is a proof of s from $((n/n) \otimes n) \otimes (n\backslash s)$.

Theorem 13.1. *There is an arrow $h : ((n/n) \otimes n) \otimes (n\backslash s) \longrightarrow s$, namely the arrow $[(({}^+(1_{n\backslash s}))^*)(1_{n/n})^+]^+$.*

$$\dfrac{\dfrac{1_{n/n} : n/n \longrightarrow n/n}{(1_{n/n})^+ : n/n \otimes n \longrightarrow n} \qquad \dfrac{\dfrac{1_{n\backslash s} : n\backslash s \longrightarrow n\backslash s}{{}^+(1_{n\backslash s}) : n \otimes (n\backslash s) \longrightarrow s}}{({}^+(1_{n\backslash s}))^* : n \longrightarrow s/(n\backslash s)}}{\dfrac{(({}^+(1_{n\backslash s}))^*)(1_{n/n})^+ : n/n \otimes n \longrightarrow s/(n\backslash s)}{[(({}^+(1_{n\backslash s}))^*)(1_{n/n})^+]^+ : n/n \otimes n \otimes n\backslash s \longrightarrow s}}$$

Note that one could have found an easier way to obtain that arrow. Intuitively, one could be tempted to assume that if it is possible to prove that there is an arrow $f : (n/n) \otimes n \longrightarrow n$, and then another arrow $g : n \otimes (n\backslash s) \longrightarrow s$, then one would obtain a proof from $((n/n) \otimes n) \otimes (n\backslash s)$ to s. The intuition would be to go from $((n/n) \otimes n)$ to n, and then from $n \otimes (n\backslash s)$ to s, as represented in the following tree.

[6] Note that 'closed' is more common than 'biclosed'.



But this is not how it works. Indeed, the composite arrow gf would not be defined in that case since g's domain would be different from f's codomain. In order to apply transitivity, one would first need to prove that there is $h : ((n/n) \otimes n) \otimes (n\backslash s) \longrightarrow n \otimes (n\backslash s)$ and then show $i : n \otimes (n\backslash s) \longrightarrow s$, which is quite different from our starting intuition. Nonetheless, it is still possible to show that there is a derived rule such that if one assumes that there are arrows $f : A \otimes B \longrightarrow C$ and $g : C \otimes D \longrightarrow E$, then one can conclude that there is an arrow $h : (A \otimes B) \otimes D \longrightarrow E$.[7]

$$\dfrac{\dfrac{\text{(H) } f : A \otimes B \longrightarrow C \qquad \dfrac{\text{(H) } g : C \otimes D \longrightarrow E}{g^* : C \longrightarrow E/D}}{(g^*)f : A \otimes B \longrightarrow E/D}}{((g^*)f)^+ : (A \otimes B) \otimes D \longrightarrow E} \qquad (\beta)$$

This makes the proof more intuitive insofar as it is easier to show the two following arrows.

$$\dfrac{1_{n/n} : n/n \longrightarrow n/n}{(1_{n/n})^+ : (n/n) \otimes n \longrightarrow n}$$

$$\dfrac{1_{n\backslash s} : n\backslash s \longrightarrow n\backslash s}{{}^+(1_{n\backslash s}) : n \otimes n\backslash s \longrightarrow s}$$

Hence, the aforementioned derived rule can be applied to obtain the following proof.

$$\dfrac{(1_{n/n})^+ : (n/n) \otimes n \longrightarrow n \qquad {}^+(1_{n\backslash s}) : n \otimes n\backslash s \longrightarrow s}{[(({}^+(1_{n\backslash s}))^*)1_{n/n}]^+ : ((n/n) \otimes n) \otimes (n\backslash s) \longrightarrow s}$$

Summing up, Lambek's syntactic calculus can be defined through category theory as a biclosed monoidal category since the operations respect the minimal properties of a tensor product with its adjoint(s). From a semantical point of view, Lambek interpreted the syntactic calculus within a topos semantics via a structure preserving functor (see Lambek 1988, p.313 for details).[8] Note, however, that the syntactic calculus can be interpreted in

[7] I am indebted to Marc-Kevin Daoust for that idea.

[8] This requires the introduction of the usual lattice operations of a Heyting algebra (cf. Lambek 1999, p.281).

some semantics that differ from toposes. The advances that have been made in category theory showed that monoidal categories, and not just toposes, play an important role in the development of logical frameworks. There are actually many semantical interpretations available in structures *weaker* than toposes. For instance, various substructural logics are extension of the syntactic calculus and can be interpreted within the framework of residuated lattices (cf. Galatos et al. 2007). Already in 1968, Lambek (1968, p.292) defined his syntactic calculus as a *residuated category*, which was defined a year later as a bi-closed monoidal category (Lambek 1969, p.98). As such, one can see why / and \ are sometimes named the left and right *residuals* of $\otimes$.

While linear logic was introduced by Girard in 1987, Abrusci (1990a; 1990b) showed how Lambek's syntactic calculus is actually a fragment of non-commutative intuitionnistic linear logic (i.e., **NILL** without intuitionnistic negation).[9] The non-commutativity of the syntactic calculus comes from the fact that it is defined as a bi-closed monoidal category instead of a symmetric (or braided) monoidal closed category. Casadio and Lambek (2002) then showed how different categorial grammars can be obtained through the extension of the calculus via some additional properties, depending on whether one wants commutativity, compact duality or involutive duality. In these cases, one would respectively obtain classical bilinear logic, compact bilinear logic or Curry's semantic calculus. The reader may consult Casadio and Lambek (2002) for details. See also Moortgat (2009) for symmetric categorial grammar, which would be an extension of classical bilinear logic (where the tensor is non-commutative).

SDL and Forrester's paradox

Leaving aside the categorical framework of the syntactic calculus, we now turn our attention to modal logic. It is worth mentioning that the term 'standard system' in the deontic logic literature is sometimes used ambiguously. A distinction must be made between *the* standard system and *a* standard system (or *the* standard *systems*). *The* standard system, SDL, is equivalent to the modal logic KD. Following Åqvist (2002, p.205), SDL is the smallest set constructed from the language $\mathcal{L} = \{(,), Prop, \neg, \supset, O\}$ (with the usual definitions for the other logical connectives, interdiction $FA =_{def} O\neg A$ and weak permission $PA =_{def} \neg O\neg A$, and $Prop$ a denumerable set of atomic propositions) which is closed under the rules *modus ponens*, O-necessitation and contains the axiom schemata (A1)–(A4).

$$\frac{\vdash A \supset B \qquad \vdash A}{\vdash B} \qquad (\textit{modus ponens})$$

$$\frac{\vdash A}{\vdash OA} \qquad (\text{O-necessitation})$$

[9] See also Casadio et al. (2004) for the comparison between the syntactic calculus and **NILL**.

$$
\begin{aligned}
&\text{every propositional tautology of } \mathcal{L} && \text{(A1)}\\
&PA \equiv \neg O \neg A && \text{(A2)}\\
&O(A \supset B) \supset (OA \supset OB) && \text{(A3)}\\
&OA \supset PA && \text{(A4)}
\end{aligned}
$$

While (A1) implies that SDL is an extension of the (classical) propositional calculus, (A2) assumes that weak permission is the dual operator of O, (A3) is the K axiom of distribution and (A4) is the well-known axiom schema (**D**). The set of well-formed formulas is defined recursively by:

$$\varphi := p_i \mid \neg\varphi \mid \varphi \supset \psi \mid O\varphi$$

The modal logic KD is the smallest set which satisfies these conditions. On the other hand, *a* standard system is a conservative extension of KD. Following Åqvist (2002, p.155), Δ is a normal system of deontic logic if $K \subseteq \Delta$, and a strongly normal deontic logic if $KD \subseteq \Delta$. Thus, *a* standard system is a strongly normal deontic logic or, put differently, the class of standard systems contains extensions of KD.

Forrester's paradox aims to show that the rule (ROM), a derived rule in KD, fails to represent the intuitive notion of deontic entailment in *the* standard system.

$$\frac{\vdash A \supset B}{\vdash OA \supset OB} \qquad \text{(ROM)}$$

The paradox can be formulated as following (see Peterson and Marquis 2012):

(1) Jones murders Smith (p)

(2) Jones ought to not murder Smith ($O\neg p$)

(3) If Jones murders Smith, then Jones ought to murder Smith gently ($p \supset Oq$)

(4) If Jones murders Smith gently, then Jones murders Smith ($q \supset p$)

By modus ponens between (1) and (3) we get (5).

(5) Jones ought to murder Smith gently (Oq)

By (ROM) and *modus ponens*, (6) follows from (4) and (5).

(6) Jones ought to murder Smith (Op)

But from (A4) and (6) we can derive (7), which contradicts (2).

(7) it is false that Jones ought to not murder Smith ($\neg O \neg p$)

We follow Åqvist's (2002, pp.161-173) presentation of the paradoxes. Let $\tau : NDL \longrightarrow KD$ be a translation function from the *Natural Deontic Language* to KD's language, and τ^{-1} be the inverse map of τ from KD to NDL. Paradoxical results arise in two situations. Either there is a formula $\varphi \in KD$ such that $\tau^{-1}(\varphi) \notin NDL$ because intuitively $\tau^{-1}(\varphi)$ seems invalid, or there is a formula $\varphi \notin KD$ such that $\tau^{-1}(\varphi) \in NDL$ because $\tau^{-1}(\varphi)$ seems intuitively valid. These are respectively a *right-to-left inadequacy* and a *left-to-right inadequacy*. Let:

$$\phi = p \wedge [(p \supset Oq) \wedge (q \supset p)]$$
$$\pi = O\neg p \wedge \phi$$

Forrester's paradox states two right-to-left inadequacies, that is:

$$\phi \supset Op \in KD \qquad \tau^{-1}(\phi \supset Op) \notin NDL$$
$$\pi \supset \bot \in KD \qquad \tau^{-1}(\pi \supset \bot) \notin NDL$$

It was shown in Peterson and Marquis (2012) that Forrester's paradox is actually *not* a paradox for *the* standard system, although it *is* a paradox for *a* standard system, which either admits $q \supset p$ as an axiom or includes a necessity operator of type $\Box$ and can translate (4) by $\Box(q \supset p)$. Forrester's paradox is not a paradox for SDL per se since:

$$\phi \supset Op \notin KD$$
$$\pi \supset \bot \notin KD$$

Nonetheless, Forrester's paradox can be objected to some extensions of KD.

We chose to analyze Forrester's paradox insofar as it incorporates three different problems for deontic logic. First, given that the paradox uses an obligation that arises in a context of violation, it incorporates Chisholm's (1963) paradox and shows that the standard systems cannot deal appropriately with conditional obligations. Secondly, the step from (4), (5) and (ROM) to (6) is an instance of Prior's (1958) good Samaritan paradox, or of Nozick's and Routley's (1962) robbery paradox.[10] In addition to these two features, Castañeda (1986) pointed out that Forrester's paradox brought to light a more subtle, linguistic problem for deontic logic. Indeed, Castañeda (1986) argued that the paradox appears when one does not distinguish between *propositions* and *practitions*. In other words, if one wants to avoid the paradox, then one must distinguish between the *context* in which an obligation arises (described by a proposition) and the *action* which is in the scope of the deontic operator. While a *proposition* can be true or false (i.e., it is declarative), a *practition* cannot (Castañeda 1986, p.37).[11] We chose to analyze Forrester's

[10] This step cannot be done in KD, but assume a standard system in which it can.

[11] Note that his formal framework was not proposed as a tool that was meant to resolve Forrester's paradox but rather that it was a happy consequence that the framework provided the material to deal with the paradox. Indeed, Forrerster's paradox was published in 1984 (although Castañeda became aware of the paradox in 1982), and the solution proposed by Castañeda (1986) is based upon the formal framework presented in Castañeda (1981), which is the result of his work done in Castañeda (1959, 1968, 1970, 1977).

paradox through the framework of categorial grammar via category theory insofar as it involves:

1. contrary-to-duties (cf. Chisholm's paradox);
2. the notion of deontic entailment (cf. the good Samaritan paradox);
3. a syntactical distinction between propositions and practitions (cf. Castañeda's analysis of Forrester's paradox).

As we will see, these three aspects can be dealt with by using a typed syntax.

The paradox revisited

Deontic logic within a typed syntax

Deontic logic is usually interpreted as a modal logic. It should be noted, however, that our approach is neither oblivious to nor incompatible with the modal tradition. Indeed, there have been many approaches that developed modal logics within categorical settings.[12] Even though these approaches mostly dealt with $S4$ modalities, it remains that modal logic can be incorporated within a categorical framework. We wish to show how this could be beneficial for deontic logic. In what follows, we illustrate how category theory provides an interesting framework to analyze many problems that deontic logic faces. We show that most of these problems can be solved using a typed theory. Since categorical logic, once the quantifiers are introduced, is mainly a typed theory, our analysis suggests that category theory is a likely candidate for the formalization of deontic logic.

Obviously, the application of categorial grammar to deontic logic requires more primitive types than only s and n. In his 1958 paper, Lambek introduced n_c, n_s and n_p for count nouns, substance nouns and plural nouns. A year after, Lambek (1959) introduced more syntactic types to analyze the English verb phrase. For instance, he augmented $\mathcal{T}$ with i, p and q, standing respectively for the types of infinitive of intransitive verb, present participle of intransitive verb and past of intransitive verb. Of course, more syntactical types can be introduced, as it is done in Casadio and Lambek (2002), where one can find types for past and present questions, past and present declarative sentences and objects. Although our analysis does not require that we burden ourselves with types for questions, the analysis of Forrester's paradox requires that we introduce types for tensed declarative sentences. As was noted by Lambek (1988, p.315), thus constructed, the syntactic calculus does not account for *imperative* sentences, which we often distinguish from factual propositions as a result of Hume's is-ought thesis.

[12] See for instance Reyes and Zolfaghari (1991) and Makkai and Reyes (1995). See also the unpublished work of Awodey and Kishida (2007).

It is worth pausing here to consider carefully the type we wish to introduce for imperative sentences (or, more accurately, normative propositions).[13] Hume's semantical dichotomy between *facts* and *norms* states that one cannot infer a descriptive sentence from a set constructed only of normative premises and, vice versa, that one cannot infer a normative sentence from a set of descriptive premises. This semantical dichotomy causes some issues when one considers Jørgensen's (1937) dilemma. In a nutshell, the dilemma arises if one assumes that the truth value of a declarative sentence depends upon the world, that is, if one adopts a (naive) correspondence theory of truth. Since there is a semantical dichotomy between facts and norms, one can safely assume that if normative propositions can be true or false, then it is not in the same conditions that descriptive propositions do. But if one assumes that the truth value of a sentence depends upon its correspondence with reality, then, assuming the semantical dichotomy, normative propositions are not declarative given that normative propositions cannot be verified empirically. Thus the dilemma: normative inferences seem to follow some logical rules, but since the object of logic (when applied to inferences) is *declarative* sentences, it seems that logic cannot deal with normative inferences seeing that normative propositions are not declarative (assuming the correspondence theory of truth).[14]

Our solution to the dilemma is to reject the correspondence theory of truth. Normative propositions *are* declarative. For example, it is true that one should not steal since there are legal norms that forbid stealing. However, we nonetheless assume Hume's semantical dichotomy since normative propositions are not true in the same conditions that descriptive propositions are.[15] But still, both are declarative. This leads us to distinguish between two syntactic types. While s_d is the syntactic type of descriptive (declarative) sentences, s_n is the syntactic type of normative (declarative) sentences.

Thus, we assume a language $\mathcal{L}$ as defined earlier but constructed from the set of primitive types $\mathcal{T} = \{s_d, s_n, n, i, p, q\}$. As it was mentioned in Lambek (1999, p.281), one must assign to each word at least one syntactical type by means of a dictionary. This begs the question of the syntactical types of words such as *ought, should, can* or *must* in a normative context, but also of words such as *obligatory, permitted* or *forbidden.* Since von Wright, the tradition has been to interpret these words as modalities that influence declarative sentences. For instance, the following sentence would be translated in SDL by Op, with p being 'John eats'.

John ought to eat

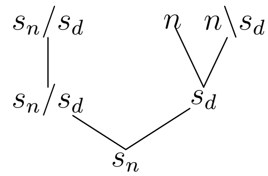
[13] We distinguish between a norm (or an imperative) and a normative proposition. While the latter is declarative, the former is not.

[14] See Peterson (2011) for an analysis of the dilemma.

[15] Following Alchourrón and Bulygin (1981), the truth value of a normative proposition depends upon a norm.

From the point of view of syntactical types, this would require that the sentence be written as

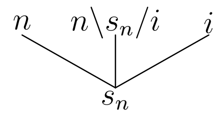
$$\text{ought (John eats)} \tag{8}$$



In this case, 'ought' would be of type s_n/s_d, that is, of the type that takes a descriptive sentence and transforms it into a normative one from the left.

Another option would be to consider the term 'ought' as it was written in the first sentence:

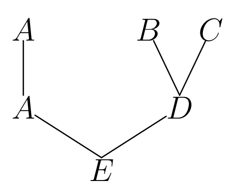
$$\text{John ought (to eat)} \tag{9}$$



In both cases, it can easily be proven that the sentences are of type s_n. The proof of (8) requires (but not necessarily) a derived rule similar to β as presented previously. The proof (left to the reader) becomes obvious when one can prove (δ):

$$\dfrac{\dfrac{(\mathrm{H})\ g : B \otimes C \longrightarrow D \qquad \dfrac{(\mathrm{H})\ f : A \otimes D \longrightarrow E}{{}^{*}f : D \longrightarrow A\backslash E}}{({}^{*}f)g : B \otimes C \longrightarrow A\backslash E}}{{}^{+}[({}^{*}f)g] : A \otimes (B \otimes C) \longrightarrow E} \tag{δ}$$

The proof of (δ) can be visualized by the following tree:



The proof of (9) is straightforward:

$$\dfrac{\dfrac{1_{n\backslash s_n/i} : n\backslash s_n/i \longrightarrow n\backslash s_n/i}{(1_{n\backslash s_n/i})^{+} : (n\backslash s_n/i) \otimes i \longrightarrow n\backslash s_n}}{{}^{+}[(1_{n\backslash s_n/i})^{+}] : n \otimes ((n\backslash s_n/i) \otimes i) \longrightarrow s_n}$$

Now, an obvious question would be to determine whether these two types are equivalent, that is, if there are arrows:

$$\gamma : s_n/s_d \longrightarrow n\backslash s_n/i$$
$$\gamma^{-1} : n\backslash s_n/i \longrightarrow s_n/s_d$$

Without pretending that such arrows do not exist, we did not find any (although this may be unsatisfactory for the reader). For now, we assume that the distinction between the two different syntactical types is sound.

With that in mind, we can revisit Forrester's paradox and see which type each sentence is. Let us consider the first type for 'ought' (hereafter the a-translations):

(1)	Jones	murders	Smith
	n	$n\backslash s_d/n$	n

(2a)	ought	(not	(Jones	murders	Smith))
	s_n/s_d	s_d/s_d	n	$n\backslash s_d/n$	n

(3a)	Jones	murders	Smith	implies	ought	(Jones
	n	$n\backslash s_d/n$	n	$\mathbf{s_d\backslash s_n/s_n}$	s_n/s_d	n

murders	Smith	gently)
$n\backslash s_d/n$	n	$s_d\backslash s_d$

(4)	Jones	murders	Smith	gently	implies	Jones	murders	Smith
	n	$n\backslash s_d/n$	n	$s_d\backslash s_d$	$s_d\backslash s_d/s_d$	n	$n\backslash s_d/n$	n

(5a)	ought	(Jones	murders	Smith	gently)
	s_n/s_d	n	$n\backslash s_d/n$	n	$s_d\backslash s_d$

(6a)	ought	(Jones	murders	Smith)
	s_n/s_d	n	$n\backslash s_d/n$	n

(7a)	not	(ought	(not	(Jones	murders	Smith)))
	$\mathbf{s_n/s_n}$	s_n/s_d	$\mathbf{s_d/s_d}$	n	$n\backslash s_d/n$	n

The proofs of the types of these sentences are relatively easy and so are left to the reader.[16] Let us now look at the translations for the second interpretation of 'ought' (hereafter the b-translations):

(2b)	Jones	ought	to not	murder	Smith
	n	$n\backslash s_n/i$	i/i	i /n	n

[16] Hint: construct a tree and then use (β) and (δ).

(3b)	Jones	murders	Smith	implies	Jones	ought	
	n	$n\backslash s_d/n$	n	$\mathbf{s_d\backslash s_n/s_n}$	n	$n\backslash s_n/i$	
	to murder	Smith	gently				
	i/n	n	$i\backslash i$				

(5b)	Jones	ought	to murder	Smith	gently
	n	$n\backslash s_n/i$	i/n	n	$i\backslash i$

(6b)	Jones	ought	to murder	Smith
	n	$n\backslash s_n/i$	i/n	n

(7b)	not	(Jones	ought	to not	murder	Smith)
	$\mathbf{s_n/s_n}$	n	$n\backslash s_n/i$	$\mathbf{i/i}$	i/n	n

Ought-to-be/Ought-to-do

The reader familiar with the deontic logic literature will in all likelihood have noticed that the first interpretation of 'ought' is related to the *Ought-to-be* interpretation of normative sentences, while the second interpretation refers to the *Ought-to-do* (see von Wright 1999b). While the *Ought-to-be* interpretation takes 'ought' as a modality that changes the truth value of a descriptive proposition (e.g., modal deontic logic within possible worlds semantics), the *Ought-to-do* interpretation considers that there are actions (rather than propositions) in the scope of the deontic operator (as in dynamic deontic logic and deontic logic with algebras for actions).

It is noteworthy that from the perspective of categorial grammar, the *Ought-to-do* interpretation of the operator O allows us to construct normative sentences that are more closely related to the English language. Indeed, the compositionality of $n\backslash s_n/i$ seems to be naturally related to the natural language. When interpreted in terms of *Ought-to-be*, 'ought' is taken as a modality that modifies the truth value of a descriptive statement, and thus the formal translation requires that we rearrange the phrase's structure. This results in an asymmetry between the formal language and the English language. Interpreting 'ought' in terms of *Ought-to-do* allows us to preserve the phrase's structure within the formal language of the syntactic calculus, allowing for a more faithful representation of the sentence into the formal language. Hence, the analysis of normative sentences from the point of view of syntactical types suggests that normative (declarative) sentences should be interpreted as being of the *Ought-to-do* type rather than of the *Ought-to-be*, meaning that the deontic operator O is of type $n\backslash s_n/i$ rather than s_n/s_d.

That being said, if one wants an *Ought-to-do* interpretation of O, then one might want to replace the primitive type i by a primitive type a for actions. Understood this way, the deontic operator O would be of type $n\backslash s_n/a$ rather than $n\backslash s_n/i$. This would mean that O is of the type that takes a noun and an action and transforms it into a normative sentence. This view is consistent with Hume's semantical dichotomy given that it implies that the truth value of a normative proposition does not rely upon the descriptive proposition in the scope of O. Thus, the *Ought-to-do* interpretation should be preferred to the *Ought-to-be* insofar as it preserves the syntactical structure of normative sentences within the English language and that it is consistent, contra the modal interpretation, with Hume's is-ought thesis.

Conditional obligations and detachment

In addition to revealing the *Ought-to-do* structure of normative sentences, the analysis of Forrester's paradox through the framework of the syntactic calculus allows us to provide a formal explanation to the problem that is implicit to Chisholm's paradox, i.e., that a material conditional is inappropriate to deal with conditional obligations. Indeed, one can see in both the a-translations and the b-translations that the material conditional in (3) is of the type that takes a normative statement and a declarative statement and turns it into a normative statement (a conditional obligation). Our analysis provides a formal explanation as to why the material conditional is not appropriate to translate conditional obligations: the conditional $\supset$ should not be interpreted as a connective of type $s_d \backslash s_n / s_n$. As such, Chisholm's paradox arises when a deontic conditional is interpreted as a connective of the type that takes a descriptive sentence (a context) and a normative sentence and transforms it into a normative sentence.

Many deontic logicians think that material conditionals are not sufficient to represent the formal properties of conditional obligations. We agree.[17] In a nutshell, to model conditional obligations, one requires a primitive operator that can represent situations where the initial conditions for the obligation are encountered.[18] In a possible worlds semantics, this is usually done in minimal models, where one can define the accessibility relation for the primitive operator in terms of a subset of a more general accessibility relation (see for example Chellas 1974). Another way of answering this problem is to introduce temporal deontic logics. Decew (1981) for instance argued that the solution to the conditional obligations problem requires temporal modalities. The main objection is that dyadic systems fail to represent the temporal character implicit to conditional obligations. An interesting solution was proposed by van Eck (1982a; 1982b), who introduced a quantified temporal deontic logic.

With that being said, one can see that there are (at least) two things going on with conditional obligations. First, it is clear that a material conditional such as $\supset$ is not of type $s_d \backslash s_n / s_n$, and thus neither (3a) nor (3b) are appropriate translations. This, which begs the question of the syntactical type of a conditional obligation, brings us to the next point: some might argue that there is a temporal parameter implicit to conditional obligations. Hence, using $\triangleright$ to represent conditional obligations, one could try to define the type of $\triangleright$ with the help of primitive types that are indexed to some states, allowing the

[17] As it is shown in Peterson (2014d), it is nonetheless possible to model conditional obligations with a monadic O and a connective similar to the linear implication.

[18] There are however problems regarding the detachment of conditional obligations. See Vorobej (1986) for an overview.

representation of the temporal parameter. A possible solution would be to consider the conditional $\triangleright$ as of the syntactical type that takes a descriptive proposition in the past (or the present) and makes it into a normative proposition in the present (or in some possible future related to that present). Put differently, if one considers that the world is in a specific state at a moment i, then $\triangleright$ is of the type that takes a descriptive sentence at i and a normative sentence in a state $i+1$ (accessible from i) and makes it into a normative conditional in state i. Hence, a solution is to consider that $\triangleright$ is of the type $s_d^{(i)}\backslash s_n^{(i)}/s_n^{(i+1)}$, where the set of primitive types is augmented with types that are indexed to states.

Issues concerning conditional obligations can also be seen through the problem of detachment. Following van Eck (1982a, p.263), the problem of detachment comes from the fact that although sometimes one might want to conclude the normative consequent from the conjunction of the descriptive antecedent and the normative conditional, as in the following reasoning, it remains that sometimes the addition of other conditions results in situations where one does not want to detach the normative consequent.

P1	If Jones drinks alcohol, then he must not drive.
P2	Jones drinks alcohol.
∴	Jones must not drive.

From a categorical perspective, the problem of detachment can be reduced to the fact that there is not always a proof of the normative consequent from the conjunction of the conditional obligation and the descriptive antecedent. Put differently, we do not have the following arrow for every A and B.[19]

$$f : A \wedge (A \triangleright B) \longrightarrow B$$

From a categorical point of view, the classical connectives $\wedge$, $\neg$ and $\supset$ can be defined on the basis of a deductive system (i.e., a category) where arrows are proofs and objects are propositions. In doing so, one obtains a Cartesian closed category with conjunction commutative and idempotent, respecting the triangle and pentagon identities and the associative law (with $\neg$ defined as $\neg A =_{def} A \supset \bot$ and $\bot$ an initial object). From this perspective, $\wedge$ and $\supset$ are adjoint. This can easily be seen from the fact that

$$\frac{g : A \wedge B \longrightarrow C}{g' : A \longrightarrow B \supset C}$$

$$\frac{h : A \longrightarrow B \supset C}{h' : A \wedge B \longrightarrow C}$$

Using these rules, one can easily show that:

[19] Note that we are looking at the logical connectives from a categorical perspective rather than from the point of view of the syntactic calculus.

$$\frac{1_{A\supset B} : A \supset B \longrightarrow A \supset B}{g'(1_{A\supset B}) : (A \supset B) \wedge A \longrightarrow B}$$

In other words, there is a proof from A implies B and A to B. The problem of detachment follows from the fact that it might happen that one does *not* want to conclude the normative consequent from the descriptive antecedent in a normative conditional. When modelling a conditional obligation through a material conditional, the problem of detachment arises from the fact that $\supset$ is the adjoint of $\wedge$. As such, to solve this problem, one would need to define a connective $\triangleright$ for conditional obligations which is not an adjoint to classical conjunction.[20]

All things considered, our analysis provides two arguments to support that $\supset$ is not an adequate connective to model conditional obligations. On the one hand, a material conditional is not sufficient to represent conditional obligations insofar as a connective for conditional obligations is not of the same syntactical type as a material conditional. On the other hand, the detachment problem shows that the connective for conditional obligations is not an adjoint to conjunction ($\wedge$), and hence that it cannot be represented by a material conditional.

Action negation

Returning to the syntactic calculus, the analysis of the paradox enables us to specifically formulate one of the problem that dynamic deontic logic (and more generally action logics) faces. Consider the logical connective 'not', as it is used in (7a) and (7b). In the a-translations, 'not' can be considered as of two types: either it takes a (declarative) descriptive sentence and transforms it into a descriptive sentence or it takes a (declarative) normative sentence and turns it into a normative one. This is perfectly consistent with our assumption that normative inferences are governed by logical rules and that normative sentences are declarative. In this case, negation can be applied to both descriptive and normative statement. An interesting fact is that the b-translation admits a third type for 'not'. Indeed, 'not' can be seen as something which takes an action (or an intransitive verb) and turns it into another action (or intransitive verb).

It seems fairly uncontroversial to assume that the 'nots' of types s_d/s_d and s_n/s_n behave according to the same logical rule inasmuch as they both apply to declarative statements. In other words, they are propositional negations that can be applied to declarative statements. But is this also the case for the 'not' of type a/a (or i/i)?

The work of Kanger (1957) and Pörn (1970) lead, with the help of Belnap and Perloff (1988) and Xu (1995), to a broad class of approaches that can be regrouped under the name *action logics*. Action logics in philosophy can be divided in two main groups, namely the STIT logics (where an agent *sees to it that...*) and dynamic logic, introduced in the deontic logic literature by Meyer (1988). The main difference between these approaches is that dynamic logic requires an algebra for actions and another for propositions, while STIT logics

[20] On this subject, see Peterson (2014d).

only need truth functional connectives (dynamic deontic logics are of the *Ought-to-do* type while STIT logics are of the *Ought-to-be*). However, there are still problems regarding the behaviour of logical connectives when applied to actions (or descriptions of actions in the case of STIT logics). While the tradition in deontic logic is to use Boolean algebras for actions (see for example Segerberg 1982b), it is noteworthy that there are still problems regarding the formalization of *action negation*.[21]

In short, the problem revolves around the interpretation of the negation of an action: must the negation of an action be considered as something which is simply not done or as something which is deliberately not done? This problem becomes clear when one introduces a deontic operator. There is a distinction between not doing an action and doing the negation of an action, and so $O\neg p$ will mean two completely different things depending on how one interprets the action negation $\neg p$. Does $O\neg p$ mean that one must do $\neg p$ or does it mean that one must not do p? Although it is not the aim of the present paper to argue in favour of that point (this is the subject of another paper, see Peterson 2014a), we think that a monoidal deductive system weaker than intuitionistic logic would be better suited than a Boolean algebra to represent the formal behaviour of complex actions. The main reason for this assumption is that the equivalence $a \equiv \neg\neg a$ is dubious when a is considered as an action proposition. Is doing a equivalent to refraining from refraining from doing a? If it is false that one did not do a, then can we conclude that one actually did a? In a nutshell, the negation of an action proposition does not seem to have a classical behaviour. As a result, the negation of a declarative sentence and the negation of an action proposition should not be considered as of the same syntactical type. While the 'not' of types s_d/s_d or s_n/s_n are truth functional and classical, it is unclear whether or not the 'not' of type a/a is truth functional or classical.[22]

Deontic entailment

So far, we have seen how our analysis sheds light upon Chisholm's paradox and the *Ought-to-be/Ought-to-do* distinction, and moreover it has shown us how to formulate precisely the problems of action negation and detachment of conditional obligations. In addition to these points, our analysis also enables us to see the good Samaritan paradox from a different perspective. The paradox of the good Samaritan in Forrester's argument can be seen through the step from (4) to (★).

(★)	Jones	ought	to murder	Smith	gently	implies
	n	$n\backslash s_n/i$	i/n	n	$i\backslash i$	$s_n\backslash s_n/s_n$
	Jones	ought	to murder	Smith		
	n	$n\backslash s_n/i$	i/n	n		

Although this is not mentioned in Forrester (1984), the step from (4) to (★) is an

[21] See Broersen (2004) for a definition of action negation within the framework of dynamic deontic logic.
[22] See Peterson (2014a) for a discussion.

instance of the principle of deontic consequences, as stated in Castañeda (1968, p.13):

> If an act A entails an act B, then (1) the obligatoriness of A entails the obligatoriness of B and (2) the forbiddenness of B entails the forbiddenness of A.

Clearly, if Jones ought to murder Smith in a gentle fashion, then obviously he also ought to murder Smith. We argued elsewhere (cf. Peterson and Marquis 2012) that, contra Forrester, his paradox is not problematic at all since despite its validity, the argument is not sound. But still, one could wonder whether there is something that can be done via the syntactical types to solve that problem. La Palme Reyes et al. (1994) used category theory to analyze the relations between types, nouns and properties. One of their example shows how some properties cannot be passed between two nouns, although these two nouns are related somehow. For instance, although a baby is a person, a big baby is not a big person. So perhaps the same kind of phenomenon is at work here. Although a gentle murder is a murder, it does not necessarily follows that a gentle murder which is obligatory is also a murder which is obligatory. The question would thus be to determine how one can represent this phenomenon through the syntactic calculus. One way of answering this problem is to fix a primitive *class* of types $\mathcal{A} = \{a_1, ..., a_n\}$ (replacing i) which would be composed of a finite number of types of actions. It is thus possible to obtain $(\star)$.[23]

$(\star)$	Jones	ought	to murder	Smith	gently
	n	$n\backslash s_n/a_2$	a_1/n	n	$a_1\backslash a_2$

On these grounds, one could then argue that the step from (4) to (★) requires that both actions in the scope of 'ought' be of the same type, which is not the case in Forrester's paradox. Indeed, the action of 'murdering Smith' is of type a_1 while 'murdering Smith gently' is of type a_2. Therefore, the step from (4) to (★) is illegitimate insofar as the property 'obligatory' cannot be passed between actions that are not of the same type.

The same strategy can be applied to the good Samaritan paradox, which goes from (10) to (11).[24]

(10) The good Samaritan binds the traveler's wounds implies
the traveler is wounded

(11) The good Samaritan ought to bind the traveler's wounds implies
the traveler ought to be wounded

In this case, it is easy to argue that the action 'binding the traveler's wound' is not of the same type as 'the traveler is wounded'. Indeed, the action of being wounded only concerns the traveler while the action of binding the traveler's wound is performed by the good Samaritan. Hence, this use of (ROM) is illegitimate: in order to be correct, its

[23] Here, we are not following the solution proposed by La Palme Reyes et al. (1994).

[24] This version is taken from Garson (2006, p.46). Note, however, that thus formulated it is not derivable neither in K or KD. That said, the formulation of Åqvist (1967) can be derived.

use must be restricted to actions that are of the same type in both the antecedent and the consequent of the conditional. As such, the application of (ROM) requires that the actions in the scope of the deontic operator are of the same syntactic type, otherwise the application is illegitimate.

Propositions and actions

Finally, it is worth mentioning that our approach implicitly incorporates Castañeda's (1986) distinction between propositions and practitions. In showing that the syntactical type of the 'ought' operator in the English language behaves according to the *Ought-to-do* interpretation, our approach implies a distinction between declarative statements and action statements. Since the syntactical type of 'ought' is $n\backslash s_n/a_i$ rather than s_n/s_d, it follows that Castañeda's distinction is implicit to our framework. However, there is more to our approach than that. Indeed, in addition to distinguishing between declarative statements and action statements, our approach also allows us to distinguish between different types of declarative statements, namely descriptive propositions of type s_d and normative propositions of type s_n. To the best of our knowledge, this is something that has not been previously done within the literature. Thus, our framework incorporates Castañeda's distinction and goes beyond: it also incorporates Hume's semantical dichotomy between facts and norms, which is important if one wants to apply logic to normative inferences. Furthermore, although Castañeda distinguishes between propositions and practitions, he still uses the same logical rules to govern both, and as we saw this is problematic if we consider action negation and conditional obligations.

Conclusion

Summing up, the contribution of the present paper was to open the dialog between three disciplines that were otherwise blind to each other. Although the bridge between category theory and linguistics was made by Lambek and further developed by Moot and Retoré (2012)[25], category theory and categorial grammar remained unrelated to deontic logic. The present paper contributes to the literature given that it provides new conceptual insights on many important features of deontic logic. First, we showed how one can formally explain the problems regarding the *Ought-to-be/Ought-to-do* distinction, the connective for conditional obligations and action negation. In a nutshell, these issues arise since they involve different syntactical types. We showed that the *Ought-to-do* interpretation of normative sentences is a more faithful representation of the English language, and thus that it is the one that should be adopted. Moreover, we showed how our approach implicitly involves Castañeda's distinction between propositions and practitions, but goes further insofar as it distinguishes between different syntactical types for declarative statements. Incidentally, it enables us to give a formal representation of Hume's semantical dichotomy.

[25] The use of deductive systems in categorial grammar is actually an active field of study and developments were made by many authors. See Moortgat (2012) for an overview.

Our framework also provided a different perspective on Forrester's and the good Samaritan paradoxes, showing that they only arise when (ROM) is applied to a conditional that contains different types of actions. Hence, the analysis of deontic logic via Lambek's syntactic calculus shed light upon many issues that it faces.

As a result of our analysis, we can see that the problems concerning conditional obligations and action negation could benefit from category theory. For instance, we saw that the detachment problem happens when the connective for conditional obligations is considered as an adjoint to conjunction. Although the main aim of this paper was to expose that deontic logic can benefit from Lambek's type theory, we also saw a glimpse of how deontic logic could benefit from a categorical analysis. As such, it opens the door to the application of category theory to deontic logic. Since categorical logic is in the first place a theory of types, our analysis suggests that deontic logic could benefit from the formal framework of category theory. For future research, we intend to provide a categorical analysis of deontic and action logics in order to precisely define the categories within which their connectives behave. Furthermore, we intend to analyze the problems regarding conditional normative inferences through the framework of categorical logic. Now that the dialog between deontic logic and category theory is open, we face a broad, rich and interesting research avenue for deontic logic.

* * *

Cet article montrait la pertinence d'analyser la logique déontique selon une théorie des types. En plus de nous permettre d'introduire la notion de système déductif, qui nous sera utile tout au long des prochains chapitres, nous avons aussi vu comment une logique typée peut résoudre certains problèmes et paradoxes de la logique déontique. Pour l'heure, nous allons laisser de côté la théorie des types de Lambek jusqu'au chapitre 16 afin de porter notre attention sur les deuxième et troisième objectifs de cette thèse. Nous allons d'abord traiter des différentes logiques de l'action du point de vue de la théorie des catégories, et ensuite nous analyserons les inférences normatives conditionnelles et les conflits d'obligations. La notion de système déductif sera reprise tout au long des prochains chapitres et notre cadre d'analyse de la logique selon la théorie des catégories sera explicité.

Chapitre 14

A categorical analysis of action logics

Abstract

The present paper provides a categorical analysis of action logics using the proof-theoretical perspective of category theory. Action logics are compared through the properties of their logical connectives and are classified as different variations of monoidal deductive systems. A philosophical analysis of action connectives is provided and we propose, according to the usual distinction between *actions* and *propositions* in dynamic logic, to distinguish between an Action Logic, representing the formal structure of actions, and a Propositional Action Logic, expressing the formal structure of the language we use to talk about actions. Our conclusion is that an Action Logic should be defined as a compact symmetric monoidal closed deductive system with sequence while a Propositional Action Logic should be defined as an intuitionistic deductive system.

Keywords: Categorical logic, Deductive system, Monoidal category, Intuitionistic logic, Action negation, Linear logic, Kleene algebra, Boolean algebra

Introduction

The concept of *action* is of obvious philosophical significance given its place in many disciplines, including law, ethics, psychology, computer science, physics and metaphysics. As propositional logic and predicate logic are meant to represent the formal structure of sentences, philosophers wondered whether it was possible to develop some kind of *logic of actions* that could represent the formal structure of actions. *Action logics* began with the work of Kanger (1957) and Pörn (1970), but the notion of action in deontic contexts was also studied at length by von Wright (1963). Following Segerberg et al. (2009, p.1), *action logics* in philosophy can roughly be divided in two broad classes. On the one hand, the work of Kanger (1957) and Pörn (1970) lead, with the help of Belnap and Perloff (1988) and Xu (1995), to what is nowadays known as STIT logics. Dynamic logic on the other hand was introduced by Pratt (for an introduction and a historical sketch, see Balbiani 2008) and was introduced in deontic contexts by Meyer (1988). Common to both dynamic and STIT logics is the fact that these frameworks have often been presented as tools that can help to analyze and clarify some notions that are relevant to the legal and the ethical discourses, either from a philosophical or a computational point of view.

The aim of the present paper is to provide a categorical analysis of action logics used in deontic contexts. Category theory will be used as a foundational framework to analyze the proof theory of action logics. This analysis will be made from a philosophical standpoint rather than a computational one. Our main objective is to develop an action logic that is relevant to human actions, and this is not, as Segerberg (1992, p.377) pointed out, the aim of the logics that are used for programming. From an epistemological perspective, category theory offers a rich and insightful foundational framework that allows one to see how different logics are identical up to equivalence.[1] In this paper, our aim will first be to show how different fragments of action logics can be compared insofar as they are different variations of monoidal categories, or, more accurately in our terms, different variations of monoidal deductive systems. Then, we propose a new action logic $\mathcal{AL}$, defined as a compact symmetric monoidal closed deductive system with sequence, but where the adjoint to conjunction is not understood as entailment, contra the usual action logics one will find within the literature.

In the first section, we begin by presenting the categorical framework, exploiting the intimate relations between monodial categories and deductive systems. This is followed by a review of the main action logics one can find within the literature, including dynamic logic, Kleene algebras, Pratt's action logic, linear logic, Boolean algebras of actions, STIT logics and Lucas's formalization of von Wright's conception of actions. We will bring to light the relations between these logics by using the properties of their action connectives and show how they are different variations of monoidal deductive systems. Then, we provide a philosophical analysis of action connectives and propose to distinguish between an Action Logic, which represents the structure of actions, and a Propositional Action Logic, which expresses the propositional structure of the language we use to talk about actions. A Propositional Action Logic $\mathcal{PAL}$ will be defined as an intuitionistic deductive system, while an Action Logic $\mathcal{AL}$ will be defined as a compact one.

Logic categorically conceived

From a categorical point of view, logical systems can be defined through the properties of their connectives. The starting point of our analysis is the proof-theoretical perspective of category theory, where deductive systems are defined as monoidal categories that respect specific criteria. Following Mac Lane (1971, pp.7-8), a category $\mathcal{C}$ is composed of

1. $\mathcal{C}$-*objects*;

[1] When deductive systems are presented as categories, the criterion of identity is *equivalence* between deductive systems, as opposed to *isomorphism* between propositions. While equivalence is the criterion of identity *between* categories, isomorphism is the criterion of identity *within* a category. For arguments in favour of the epistemological virtues of category theory, see Marquis (2009).

2. $\mathcal{C}$-*arrows*;

3. an operation assigning to each arrow a domain and a codomain within the $\mathcal{C}$-objects;

4. composition of arrow gf for each pair $f : x \longrightarrow y$ and $g : y \longrightarrow z$ that respects associativity, i.e. $h(gf) = (hg)f$;

5. an identity arrow 1_y for each $\mathcal{C}$-object y such that $1_y f = f$ and $g1_y = g$ for each pair $f : x \longrightarrow y$ and $g : y \longrightarrow z$ (the identity law).

The definition of a deductive system is intimately related to the definition of a category. Indeed, one simply has to see that the notion of *proof* (or *deduction*) used in logical systems possesses two specific properties, namely the reflexivity and transitivity of the consequence relation. From φ one can prove φ, and if from φ one can prove ψ and from ψ one can prove ρ, then from φ one can prove ρ. Hence, a *deductive system* can be defined as a category $\mathcal{C}$, where objects are propositions and arrows are proofs. An arrow $f : \varphi \longrightarrow \psi$ represents a proof (a deduction) of ψ from φ. If there is such an arrow $f : \varphi \longrightarrow \psi$, then ψ is a logical consequence of φ within the deductive system. The reflexivity of the consequence relation is captured via the identity arrow (1) for each formula φ:

$$\frac{}{1_\varphi : \varphi \longrightarrow \varphi} \qquad (1)$$

The identity arrow is an axiom (schema) of the deductive system and it respects the identity law. The transitivity of the consequence relation is represented by the rule (cut) and composition of arrows respects associativity:

$$\frac{f : \varphi \longrightarrow \psi \qquad g : \psi \longrightarrow \rho}{gf : \varphi \longrightarrow \rho} \qquad \text{(cut)}$$

So far, only the consequence relation is defined and there are no logical connectives. But one can easily add some via the definition of a monoidal category. Following Mac Lane (1971, pp.161-2), a *monoidal category* is a category equipped with a tensor product $\otimes$ and a unit object I. The tensor product is associative, meaning that there is an arrow $a_{x,y,z}$ for each $\mathcal{C}$-objects x, y and z, and the unit object respects l_x and r_x for each $\mathcal{C}$-objects x:

$$\begin{aligned}
&a_{x,y,z} : (x \otimes y) \otimes z \longrightarrow x \otimes (y \otimes z) \\
&l_x : I \otimes x \longrightarrow x \\
&r_x : x \otimes I \longrightarrow x
\end{aligned}$$

The arrows $\alpha_{x,y,z}$, l_x and r_x are natural isomorphisms.[2] Moreover, the tensor product and the unit object have to respect the triangle and pentagon identities, expressed by the two following commutative diagrams:

[2] Hence we also have $a^{-1}_{x,y,z} : x \otimes (y \otimes z) \longrightarrow (x \otimes y) \otimes z$, $l^{-1}_x : x \longrightarrow 1 \otimes x$ and $r^{-1}_x : x \longrightarrow x \otimes 1$. See also the nlab entry on monoidal category http://ncatlab.org/nlab/show/monoidal+category.

$$
\begin{tikzcd}
(x \otimes I) \otimes y \arrow[rr, "a_{x,I,y}"] \arrow[dr] & & x \otimes (I \otimes y) \arrow[dl] \\
& x \otimes y &
\end{tikzcd}
$$

$$
\begin{tikzcd}
& (w \otimes x) \otimes (y \otimes z) \arrow[r] & w \otimes (x \otimes (y \otimes z)) \\
((w \otimes x) \otimes y) \otimes z \arrow[ur] \arrow[dr] & & \\
& (w \otimes (x \otimes y)) \otimes z \arrow[r] & w \otimes ((x \otimes y) \otimes z) \arrow[uu]
\end{tikzcd}
$$

As a category can be represented via the logical rules and axioms of a deductive system, a monoidal category can be expressed by a deductive system in additions with some rules to govern the tensor product and the unit object. Following Baez and Stay (2011, p.140), these rules are (t), (a), (l) and (r). The double line means that the rule can be applied both ways (with a, l and r natural transformations and the tensor product respecting the triangle and pentagon identities).

$$
\frac{f : \varphi \longrightarrow \psi \qquad g : \rho \longrightarrow \tau}{h : \varphi \otimes \rho \longrightarrow \psi \otimes \tau} \qquad \text{(t)}
$$

$$
\begin{array}{c} f : \tau \longrightarrow (\varphi \otimes \psi) \otimes \rho \\ \hline\hline g : \tau \longrightarrow \varphi \otimes (\psi \otimes \rho) \end{array} \qquad \text{(a)}
$$

$$
\begin{array}{c} f : \varphi \longrightarrow I \otimes \psi \\ \hline\hline g : \varphi \longrightarrow \psi \end{array} \qquad \text{(l)}
$$

$$
\begin{array}{c} f : \varphi \longrightarrow \psi \otimes I \\ \hline\hline g : \varphi \longrightarrow \psi \end{array} \qquad \text{(r)}
$$

In addition to the associativity of the tensor product and the properties of the unit object, one can augment a monoidal category with a commutative tensor product. A *braided monoidal category* is a monoidal category obtained through the addition of a natural isomorphism $\beta_{x,y}$, with the tensor product respecting the two hexagon identities (cf. Mac Lane 1971, pp.251-3).[3]

$$
\beta_{x,y} : x \otimes y \longrightarrow y \otimes x
$$

[3] See also Baez and Stay (2011, p.115).

$$
\begin{tikzcd}
x \otimes (y \otimes z) \arrow[r] \arrow[d] & (x \otimes y) \otimes z \arrow[r] & (y \otimes x) \otimes z \arrow[d] \\
(y \otimes z) \otimes x & y \otimes (z \otimes x) \arrow[l] & y \otimes (x \otimes z) \arrow[l]
\end{tikzcd}
$$

$$
\begin{tikzcd}
(x \otimes y) \otimes z \arrow[r] \arrow[d] & x \otimes (y \otimes z) \arrow[r] & x \otimes (z \otimes y) \arrow[d] \\
z \otimes (x \otimes y) & (z \otimes x) \otimes y \arrow[l] & (x \otimes z) \otimes y \arrow[l]
\end{tikzcd}
$$

The tensor product in a braided monoidal category is commutative up to isomorphism. A *symmetric monoidal category* is a braided monoidal category which satisfies $\beta_{y,x}\beta_{x,y} = 1_{x\otimes y}$, and hence the tensor product is commutative (in which case one of the hexagon implies the other, see Mac Lane 1971, p.253).

Following Baez and Stay (2011, p.130), a braided monoidal category in terms of deductive systems is obtained when one adds the rule (b) to the rules for a monoidal category (with (b) a natural transformation satisfying the hexagon identities).

$$
\frac{f : \varphi \longrightarrow \psi \otimes \tau}{g : \varphi \longrightarrow \tau \otimes \psi} \qquad \text{(b)}
$$

In order to obtain a symmetric monoidal category, one simply has to assume that (b) is its own inverse, hence $\beta_{\psi,\varphi}\beta_{\varphi,\psi} \cong 1_{\varphi\otimes\psi}$.

When the unit object is terminal and binary products are defined for each $\mathcal{C}$-objects x and y, a symmetric monoidal category becomes a Cartesian monoidal category. Put differently, a monoidal category is Cartesian if the product respects the universal property defined by the following commutative diagram.

$$
\begin{tikzcd}
 & z \arrow[dl, "f"'] \arrow[d, "\langle f,g \rangle"] \arrow[dr, "g"] & \\
x & x \otimes y \arrow[l, "pr_x"] \arrow[r, "pr_y"'] & y
\end{tikzcd}
$$

This diagram must be interpreted as follows: for all f and g such that $f : z \longrightarrow x$ and $g : z \longrightarrow y$ there is one and only one arrow $\langle f, g\rangle : z \longrightarrow x \otimes y$ making the diagram commute.

From a proof-theoretical point of view, a Cartesian category is obtained by adding the axiom schema (!) and the rule (Cart) to those of a monoidal category. Actually, since (b), (t), (a), (l) and (r) can be proven from (Cart) and (!), it can be obtained by adding them to (1) and (cut).

$$\frac{}{!_\varphi : \varphi \longrightarrow I} \quad (!)$$

$$\frac{f : \varphi \longrightarrow \psi \qquad g : \varphi \longrightarrow \rho}{\langle f, g \rangle : \varphi \longrightarrow \psi \otimes \rho} \quad \text{(Cart)}$$

An interesting property of a Cartesian category is that it satisfies idempotence for products. An operation $\circ$ is said to be *idempotent* when $x \circ x \cong x$. One can see that product satisfies idempotence insofar as two special cases of (Cart) are:

$$\frac{1_\varphi : \varphi \longrightarrow \varphi \qquad 1_\varphi : \varphi \longrightarrow \varphi}{\langle 1_\varphi, 1_\varphi \rangle : \varphi \longrightarrow \varphi \otimes \varphi}$$

$$\frac{pr_\varphi : \varphi \otimes \varphi \longrightarrow \varphi \qquad pr_\varphi : \varphi \otimes \varphi \longrightarrow \varphi}{1_{\varphi \otimes \varphi} : \varphi \otimes \varphi \longrightarrow \varphi \otimes \varphi}$$

Therefore, $\varphi \cong \varphi \otimes \varphi$. Note that in the first case the rule is applied top-down while in the second one it is applied bottom-up.

So far, we have seen how one can build a logical system as a monoidal category by adding some rules to represent the properties of the tensor product. That being said, each of these categories can be transformed into a closed category by adding the appropriate adjoint connective. A *closed monoidal category* is a monoidal category for which the tensor product has right adjoint. Note, however, that when the tensor is not commutative (i.e., when $\mathcal{C}$ is not symmetric), the tensor can have two different right adjoints. Following Baez and Stay (2011, p.120), a monoidal category is said to be *left closed* when there is an isomorphism between the class of morphisms (of $\mathcal{C}$) from $x \otimes y$ to z and the class of morphisms from x to $[y, z]$, and it is *right closed* when there is an isomorphism between the class of morphisms from $x \otimes y$ to z and the class of morphisms from y to $[x, z]$.[4] In other words, a monoidal category is left closed when:

$$\mathrm{Hom}_{\mathcal{C}}(x \otimes y, z) \cong \mathrm{Hom}_{\mathcal{C}}(x, [y, z])$$

As such, for each arrow $f : x \otimes y \longrightarrow z$ there is an arrow $g : x \longrightarrow [y, z]$, and vice versa. Similarly, a monoidal category is right closed when:

$$\mathrm{Hom}_{\mathcal{C}}(x \otimes y, z) \cong \mathrm{Hom}_{\mathcal{C}}(y, [x, z])$$

Following the nlab entry on closed monoidal categories, if $\mathcal{C}$ is symmetric, then both adjoints are the same since $x \otimes y \cong y \otimes x$. However, if $\mathcal{C}$ is not symmetric, then $x \otimes y$ is

[4] See also http://ncatlab.org/nlab/show/closed+monoidal+category.

not the same functor as $y \otimes x$, and hence each can have a right adjoint. In such a case, the monoidal category can be left closed, right closed, or biclosed.

Identifying the internal Hom $[\varphi, \psi]$ with $\varphi \rhd \psi$, one obtains a (left) closed monoidal category by adding the rule (cl).

$$\begin{array}{c} f : \varphi \otimes \psi \longrightarrow \rho \\ \hline\hline h : \varphi \longrightarrow \psi \rhd \rho \end{array} \qquad (\text{cl})$$

If $\mathcal{C}$ is not symmetric and one also wants it to be right closed, then one simply has to add the rule (cl').

$$\begin{array}{c} f : \varphi \otimes \psi \longrightarrow \rho \\ \hline\hline h : \psi \longrightarrow \varphi \blacktriangleright \rho \end{array} \qquad (\text{cl'})$$

At this point, we have only considered monoidal categories with a unit object, a tensor product and its adjoint(s). That being said, one can easily introduce dual objects on the ground of a closed monoidal deductive system by defining $x^* =_{def} x \rhd *$, in which case dual objects satisfy $*_x$.

$$*_x : x \longrightarrow x^{**}$$

In the case of deductive systems, the morphism $*_\varphi : \varphi \longrightarrow \varphi^{**}$ follows from:

$$(\text{b})\ \dfrac{\dfrac{\dfrac{1_{\varphi\otimes\varphi^*} : \varphi \otimes \varphi^* \longrightarrow \varphi \otimes \varphi^*}{b(1_{\varphi\otimes\varphi^*}) : \varphi \otimes \varphi^* \longrightarrow \varphi^* \otimes \varphi} \qquad \dfrac{\dfrac{1_{\varphi^*} : \varphi^* \longrightarrow \varphi^*}{1_{\varphi^*} : \varphi^* \longrightarrow \varphi \rhd *}\,(\text{def } *)}{cl(1_{\varphi^*}) : \varphi^* \otimes \varphi \longrightarrow *}\,(\text{cl})}{\dfrac{cl(1_{\varphi^*})b(1_{\varphi\otimes\varphi^*}) : \varphi \otimes \varphi^* \longrightarrow *}{cl(cl(1_{\varphi^*})b(1_{\varphi\otimes\varphi^*})) : \varphi \longrightarrow \varphi^* \rhd *}\,(\text{cl})}\,(\text{cut})}{*_\varphi : \varphi \longrightarrow \varphi^{**}}\,(\text{def } *)$$

Moreover, if $*$ is a dualizing object, then $*_x$ becomes a natural isomorphism and one gets $x \cong x^{**}$.[5]

Now, let us return to the definition of a symmetric monoidal category. Following Blute and Scott (2004, p.24), if one adds a dualizing object $*$ and defines the dual object by $x^* =_{def} x \rhd *$, then one obtains a *-autonomous category, satisfying the two following isomorphisms (cf. Barr 1979).

$$x \rhd y \cong y^* \rhd x^*$$
$$x \otimes y \cong (x \rhd y^*)^*$$

If one goes further and adds the following natural isomorphism to a *-autonomous category, then one gets a compact symmetric monoidal closed category (cf. Barr 1991).

$$x \rhd y \cong x^* \otimes y$$

[5] Following Barr (1991, p.161), $*$ is a dualizing object specifically when $x \cong x^{**}$ for each x.

Following lemma 2.18 of Blute and Scott (2004, p.28), compact symmetric monoidal closed categories are *-autonomous categories where the unit is a dualizing object. Note, however, that the converse is not true: a *-autonomous category where the unit is a dualizing object is not necessarily a compact closed category.[6] From the point of view of deductive systems, one obtains the tensor's fragment of a *-autonomous category by defining $x^* =_{def} x \triangleright *$ and by adding the axiom (**), which is a natural isomorphism, to the rules and axioms for a symmetric monoidal closed deductive system.

$$\frac{}{**_\varphi : \varphi^{**} \longrightarrow \varphi} \qquad (**)$$

A compact symmetric monoidal closed deductive system is obtained by the addition of the following axioms.

$$\frac{}{cpt1_{\varphi,\psi} : \varphi \triangleright \psi \longrightarrow \varphi^* \otimes \psi} \qquad \text{(cpt1)}$$

$$\frac{}{cpt2_{\varphi,\psi} : \varphi^* \otimes \psi \longrightarrow \varphi \triangleright \psi} \qquad \text{(cpt2)}$$

Returning to monoidal categories, one may want to consider a larger set of connective through duality. Having a category $\mathcal{C}$, one can define the dual (opposite) category $\mathcal{C}^{op}$ of a category $\mathcal{C}$, where the objects of $\mathcal{C}^{op}$ are the objects of $\mathcal{C}$ but where the arrows are reversed (see Mac Lane 1971, p.33). In the case of Cartesian categories, the co-tensor is defined by the universal property of the co-product. A *co-Cartesian category* $\mathcal{C}^{op}$ is obtained by defining a coproduct $\oslash$ via the universal property expressed by the following commutative diagram and by adding an initial object 0.

$$\begin{tikzcd} x \arrow[r, "i_x"] \arrow[dr, "f"'] & x \oslash y \arrow[d, "{[f,g]}"] & y \arrow[l, "i_y"'] \arrow[dl, "g"] \\ & z & \end{tikzcd}$$

The universal property of the coproduct means that for all $f : x \longrightarrow z$ and $g : y \longrightarrow z$ there is one and only one arrow $x \oslash y \longrightarrow z$ making the above diagram commute. In this case, one can add the rule (co-Cart) and the axiom schema (0) to a Cartesian category. In this case, we will speak of a *bi-Cartesian closed deductive system.*

$$\frac{f : \varphi \longrightarrow \rho \qquad g : \psi \longrightarrow \rho}{[f, g] : \varphi \oslash \psi \longrightarrow \rho} \qquad \text{(co-Cart)}$$

$$\frac{}{0_\varphi : 0 \longrightarrow \varphi} \qquad (0)$$

[6] Thanks to Micheal Barr for this clarification.

An interesting case of (co-Cart) is that similarly to (Cart), it satisfies idempotence for coproduct. Indeed, we have:

$$\frac{1_\varphi : \varphi \longrightarrow \varphi \qquad 1_\varphi : \varphi \longrightarrow \varphi}{[1_\varphi, 1_\varphi] : \varphi \oslash \varphi \longrightarrow \varphi}$$

$$\frac{i_\varphi : \varphi \longrightarrow \varphi \oslash \varphi \qquad i_\varphi : \varphi \longrightarrow \varphi \oslash \varphi}{1_{\varphi \oslash \varphi} : \varphi \oslash \varphi \longrightarrow \varphi \oslash \varphi}$$

Hence $\varphi \cong \varphi \oslash \varphi$.

If one adds duals to a bi-Cartesian closed deductive system by defining $x^* =_{def} x \triangleright *$, then one will obtain the proof theory of intuitionistic logic, with the symbols $\{\otimes, \triangleright, I, \oslash, 0\}$ corresponding respectively to $\{\wedge, \Rightarrow, \top, \vee, \perp\}$.[7] Classical propositional logic is obtained on the ground of a bi-Cartesian closed deductive system (with dualizing object) via the addition of the axiom (**), which satisfies the natural isomorphism $x \cong x^{**}$. Hence, from a proof-theoretical perspective, both intuitionistic logic and classical logic take place within a bi-Cartesian closed deductive system, the latter obtained with the addition of involutive negation.

We conclude this section with a brief recapitulation of what was presented so far. Figure 14.1, which is inspired by Baez's and Stay's (2011, p.99) presentation, shows the relations between the tensor's fragment of the deductive systems. From a proof-theoretical standpoint, logical systems can be defined as monoidal categories that satisfy specific rules and axioms, depending on the properties of their connectives. Table 14.1 provides a list of the main deductive systems that were presented using the tensor's fragment $\{\otimes, \triangleright, I\}$. The upshot of this mode of presentation is that the categorical framework enables us to see that different logical systems can be built upon weaker systems by focusing on specific properties of the logical connectives. Logical systems built in the opposite categories from the co-tensor's fragment are obtained by reversing the arrows, in which case one obtains a co-monoidal deductive system, a co-symmetric deductive system, a co-Cartesian deductive system, etc. If one wants to consider a full set of connectives with both $\mathcal{C}$ and $\mathcal{C}^{op}$, then one gets a bi-monoidal deductive system, a bi-symmetric deductive system, a bi-Cartesian deductive system, etc. According to our appellation, a co-symmetric Cartesian deductive system for example would refer to a deductive system in which the co-tensor is axiomatized via a co-symmetric deductive system and the tensor is axiomatized via a Cartesian deductive system.

Our analysis is based upon the intimate relations that can be found between categories and deductive systems. On the one hand, it is possible to construct a category which will have specific properties from a deductive system within which the consequence relation is reflexive and transitive and where the tensor respects some specific rules. This way of thinking was advocated namely by Lambek, who explicitly defined categorical notions on the grounds of deductive systems (see for instance Lambek 1968, 1969). On the

[7] See Blute and Scott (2004, p.21).

other hand, it is also possible to define a deductive system as a category (as opposed to defining a category as a deductive system), where the tensor will have some specific properties depending on the characteristics of the category. Before introducing some terminology, let us clarify our epistemic standpoint. It is our view that category theory is somewhat epistemically superior to logic (for arguments in favour of this point, we refer the reader to Marquis 2009). Indeed, category theory provides a rich, powerful, but more importantly *abstract* framework that can be used to analyze a variety of domains, including logic. Hence, reducing a category to a deductive system would be too restrictive. As such, although the terminology we introduce and the presentation we made suggest a reduction of categories to deductive systems (or a conceptual equivalence between these two notions), let us emphasize that we understand a deductive system as a particular instance of some category. Put differently, despite the fact that one can find a conceptual equivalence between categories and deductive systems within the literature, it is our view that deductive systems should be defined as categories that respect specific criteria (rather than the converse). Keeping these remarks in mind, we propose the following terminology.

Terminology 1. *A* Classical deductive system *is a bi-Cartesian closed deductive system with dualizing object.*

Terminology 2. *An* Intuitionistic deductive system *is a bi-Cartesian closed deductive system with negation.*

Terminology 3. *A* *-autonomous deductive system *is a bi-symmetric closed deductive system with a dualizing object.*

Terminology 4. *A* compact deductive system *is a compact symmetric monoidal closed deductive system such that* $\varphi \triangleright \psi \cong \varphi^* \otimes \psi$.

Terminology 5. *A* distributive deductive system *is a deductive system where the tensor distributes over the co-tensor and* $\alpha \otimes 0 \cong 0 \otimes \alpha \cong 0$.[8]

[8] Not to be confused with a *distributive category* (see the nlab entry http://ncatlab.org/nlab/show/distributive+category for the latter).

$$
\begin{tikzcd}
\textbf{Deductive system} \arrow[d] & & & & \\
\textbf{Monoidal} \arrow[r] \arrow[d] & \textbf{Monoidal closed} \arrow[r] \arrow[d] & \begin{array}{c}\textbf{Monoidal closed}\\ \textbf{with negation}\end{array} \arrow[r] \arrow[d] & \begin{array}{c}\textbf{Monoidal closed}\\ \textbf{with dualizing object}\end{array} \arrow[d] & \\
\textbf{Symmetric} \arrow[r] \arrow[d] & \textbf{Symmetric closed} \arrow[r] \arrow[d] & \begin{array}{c}\textbf{Symmetric closed}\\ \textbf{with negation}\end{array} \arrow[r] \arrow[d] & \begin{array}{c}\textbf{Symmetric closed}\\ \textbf{with dualizing object}\end{array} \arrow[r] \arrow[d] & \textbf{Compact} \\
\textbf{Cartesian} \arrow[r] & \textbf{Cartesian closed} \arrow[r] & \begin{array}{c}\textbf{Cartesian closed}\\ \textbf{with negation}\end{array} \arrow[r] & \begin{array}{c}\textbf{Cartesian closed}\\ \textbf{with dualizing object}\end{array} &
\end{tikzcd}
$$

Figure 14.1: Relations between deductive systems

Logical systems	Logical rules and axioms
Deductive system reflexivity and transitivity of deduction	(1), (cut)
Monoidal deductive system associativity of $\otimes$, absorption of $\top$	(1), (cut) (t), (a), (l), (r)
Monoidal closed deductive system $\otimes$ possesses adjoint(s)	(1), (cut), (t), (a), (l), (r) (cl), (cl')
Symmetric deductive system commutativity of $\otimes$	(1), (cut), (t), (a), (l), (r) (b)
Symmetric closed deductive system only one adjoint	(1), (cut), (t), (a), (l), (r) (b), (cl)
Symmetric closed deductive system with dualizing object	(1), (cut), (t), (a), (l), (r) (b), (cl), (**)
Compact deductive system $\varphi \triangleright \psi \cong \varphi^* \otimes \psi$	(1), (cut), (t), (a), (l), (r) (b), (cl), (**), (cpt1), (cpt2)
Cartesian deductive system idempotence, $\top$ is terminal	(1), (cut), (Cart), (!)
Cartesian closed deductive system with negation $x^* =_{def} x \triangleright *$	(1), (cut), (Cart), (!), (cl)
Cartesian closed deductive system with dualizing object	(1), (cut), (Cart), (!), (cl), (**)

Table 14.1: Deductive systems with their corresponding rules and axioms

A survey of action logics

Now turning our attention to the action logic literature, our aim is to illustrate how different approaches can be compared through the properties of their tensor product. We will briefly present some of the main approaches one can find within the literature and show how each can be viewed as a variation of a monoidal deductive system. We begin by presenting dynamic logic and Kleene algebras, which will be followed by Pratt's action logic, linear logic, Boolean algebras of actions, STIT logics and Lucas's formalization of von Wright's conception of actions. We conclude this section with a comparison of these systems.

Dynamic logic

Dynamic logic was developed by Vaughan Pratt (cf. Segerberg 1992, p.375) and was meant to represent the evolution of a scenario after the execution of some *actions*, or, in Pratt's case, after the execution of some computer *program* (see Pratt 1976, 1980). The main characteristic of dynamic logic is that it distinguishes between *propositions* (or *assertions*) and *actions* (or *programs*). It began with the introduction of a modal operator $[\alpha]\varphi$, which is read *after the execution of* α, φ *holds*, and complex actions were composed from the connectives $\cup$, ; and *, representing respectively the choice action, the sequence of action and iteration of action (for an introduction, see Harel et al. 2000). The set of well-formed formulas is defined from the language $\mathcal{L} = \{Prop, Act, (,), \cup, ;, ^*, \neg, \supset, [\,]\}$ (with $Prop$ and Act denumerable sets of atomic propositions and actions) by:

$$\varphi := \ p_i \mid \neg\varphi \mid \varphi \supset \psi \mid [\alpha]\varphi$$

The set of complex actions $Act*$ defined by (see Harel et al. 2000, p.165):

$$\alpha := \ a_i \mid \alpha \cup \beta \mid \alpha;\beta \mid \alpha^*$$

The $[\,]$ modality is axiomatized as a normal $\Box$ modality within the modal system K. It is known that the action structure inherent to dynamic logic can be expressed by a Kleene algebra (Harel et al. 2000, p.389). A *Kleene algebra* $\mathcal{KA}* = (K, \cup, ;, 0, 1, *)$ is an idempotent semiring which satisfies the three following equations (Harel et al. 2000, p.419).

$$1 \cup (\alpha;\alpha^*) = 1 \cup (\alpha^*;\alpha) = \alpha^* \tag{K*1}$$

$$\text{if } \beta \cup (\alpha;\gamma) \leq \gamma, \text{ then } \alpha^*;\beta \leq \gamma \tag{K*2}$$

$$\text{if } \beta \cup (\gamma;\alpha) \leq \gamma, \text{ then } \beta;\alpha^* \leq \gamma \tag{K*3}$$

The definition of $\alpha \leq \beta$ is $\alpha \leq \beta =_{def} \alpha \cup \beta = \beta$. Concentrating only on the fragment for choice and sequence, let $\mathcal{KA} = (K, \cup, ;, 0, 1)$. The logic of programs implicit to propositional dynamic logic can be expressed via an idempotent semiring. Indeed, a $\mathcal{KA}$ has to satisfy the axioms:

$$\alpha \cup (\beta \cup \gamma) = (\alpha \cup \beta) \cup \gamma \tag{K1}$$
$$\alpha \cup \beta = \beta \cup \alpha \tag{K2}$$
$$\alpha \cup 0 = \alpha = \alpha \cup \alpha \tag{K3}$$
$$\alpha; (\beta; \gamma) = (\alpha; \beta); \gamma \tag{K4}$$
$$\alpha; 1 = \alpha = 1; \alpha \tag{K5}$$
$$\alpha; 0 = 0 = 0; \alpha \tag{K6}$$
$$\alpha; (\beta \cup \gamma) = (\alpha; \beta) \cup (\alpha; \gamma) \tag{K7}$$
$$(\alpha \cup \beta); \gamma = (\alpha; \gamma) \cup (\beta; \gamma) \tag{K8}$$

In short, 0 is the unit of $\cup$, which is associative (K1), commutative (K2) and idempotent (K3). Hence, $(K, \cup, 0)$ can be seen as an idempotent symmetric monoid. As a result, $\cup$ can be defined via the universal property of a co-product within a co-Cartesian category. This follows from the fact that a Kleene algebra is an idempotent semiring. Since a semiring is by definition a semigroup[9], it follows that $(K, \cup, 0)$ is a join-semilattice insofar as any commutative idempotent semigroup is a semilattice.[10] Thus, a Kleene algebra is a join-semilattice. This can also be seen insofar as (K1), (K2) and (K3) give us the equations (L2), (L4), (L6) and (L8) for a join-semilattice (see the section on Boolean algebras below). Hence, $(K, \cup, 0)$ possesses all finite joins, meaning that for all α and β we have the injection maps:

$$\alpha \leq \alpha \cup \beta$$
$$\beta \leq \alpha \cup \beta$$

Moreover, it means that if $\alpha \leq \gamma$ and $\beta \leq \gamma$, then $\alpha \cup \beta \leq \gamma$. Replacing '$\leq$' by a deduction arrow '$\longrightarrow$', one can easily derive (co-Cart). Indeed, the latter condition gives us:

$$\frac{\alpha \leq \gamma \qquad \beta \leq \gamma}{\alpha \cup \beta \leq \gamma}$$

The other direction of (co-Cart) follows directly from the transitivity of the preorder: assume $\alpha \cup \beta \leq \gamma$, then we have:

$$\alpha \leq \alpha \cup \beta \leq \gamma$$
$$\beta \leq \alpha \cup \beta \leq \gamma$$

Similarly, if one assumes that $(K, \cup, 0)$ is axiomatized by a co-Cartesian deductive system, then one can show that K possesses all finite joins (the proof is left to the reader).

[9] See http://ncatlab.org/nlab/show/rig.
[10] See http://ncatlab.org/nlab/show/semilattice.

As a result, the logic of $(K, \cup, 0)$ can be axiomatized within a co-Cartesian deductive system. In the case of $(K, ;, 1)$, one can see that 1 is the unit of ;, which is associative, and thus $(K, ;, 1)$ can be seen as a monoid, implying that the logic of $(K, ;, 1)$ can be represented by a monoidal deductive system. Finally, ; distributes over $\cup$ and (K6) gives us

$$\alpha; 0 = 0 = 0; \alpha$$

Therefore, a Kleene algebra $\mathcal{KA} = (K, \cup, ;, 0, 1)$ can be seen as a distributive co-Cartesian monoidal deductive system.

After Pratt's developments, dynamic logic was introduced in the context of deontic logic by Meyer (1988), who introduced the connectives $\bar{\alpha}$, $\alpha \cap \beta$ and $\varphi \to \alpha/\beta$, respectively representing action negation, joint action and the conditional action 'if φ, then α else β'. In addition to a denumerable set of atomic actions $Act = \{\varnothing, U, a_1, ..., a_n, ...\}$, Meyer includes the action which *fails* $\varnothing$ and the *universal action* U. The modality $[\alpha]$ is axiomatized as a K modality with the necessitation rule and the K axiom for distribution together with the propositional axioms and the *modus ponens* rule. The set of well-formed formulas of propositional dynamic logic is defined recursively from the language $\mathcal{L} = \{Prop, Act, (,), ^-, \cup, \cap, ;, \neg, \supset, [\,]\}$ (we leave aside the connective $\to /$ since it has not been considered further) by

$$\varphi := \ p_i \mid \neg\varphi \mid \varphi \supset \psi \mid [\alpha]\varphi$$

The set of complex actions $Act*$ defined by

$$\alpha := \ a_i \mid \bar{\alpha} \mid \alpha \cap \beta \mid \alpha \cup \beta \mid \alpha; \beta$$

The other logical connectives are defined as usual. The axioms for actions are

$$[\overline{\overline{\alpha}}]\varphi \equiv [\alpha]\varphi \tag{A1}$$
$$[\alpha; \beta]\varphi \equiv [\alpha]([\beta]\varphi) \tag{A2}$$
$$[\overline{\alpha; \beta}]\varphi \equiv [\overline{\alpha}]\varphi \wedge [\alpha]([\overline{\beta}]\varphi) \tag{A3}$$
$$[\alpha \cup \beta]\varphi \equiv [\alpha]\varphi \wedge [\beta]\varphi \tag{A4}$$
$$[\overline{\alpha}]\varphi \vee [\overline{\beta}]\varphi \supset [\overline{\alpha \cup \beta}]\varphi \tag{A5}$$
$$[\alpha]\varphi \vee [\beta]\varphi \supset [\alpha \cap \beta]\varphi \tag{A6}$$
$$[\overline{\alpha}]\varphi \wedge [\overline{\beta}]\varphi \equiv [\overline{\alpha \cap \beta}]\varphi \tag{A7}$$
$$[\varnothing]\varphi \tag{A9}$$

It is noteworthy that (A2) implies that sequence is not primitive. Hence, let us concentrate on the fragment constructed from $^-$, $\cap$ and $\cup$. Without explicitly defining the algebra of actions inherent to his approach, Meyer (1988, p.132) asserts that the following equations are satisfied.

$$\overline{\overline{\alpha}} = \alpha \tag{DL1}$$
$$\overline{\alpha \cup \beta} = \overline{\alpha} \cap \overline{\beta} \tag{DL2}$$
$$\overline{\alpha \cap \beta} = \overline{\alpha} \cup \overline{\beta} \tag{DL3}$$
$$\alpha \cup \varnothing = \alpha \tag{DL4}$$
$$\alpha \cup \alpha = \alpha \tag{DL5}$$
$$\alpha \cap \alpha = \alpha \tag{DL6}$$
$$\alpha \cap \overline{\alpha} = \varnothing \tag{DL7}$$
$$\alpha \cap (\beta \cup \gamma) = (\alpha \cap \beta) \cup (\alpha \cap \gamma) \tag{DL8}$$
$$\alpha \cup (\beta \cap \gamma) = (\alpha \cup \beta) \cap (\alpha \cup \gamma) \tag{DL9}$$

When writing on the algebraic properties of actions, Meyer did not refer to a Boolean algebra of actions, and hence it is fair to assume that the algebra of actions he had in mind for his approach was somewhat weaker than a Boolean algebra. But he did not either refer explicitly to a Kleene algebra, and so it is difficult to see exactly what he had in mind regarding the action logic inherent to his approach. Let DDL refer to the (fragment of the) action logic inherent to the Dynamic Deontic Logic presented in Meyer (1988). In what follows, we will first expose without any extra assumption the properties of the action connectives we can infer from Meyer's presentation. We will compare his approach to a Kleene algebra, then we will examine what needs to be done if one wants to define explicitly DDL as a Boolean algebra, and finally we will show that DDL actually *is* a Boolean algebra.

Although Meyer did not mention whether $\cap$ and $\cup$ were associative or commutative, associativity and commutativity of $\cup$ can be recovered from (A4). Idempotence of both $\cap$ and $\cup$ is given by (DL5) and (DL6), and distributivity of $\cap$ over $\cup$ is given by (DL8). Moreover, distributivity of $\cup$ over $\cap$ is given by (DL9). Finally, (DL1) gives us involutive negation.

Now, let us compare DDL and $\mathcal{KA}$. Associativity and commutativity of choice can be recovered in DDL from (A4), and hence (K1) and (K2) can also be recovered. (K3) is given by (DL4) and (DL5). (K7) is given by (DL8). There are no references to (K4), (K5), (K6) or (K8).

When one looks at (DL1), (DL2) and (DL3), one gets the feeling that Meyer wanted the structure of his action algebra to be classical. Anticipating on the following sections, let us note that Meyer's algebra of actions must be stronger than the multiplicative fragment **MLL** of linear logic insofar as he assumes the distributivity of $\cap$ over $\cup$. Hence, since Meyer's action algebra cannot be reduced to **MLL**, it seems that the only way to obtain both De Morgan dualities and involutive negation is to assume a Boolean algebra for actions.

The question we now want to examine is what extra assumptions are required to go from DDL to a Boolean algebra $\mathcal{BA}$. By definition, a $\mathcal{BA}$ is a distributive complemented lattice (see the section below for the definition). From (DL8) and (DL9) we obtain the two conditions for distributivity, and from (DL7) we obtain one part of the condition for complementarity. Therefore, the extra assumptions one needs in order to obtain a $\mathcal{BA}$ from DDL are that it is a lattice and that $\alpha \cup \overline{\alpha} = U$.

Looking at Meyer's paper, it is hard definitively to assert (or deny) that he assumed these two conditions. Meyer's approach takes place within the dynamic logic tradition, and hence it is reasonable to assume that he minimally considers $\cap$ to be associative and U to be its unit, which would enable him to recover (K4) and (K5). Moreover, (K6) can be recovered from distributivity and (DL7). Indeed, we have

$$\alpha \cap \overline{\alpha} = \varnothing \tag{DL7}$$

Hence from (DL4) we have

$$(\alpha \cap \overline{\alpha}) \cup \varnothing = \varnothing$$

By commutativity of $\cup$ and (DL9) we obtain

$$(\varnothing \cup \alpha) \cap (\varnothing \cup \overline{\alpha}) = \varnothing$$

Since $\alpha \cup \varnothing = \alpha$ we get

$$\alpha \cap (\varnothing \cup \overline{\alpha}) = \varnothing$$

By (DL8) we have

$$(\alpha \cap \varnothing) \cup (\alpha \cap \overline{\alpha}) = \varnothing$$

Since $\alpha \cap \overline{\alpha} = \varnothing$ it gives us

$$(\alpha \cap \varnothing) \cup \varnothing = \varnothing$$

Hence from (DL4) we obtain

$$\alpha \cap \varnothing = \varnothing$$

Finally, the commutativity of $\cup$ and the De Morgan dualities suggest that Meyer assumed $\cap$ to be commutative. This would explain the absence of any reference to (K8), which would hence be derivable from (DL8) and commutativity. As a result, it is quite reasonable to assume that Meyer minimally presupposed the structure of a Kleene algebra for DDL.

We saw that a Kleene algebra is an idempotent semiring, hence a join-semilattice. Hence, to define explicitly DDL as a $\mathcal{BA}$, one would need to assume that it is also a meet-semilattice (hence a lattice) and that it is complemented. It is noteworthy that a Kleene algebra is not a meet-semilattice insofar as the sequence operator is neither commutative nor idempotent. In Meyer's case, though, we saw that it would be safe to assume commutativity of $\cap$ and we already have idempotence, and as such assuming that DDL is a lattice would be consistent with his position. Moreover, the complementarity could easily be recovered. Indeed, from (DL4) and (DL7) we have:

$$\alpha \cup (\alpha \cap \overline{\alpha}) = \alpha$$

By (DL9) we obtain:

$$(\alpha \cup \alpha) \cap (\alpha \cup \overline{\alpha}) = \alpha$$

From (DL5) we know that $\alpha \cup \alpha = \alpha$, and therefore we have:

$$\alpha \cap (\alpha \cup \overline{\alpha}) = \alpha$$

This suggests that $\alpha \cup \overline{\alpha}$ is the unit of $\cap$, and hence that $\alpha \cup \overline{\alpha} = U$ if we assume that U is the unit of $\cap$.

So far, we have seen how one could (explicitly) define DDL as a $\mathcal{BA}$ from Meyer's initial remarks. Now we show that DDL is actually implicitly a $\mathcal{BA}$. To do so, we need to show that DDL satisfies the equations (L1)–(L14) that define a distributive complemented lattice (see the section below). We already know that (L2) and (L4), which express the associativity and commutativity of $\cup$ can be recovered from (A4), and that (L5) and (L6) for idempotence are given by (DL5) and (DL6). The distributivity conditions (L11) and (L12) are given by (DL8) and (DL9) and the first part of complementarity (L13) is given by (DL7). (DL4) yields (L8). As such, we only need to show that DDL satisfies

$$\alpha \frown (\beta \frown \gamma) = (\alpha \frown \beta) \frown \gamma \tag{L1}$$

$$\alpha \frown \beta = \beta \frown \alpha \tag{L3}$$

$$\alpha \frown 1 = \alpha \tag{L7}$$

$$\alpha \smile (\alpha \frown \beta) = \alpha \tag{L9}$$

$$\alpha \frown (\alpha \smile \beta) = \alpha \tag{L10}$$

$$\alpha \smile \bar{\alpha} = 1 \tag{L14}$$

Associativity and commutativity of $\cap$ expressed by (L1) and (L3) are recovered by:

$$\begin{aligned}
\overline{\alpha \cap (\beta \cap \gamma)} &= \overline{\alpha} \cup \overline{\beta \cap \gamma} && \text{(by DL3)}\\
&= \overline{\alpha} \cup (\overline{\beta} \cup \overline{\gamma}) && \text{(by DL3)}\\
&= (\overline{\alpha} \cup \overline{\beta}) \cup \overline{\gamma} && \text{(by associativity)}\\
&= \overline{\alpha \cap \beta} \cup \overline{\gamma} && \text{(by DL3)}\\
&= \overline{(\alpha \cap \beta) \cap \gamma} && \text{(by DL3)}
\end{aligned}$$

$$
\begin{aligned}
\overline{\alpha \cap \beta} &= \overline{\alpha} \cup \overline{\beta} && \text{(by DL3)}\\
&= \overline{\beta} \cup \overline{\alpha} && \text{(by commutativity)}\\
&= \overline{\beta \cap \alpha} && \text{(by DL3)}
\end{aligned}
$$

(L7) can be derived from the fact that:

$$\overline{\alpha} = \overline{\alpha} \cup \varnothing \qquad \text{(by DL4)}$$

Therefore, we can derive:

$$
\begin{aligned}
\alpha &= \overline{\overline{\alpha}} && \text{(by DL1)}\\
&= \overline{\overline{\alpha} \cup \varnothing} && \\
&= \overline{\overline{\alpha}} \cap \overline{\varnothing} && \text{(by DL2)}\\
&= \alpha \cap \overline{\varnothing} && \text{(by DL1)}
\end{aligned}
$$

Hence, $\overline{\varnothing}$ is the unit of $\cap$. Moreover, since (K6) can be recovered in DDL, (DL3) gives us:

$$\overline{\varnothing} = \alpha \cup \overline{\varnothing}$$

Thus, it follows that the absorption laws (L9) and (L10) can be derived:

$$
\begin{aligned}
\alpha &= \alpha \cap \overline{\varnothing} && \text{(by L7)}\\
&= \alpha \cap (\beta \cup \overline{\varnothing}) && \text{(by K6 and DL3)}\\
&= (\alpha \cap \beta) \cup (\alpha \cap \overline{\varnothing}) && \text{(by DL8)}\\
&= (\alpha \cap \beta) \cup \alpha && \text{(by L7)}\\
&= \alpha \cup (\alpha \cap \beta) && \text{(by commutativity)}
\end{aligned}
$$

$$
\begin{aligned}
\alpha &= \alpha \cup \varnothing && \text{(by DL4)}\\
&= \alpha \cup (\varnothing \cap \beta) && \text{(by K6)}\\
&= (\alpha \cup \varnothing) \cap (\alpha \cup \beta) && \text{(by DL9)}\\
&= \alpha \cap (\alpha \cup \beta) && \text{(by DL4)}
\end{aligned}
$$

Finally, from (DL7) and (DL3) we have:

$$\overline{\varnothing} = \overline{\alpha} \cup \overline{\overline{\alpha}}$$

Thus, by commutativity and (DL1) we have (L14):

$$\overline{\varnothing} = \alpha \cup \overline{\alpha}$$

As a result DDL satisfies the equations of a distributive complemented lattice.

The reader can compare Meyer's definition of DDL and the properties of a Boolean algebra in table 14.2. When compared through the properties of their connectives, the only difference between them is that $\cap$ is not explicitly said to be commutative and associative within DDL. Otherwise, the result is the same, and as we saw Meyer's definition yields implicitly a $\mathcal{BA}$. As such, it is noteworthy that if Meyer explicitly assumed that DDL is a lattice and that $\alpha \cup \overline{\alpha} = U$, then DDL would have been seen as a precursor of what Kozen (1997) defined as a *Kleene algebra with test*, that is, a Kleene algebra embedded in a Boolean algebra to model negation of programs (see Harel et al. 2000, p.421).

To sum up, from a categorical perspective, a Kleene algebra $\mathcal{KA}$ (without the Kleene star *) can be seen as a distributive co-Cartesian monoidal deductive system and DDL as a Classical deductive system, as we will see in the section below. Let us conclude this section by mentioning that in addition to $\mathcal{KA}$ and DDL, there are other approaches in dynamic logic which have explicitly used Boolean algebras for actions. For instance the reader may look at the work of Castro and Maibaum (2009), who develop a dynamic logic as an extension of a Boolean algebra of actions.[11]

Pratt

In addition to the introduction of dynamic logic, Pratt (1991) also introduced another form of action logic, which looses the distinction between actions and assertions. Pratt's action logic is presented as a logic for programming that includes the three basic control elements *choice*, *sequence* and *iteration* (Pratt 1991, p.100). The choice $\alpha + \beta$ between α and β is introduced similarly to the additive disjunction of linear logic, where the choice implies the performance of either one or the other, but not both. Sequence $\alpha; \beta$ is introduced as a non-commutative conjunction where α can be performed before β or simultaneously (hence the non commutativity), and iteration α^* is the performance of α an arbitrary number of times (Pratt 1991, p.100). The constant 0 and 1 are introduced as the *empty* and the *skip* (or universal) actions respectively.

Following Pratt (1991, p.102), an Action algebra $\mathcal{A}* = \langle A, +, 0, ;, 1, \rightarrow, \leftarrow, ^* \rangle$ is defined by two monoids $(A, +, 0)$ and $(A, ;, 1)$ with $+$ commutative and idempotent, satisfying the following rules.

$$\frac{\alpha \leq \gamma \leftarrow \beta}{\overline{\alpha; \beta \leq \gamma}}$$

$$\frac{\beta \leq \alpha \rightarrow \gamma}{\overline{\alpha; \beta \leq \gamma}}$$

[11] For more information on categories and dynamic logic, the reader can consult van Benthem (1991), and for other approaches in dynamic deontic logic, see van der Meyden (1996) for a dynamic logic of permission, Royakkers (1998) for dynamic logic and agency and Broersen (2004) for the analysis of action negation.

As such, $\rightarrow$ and $\leftarrow$ are the right adjoints of ;. Note that these rules are sometimes called the *residuation laws* in the literature (see Galatos et al. 2007). An action algebra also has to satisfy:

$$\frac{1+\beta;\beta+\alpha \leq \beta}{\alpha^* \leq \beta}$$

$$1+\alpha^*;\alpha^*+\alpha \leq \alpha^*$$

The symbols $\rightarrow$ and $\leftarrow$ stand respectively for pre-implication (had α, then β) and post-implication (β if ever α). We concentrate only on the fragment for conjunctive action and disjunctive action without iteration, hence let $\mathcal{A} = \langle A, +, 0, ;, 1, \rightarrow, \leftarrow \rangle$.

An interesting result is that $\mathcal{A}* \subseteq \mathcal{KA}*$ (see Pratt 1991, p.109). Hence, $\mathcal{A} \subseteq \mathcal{KA}$. It trivially follows from the definition of an Action algebra that $\mathcal{A}$ satisfies (K1)-(K5). Following Pratt (1991, p.104), since $\rightarrow$ and $\leftarrow$ are adjoints of ;, it follows that one can also derive (K6), (K7) and (K8). Pratt actually shows that in finite cases, every $\mathcal{KA}$ expands to an Action algebra $\mathcal{A}$. This results does not, however, hold for infinite cases (see Pratt 1991, p.109). Hence, there are some Kleene algebras that are *not* Action algebras (see Pratt 1991, pp.109-10), and thus $\mathcal{KA}$ is a little stronger than $\mathcal{A}$, although both can be seen as idempotent semirings. From a proof-theoretical perspective, an Action algebra $\mathcal{A}$ can be seen as a co-Cartesian monoidal closed deductive system.[12]

Linear logic

Linear logic was introduced by Girard (1987) and was meant to provide a constructive framework other than intuitionistic logic. The main objective of linear logic is to be able to deal with limited resources, and this led to the introduction of two types of conjunction (with respect to their dual disjunction). The syntax of linear logic is built from the language $\mathcal{L} = \{Prop, {}^{\perp}, \top, \bot, \mathbf{1}, \mathbf{0}, \otimes, ⅋, \&, \oplus, !, ?, \forall, \exists\}$, with *Prop* a denumerable set of atoms, !/? operators to add and subtract resources and $\forall/\exists$ the usual quantifiers. For the sake of our analysis, we will only present the **MALL** fragment of linear logic, which is the fragment constructed upon $\mathcal{L} = \{Prop, {}^{\perp}, \top, \bot, \mathbf{1}, \mathbf{0}, \otimes, ⅋, \&, \oplus\}$. The main feature of linear logic is that it incorporates two pairs of dual connectives, namely $\otimes$, ⅋ and $\&, \oplus$, with a linear implication $\multimap$ defined as the adjoint of $\otimes$.

Before providing the categorical presentation of **MALL**, let us have an example to illustrate the difference between the additive and the multiplicative fragments. Let:

p: Paul is at a T intersection
q: Paul turns right
r: Paul turns left

[12] See also Bimbó et Dunn (2005) for Kleene algebras and Pratt's action logic.

Now, consider the four following formulas:

$$p \multimap q \otimes r \tag{1}$$
$$p \multimap q \parr r \tag{2}$$
$$p \multimap q \& r \tag{3}$$
$$p \multimap q \oplus r \tag{4}$$

The conjunction $\otimes$ requires that both actions are done. In computer science, it means parallel execution. Thus, (1) signifies that if Paul is at a T intersection, then he turns both left and right, which is obviously impossible. On the other hand, & is a conjunction that implies a choice where only one of the actions will be done. Hence, (3) makes sense insofar as it means that if Paul is at a T intersection, then he turns right, else he turns left. As Girard (1995, p.2) points out, this is the analogue of 'if ... then ... else ... ' in computer science. The dual $\oplus$ of & is a disjunction that respects this reading. While & implies that one has to decide which action in the conjunction will be done, $\oplus$ is a 'choice' without decision, Paul will turn right or left, but his action will be determined via external causes. In other words, with & the choice is left to the agent but with $\oplus$ it is not, which is the distinction between outer and inner non determinism (Girard 1995, p.3). Finally, as Girard (1995, p.3) says, the disjunction $\parr$ expresses a dependency between two actions according to the reading of $\otimes$, meaning that both action could be done at the same time.

Following Blute and Scott (2004), the proof theory of **MALL** can be reproduced by a *-autonomus category with products and coproducts. This result was actually provided by Seely (1989), who showed how the deductive system for the multiplicative fragment **MLL** can be obtained through a *-autonomous category and the **MALL** fragment results by adding products and coproducts. The **MALL** fragment of linear logic is constructed from the multiplicative $\{\otimes, \parr, \multimap, \mathbf{1}, \bot\}$ and additive $\{\&, \oplus, \mathbf{0}, \top\}$ fragments (see Girard 1995, p.10 for the definitions). Recall the definition of a *-autonomous category: it is a symmetric closed monoidal category with a dualizing object. Thus, the multiplicative fragment **MLL** can be obtained through the axioms (1), (**) and the rules (cut), (t), (a), (l), (r), (b), (cl) for a symmetric closed monoidal category, with $\parr$, $\multimap$ and $\mathbf{1}$ being respectively the dual, adjoint and unit of $\otimes$ and $\bot$ the dualizing object satisfying $\varphi \cong \varphi^{\bot\bot}$. The **MALL** fragment is obtained by adding the rules and axioms of a bi-Cartesian deductive system for the additive fragment, with $\mathbf{0}$ the initial object and $\top$ the terminal. In short, **MALL** is a *-autonomous deductive system with products and co-products.[13]

The multiplicative fragment of linear logic **MLL** is a *-autonomous deductive system, that is, a bi-symmetric closed monoidal deductive system with a dualizing object. Hence, the tensor product and the co-tensor are both associative and commutative, but neither are idempotent. Note, however, that although closed, the multiplicative fragment is not distributive (in the sense of terminology 5).[14] There are, however, other variations of linear logic that involve different properties for the multiplicative connectives. First, following

[13] For a more succinct presentation of the **MALL** fragment in terms of sequent calculus than the one of Girard, see Schalk (2004).

[14] **MLL** is, however, *weakly* or *linearly* distributive, i.e., there are natural transformations

Casadio (2001, p.167), there is the multiplicative (and exponential free) fragment of the non-commutative intuitionistic linear logic **NILL** introduced by Abrusci (1990b), which corresponds to Lambek's (1958) Syntactic Calculus with negation (see also Abrusci 1990a). Hence, the multiplicative fragment of **NILL** built from $\{\otimes, \multimap, \circ\!\!-, \mathbf{1}\}$ corresponds to what Lambek (1968) defined as a residuated category and can be axiomatized by a monoidal bi-closed deductive system. **NILL** without negation (i.e., the Syntactic Calculus) also corresponds to the $(A, ;, 1)$ fragment of Pratt's Action algebra $\mathcal{A}$. Secondly, there is the multiplicative fragment **MNLL** of the non-commutative linear logic **NLL** introduced by Abrusci (1991). It can be obtained from a syntactical point of view when one introduces involutive negation (see also Abrusci and Ruet 2000), and thus it can be seen as a monoidal closed deductive system with a dualizing object.

Boolean algebras of actions

We now turn our attention to Boolean algebras of actions. The classical structure of action propositions can be fully expressed when one uses a Boolean algebra of actions rather than the usual propositional dynamic logic. Segerberg (1982b), for instance, proposed to define a Deontic Action algebra on the ground of a Boolean algebra of actions. Also distinguishing between *actions* and *propositions*, he assumes a language $\mathcal{L} = \{Act, \smile, \frown, ^-\}$ with Act a denumerable set of atomic actions together with the impossible action $\mathbf{0}$ and the universal action $\mathbf{1}$. Segerberg (1982b, p.269) uses $\smile$, $\frown$ and $^-$ to denote the union, the conjunction and the negation of actions respectively. The set A of actions is recursively defined by:

$$\alpha := \mathbf{0} \mid \mathbf{1} \mid a_i \mid \bar{\alpha} \mid \alpha \frown \beta \mid \alpha \smile \beta$$

A logic of action is understood as a Boolean algebra. Following Goldblatt (2006, p.134), a Boolean algebra $\mathcal{BA} = (A, \frown, \smile, ^-, \mathbf{0}, \mathbf{1})$ is by definition a distributive complemented lattice. A lattice can algebraically be defined by the following equations.[15]

$$\alpha \frown (\beta \frown \gamma) = (\alpha \frown \beta) \frown \gamma \tag{L1}$$
$$\alpha \smile (\beta \smile \gamma) = (\alpha \smile \beta) \smile \gamma \tag{L2}$$
$$\alpha \frown \beta = \beta \frown \alpha \tag{L3}$$
$$\alpha \smile \beta = \beta \smile \alpha \tag{L4}$$
$$\alpha \frown \alpha = \alpha \tag{L5}$$
$$\alpha \smile \alpha = \alpha \tag{L6}$$
$$\alpha \frown 1 = \alpha \tag{L7}$$
$$\alpha \smile 0 = \alpha \tag{L8}$$
$$\alpha \smile (\alpha \frown \beta) = \alpha \tag{L9}$$
$$\alpha \frown (\alpha \smile \beta) = \alpha \tag{L10}$$

$\varphi \otimes (\psi \parr \rho) \longrightarrow (\varphi \otimes \psi) \parr \rho$ and $(\varphi \parr \psi) \otimes \rho \longrightarrow \varphi \parr (\psi \otimes \rho)$. See Blute and Scott (2004, p.28) or the nlab entry http://ncatlab.org/nlab/show/linearly+distributive+category.

[15] See http://ncatlab.org/nlab/show/lattice.

(L1) and (L2) express associativity of $\smile$ and $\frown$, as (L3)-(L4) and (L5)-(L6) represent commutativity and idempotence. (L7) and (L8) make 1 and 0 into the respective unit of $\frown$ and $\smile$, and (L9)-(L10) are the absorption laws.

Equivalently, a lattice can be defined as a a poset $\mathcal{P} = (A, \leq)$ which possesses meets and joins, that is, for each $\alpha, \beta, \gamma \in A$:

$$\begin{aligned} &\alpha \frown \beta \leq \beta \text{ and } \alpha \frown \beta \leq \beta \\ &\text{if } \gamma \leq \alpha \text{ and } \gamma \leq \beta, \text{ then } \gamma \leq \alpha \frown \beta \end{aligned} \tag{meets}$$

$$\begin{aligned} &\alpha \leq \alpha \smile \beta \text{ and } \beta \leq \alpha \smile \beta \\ &\text{if } \alpha \leq \gamma \text{ and } \beta \leq \gamma, \text{ then } \alpha \smile \beta \leq \gamma \end{aligned} \tag{joins}$$

Moreover, for each $\alpha \in A$:

$$\begin{aligned} \alpha &\leq \mathbf{1} \\ \mathbf{0} &\leq \alpha \end{aligned}$$

Hence, $\alpha \frown \beta$ is a greatest lower bound and $\alpha \smile \beta$ a least upper bound.

A lattice is said to be *distributive* when:

$$\alpha \smile (\beta \frown \gamma) = (\alpha \smile \beta) \frown (\alpha \smile \gamma) \tag{L11}$$
$$\alpha \frown (\beta \smile \gamma) = (\alpha \frown \beta) \smile (\alpha \frown \gamma) \tag{L12}$$

It is complemented when:

$$\alpha \frown \bar{\alpha} = \mathbf{0} \tag{L13}$$
$$\alpha \smile \bar{\alpha} = \mathbf{1} \tag{L14}$$

It trivially follows from these definitions that a lattice can be viewed as a Cartesian monoidal category, with $\frown$ the product, $\smile$ the co-product, $\mathbf{0}$ an initial object and $\mathbf{1}$ a terminal object, the meets and the joins giving us (Cart) and (co-Cart).[16]

That being said, a $\mathcal{BA}$ can also be seen as a Cartesian *closed* poset. Indeed, one simply has to define $\alpha \triangleright \beta$ by $\bar{\alpha} \smile \beta$, and distributivity will imply that $\triangleright$ is the adjoint of $\frown$ (see Awodey 2006, p.129).

A logic of action which is understood as a Boolean algebra can thus be axiomatized as a Classical deductive system. For approaches that consider the logic of action as a Boolean algebra of actions, see for instance Lindahl and Odelstad (2011), Trypuz and Kulicki (2010) or Castro and Maibaum (2009).

[16] Note that when a poset is seen as a category, $\leq$ must be viewed as an arrow between equivalence classes of actions rather than actions per se insofar as $\leq$ is an antisymmetric relation.

Stit logics

The distinction between propositions and assertions is not always taken for granted. Following Segerberg et al. (2009, p.1), the banner of 'action logics' can be divided in two main groups in philosophy, namely the STIT logics and dynamic logic. Although the origin of STIT logic is usually attributed to the work of Belnap and Perloff (1988), who introduced the modal operator $[i\ stit : \varphi]$ (where STIT stands for *seeing to it that*), this variant of action logic can actually be traced back to the work of Kanger (1957).[17] The first formal semantics for STIT logics was proposed by Chellas (1969) (cf. Segerberg 1992, p.369) and the modal operator E_i (originally written D_i), which is another version of the STIT operator and was meant to be read '*agent i brings it about that*', was introduced in the work of Pörn (1970). The work of Belnap and Perloff (1988) was further developed by Xu (1995), and nowadays one can find STIT logics for instance in the work of Horty and Belnap (1995), Horty (2001), Pacheco and Carmo (2003) and Broersen (2011a).

STIT logics are modal logics that use a K modality indexed to an agent (or a set of agents). The STIT modality is usually axiomatized by KM, as in the work of Pörn (1970) and Pacheco and Carmo (2003), but can also be axiomatized via an $S5$ modality (cf. Horty 2001) or by a KD modality as in the XSTIT framework of Broersen (2011a). Assuming a language $\mathcal{L} = \{Prop, (,), \neg, \supset, [i\ stit]\}$ (with $Prop$ a denumerable set of atomic sentences), one can recursively define the set of well-formed formulas by:

$$\varphi := p_i \mid \neg\varphi \mid \varphi \supset \psi \mid [i\ stit]\varphi$$

The other connectives are defined as usual. The $[i\ stit]$ modality is axiomatized through the usual axioms and rules of modal logic together with suitable axiom(s), depending on the framework adopted. If $[i\ stit]$ behaves according to (**M**), then one can infer from $[i\ stit]\varphi$ that φ is true, which will be accompanied by a reflexive relation R on the semantical model. Likewise, truth conditions for the $[i\ stit]$ modality will be usually defined via the accessibility relation on the model $\mathcal{M}$, the only difference between the usual model of modal logic being that $\mathcal{M}$ is augmented with a set of agent Ag and the truth condition for $[i\ stit]$ is indexed to a relation R_i, and thus one obtains a class of accessibility relations indexed to some agents.[18]

One interest of STIT logic is that it enables one to distinguish between doing a negative action and simply not doing an action. Indeed, it enables one to distinguish between not doing φ, represented by $\neg[i\ stit]\varphi$, and consciously doing $\neg\varphi$, expressed by $[i\ stit]\neg\varphi$. As the distinction between propositions and assertions can be seen as an argument in favour of dynamic logic over STIT logics, the fact that the latter allows to distinguish between '*not doing* φ' and '*doing not* φ' can be seen as an argument in its favour.

In a STIT framework one considers descriptions of states rather than actions in themselves. The modal operator $[i\ stit]\varphi$ means that the agent i brings about (makes true)

[17] This was reprinted in Kanger (1971) and further developed in Kanger and Kanger (1966) and Kanger (1972).

[18] See Chellas (1980) or Garson (2006) for an introduction to modal logics.

some description of the world φ. Therefore, the emphasis is not put on the action in itself but is rather put upon the *result* of the action (i.e., its outcome), that is, some specific description of the world. That being said, since STIT logics are modal logics, it follows that their structure is fundamentally classical (Boolean), and hence that the logic of the connective $\wedge$, $\vee$, $\neg$ and $\supset$ can be axiomatized via a Classical deductive system, as it is the case for Boolean algebras. Since the actions are viewed as descriptions within a STIT framework and the logic of description is Boolean, it follows that the logic of actions inherent to the STIT frameworks is classical.

Thierry Lucas

Lucas's work stands among the few approaches that explicitly applied category theory to deontic logic. His logic of action, inspired by the work of von Wright, is meant to express the distinction between the algebraic structure of action sentences and the propositional structure of normative sentences. In the first part of his paper, Lucas (2006) reconstructs von Wright's logic of actions using action operators and a modality for necessity. However, for the sake of our analysis, we will only concentrate on the second part of the 2006 paper, where Lucas emphasizes the algebraic structure of an action logic, which was further analyzed in Lucas (2007, 2008).

The originality of Lucas's (2006, p.101) approach is that he considers actions as *morphisms* in a category-theoretic framework. Indeed, he understands actions as mapping from a set of initial conditions to a set of results. Lucas (2008, p.370) defines an algebra of action as a triple $\langle B, \mathcal{ACT}, C\rangle$ with B and C two Boolean algebras (that may be different from each other) and $\mathcal{ACT} = \langle A, 0, 1, \neg, \wedge, \vee, C_0, \rangle$ a poset that behaves as a bi-intuitionistic logic (Lucas 2006, p.86) (the notion of cosupport C_0 is introduced to represent the conditions under which the negation of the action is obtained). Assuming that $\mathcal{F}_1$ and $\mathcal{F}_2$ are respectively the set of well-formed formulas of B and C, Lucas (2006, p.102) views an action as a morphism $a : \Sigma \longrightarrow \mathcal{F}_2$, with $\Sigma \subseteq \mathcal{F}_1$. Hence, an action is considered as a mapping from a set of conditions to a set of results (Lucas 2006, p.101). Actions have to satisfy a coherence condition such that $a(\sigma) = a(\sigma')$ for $\sigma, \sigma' \in \Sigma$, meaning that the action a yields the same result for every sub-condition $\sigma \in \Sigma$. In other words, the sub-conditions are consistent with each other, that is, the action regarding two sub-conditions of a set Σ of conditions yields the same result. Actions are pre-ordered by $\leq$ (Lucas 2008, p.369), 1 is the *empty* action and 0 the *total* (or *null*) action. The empty action is the equivalent of doing nothing (the trivial action) while the null action is impossible to be performed.

On the ground of $(A, 0, 1, \wedge, \vee)$, Lucas (2006, p.106) defines $(A, 0, 1, \wedge, \vee, \sim, \rightarrow)$ as an intuitionistic deductive system, with $\rightarrow$ the adjoint of $\wedge$ and $\sim$ the intuitionistic negation defined by $\sim \alpha =_{def} \alpha \rightarrow 0$ (see also Lucas 2008, p.383). However, there is more to it: he also defines an adjoint to disjunction, together with another negation v. The adjoint to disjunction is governed by the rule (cl'), and negation is defined by $v\alpha = 1\backslash\alpha$.

$$\frac{\gamma \leq \alpha \vee \beta}{\overline{\gamma\backslash\beta \leq \alpha}} \qquad \text{(cl')}$$

This is why Lucas defines $\mathcal{ACT}$ as a bi-intuitionistic logic. A bi-intuitionistic logic is understood as a bi-Heyting algebra, that is, a bounded distributive lattice which is a Heyting and a co-Heyting algebra.[19]

Although we will restrict our attention only to the $(A, 0, 1, \wedge, \vee)$ fragment of $\mathcal{ACT}$, let us note that Lucas (2008, p.370) uses $\neg$ to dualize the connectives, introducing the dual connectives $(-)^*$ defined by the following equations.

$$
\begin{aligned}
&\alpha \leq^* \beta \Leftrightarrow \neg\beta \leq \neg\alpha \\
&\alpha \leq \beta \Leftrightarrow \neg\beta \leq^* \neg\alpha \\
&\alpha \wedge^* \beta = \neg(\neg\alpha \vee \neg\beta) \\
&\alpha \vee^* \beta = \neg(\neg\alpha \wedge \neg\beta) \\
&1^* = \neg 0 \\
&0^* = \neg 1 \\
&\alpha \backslash^* \beta = \neg(\neg\beta \rightarrow \neg\alpha) \\
&v^*\alpha = \neg \sim \neg\alpha \\
&\alpha \rightarrow^* \beta = \neg(\neg\beta \backslash \neg\alpha) \\
&\sim^* \alpha = \neg v \neg\alpha
\end{aligned}
$$

Thus defined, both $\wedge$ and $\wedge^*$ distribute over $\vee$. Returning to the fragment that concerns us, the logic of $(A, 0, 1, \wedge, \vee)$ can be seen as an intuitionistic deductive system with an adjoint to disjunction and another negation.

Summing up

All of the aforementioned approaches can be compared on the ground that they are all variations of monoidal deductive systems. Table 14.2 provides a comparison between the systems discussed so far, concentrating only upon the fragments including tensor $\otimes$, co-tensor $\oslash$, adjoint(s) $\triangleright$, $\blacktriangleright$ and negation * when they are defined (with only the multiplicative fragments for variations of linear logic). The notation has been made uniform according to the categorical analysis provided at the beginning of this paper. We also introduced a Heyting algebra $\mathcal{HA}$ (which corresponds to an Intuitionistic deductive system) and Lambek's Syntactic Calculus $\mathcal{SC}$ for comparison.[20] Although we will not explicit each consequence of the comparison made in table 14.2, let us note that:

1. $\otimes$ is non-commutative for **NILL**, **MNLL**, $\mathcal{A}$ and $\mathcal{KA}$;

2. in every case except the linear fragments, $\otimes$ distributes over $\oslash$;

[19] Note that 'bi-intuitionistic' here refers to an intuitionistic logic with two negations rather than a 'bi-' deductive system as presented earlier. The relationship between bi-Heyting algebras and modal logic has been discussed by Reyes and Zolfaghari (1996).

[20] The abbreviations in table 14.2 stand for associativity, commutativity, idempotence and distributivity.

3. negation is involutive in **MNLL**, DDL, **MLL** and $\mathcal{BA}$;
4. $\oslash$ is commutative and idempotent in $\mathcal{A}$, $\mathcal{KA}$, DDL, $\mathcal{HA}$, $\mathcal{ACT}$ and $\mathcal{BA}$;
5. $\otimes$ is idempotent in DDL, $\mathcal{HA}$, $\mathcal{ACT}$ and $\mathcal{BA}$;
6. only **NILL**, $\mathcal{HA}$ and $\mathcal{ACT}$ provide a non-involutive negation.

	Ass.	Comm.	Idem.	Dist.	Adjoint(s)	Negation
$\mathcal{SC}$	$\otimes$				$\triangleright$ $\blacktriangleright$	
NILL	$\otimes$ $\oslash$				$\triangleright$ $\blacktriangleright$	$\alpha \leq \alpha^{**}$
MNLL	$\otimes$ $\oslash$				$\triangleright$ $\blacktriangleright$	$\alpha \cong \alpha^{**}$
$\mathcal{A}$	$\otimes$ $\oslash$	$\oslash$	$\oslash$	$\otimes$ over $\oslash$	$\triangleright$ $\blacktriangleright$	
$\mathcal{KA}$	$\otimes$ $\oslash$	$\oslash$	$\oslash$	$\otimes$ over $\oslash$		
DDL	$\oslash$	$\oslash$	$\otimes$ $\oslash$	$\otimes$ over $\oslash$ $\oslash$ over $\otimes$		$\alpha \cong \alpha^{**}$
MLL	$\otimes$ $\oslash$	$\otimes$ $\oslash$			$\triangleright$	$\alpha \cong \alpha^{**}$
$\mathcal{HA}$	$\otimes$ $\oslash$	$\otimes$ $\oslash$	$\otimes$ $\oslash$	$\otimes$ over $\oslash$ $\oslash$ over $\otimes$	$\triangleright$	$\alpha \leq \alpha^{**}$
$\mathcal{ACT}$	$\otimes$ $\oslash$	$\otimes$ $\oslash$	$\otimes$ $\oslash$	$\otimes$ over $\oslash$ $\oslash$ over $\otimes$	$\triangleright$ $\backslash$	$\alpha \leq \alpha^{**}$
$\mathcal{BA}$	$\otimes$ $\oslash$	$\otimes$ $\oslash$	$\otimes$ $\oslash$	$\otimes$ over $\oslash$ $\oslash$ over $\otimes$	$\triangleright$	$\alpha \cong \alpha^{**}$

Table 14.2: Properties of action connectives

Categorical action logic

The starting point of our investigation is the distinction between *actions* and *assertions*. It was already pointed out by Goldman (1970, p.47) that it is unclear whether or not propositional connectives could or should be applied to actions. It is our view that *actions* and *propositions* have different structures, and as such they do not share the same connectives. Following the dynamic logic tradition but also Lucas (2006), we assume that the algebraic structure of actions is fundamentally different from the structure of propositions.[21] In what follows, we propose to define an Action Logic as a logic that can represent the formal structure of actions, while a Propositional Action Logic is a logic that expresses the structure of the language we use to talk about actions. Hence, actions are not understood as declarative sentences in an Action Logic. What we mean by an *action* will be discussed further below in the section on action negation. For now, it suffices to see an action as something which can be referred to. We begin by providing a philosophical analysis of action connectives and then propose to define an Action Logic and a Propositional Action Logic respectively as a compact deductive system and an intuitionistic deductive system.

[21] Segerberg (1982b) distinguishes between actions and assertions but uses a Boolean algebra of action and Boolean propositional connectives.

A logic of actions

The first thing that needs to be done is to determine which types of actions we wish to consider for our Action Logic. Assuming that there are such things as *atomic actions*, we must determine the action connectives that can be used to construct complex actions. It was mentioned at the beginning of this paper that the aim is to provide an analysis of action logics from a philosophical perspective rather than a computational one. Although some action connectives are relevant for programming, we wish to develop an Action Logic that is relevant for human actions in deontic contexts instead of a logic that is useful for computer science.

Following Pratt (1991, p.99), the three basic control elements of imperative programming languages are choice, sequence, and iteration. The first consequence of our philosophical standpoint is that we will not consider iteration of action insofar as it is not relevant for human actions. Iteration and recursion are fundamental for programming languages and algorithms, but they are not part of *human* actions. Of course, there are actions that require some form of repetition (e.g., a work chain or a serial killer), but these actions can be seen as sequences rather than pure recursive actions. A second consequence of our approach is that *choice* will not be considered as a connective of our Action Logic but will rather be considered as a connective of the Propositional Action Logic. From the perspective of human actions, *choice* is not an action in itself. Although one can have a choice between two actions, in which case one will (ideally) perform one or the other, one will not however be *performing* the action '*choice*'. The third connective, sequence, can be used to construct complex human actions, but it will not be considered as the primary connective.

Action conjunction and sequence

The primary action connective we consider is the conjunction of action $\otimes$, where $\alpha \otimes \beta$ means the action α *together with* β. The conjunction of actions expressed by $\otimes$ implies that both actions are done simultaneously. It implies that α is done in conjunction with β (hence the use of $\otimes$ to suggest simultaneity). Action conjunction is introduced similarly to the multiplicative conjunction of linear logic. Since the 'with' operator implies simultaneity and sequence implies a temporal order, it follows that action conjunction and action sequence are two different action connectives. While $\otimes$ is commutative and associative, the sequence connective $\curvearrowright$ is only associative (and not commutative, for obvious reasons). A sequence action $\alpha \curvearrowright \beta$ means α *and then* β. We assume that the special action $*$ is the unit of $\otimes$. $*$ is interpreted as the *no change* or the *nothing* action. Doing α together with *no change* is the same as doing α, and vice versa. The special action $*$ is also the unit of sequence. Doing α and then *no change* is the same as doing α, as it is the case for doing *no change* and then α. Considering that we do not want *choice* to be included within our Action Logic, we do not define the co-tensor.

Now, the question that needs to be asked is whether $\otimes$ possesses an adjoint, and if so, then what is its philosophical interpretation. Since implication is usually the adjoint to conjunction, it seems natural to first ask if $\Rightarrow$ is the adjoint of $\otimes$. While we answer positively the first question ($\otimes$ possesses an adjoint), its adjoint is not implication. A conditional 'if α, then β' can have one of two meanings within the natural language when talking about actions. In oder to illustrate this, consider the two following conditionals:

(1) If Jones murders Smith gently, then Jones murders Smith.

(2) If Jones murders Smith, then he will try to hide the body.

The first conditional expresses a semantical entailment. However, this kind of entailment does not produce an action per se. Representing (1) by $\alpha \Rightarrow \beta$, '$\Rightarrow$' is a propositional connective (i.e., a connective that can be applied to declarative statements) rather than an action connective. Hence, (1) expresses a semantical entailment between propositions rather than actions. The conditional 'If Jones murders Smith gently, then Jones murders Smith' is not an action in itself but is rather a semantical entailment between 'murdering Smith gently' and 'murdering Smith'. Rather than pertaining to our Action Logic, the entailment expressed by (1) will be represented within the Propositional Action Logic.

The second conditional expresses that an action is conditional to another one. That Jones tries to hide the body of Smith requires that he murdered Smith in the first place. A conditional action happens when an action cannot be done unless some specific requirements have been met. In order to checkmate one's opponent, the white must first open, and then its opponent must do the same, then he will move another piece, etc. When building a house, building the roof is conditional to building the foundations and the walls. Similarly, do not try to frost a cupcake if it has not been done yet. Conditional actions are not, however, actions per se. Whether an action is conditional or not can be expressed when *talking* about actions. We can speak of conditional actions using $\Rightarrow$ but it does not produce a complex action. The fact that an action α is conditional to other actions implies that α can only be accomplished when some requirements have been met within a specific order. Hence, when saying that an action is conditional, we are saying that this action can only be realized through some specific sequences.

The connective $\Rightarrow$ is a propositional one and hence does not pertain to an Action Logic. Thus, it is not the adjoint to action conjunction. Is the sequence connective an adjoint to conjunctive action? If so, we would obtain:

$$\frac{\alpha \otimes \beta \longrightarrow \gamma}{\alpha \longrightarrow \beta \curvearrowright \gamma} \qquad \text{(cl)}$$

However, the interpretation of this rule would not be philosophically accurate. Indeed, an arrow between two actions represents logical entailment. Thus, (cl) would imply that if γ is a deducible from $\alpha \otimes \beta$, then the sequence $\beta \curvearrowright \gamma$ is deducible from α. But γ is not something which would be done subsequently after β, it is only a logical consequence of the conjunctive action. The distinction between sequence and conjunction is meant to distinguish between simultaneity and succession. Assuming that sequence is the adjoint of conjunction would blur that distinction. It is not because the action α together with β logically implies the action γ that α logically implies that γ will be done after β.

There is, however, an action connective than can be seen as the adjoint of $\otimes$. Indeed, it is our view that the appropriate adjoint for the *together with* connective is the *without* connective $\ominus$. To the best of our knowledge, the action connective 'without' has never been considered within the action logic literature. The behaviour of $\ominus$ is represented by (cl).

$$\frac{\alpha \otimes \beta \longrightarrow \gamma}{\beta \longrightarrow \gamma \ominus \alpha} \quad \text{(cl)}$$

We wrote $\gamma \ominus \alpha$ instead of $\alpha \ominus \gamma$ to keep the intuitive reading of γ without α, hence the notation reminding us of the subtraction γ minus α. In other words, if γ is a consequence of α together with β, then γ without α is a consequence of β (and γ without β is a consequence of α since we assumed that conjunction is commutative). That is to say, if adding β to α gives us the action γ, then subtracting β to γ gives us α back. A particular instance of (cl) is:

$$\frac{\alpha \otimes \beta \longrightarrow \alpha \otimes \beta}{\beta \longrightarrow (\alpha \otimes \beta) \ominus \alpha} \quad \text{(cl)}$$

For example, from the action of *walking together with singing* one can conclude that the action of *walking* yields the action of *walking and singing but without singing*.

Having introduced the conjunction and its adjoint, we can now see that we do not want $\otimes$ to be idempotent. First, if $\otimes$ is idempotent, then we obtain both $\alpha \longrightarrow \alpha \otimes \alpha$ and $\alpha \otimes \alpha \longrightarrow \alpha$. Although the latter might seem desirable, i.e., if one does α together with α it logically follows that one does α, the former is not. One of the main concern of Girard when developing linear logic was to be able to deal with limited resources. That said, $\alpha \longrightarrow \alpha \otimes \alpha$ violates this principle, and hence idempotence must be rejected. But moreover, since $\ominus$ is the adjoint of $\otimes$, idempotence would give us that from $\alpha \otimes \alpha \longrightarrow \alpha$ (which seemed a desirable consequence) one can infer $\alpha \longrightarrow \alpha \ominus \alpha$, which is clearly undesirable from a philosophical perspective: doing the action α logically implies doing α without doing α? As a result, it is our view that a proper Action Logic should not be constructed from a Cartesian closed deductive system but should rather be constructed (minimally) upon a symmetric monoidal closed deductive system, with sequence axiomatized by a monoidal deductive system.[22] Thus constructed, the logic of action conjunction can be compared to the tensor's fragment of **MLL** given that they both share a $*$-autonomous structure. However, as we will see, $\ominus$ is not reducible to a linear implication $\multimap$.

[22] By the same reasoning, if we want to be able to deal with limited resources, sequence must not be idempotent.

Action negation

The last action connective we wish to investigate is also the most intriguing from a philosophical perspective. When applied to deontic logic, action logics usually include an action connective for negative actions. But what does it mean, if it means anything, to say that one is *doing* a *negative action*? It might make sense to use negation when talking about actions[23], but as Walton (1980, p.321) pointed out one must be careful when trying to interpret action negation ontologically. Although action negation has a meaning when we are *talking* about actions, it is unclear whether or not it makes any sense to speak of an *action negation* within an Action Logic per se.

Negation of actions (or programs) are often introduced within the logic literature without further ado. Wiseman (1970), for example, introduced a connective for action negation but did not explain its meaning. He sees a problem with action logics but only because they presuppose the existence of *acts*, which he finds dubious. In the computer science literature, Meyer (1988) introduced action negation without explaining its meaning, and negation of programs were introduced by Kozen (1997) to model tests, but the significance of a *program negation* was left unexplained. Similarly, when thinking about programs, Solin (2012) introduces the negation of a program as a mere useful fictional technical tool, without saying more about what it means for a program to be *negated*.

All of these interrogations regarding action negation rely upon a much more fundamental question: what is an *action*? What is it to *act*? Von Wright (1968, p.38) answers the latter by saying that acting is to intentionally bring about or prevent a change in the world. This understanding is however not accepted unanimously. Davidson (1967, p.50) for instance rather says that acting is to bring about a state of affair, but is "neutral with respect to the question whether the action was intentional" or not. The main characteristic of an action is that it (causally) brings change in the world. This change, however, need not be voluntary: to act is not to be responsible for acting. Goldman (1970, p.46) would agree with this understanding. These conceptions are also in accordance with Pörn (1970, p.5), for whom acting is to bring about a state of affair, although this state of affair can either be the end of the action, its result or one of its consequences.

Philosophers agree that there are such things as *actions*. These *actions* are characterized by the fact that they have causal repercussions in the world. As such, actions are often associated with the description of the state they produce. This is not to say that an action is a mere change in the world (von Wright 1963, p.36), or that it can be reduced to a description. Rather, it is to say that there *are* actions, and one way to see them is to look at the changes they produce. In the STIT literature, actions are associated with the description of the state they bring about. It is presupposed that there are such things as *actions*, and these actions can be performed to obtain results in the world. For Lucas, actions are considered as mappings from the description of a state to another description. An action is thus something which can influence the evolution of the world. Even though dynamic logic tends to distinguish actions from assertions, its ontology of actions is un-

[23] Walton (1980, p.320) for example distinguished between fourteen different locutions of action negation within the language.

derstood similarly. When looking at the semantics of Segerberg (1982b) or Meyer (1988), one sees that actions are considered as sequences of events (or points in time). Actions are seen, in this case, as sequences of descriptions.

One can thus safely assume that an action has causal power and that it brings changes in the world. This change may be done actively or passively, and it might also result from *not doing*. As Goldman (1970, p.10) distinguishes between *action-types* and *action-tokens*, von Wright (1963, p.36) speaks of *act-categories* and *act-individuals*. Act-categories are classes of particular actions. For Goldman (1970, p.10), an act-token is an instance of an act-type. Since the negation of an act-type is also an act-type (Goldman 1970, p.18), it follows that an instance of the negation of an act-type is an act-token. Hence, Goldman (1970, pp.47-8) admits that there are such things as *negative acts*, but he leaves the matter open since it brings difficult problems. Von Wright on the other hand always sees an action as *positive*. He distinguishes between *productive* and *preventive* actions (von Wright 1968, p.38). One can either produce by bringing about that the state stays the same or by bringing about a change, and one can prevent by leaving the state unchanged or by letting a change happen.

This brings us to the notions of *omission*, *refraining* and *forbearance*. Although this might seem paradoxical, the doing of a negative action is usually understood as not doing. One *performs* a negative action by not performing said action. This non-performance can be intentional or not. If it is, we speak of a forbearance, and if not, an omission. An omission is thus understood as an intentional or unintentional not-doing (i.e., consciously not-doing or unconsciously not-doing). Mossel (2009) presented a nice review of the literature on negative actions. He argued that there are such things as *negative acts*, and they are characterized by the fact that they have causal power. Moreover, he argues that we are *responsible* for our negative actions (Mossel 2009, p.330). While an *omission* is not necessarily voluntary, *forbearance* (*refraining*) is. Hence, a negative action in Mossel's view is a form of forbearance, which is intentional. According to his understanding, omissions are not negative acts.

For von Wright (1963, p.36), *acting* is to bring about 'at will' a change in the world. On Davidson's (1967, p.50) and Goldman's (1970, p.46) accounts, *acting* should be understood independently from *intention* and *responsibility*. It is peculiar though that while Goldman asserts that the notion of action should be analyzed independently from *intentionality*, he also says (Goldman 1970, p.48) that *negative actions* are characterized by the fact that they are *intentional* omissions (i.e., forbearances). Similarly, although von Wright asserts that to *act* is to bring about a state of affair intentionally, he also says that forbearing to do something does not presuppose awareness (von Wright 1963, pp.45-6). In both cases, the authors seem to assume some form of duality between *acting* and *not-acting*. One considers that *acting* presupposes intentionality but *not-acting* does not, while the other assumes that *acting* does not presuppose intentionality but *not-acting* does.

It is noteworthy that Davidson's and Goldman's accounts are consistent with the legal understanding of *actions*. In criminal law, liability is determined on two grounds:

the *mens rea* and the *actus reus*. The *actus reus* corresponds to the action *per se* while the *mens rea* refers to the agent's mind, to his intention before and while acting. A convict can only be held criminally responsible if he *did* indeed commit the crime and it was done *voluntarily*.[24] Hence, from a legal point of view, *action* and *voluntariness* are two notions that should be treated separately.

But if *acting* does not presuppose any form of intentionality, and negative actions are a form of acting, how can one be justified to assume that negative actions are intentional *per se*? These dualities are quite puzzling. An action does not presuppose intentionality, yet not-acting, which is also a form of action, does? Are these duality inherent to the structure of actions?

In our view, no. Even though there are cases within the natural language where these dualities are at play, there are also cases where they are not. In criminal law for instance, if an agent *did* indeed commit a crime, then he did it (*actus reus*) intentionally (*mens rea*). The *actus reus* is an action that does not presupposes intentionality, but the action of committing the crime does. If the agent *did not* commit the crime in question, then either he did it unintentionally, or he did not do it, whether it was intentional or not. Thus, when one commits a crime it is intentional, but when one does not it may be intentional or not. As such, assuming that acting is intentional does not necessarily entail that not-acting is not. Similarly, if an agent respects the law, it needs not be intentional. Consider 'it is forbidden to steal'. One respects this norm either by omitting or forbearing to steal (i.e., one respects the norm by *not stealing*). But if an agent breaks the law (does not respect the law), it can be with or without intention: it is not because an agent broke the law (*actus reus*) that he will be held responsible for it. In a nutshell, one can respect the law intentionally or not, as one can *not* respect the law intentionally or not. Hence, it is not a necessary assumption that *acting* (or *not-acting*) presupposes intentionality, while *not-acting* (or *acting*) does not.[25]

It would be a mistake to presuppose that negative actions are intentional *per se*. A negative action can either be a forbearance or an omission. Otherwise, if omission is neither an act nor a negative act, how could we say that someone committed a crime by omission? There are situations where one can break the law by omission, and thus *omission* must be considered as some sort of negative action. We think that Goldman's and Davidson's conceptions of action without intentionality should be preferred to von Wright's insofar as it is more faithful to our understanding of action in criminal law, where one can act without being responsible for ones actions. This, however, also implies that negative actions are not necessarily intentional.

[24] To be precise, the *mens rea* refers to the agent's state of mind. Acting *volontarily* does not presuppose an explicit intention, but only means that the agent is in a state of mind such that he has the capacity to understand what he is doing, even though he cannot anticipate every possile outcome of his actions.

[25] Although criminal liability usually requires both an illegal act and a criminal mind, it should be noted that there are exceptions within Canadian law. First, some cases involve the notion of *absolute* liability, which depends on the act alone (see *R. v. Canning, 2004 SKPC 13* for the definition). Second, there are specific cases where murders are straightaway considered as first degree murders (e.g., the killing of a police officer or a warden acting in the course of their duties), notwithstanding the intention of the person perpetrating the act (see for instance the *Criminal Code, R.S.C., 1985, c. C-46, art. 231.4-231.6*.

Actions are thus understood minimally as things that have causal powers. An action is characterized by the fact that it affects the world. This is the minimal understanding of *action* we adopt. To *act* is to affect the world, either by changing it or by not changing it, intentionally or not. To act is to play a role within the evolution of the world. It might seem paradoxical to say that one can play a role by not playing a role, but it actually is not. By not acting, either intentionally or not, one influences (directly or not) how the world changes. This understanding allows us to speak of both positive and negative actions, the latter including omissions and forbearances.

It is sometimes argued within the literature that STIT logics should be preferred to dynamic logic insofar as they enable us to distinguish between doing the negation of an action ($[i\ stit]\neg\varphi$) and not doing an action ($\neg[i\ stit]\varphi$), which are not (and should not) be equivalent. In the STIT frameworks, we usually have $[i\ stit]\neg\varphi \supset \neg[i\ stit]\varphi$ (unless the STIT modality is not axiomatized by KD or KM), but the converse does not hold. Omission does not imply forbearance, but forbearance does imply omission. If it is true that the agent saw to it that not-α was done, then it is true that he did not see to it that α was. Note that the aforementioned duality is at play in this conditional. While the $[i\ stit]$ modality presupposes intentionality, its negation does not. It is noteworthy that this duality appears when we are *talking* about actions (i.e., when we use declarative statements). As such, this duality must be taken care of within a Propositional Action Logic.

Lenk (1977) was, to the best of our knowledge, the first to explore the connections between different lattice structures and action negation. Lenk distinguishes between three types of negations: non-execution, omission and forbearance. He argued that Boolean negation roughly represents non-execution (Lenk 1977, p.258) and that the negation within a Heyting algebra is better suited to represent forbearance (Lenk 1977, p.260). He thus concludes that a logic of action should include at least two negations, a classical and an intuitionistic one, to represent respectively non-execution (not-doing) and forbearance. The negation for non-execution is classical insofar as it respects the law of excluded middle: either an action is performed or it is not. However, it is not because an action is not performed that its negation is. Hence, negation representing forbearance is intuitionistic since to refrain from forbearing to do α is different then doing α. Doing α implies not refraining from doing it, but it is not because one is not refraining from doing α that one is doing it (since one might simply not do α).

Lenk (1977, p.262) then suggests that if one wants to represent *omission*, then one should look at a structure which is neither a Boolean nor a Heyting algebra. The distinction between an *omission* and *non-execution* is meant to distinguish between actions that are in one's power to *not do* from actions that are not. For example, Paul cannot *omit* or *refrain* to fly into the Sun, but Superman can. What is of interest to our approach is that Lenk (1977, p.263) actually suggests that an action logic could be inspired by the logic of quantum mechanics. Indeed, he hinted at the parallel that can be made between quantum mechanics and action logics, where actions operators are not distributive but are linearly (or weakly) distributive, and where actions have some form of complement.

Neither the logical tools nor the terminology were available for him at the time, but what Lenk hinted at is that one should look at a logic that has the structure of a compact closed category to model the structure of actions. This is the basic idea that guides our conception of an Action Logic, which will be obtained by making some slight modifications to the $*$-autonomous' structure.[26]

Although we came by this idea independently, Lenk's suggestion that an action logic could perhaps be constructed in the light of quantum mechanics and the properties of particles is at the core of our approach. As Lenk suggested, we will use a structure that is neither a Boolean nor a Heyting algebra to model negative actions that can be omissions. In what follows, we understand an omission as an action which is not done, whether it is intentional or not. Hence, an omission encompasses forbearance and is seen as the complement of an action. Moreover, following Lenk's suggestion, we will use an intuitionistic negation to model forbearance in a propositional action logic.

It is our view that an action possesses a *dual* or a *complement*, which can be referred to as a *negative action*. Since an action does not presuppose any intentionality, its dual also does not, and hence a negative action is understood as either an omission or a forbearance. Let α^* be the *dual* (or the *complement*) of an action α. It can be understood as its *negation*, although *dual* might be a proper unambiguous term. Dual actions are defined by:

$$\alpha^* =_{def} * \ominus \alpha$$

The dual of α is understood as *no change without* α. Consider the 'no change' situation, subtract α to it and one gets its dual. Four consequences of this definition are:

$$\alpha \otimes \alpha^* \longrightarrow * \tag{14.1}$$

$$\alpha \longrightarrow \alpha^{**} \tag{14.2}$$

$$\beta \ominus \alpha \cong \alpha^* \ominus \beta^* \tag{14.3}$$

$$\alpha \otimes \beta \cong (\beta^* \ominus \alpha)^* \tag{14.4}$$

Since neither α nor α^* presuppose intentionality, α can be seen as an action which is *present* (intentionally or not) while α^* refers to the absence of α (intentionally or not). The first consequence means that the action α together with its dual leads to *no change*. That is, they cancel each other. This is consistent with Lenk's suggestion. The second consequence means that if an action is present, then its absence is absent. The third consequence implies that β without α is isomorphic to α^* without β^*. Equation 14.4 follows from the fact that $\ominus$ is the adjoint of $\otimes$ (see Barr 1991, p.161).

Let us pursue the analogy between quantum mechanics (Hilbert spaces) and an Action Logic by considering the following property. An action and its dual in an Action Logic can be compared to a particle and its corresponding antiparticle in physics.

$$\alpha^* \otimes \beta \cong \beta \ominus \alpha \tag{14.5}$$

[26] The co-tensor will not be defined, hence we cannot speak of linear distributivity. Moreover, the adjoint to conjunctive action is not linear implication. An Action Logic will be defined as a compact deductive system (with sequence) rather than a symmetric monoidal closed deductive system.

Do we wish that our Action Logic satisfies that β without α is isomorphic to β together with α^*? This principle is actually quite plausible. It means, for example, that the action *driving without drinking a bottle of scotch* is isomorphic to *driving together with not drinking a bottle of scotch*.

There is thus a strong connection between the structure of an Action Logic and quantum mechanics since as an Hilbert space, an Action Logic can be seen as an instance of a compact closed category in virtue of equation 14.5. Some consequences of adopting 14.5 are:

$$* \longrightarrow \alpha \otimes \alpha^* \tag{14.6}$$
$$\alpha \cong \alpha^{**} \tag{14.7}$$
$$(\alpha \otimes \beta)^* \cong \alpha^* \otimes \beta^* \tag{14.8}$$
$$\alpha \otimes \beta \cong (\beta^* \otimes \alpha^*)^* \tag{14.9}$$

Equations 14.6 and 14.7 follow from the fact that the unit dualizes, equation 14.8 is a consequence of equation 14.5 (see Kelly and Laplaza 1980, p.194) and equation 14.9 follows from 14.7, 14.8 and the fact that if $\alpha \cong \beta$, then $\alpha^* \cong \beta^*$.[27]

To sum up, our proposal is to define an Action Logic where:

1. action conjunction $\otimes$ is commutative, associative, and has a unit $*$;
2. $\otimes$ possesses an adjoint $\ominus$;
3. action negation is defined by $\alpha^* =_{def} * \ominus \alpha$;
4. action negation is involutive (from equation 14.7);
5. $\alpha^* \otimes \beta \cong \beta \ominus \alpha$;
6. sequence is associative and has a unit $*$.

While sequence can be axiomatized within a monoidal deductive system, the logic of the fragment $\{\otimes, \ominus, *\}$ possesses the characteristics of a compact deductive system.

A last thing to point out is that although one might argue that the *without* connective should respect 14.10, this would violate our interpretation of action conjunction.

$$\alpha \ominus \beta \longrightarrow \alpha \tag{14.10}$$

To see why one might want the *without* connective to respect 14.10, consider the following sentence.

(a) If Paul drives his car without drinking a bottle of scotch, then he drives his car.

[27] This is quite straightforward assuming that $\alpha \cong \beta$ since $\alpha \cong \beta \cong \beta^{**} \cong * \ominus \beta^*$, and hence $\alpha \otimes \beta^* \cong *$ by the fact that $\ominus$ is the adjoint of $\otimes$, and by the same reasoning one obtains $\beta^* \cong * \ominus \alpha \cong \alpha^*$.

Despite that this might seem desirable within the natural language, recall that action conjunction is understood similarly to the multiplicative conjunction. Hence, we do not want to satisfy equation 14.11.

$$\alpha^* \otimes \beta \longrightarrow \beta \ominus \alpha \longrightarrow \beta \tag{14.11}$$

A logic of actions

As a result of our analysis, we propose to define an Action Logic as a compact deductive system with sequence. More specifically, assume a category $\mathcal{C}$ where $\mathcal{C}$-objects are actions a_i and $\mathcal{C}$-arrows are proofs (deductions). Let Act stands for the collection of all atomic actions. The language of an Action Logic $\mathcal{AL}$ is $\mathcal{L}_{\mathcal{AL}} = \{Act, \otimes, \ominus, \curvearrowright, *, (,)\}$. Well-formed formulas ($WFF_{\mathcal{AL}}$) are defined recursively by:

$$\alpha := \ a_i \mid \ * \ \mid \alpha \otimes \beta \mid \alpha \ominus \beta \mid \alpha \curvearrowright \beta$$

Definition 14.1. Sequence *is axiomatized by a monoidal deductive system from* $\mathcal{L}_{\curvearrowright} = \{Act, \curvearrowright, *, (,)\}$ *and* $WFF_{\mathcal{AL}}$.

Definition 14.2. *An* Action Logic $\mathcal{AL}$ *is defined on the grounds of* $\mathcal{L}_{\mathcal{AL}}$ *and* $WFF_{\mathcal{AL}}$ *by:*

1. *Sequence;*
2. *The fragment* $\{Act, \otimes, \ominus, *, (,)\}$ *is axiomatized by a compact deductive system, with:*
   (a) $*$ *the unit of* $\otimes$ *(which also dualizes);*
   (b) $\alpha^* =_{def} * \ominus \alpha$.

Hence, an Action Logic is understood as a compact deductive system with sequence.

A logic for action propositions

We can now define a logic for action propositions. The idea is to develop a two-sorted language that includes both descriptive propositions and action propositions. To avoid confusion, let us define the set AP of action propositions by

$$\text{if } \alpha \in WFF_{AL}, \text{ then } \boldsymbol{\alpha} \in \mathsf{AP}$$

The bold notation $\boldsymbol{\alpha}$ is used when actions are understood as declarative statements. Hence, $\boldsymbol{\alpha}$ is the proposition one uses to refer to the action α. A Propositional Action Logic is understood as a two sorted deductive system where the collection of objects is composed of members of AP and of a collection $Prop$ of descriptive atoms p_i (and where arrows are proofs). Let $\mathcal{L}_{\mathcal{PAL}} = \{\mathsf{AP}, Prop, \wedge, \vee, \Rightarrow, \top, \bot, (,)\}$ be the language of a Propositional Action Logic. Well-formed formulas ($WFF_{\mathcal{PAL}}$) are defined recursively by:

$$\varphi := \ \boldsymbol{\alpha} \mid p_i \mid \top \mid \bot \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \Rightarrow \psi$$

Definition 14.3. *A* Propositional Action Logic $\mathcal{PAL}$ *is an intuitionistic deductive system built from* $\mathcal{L}_{\mathcal{PAL}}$ *and* $WFF_{\mathcal{PAL}}$.

It goes without justification that propositional conjunction, disjunction and implication can be axiomatized within a bi-Cartesian closed deductive system. The only thing we need to show is that negation, when applied to action propositions, is not involutive. The underlying issue is whether or not we want $\neg\neg\boldsymbol{\alpha} \longrightarrow \boldsymbol{\alpha}$. Consider the action α and the action proposition $\boldsymbol{\alpha}$:

α: Paul drives his car

Saying that $\boldsymbol{\alpha}$ is false can have one of two meanings:

(1) Paul forbears to drive his car;

(2) Paul omits to drive his car.

Each time one adds a propositional negation before an action proposition, one will obtain these two possibilities. As Lenk suggested, saying that it is false that an agent forbears does not necessarily imply that the action was done since it might be false in virtue of a simple omission. Hence, the negation of the negation of $\boldsymbol{\alpha}$ will not necessarily give us $\boldsymbol{\alpha}$ back.

These two meanings can be exemplified using the distinction made within STIT frameworks between $[i\ stit]\neg\varphi$ and $\neg[i\ stit]\varphi$. As in the STIT frameworks where one wants $[i\ stit]\neg\varphi \supset \neg[i\ stit]\varphi$ but does not want the converse, one should want $\boldsymbol{\alpha} \longrightarrow \neg\neg\boldsymbol{\alpha}$ in $\mathcal{PAL}$ but not $\neg\neg\boldsymbol{\alpha} \longrightarrow \boldsymbol{\alpha}$. When understanding actions as propositions, we tend to associate actions with some descriptions of the world. However, there is a difference between the truth of a description and the fact that the truth of the description is a consequence of one's actions. As a result, the law of excluded middle does not hold when talking about actions insofar as there *is* a third possibility, namely omission. It is entirely possible that an agent omits to both α and α^*. The action $\alpha^* \otimes \alpha^{**}$ simply leads to *no change*. The propositional negation thus has an intuitionistic behaviour since we do not necessarily have that either $\neg\boldsymbol{\alpha}$ or $\neg\neg\boldsymbol{\alpha}$.

Conclusion

Summing up, we provided a categorical analysis of action logics by showing how action logics within the literature can be compared through the properties of their connectives. We began by presenting the categorical proof-theoretic framework and then showed how action logics can be seen as variations of monoidal deductive systems. As a result of our analysis, we saw that:

1. **NILL** is a bi-monoidal closed deductive system with negation;

2. **MNLL** is a bi-monoidal closed deductive system with a dualizing object;

3. Pratt's action logic $\mathcal{A}$ is a co-Cartesian monoidal closed deductive system;

4. A Kleene algebra $\mathcal{KA}$ is a distributive co-Cartesian monoidal deductive system;

5. The logic of action inherent DDL can be represented via a Classical deductive system;

6. **MLL** is a *-autonomous deductive system;

7. $\mathcal{BA}$ is a Classical deductive system.

We proposed to distinguish between an Action Logic and a Propositional Action Logic. The former represents the logical structure of actions while the latter expresses the propositional structure of the sentences we use to talk about actions. An Action Logic was defined as a compact symmetric monoidal closed deductive system with sequence (in short, a compact deductive system with sequence), and a Propositional Action Logic was defined as an intuitionistic deductive system, which is consistent with our assumption that the structure of actions is fundamentally different from the structure of the language we use to talk about actions. The reader can compare tables 14.2 and 14.3 to see the differences between our proposal and the action logics one can find within the literature.

	Ass.	Comm.	Idem.	Dist.	Adjoint(s)	Negation
$\mathcal{AL}$	$\otimes$ $\quad\curvearrowright$	$\otimes$			$\ominus$	$\alpha \cong \alpha^{**}$
$\mathcal{PAL}$	$\wedge$ $\quad\vee$	$\wedge$ $\quad\vee$	$\wedge$ $\quad\vee$	$\wedge$ over $\vee$	$\Rightarrow$	$\alpha \leq \neg\neg\alpha$

Table 14.3: Properties of action connectives (2)

1. *Choice* is not an action but can be expressed in $\mathcal{PAL}$ via $\vee$.

2. Semantical entailment and conditional actions are represented in $\mathcal{PAL}$ by $\Rightarrow$.

3. The tensor in $\mathcal{AL}$ is stronger than in **NILL**, **MNLL**, $\mathcal{A}$ and $\mathcal{KA}$.

4. The tensor in $\mathcal{AL}$ is weaker than in DDL, $\mathcal{HA}$, $\mathcal{ACT}$ and $\mathcal{BA}$.

5. There is no co-tensor in $\mathcal{AL}$.

6. As in **NILL**, **MNLL**, **MLL**, $\mathcal{ACT}$, $\mathcal{HA}$ and $\mathcal{BA}$ the tensor in $\mathcal{AL}$ possesses an adjoint.

7. Contrary to **NILL**, **MNLL**, **MLL**, $\mathcal{ACT}$, $\mathcal{HA}$ and $\mathcal{BA}$, the adjoint in $\mathcal{AL}$ is not implication.

8. Action negation in $\mathcal{AL}$ is involutive, as in **MNLL**, DDL, **MLL** and $\mathcal{BA}$.

9. Propositional negation of action is not involutive.

The closer system to our approach would be the tensor's fragment of **MLL** but without linear implication and with a radically different philosophical interpretation. Our approach contributes to the literature insofar as the Action Logic $\mathcal{AL}$ provides a new system of logic that has, to the best of our knowledge, never been explored by philosophers.[28] In addition to the fact that we provided a framework that can help to analyze and compare the different action logics, our proposal brings to light the differences between the structure of actions and the structure of action propositions by providing a structure which is radically *not* propositional for the Action Logic.

Many consequences of this approach are yet to be explored and the philosophical interpretation of these consequences are yet to be done. For future research, we intend to extend this approach to deontic logic on the ground of a three-sorted language that includes both descriptive and action propositions but also normative propositions. We will need to explore the relations between $\mathcal{AL}$ and $\mathcal{PAL}$ and see how one can be embedded into the other. We will also need to provide a proper semantics for this system. Another avenue for research will be to determine how to interpret the co-tensor within an Action Logic built upon a bi-compact deductive system. Finally, the framework will be used to analyze the paradoxes of deontic logic and it will also be applied to legal discourse to see if it can shed some light upon the validity of normative inferences.

* * *

Le présent chapitre répond au second objectif de la thèse, à savoir fournir une logique de l'action capable de représenter l'action humaine. L'intérêt d'une telle approche s'apercevra notamment lors de la construction d'un système déontic déductif au chapitre 16, où la logique d'action sera utilisée en conjonction avec les opérateurs déontiques. Pour l'instant, nous allons aborder le chapitre qui répond au troisième objectif de thèse et montre comment notre cadre de travail permet de modéliser adéquatement les inférences normatives conditionnelles.

[28] It should be noted that logics that are analogues to compact deductive systems are used in quantum physics and quantum logic. See for instance Duncan's (2006) Ph.D. thesis and further work.

Chapitre 15

Contrary-to-duty reasoning: A categorical approach

Abstract

This paper provides an analysis of contrary-to-duty reasoning from the proof-theoretical perspective of category theory. While Chisholm's paradox hints at the need of dyadic deontic logic by showing that monadic deontic logics are not able adequately to model conditional obligations and contrary-to-duties, other arguments can be objected to dyadic approaches in favour of non-monotonic foundations. We show that all these objections can be answered at once by modeling conditional obligations within a deductive system defined as a symmetric monoidal closed category. Using category theory as a foundational framework for logic, we show that it is possible to model conditional normative reasoning and conflicting obligations within a monadic approach without adding further operators or considering deontic conditionals as primitive.

Keywords: Conditional obligations, Factual detachment, Augmentation, Deontic explosion, Normative consistency, Categorical logic

Introduction

Many paradoxes have impeded deontic logic since the work of von Wright (1951). The first objection was formulated by Prior (1954) and is known as the *paradox of derived obligations*. It intended to show that von Wright's conception of commitment fails when analyzed in terms of derived obligations. In response to Prior, von Wright (1956) introduced what would later be known as *dyadic deontic logic*. Other paradoxes were formulated against von Wright's initial monadic approach[1], but the most important of them all was presented in 1963 by Chisholm. Chisholm's contrary-to-duty paradox shows that monadic frameworks cannot deal properly with conditional obligations and contrary-to-duties. As a result, dydadic deontic logics were introduced by von Wright (1967) to model condi-

[1] See McNamara (2010) for an overview.

tional normative reasoning. But even though dyadic deontic logics can solve Chisholm's puzzle, they still face important problems when trying to deal with conflicting obligations and detachment. Non-monotonic deontic logics were introduced as solutions to all these criticisms.

Since Chisholm's paradox, it is generally admitted by scholars that a monadic deontic framework without further operators cannot properly model conditional obligations without considering them as primitive. In this paper, we show that it is actually possible, although this possibility comes at a cost. Concentrating on the paradoxes that arise from the properties of the consequence relation, as opposed to the paradoxes that come from the rules governing the behaviour of the deontic operators, we show that all the usual objections against standard or dyadic deontic logic in favour of non-monotonic foundations can be met simply by removing two structural properties of the classical consequence relation. To see this, however, one needs to look at deontic logic from a categorical perspective. We thus assume category theory as a foundational framework for logic and show how a categorical analysis of the proof theory of deontic logic can help to answer these objections.

In the first section, we expose how Chisholm's paradox raises a difficulty to standard deontic logic and then show how dyadic deontic logic answers the problem. The necessity of considering the dyadic operator $O(/)$ as a primitive is discussed in conjunction with the problems of augmentation and detachment of deontic conditionals. Then, we provide an overview of the main arguments in favour of non-monotonic foundations for deontic logic. Horty's (1997) paper is taken as an example and we show how non-monotonic deontic logic can deal with these challenges. The different forms of defeasibility in contrary-to-duty reasoning introduced by van der Torre and Tan (1997) are discussed. Our conceptual framework is then presented. Category theory is taken as a foundational framework and logical systems are analyzed through the properties of their connectives, which allows us to provide a classification for different types of deductive systems. Finally, we show that all the objections against standard deontic logic and dyadic deontic logic can be met within a monadic system $\mathcal{CNR}$ defined as a symmetric monoidal closed category. We analyze some objections to standard deontic logic and show how $\mathcal{CNR}$ can handle conditional normative reasoning. We briefly compare our approach to the input-output logics of Makinson and van der Torre (2000) and to the deontic linear logic proposed by Lokhorst (1997), and we conclude with remarks for future research in the last section.

Contrary-to-duties and dyadic deontic logic

Dyadic deontic logic was introduced by von Wright (1956) as a response to Prior's (1954) criticism. Although conditional permission drew the attention of Rescher (1958, 1962), the first well defined dyadic deontic logic appeared with von Wright (1967).[2] The need for a dyadic deontic logic arises when we consider Chisholm's objection against a monadic deontic logic such as the standard system. The standard system of deontic logic (i.e., the

[2] Rescher's approach was seriously undermined by Anderson's (1959; 1962) and Robison's (1967) criticisms.

modal logic KD, see Åqvist 2002, p.155), is a refinement of von Wright's initial approach. It can be defined from the language $\mathcal{L} = \{(,), Prop, \neg, \supset, O\}$, with $Prop = \{p_1, ..., p_n, ...\}$ a denumerable set of atomic sentences and the set of well-formed formulas defined recursively by:

$$\varphi := p_i \mid \neg\varphi \mid \varphi \supset \psi \mid O\varphi$$

The modal logic KD is an extension of the classical propositional calculus, to which we add the K axiom schema for distribution, the O-necessitation rule and the axiom schema (**D**) for normative consistency.

$$\vdash_{KD} O(\varphi \supset \psi) \supset (O\varphi \supset O\psi) \tag{K}$$

$$\vdash_{KD} O\varphi \supset P\varphi \tag{\textbf{D}}$$

$$\frac{\vdash_{KD} \varphi}{\vdash_{KD} O\varphi} \text{ (O-necessitation)}$$

P represents weak permission and is defined by $P\varphi =_{def} \neg O \neg \varphi$. The other logical connectives are defined as usual.[3]

Following Åqvist's (2002) presentation, a paradox arises in a deontic logic Δ either when there is a formula φ derivable in Δ but for which the translation does not seem derivable within the natural normative language, or there is a formula φ which is not derivable in Δ but for which the translation seems derivable within our natural normative language. Chisholm's contrary-to-duty paradox can be formulated as follows.

1. Paul ought to not steal.
2. It ought to be that if Paul does not steal, then he does not give back what was stolen.
3. If Paul steals, then he ought to give back what was stolen.
4. Paul steals.

A *contrary-to-duty* is an obligation that arises in a context where a previous obligation has been violated. Hence, contrary-to-duties are meant to represent what should happen in sub-ideal situations (cf. Jones and Pörn 1985). In the standard system, Chisholm's set is translated by:

$$O\neg p \tag{15.1}$$

$$O(\neg p \supset \neg q) \tag{15.2}$$

$$p \supset Oq \tag{15.3}$$

$$p \tag{15.4}$$

[3] Note that *a* standard system is a proper extension of KD, while *the* standard system is the normal modal system KD (see Åqvist 2002, p.155).

The need to formulate the paradox using both 15.2 and 15.3 follows from the fact that the propositions within Chisholm's set are independent in the natural language, but using 15.5 and 15.6 or 15.7 and 15.8 instead of the original formulation would yield that 15.6 is a consequence of 15.1, and 15.7 is a consequence of 15.4.[4]

$$O(\neg p \supset \neg q) \tag{15.5}$$
$$O(p \supset q) \tag{15.6}$$

$$\neg p \supset O\neg q \tag{15.7}$$
$$p \supset Oq \tag{15.8}$$

The Chisholm set yields a paradox for standard deontic logic given that it is inconsistent within KD while it seems consistent in the natural language. Put differently, theorem 15.9 is derivable in KD but is not within our natural normative language, hence a paradox.

$$\vdash_{KD} \neg(O\neg p \wedge (O(\neg p \supset \neg q) \wedge ((p \supset Oq) \wedge p))) \tag{15.9}$$

Chisholm's paradox entails a contradiction within the standard system since 15.1 and 15.2 yield $O\neg q$ by deontic detachment, 15.3 and 15.4 give Oq by factual detachment, and hence we have $O\neg q \wedge Oq$, which is impossible according to the axiom schema (**D**).[5]

Dyadic deontic logic was thus introduced to model deontic conditionals. There have been many developments in dyadic deontic logic since von Wright's proposal (for an overview, see Åqvist 2002), including van Fraassen's (1972), and many approaches have been introduced to deal with conditional obligations (for a discussion, see Tomberlin 1981). For simplicity, we only present the dyadic version of KD. The dyadic deontic logic KD_{dy} can be obtained by replacing the monadic operator O within KD by the dyadic $O(\varphi/\psi)$, which reads 'φ is obligatory under conditions ψ'.[6] The well-formed formulas are defined by:

$$\varphi := p_i \mid \neg\varphi \mid \varphi \supset \psi \mid O(\varphi/\psi)$$

The dyadic K_{dy} axiom schema for distribution is:

$$\vdash_{KD_{dy}} O(\varphi \supset \rho/\psi) \supset (O(\varphi/\psi) \supset O(\rho/\psi)) \tag{K_{dy}}$$

The dyadic O_{dy}-necessitation rule is:

$$\frac{\vdash_{KD_{dy}} \varphi}{\vdash_{KD_{dy}} O(\varphi/\psi)}\ (O_{dy}\text{-necessitation})$$

[4] On this point, see Åqvist (1967), Decew (1981), Hansen (1999) and Tomberlin (1981, 1983).

[5] For other discussions on Chisholm's paradox and conditional obligations, see among others Åqvist (1967), Mott (1973), Chellas (1974), Decew (1981), Tomberlin (1981) and Niles (1997).

[6] Some authors write $O_\psi\varphi$ instead.

And the dyadic axiom schema ($\mathbf{D}_{dy}$) for normative consistency is:

$$\vdash_{KD_{dy}} O(\varphi/\psi) \supset \neg O(\neg\varphi/\psi) \tag{$\mathbf{D}_{dy}$}$$

Dyadic permission is defined by $P(\varphi/\psi) =_{def} \neg O(\neg\varphi/\psi)$. The monadic operator O is defined by $O\varphi =_{def} O(\varphi/\top)$, i.e., an unconditional obligation holds under tautologous conditions.

Chisholm's set is translated in KD_{dy} either by

$$O(\neg p) \tag{15.10}$$
$$O(\neg p \supset \neg q) \tag{15.11}$$
$$O(q/p) \tag{15.12}$$
$$p \tag{15.13}$$

or

$$O(\neg p) \tag{15.14}$$
$$O(\neg q/\neg p) \tag{15.15}$$
$$O(q/p) \tag{15.16}$$
$$p \tag{15.17}$$

The paradox is avoided since both translations are consistent within KD_{dy} and in each translation the propositions are independent from each other.[7]

Nute and Yu (1997, p.6) argued that Chisholm's paradox arises in standard deontic logic since it does not discriminate between *factual* and *deontic* detachment.[8] Indeed, in KD, the paradox follows from the fact that we can deontically detach $O\neg q$ from $O\neg p$ and $O(\neg p \supset \neg q)$, but we can also factually detach Oq from p and $p \supset Oq$. The resolution of Chisholm's paradox within a dyadic framework such as KD_{dy} can be seen through the fact that it (completely) lacks factual detachment. Since an unconditional obligation Oq cannot factually be detached from the context p and the conditional obligation $O(q/p)$, it follows that there is no normative inconsistency with $O\neg q$.

That detachment cannot be performed follows from the fact that $O(/)$ is taken as a primitive. Following Chellas (1980, p.201), the available definitions for a dyadic $O(\varphi/\psi)$ in terms of a monadic operator are:

$$\psi \supset O\varphi \tag{15.18}$$
$$O(\psi \supset \varphi) \tag{15.19}$$
$$\Box(\psi \supset O\varphi) \tag{15.20}$$

[7] For the construction of a KD_{dy}-model, see Åqvist (2002, pp.234-5) and replace the shift-reflexive and transitive accessibility relation of a $\mathbf{O}_{dy}\mathbf{S4}$-model by a serial one.

[8] See Loewer and Belzer (1983) and Vorobej (1986) for a discussion on detachment.

Here, $\Box$ is a normal K-necessity operator. The first definition uses a plain material conditional to model 'if ψ, then φ ought to be the case'. Definition 15.18 is to be avoided given that it would make $O(\varphi/\psi)$ true whenever ψ is false. Hence, it would follow that if Paul did not pay his taxes, then he has a conditional obligation to commit adultery under the condition that he pays them. The second definition uses a material conditional $\psi \supset \varphi$ within the scope of a monadic O. Such a conditional was initially used by von Wright (1951) to model the notion of *commitment*, but was also used by Hintikka (1970) to model the notion of *deontic consequence*. Definition 15.19 is however to be avoided since it makes $O(\varphi/\psi)$ true whenever $O\neg\psi$ or $O\varphi$ is. For instance, it would imply that if it is true that Paul has the obligation help his neighbour, then he also has this obligation under the conditions that it threatens his life. Similarly, it would imply that if Paul has an obligation to not steal, then he has the obligation to commit adultery under the condition that he steals. Finally, all of these definitions yield that if φ ought to be under conditions ψ, then it ought also to be under conditions $\rho \wedge \psi$ for any ρ. This last problem is also known as the problem of *strengthening the antecedent* of a deontic conditional (cf. Alchourrón 1996) or the problem of *augmentation* (cf. Jones 1991). It arises whenever one tries to model deontic conditionals with a material implication.

Although dyadic frameworks can solve Chisholm's paradox by avoiding either factual or deontic detachment, it remains that dyadic deontic logics face the (factual) detachment problem (see van Eck 1982a): either a dyadic deontic logic completely lacks factual detachment, or it has unrestricted factual detachment. However, these two situations are undesirable. The (factual) detachment problem can be expressed as follows: even though there are situations where we want to avoid detachment, there are other situations where detachment seems desirable. In Chisholm's paradox for example, we want to infer that *Paul ought to give back what was stolen* from the contrary-to-duty *if Paul steals, then he ought to give back what was stolen* and the fact that *Paul steals*. As such, it seems that if there is a conditional obligation $O(\varphi/\psi)$ and the conditions ψ are met, then one can conclude that φ ought to be. But this pattern of inference can lead to undesirable consequences. In some situations there might be other conditions ρ that could block the inference. For instance, one has a conditional obligation to help his neighbour if he is drowning, but one does not ought to help him if it puts one's life at risk. Similarly, Paul has a conditional obligation to drive his pregnant wife to the hospital if she is near delivering, but if for some reason he is drunk, then he does not. This is, in a nutshell, the (factual) detachment problem for conditional obligations: we want to allow for *restricted* detachment, when we can insure that no other information can thwart the detached obligation, but we also want to avoid *unrestricted* detachment, in case of some extra relevant information. Factual detachment is thus desirable only if it can be restricted to situations in which we can guarantee that nothing else will defeat the conditional obligation.

Contrary-to-duties and monotonicity

Detachment and augmentation bring us to non-monotonic deontic logics. A consequence relation $\vdash$ is *monotonic* when it respects the inference pattern (wk). It is an analogue of Gentzen's 1934 substructural rule weakening.

$$\frac{\varphi \vdash \psi}{\rho, \varphi \vdash \psi} \quad \text{(wk)}$$

In other words, if ψ is a consequence of φ, then adding ρ to φ will not change the fact that ψ can be derived. There are two ways to argue in favour of the non-monotony of normative inferences. The first one insists upon the non-monotonic character of conditional obligations and contrary-to-duties, the other regards the capacity of non-monotonic frameworks to deal with conflicting obligations. As such, non-monotonic deontic logic were introduced as frameworks that are better suited to represent both contrary-to-duties *and* conflicting obligations.[9]

On the one hand, the non-monotonic character of deontic conditionals appears when one considers that normative inferences do not satisfy the inference pattern (wk).

$$\frac{\psi, O(\varphi/\psi) \vdash O\varphi}{\rho, \psi, O(\varphi/\psi) \vdash O\varphi} \quad \text{(wk)}$$

Since a monotonic consequence relation satisfies (wk), it follows that if normative inferences do not satisfy (wk), then they are not monotonic. Even though $O\varphi$ could be detached from $O(\varphi/\psi)$ and ψ, it does not follow that it can be detached from $O(\varphi/\psi)$, ψ and ρ, and thus reasoning with deontic conditionals is non-monotonic. The inference pattern (wk), as we will see later, is related to the problems of augmentation and (factual) detachment.

On the other hand, it can be argued that dyadic and monadic deontic logics should be rejected since they cannot allow for conflicting obligations. In both KD and KD_{dy}, it is impossible to have a situation where φ and $\neg\varphi$ are obligatory: the axiom schema for normative consistency makes it impossible to have a situation where $O\varphi \wedge O\neg\varphi$. As such, standard deontic logic (or dyadic KD) is unable to model conflicting obligations.

But more importantly, a conflict of obligations in these systems entails that anything can be seen as obligatory since we have the following theorems for any ψ.

$$\vdash_{KD} (O\varphi \wedge O\neg\varphi) \supset O\psi \tag{15.21}$$
$$\vdash_{KD_{dy}} (O(\varphi/\rho) \wedge O(\neg\varphi/\rho)) \supset O\psi \tag{15.22}$$

[9] See Nute (1997) for an introduction to non-monotonic deontic logics and Antonelli (2012) for an introduction to non-monotonic logics.

This is known as the problem of *deontic explosion*. Hence, even if one can deal with Chisholm's paradox through dyadic deontic logic, it remains unable to reason with conflicting obligations (Ryu and Lee 1997, p.123). Horty (1997) for instance argues that a non-monotonic framework should be preferred to the monotonic consequence relation of KD or KD_{dy} seeing that it does not allow for conflicting obligations.[10] Indeed, the axioms (**D**) and ($\mathbf{D}_{dy}$) are (respectively) propositionally equivalent to 15.23 and 15.24.

$$\neg(O\varphi \wedge O\neg\varphi) \tag{15.23}$$

$$\neg(O(\varphi/\psi) \wedge O(\neg\varphi/\psi)) \tag{15.24}$$

As a result, conflicting obligations are logically impossible within KD or KD_{dy}. However, there are situations within real life where conflicting obligations arise, and as such a deontic logic should be able to represent normative conflicts.

In Horty's (1997, p.22) view, the incapacity of the standard system to model conflicting obligations comes from the fact that it validates an unrestricted aggregation (in his terms, agglomeration) principle, which in turn entails deontic explosion. In KD we have:

$$\frac{O\varphi \wedge O\neg\varphi \vdash \bot \qquad \bot \vdash O\psi}{O\varphi \wedge O\neg\varphi \vdash O\psi}$$

But more generally in K we have:

$$\frac{O\varphi \wedge O\neg\varphi \vdash O(\varphi \wedge \neg\varphi) \qquad O(\varphi \wedge \neg\varphi) \vdash O\psi}{O\varphi \wedge O\neg\varphi \vdash O\psi}$$

So even though one might reject the axiom schema (**D**) to model conflicting obligations within K, it remains that aggregation together with the principle *ex falso sequitur quodlibet* (from the false anything follows) entail deontic explosion.[11] In both cases, Horty argues that these two patterns of inference are to be avoided if one wants to program a machine with automated reasoning: it would not be desirable that a machine considers anything as a goal when there is a conflict regarding what it should do (Horty 1997, p.20). His conclusion is that the aggregation principle 15.25 should be restricted to cases within which no conflict arises.

$$O\varphi \wedge O\neg\varphi \vdash O(\varphi \wedge \neg\varphi) \tag{15.25}$$

Horty's (1997, p.28) solution is to define the relation of deontic consequence through Reiter's (1980) default logic. In a nutshell, the idea is to say that $O\varphi$ is a consequence of a set of ought statements Γ if and only if the proposition φ pertains to some extension $\mathcal{E}$ of the default theory $\Delta_\Gamma = \langle \mathcal{W}, \mathcal{D} \rangle$, where the set of formulas $\mathcal{W}$ is empty and the set of defaults is defined by:

$$\mathcal{D} = \{(\top : \varphi/\varphi) : O\varphi \in \Gamma\}$$

[10] See also Horty (1994).

[11] See also Goble (2004, 2009) for a discussion of normative conflicts.

This is the *credulous* strategy, which allows to conclude $O\varphi$ as soon as φ pertains to some $\mathcal{E}$, as opposed to the *skeptical* strategy, which tells us to conclude $O\varphi$ only when φ pertains to every extension $\mathcal{E}$ of Δ_Γ (Horty 1997, p.26). Generally, a default $(\varphi : \psi/\rho)$ means that in a context where φ is the case, one can conclude ψ unless ρ is the case. Thus, a default $(\top : \varphi/\varphi)$ means that in a context where we have *truth*, one can infer φ inasmuch as it is consistent with φ. An extension $\mathcal{E}$ for a default theory Δ_Γ is defined by:

$$\begin{aligned}
&\mathcal{W} \subseteq \mathcal{E}\\
&\mathcal{E} = Cn(\mathcal{E})\\
&\text{for all } (\varphi : \psi/\rho) \in \mathcal{D},\ \varphi \in \mathcal{E} \text{ and } \neg\rho \notin \mathcal{E} \Rightarrow\ \psi \in \mathcal{E}
\end{aligned}$$

Hence, an extension $\mathcal{E}$ for a default theory Δ_Γ contains its consequences and is such that it contains every conclusion that can be drawn from the default rules when the context is present and the 'unless' clause is not satisfied. As a result, an obligation can only be derived when it is consistent with the other obligations that can be derived in some extension $\mathcal{E}$ of Δ_Γ.

As the following example shows, this strategy allows for conflicting obligations. Let the default set for the default theory Δ_Γ be:

$$\mathcal{D} = \{(\top : p/p), (\top : \neg q/\neg q), (\top : p \supset q/p \supset q)\}$$

This theory possesses many extensions, depending on which proposition we consider first. For instance, if we consider $(\top : p/p)$ first, then we have $\top \in \mathcal{E}_1$ since $\mathcal{W} = \varnothing$ and $\top \in Cn(\varnothing)$, and since $\neg p \notin \mathcal{E}_1$ we obtain $p \in \mathcal{E}_1$. Then, we can either consider $(\top : p \supset q/p \supset q)$ or $(\top : \neg q/\neg q)$. Considering the former first, we get $p \supset q \in \mathcal{E}_1$ by the same reasoning and since $\mathcal{E}_1 = Cn(\{p, p \supset q\})$, we obtain $q \in \mathcal{E}_1$ and thus also $\neg\neg q \in \mathcal{E}_1$. From the third condition we cannot conclude $\neg q \in \mathcal{E}_1$ since $\neg\neg q \in \mathcal{E}$ (to do so, we would need $\neg\neg q \notin \mathcal{E}_1$). Had we considered the latter first, we would have obtained $\neg q \in \mathcal{E}_2$ since $\top \in \mathcal{E}_2$ and $\neg\neg q \notin \mathcal{E}_2$, thus $\neg(p \supset q) \in \mathcal{E}_2$ since $\neg(p \supset q) \in Cn(\{p, \neg q\})$. Finally, if we begin by considering $(\top : p \supset q/p \supset q)$ and then $(\top : \neg q/\neg q)$, we get $\mathcal{E}_3 = Cn(\{p \supset q, \neg q\})$ for which $\neg p \in \mathcal{E}_3$. We therefore have:

$$\begin{aligned}
\mathcal{E}_1 &= Cn(\{p, p \supset q\})\\
\mathcal{E}_2 &= Cn(\{p, \neg q\})\\
\mathcal{E}_3 &= Cn(\{p \supset q, \neg q\})
\end{aligned}$$

From the credulous strategy, one could then conclude both Op and $O\neg q$ from Γ since $p \in \mathcal{E}_1$ and $\neg q \in \mathcal{E}_2$. However, since $\neg q \wedge q$ is neither in $\mathcal{E}_1$, $\mathcal{E}_2$ or $\mathcal{E}_3$, $O(\neg q \wedge q)$ cannot be deduced from Γ.

Although some authors insist on conflicting obligations to justify non-monotonic foundations for deontic logic (e.g., Ryu and Lee 1997; Horty 1997), it can be argued that the non-monotony of normative reasoning comes from contrary-to-duty reasoning and deontic conditionals (e.g., van der Torre and Tan 1997). Indeed, conflicting obligations arise in specific *contexts*, that is, under given *circumstances* or *conditions*. As such, conflicting obligations result from augmentation and detachment. Following van der Torre and Tan (1997), conflicting obligations can arise in various situations, which lead to three different types of defeasible reasoning. *Factual defeasibility* is meant to model the violation of a contrary-to-duty (van der Torre and Tan 1997, p.103). In Chisholm's paradox for instance, the conflict of obligations comes from the violation of $O\neg p$, in which case the contrary-to-duty $p \supset Oq$ dictates what should be done in the sub-optimal situation where $O\neg p$ is violated by p. The contrary-to-duty obligation Oq does not *cancel* the previous obligation but *overshadows* it, meaning that the obligation $O\neg p$ is still in force despite its violation.

Some authors (e.g., Decew 1981) argued that the solution to Chisholm's paradox requires the introduction of temporal parameters, since the obligation Oq applies only at a moment that succeeds the violation of $O\neg p$. However, there are other forms of conditional normative reasoning that do *not* involve a temporal parameter (for the original example, see Prakken and Sergot 1996), and thus temporal deontic logic is not enough to model conditional normative reasoning (Straßer 2011, p.62).

Normative conflicts do not necessarily arise from contrary-to-duties. Consider for instance the following example of Paul and the Nazi. Suppose that Paul is hiding a Jewish family in his basement and that a Nazi comes to his house and asks him if there are Jews in the neighbourhood. Paul has a *prima facie* obligation to not lie, but in the context that lying could save the family, one could easily accept that Paul has an obligation to lie. In such situation, there is a conflict between the *prima facie* obligation to tell the truth and the conditional obligation to lie under the circumstances that it can save the family. This is not a contrary-to-duty since the principle *if it can save the family, Paul ought to lie* does not tell us what Paul should do in the sub-optimal situation that he violates his ought to not lie, but it rather tells us that he ought to violate it. This is a case of *weak overridden defeasibility* (van der Torre and Tan 1997, p.117). In this example, the *prima facie* ought is not *cancelled* but is only *overshadowed*. Paul still has a *prima facie* ought to not lie but the ought to lie takes priority in the context that it can save the family.

Finally, conflicts of obligations can also arise from conflicting norms. Consider the following fence example (cf. Prakken and Sergot 1996).

1. There should be no fence.

2. But if the cottage is by the sea, then there can be a fence.

Some norms express exceptions. The first norm tells us that in general there ought to be no fence around the cottage. However, if the cottage is by the sea, then there is a more specific norm which states that there is an exception to the previous norm, and in this case there can be a fence. The fence example shows a situation in which there is *strong overridden defeasibility* (van der Torre and Tan 1997, p.115). If the cottage is by the sea, the obligation that there be no fence is *cancelled* by the more specific norm which states that there can be a fence. The first obligation is not simply overshadowed:

the obligation that there be no fence is not still in force if the cottage is by the sea.[12]

The strategy to solve conflicts of obligations by prioritizing one over the other was inspired by Alchourrón and Makinson (1981). Van der Torre and Tan (1997) showed that there are different ways to solve conflicts of obligations depending on the context in which the conflict arises. To solve a conflict, one obligation must defeat the other, either by cancelling it or by overshadowing it. That said, there can still be situations of unsolvable conflicts. For instance, think of a mother who can only save one of her children but ought to save both. In such a situation, it is unlikely that the mother will accept that the obligation to save one child defeats the other. If this happens, one must however not be able to conclude that anything follows.

Summing up, the logic of normative reasoning faces many challenges. First, it must be able to block augmentation and allow for restricted detachment, but not allow for unrestricted detachment. Second, it must be able to model conflicting obligations. Conflicts of obligations take different forms, namely conflicts that arise in situations of contrary-to-duty reasoning, conflicts between *prima facie* and conditional obligations and conflicts that result from inconsistent norms. In such cases, a proper deontic logic must not only be able to model the consistency of these inferences but must also be able to tell which obligation holds. When conflicts of obligations are unsolvable, the logic must insure that anything does *not* follow. We will see later that all these challenges can be met within the framework of categorical logic. Before doing so, let us first expose the conceptual framework we adopt.

Category theory and deductive systems

Having presented the challenges that must be met, we now turn our attention to the framework of categorical logic. The present section is based on previous work (cf. Peterson 2014a). The notion of *category* was introduced by Eilenberg and Mac Lane (1945) to define functors and natural transformations (cf. Marquis 2013). Since then, category theory has however become a distinct, successful and fruitful field of studies. A *category* is defined as a collection of objects together with a collection of arrows, the latter having objects for domain and codomain and respecting composition and identity (Mac Lane 1971, pp.7-8), i.e.,

1. for each pair $f : x \longrightarrow y$ and $g : y \longrightarrow z$ there is $gf : x \longrightarrow z$ such that $h(gf) = (hg)f$;
2. there is an identity arrow $1_y : y \longrightarrow y$ for each object y such that for each pair $f : x \longrightarrow y$ and $g : y \longrightarrow z$, $1_y f = f$ and $g 1_y = g$.

We assume category theory as a proof-theoretical foundational framework for logic. The relation between category theory and logic can be traced back to Lambek (1958), who

[12] See van der Torre and Tan (1999, p.53) for a summary of the examples used in non-monotonic frameworks and contrary-to-duty reasoning.

introduced the *syntactic calculus*, and later defined different categorical notions in terms of *deductive systems* (Lambek 1968, 1969).

Although we follow Lambek in that we assume that there is an intimate relationship between categories and deductive systems, we will not, however, define the former in terms of the latter.[13] Instead, let us define a *deductive system* as a category whose objects are *formulas* and whose arrows are *proofs* (or *deductions*). Composition of arrows is represented by the rule (cut), which respects associativity, and the axiom (1) represents the identity arrow for each formula and it respects the identity law.

$$\frac{f : \varphi \longrightarrow \psi \qquad g : \psi \longrightarrow \rho}{gf : \varphi \longrightarrow \rho}\ (\text{cut})$$

$$\frac{}{1_\varphi : \varphi \longrightarrow \varphi}\ (1)$$

Note that at this stage the definition of a deductive system is rather minimal: there are no logical connectives and only the consequence relation is defined. A *monoidal deductive system* is a deductive system to which we add a tensor product $\otimes$ together with a unit $\top$. The tensor product is not necessarily commutative, but it is minimally associative. Following Baez and Stay (2011, p.140), it is constructed from the rules (t), (a), (l) and (r).[14]

$$\frac{f : \varphi \longrightarrow \psi \qquad g : \rho \longrightarrow \tau}{h : \varphi \otimes \rho \longrightarrow \psi \otimes \tau}\ (\text{t}) \qquad\qquad \frac{f : \varphi \longrightarrow \top \otimes \psi}{\overline{g : \varphi \longrightarrow \psi}}\ (\text{l})$$

$$\frac{f : \tau \longrightarrow (\varphi \otimes \psi) \otimes \rho}{\overline{g : \tau \longrightarrow \varphi \otimes (\psi \otimes \rho)}}\ (\text{a}) \qquad\qquad \frac{f : \varphi \longrightarrow \psi \otimes \top}{\overline{g : \varphi \longrightarrow \psi}}\ (\text{r})$$

The rules (t) and (a) make the tensor product associative, while (l) and (r) make $\top$ into the unit of $\otimes$. Note that (a), (l) and (r) are natural transformations and the tensor product respects the triangle and pentagon identities. This makes a monoidal deductive system into a monoidal category (see Mac Lane 1971, pp.161-3).

If one goes further and wants a commutative tensor product, one can define a *symmetric deductive system* as a monoidal deductive system where the tensor product respects (b) and this rule is its own inverse.[15]

$$\frac{f : \varphi \longrightarrow \psi \otimes \tau}{\overline{g : \varphi \longrightarrow \tau \otimes \psi}}\ (\text{b})$$

[13] The rational behind that shift of perspective is that category theory can be seen as a foundational framework for logic, but also for other disciplines. The reader is referred to Marquis (2009) for arguments in favour of the epistemic superiority of category theory.

[14] A double line means that the rule can be applied both ways.

[15] Rejecting the condition that (b) is its own inverse would gives us a *braided deductive system*, where the tensor is only commutative up to isomorphism (see Baez and Stay 2011, p.130). In a symmetric deductive system, (l) can be proven from (r) and (b).

The rule (b) is a natural transformation satisfying the hexagon identities. In such a case, a symmetric deductive system is an instance of a symmetric monoidal category (see Mac Lane 1971, p.253).

If one goes even further and adds the conditions that the unit is a terminal object via the axiom (!) and that the tensor product is a categorical product by the rule (Cart) to a symmetric deductive system, then one obtains a *Cartesian deductive system*. Note that (b), (t), (a) and the top-down part of (r) or (l) can all be proven from (Cart), and the bottom-up part of (l) or (r) can be proven from (Cart) and (!).[16] In this case, a Cartesian deductive system is an instance of a Cartesian category (see Mac Lane 1971, p.72 for Cartesian categories, i.e., categories with finite products and a terminal object).

$$\frac{}{!_\varphi : \varphi \longrightarrow \top}\ (!)$$

$$\frac{f : \varphi \longrightarrow \psi \qquad g : \varphi \longrightarrow \rho}{\langle f, g \rangle : \varphi \longrightarrow \psi \otimes \rho}\ (\mathrm{Cart})$$

So far, we have seen how one can construct different kinds of deductive systems by specifying the properties of the tensor product. But each of these deductive system can be augmented with other logical connectives. On the one hand, it is possible in each of the aforementioned deductive systems to introduce an adjoint to the tensor product, in which case we will speak of a *closed deductive system*. To do so, one simply has to add the rule (cl), where $\triangleright$ is a right adjoint functor to the tensor product.[17]

$$\frac{f : \varphi \otimes \psi \longrightarrow \rho}{h : \varphi \longrightarrow \psi \triangleright \rho}\ (\mathrm{cl})$$

If the deductive system is not symmetric, then there can be two different (right) adjoints to the tensor product. While a monoidal deductive system which satisfies (cl) is *left* closed, a monoidal deductive system is *right* closed when it satisfies (cl').

$$\frac{f : \varphi \otimes \psi \longrightarrow \rho}{h : \psi \longrightarrow \varphi \blacktriangleright \rho}\ (\mathrm{cl'})$$

If the monoidal deductive system is both left closed and right closed, we say that it is *biclosed*.[18] That said, if the deductive system is symmetric, then both adjoints are the same, and hence it suffices to speak of a *closed* symmetric deductive system. In the case

[16] The case of (t) might be less obvious. Proof: assume both $\varphi \longrightarrow \psi$ and $\rho \longrightarrow \tau$. By (Cart) we have $\varphi \otimes \rho \longrightarrow \varphi$, hence by (cut) $\varphi \otimes \rho \longrightarrow \psi$. Similarly we obtain $\varphi \otimes \rho \longrightarrow \tau$, hence by (Cart) $\varphi \otimes \rho \longrightarrow \psi \otimes \rho$.

[17] Note that φ, ψ and ρ are understood as meta-variables. For example, we could have $\varphi = p$, $\psi = p \triangleright q$ and $\rho = q$. The rule (cl) can be compared to the deduction theorem.

[18] See the nlab entry http://ncatlab.org/nlab/show/closed+monoidal+category.

of a symmetric deductive system, we will use the symbol $\multimap$ for the adjoint to the tensor product $\otimes$, and $\supset$ for the adjoint to the product $\wedge$ in Cartesian deductive systems.

Having the adjoint to the tensor product at our disposition, it is now possible to define negation by introducing a special object $\bot$ and defining $\neg x =_{def} x \triangleright \bot$. This definition gives us both 15.26 and 15.27, which follows from the proof below. We thus speak of a *closed deductive system with negation*.

$$\bot_\varphi : \varphi \longrightarrow \neg\neg\varphi \tag{15.26}$$

$$cl(1_{\neg\varphi})b(1_{\varphi\otimes\neg\varphi}) : \varphi \otimes \neg\varphi \longrightarrow \bot \tag{15.27}$$

$$\text{(b)}\ \dfrac{\dfrac{1_{\varphi\otimes\neg\varphi} : \varphi\otimes\neg\varphi \longrightarrow \varphi\otimes\neg\varphi}{b(1_{\varphi\otimes\neg\varphi}) : \varphi\otimes\neg\varphi \longrightarrow \neg\varphi\otimes\varphi} \quad \dfrac{\dfrac{1_{\neg\varphi} : \neg\varphi \longrightarrow \neg\varphi}{1_{\neg\varphi} : \neg\varphi \longrightarrow \varphi \triangleright \bot}\,(\text{def }\neg)}{cl(1_{\neg\varphi}) : \neg\varphi\otimes\varphi \longrightarrow \bot}\,(\text{cl})}{\dfrac{\dfrac{cl(1_{\neg\varphi})b(1_{\varphi\otimes\neg\varphi}) : \varphi\otimes\neg\varphi \longrightarrow \bot}{cl(cl(1_{\neg\varphi})b(1_{\varphi\otimes\neg\varphi})) : \varphi \longrightarrow \neg\varphi \triangleright \bot}\,(\text{cl})}{\bot_\varphi : \varphi \longrightarrow \neg\neg\varphi}\,(\text{def }\neg)}\,(\text{cut})$$

If one wants the negation to have a classical behaviour, then one simply has to assume the axiom $(\neg\neg)$ to a closed deductive system with negation, in which case one gets a *closed deductive system with classical negation*, where $\varphi \cong \neg\neg\varphi$ is a natural isomorphism.

$$\dfrac{}{\neg\neg_\varphi : \neg\neg\varphi \longrightarrow \varphi}\,(\neg\neg)$$

On the other hand, one can define the co-tensor by considering a co-monoidal deductive system, which would be an instance of the opposite of a monoidal category $\mathcal{C}^{op}$ (see Mac Lane 1971, p.33). The co-tensor $\oslash$ is obtained by reversing the arrows and replacing $\otimes$ and $\top$ respectively by $\oslash$ and $\bot$ to obtain the rules (co-t), (co-a), (co-l), (co-r) and (co-b). If the co-tensor is assumed to be a categorical co-product and $\bot$ is initial, then one obtains a co-Cartesian deductive system through the addition of the axiom (0) and the rule (co-Cart). In such cases, we speak of a *deductive system with co-tensor*.[19]

$$\dfrac{}{0_\varphi : \bot \longrightarrow \varphi}\,(0)$$

$$\dfrac{f : \varphi \longrightarrow \rho \qquad g : \psi \longrightarrow \rho}{[f, g] : \varphi \oslash \psi \longrightarrow \rho}\,(\text{co-Cart})$$

[19] One could also speak of a *co-deductive system* if one only wanted to refer to the co-tensor's fragment of the deductive system.

Following Peterson (2014a), figure 15.1 sums up the relations between the deductive systems presented so far. The strength of category theory is that it provides a foundational framework that enables us to see the common structure shared by different theories. When applied to logic, the definition of a deductive system enables us to see that different logical systems share the same categorical structure. For instance, a Heyting algebra, a distributive lattice and intuitionistic logic are simply three ways to express the same structure, namely that of a Cartesian closed category. Hence, from a categorical perspective, a Heyting algebra is similar to intuitionistic logic, although both are in fact different and are formulated in different ways. Category theory allows us to see different things as *identical* from an abstract perspective. While isomorphism is the criterion of identity *within* a category, equivalence is the criterion of identity *between* categories. The proof-theoretical analysis of deductive systems from a categorical point of view allows us to compare different logics and see that despite their differences, some share the same structural properties. In what follows, we will see that the challenges that monadic and dyadic deontic logics face can be met within a monadic approach defined as a symmetric closed deductive system with classical negation and co-product.

We conclude this section by comparing the deductive systems presented so far with the literature. We use well-known systems as representatives.

1. Lambek's (1958) syntactic calculus **SC** is an instance of a (bi)closed monoidal deductive system.

2. The multiplicative fragment of the non-commutative intuitionistic linear logic **NILL** introduced by Abrusci (1990b) is an instance of a (bi)closed monoidal deductive system with negation and co-tensor.

3. The multiplicative fragment **MNLL** of the non-commutative linear logic introduced by Abrusci (1991) and further developed by Abrusci and Ruet (2000) is an instance of a (bi)closed monoidal deductive system with classical negation and co-tensor.

4. The multiplicative fragment of intuitionistic linear logic **MILL** introduced by Girard (1987) is a closed symmetric deductive system with negation and co-tensor.

5. The multiplicative fragment of linear logic **MLL** introduced by Girard (1987) is a closed symmetric deductive system with classical negation and co-tensor.[20]

6. The proof theory of intuitionistic logic **IL** (or a Heyting algebra) corresponds to a closed Cartesian deductive system with negation and co-product.

7. The proof theory of classical logic **CL** (or a Boolean algebra) corresponds to a closed Cartesian deductive system with classical negation and co-product.

[20] Seely (1989) showed that the **MALL** fragment of linear logic (i.e., the multiplicative and additive fragment) can be defined as a $*$-autonomous category with product and co-product. In terms of deductive systems, this means a symmetric closed deductive system with classical negation and co-tensor for the multiplicative fragment together with a Cartesian deductive system with co-product for the additive one. See also Blute and Scott (2004).

$$
\begin{tikzcd}
\textbf{Deductive system} \arrow[d] & & & \\
\textbf{Monoidal} \arrow[r] \arrow[d] & \textbf{Monoidal closed} \arrow[r] \arrow[d] & \substack{\textbf{Monoidal closed}\\ \textbf{with negation}} \arrow[r] \arrow[d] & \substack{\textbf{Monoidal closed}\\ \textbf{with classical negation}} \arrow[d] \\
\textbf{Symmetric} \arrow[r] \arrow[d] & \textbf{Symmetric closed} \arrow[r] \arrow[d] & \substack{\textbf{Symmetric closed}\\ \textbf{with negation}} \arrow[r] \arrow[d] & \substack{\textbf{Symmetric closed}\\ \textbf{with classical negation}} \arrow[d] \\
\textbf{Cartesian} \arrow[r] & \textbf{Cartesian closed} \arrow[r] & \substack{\textbf{Cartesian closed}\\ \textbf{with negation}} \arrow[r] & \substack{\textbf{Cartesian closed}\\ \textbf{with classical negation}}
\end{tikzcd}
$$

Figure 15.1: Relations between deductive systems

Deductive systems and monotony

Preliminaries

The main objective of this paper is to show that the factual detachment problem and deontic explosion can be solved within a monadic approach without adding further operators or considering deontic conditionals as primitive. There have been recent developments in adaptive logic to answer the objections regarding normative conflicts (cf. Beirlaen et al. 2013; Beirlaen and Straßer 2013) and the factual detachment problem (cf. Straßer 2011). Straßer (2011) actually provided a nice discussion of contrary-to-duty reasoning and conditional obligations and proposed a dyadic adaptive framework that can properly answer the (factual) detachment problem. That said, the aim of the present paper is to show that all the objections against standard deontic logic and dyadic deontic logic can be met within a monadic framework without going non-monotonic or adaptive. This is not to say that this approach is preferable to what we find within the literature, but it is rather to show that there is a possibility that has not been explored by the community yet. We will show that the factual detachment problem and deontic explosion come from the structural properties of the logic we use to reason with deontic conditionals. From a categorical perspective, removing these properties brings us to a specific type of deductive system.

What follows can thus be understood as a categorical argument in favour of non-Cartesian foundations for conditional normative reasoning. But before going further, we need to clarify the scope of our proposal. Our aim is to show that the factual detachment problem and deontic explosion, which are the two main motivations in favour of non-monotonic foundations for deontic logic, come from the properties of the consequence relation. As a result, our analysis will be restricted to the problems that arise from the properties of the consequence relation, leaving aside other principles that should or should not govern the behaviour of the deontic operators, such as the principle of aggregation. There are a few reasons for doing so. First, although it is not the aim of the present paper to discuss this issue, we do not think that deontic logic is of the type *ought-to-be* (see von Wright 1999a for the distinction), and thus we do not want to assume unnecessary conditions regarding the propositions that should or should not be in the scope of the deontic operator. Furthermore, even if we adopt an *ought-to-do* interpretation of O, we do not think that the available action logics are satisfying to model human action (see Peterson 2014a), and as such we simply do not want to assume *any* restriction on the behaviour of O regarding the formulas that are within its scope.

But more importantly, it is our view that the factual detachment problem and deontic explosion are two problems that concern mainly the structural properties of the logic we use to model conditional normative reasoning. Even though it can be argued that deontic explosion comes from aggregation (see for example Horty 1997), which is not a property of the consequence relation but is rather a principle that concerns the behaviour of O, it remains that in the end deontic explosion comes from the fact that the logic satisfies *ex falso sequitur quodlibet* (Goble 2009, p.78), which regards the structural properties of the logic. Removing this principle might be seen as a drastic solution to solve the problem of deontic explosion, but in our view it is actually the right one. Indeed, it is *false* that from a normative contradiction anything follows. As such, a logic for normative reasoning that allows for conflicting obligations must not satisfy *ex falso sequitur quodlibet*.[21]

Since we do not concern ourselves with the rules that should or should not govern the behaviour of O, the deductive system we propose in the next section should be seen as a *foundational framework* for conditional normative reasoning, to which one can add the desired rules and axioms to model the deontic operators. Hence, the following section should not be seen as an argument in favour of a specific deontic logic but should rather be seen as a categorical argument in favour of a foundational framework for conditional normative reasoning.

The roots of the problems

Although normative conflicts can be seen as the main motivation for non-monotonic foundations in deontic logic (e.g., Horty 1997; Royakkers and Dignum 1997; Ryu and Lee 1997), we think that the non-monotony of normative inferences is best argued by the fact that deontic conditionals do not satisfy (wk).

$$\frac{\varphi \wedge (\varphi \supset O\psi) \longrightarrow O\psi}{\rho \wedge (\varphi \wedge (\varphi \supset O\psi)) \longrightarrow O\psi} \quad \text{(wk)}$$

Since (wk) represents monotony and normative inferences do not satisfy (wk), it follows that the consequence relation of normative inferences is non-monotonic.

[21] It might be argued that *ex falso sequitur quodlibet* should not be rejected since doing so would imply that we also reject the disjunctive syllogism $(\varphi \vee \psi) \wedge \neg\varphi \longrightarrow \psi$ (see for instance Goble 2004, p.79). We will return on this point in the conclusion.

The two main problems one faces when trying to model deontic conditionals are the problem of *augmentation* and the problem of (factual) *detachment*.[22] The problem of augmentation operates at two levels. First, there is the problem of augmentation in regards to (factual) detachment of deontic conditionals. In this aspect, augmentation and detachment are two related issues. For simplicity, let us concentrate only upon the consequence relation of the standard system. Standard deontic logic, as a modal logic, is an extension of classical propositional logic. Therefore, its consequence relation satisfies minimally (and trivially) the axioms and rules of a closed Cartesian deductive system with classical negation and co-product. For visual clarity, we drop the names of the arrows. Factual detachment follows from (cl):

$$\dfrac{\dfrac{}{\varphi \supset O\psi \longrightarrow \varphi \supset O\psi}\ (1)}{\varphi \wedge (\varphi \supset O\psi) \longrightarrow O\psi}\ (\mathrm{cl})$$

This in turn yields augmentation insofar as:

$$\dfrac{\dfrac{\dfrac{}{\rho \wedge (\varphi \wedge (\varphi \supset O\psi)) \longrightarrow \rho \wedge (\varphi \wedge (\varphi \supset O\psi))}\ (1)}{\rho \wedge (\varphi \wedge (\varphi \supset O\psi)) \longrightarrow \varphi \wedge (\varphi \supset O\psi)}\ (\mathrm{Cart}) \qquad \dfrac{\dfrac{}{\varphi \supset O\psi \longrightarrow \varphi \supset O\psi}\ (1)}{\varphi \wedge (\varphi \supset O\psi) \longrightarrow O\psi}\ (\mathrm{cl})}{\rho \wedge (\varphi \wedge (\varphi \supset O\psi)) \longrightarrow O\psi}\ (\mathrm{cut})$$

In this regard, standard deontic logic satisfies (wk) and allows for augmentation with respect to detachment. Put differently, it allows to strengthen the premises of a normative inference, and this is related to the monotony of the consequence relation. As we can see from the proof above, this follows from the fact that the consequence relation satisfies (Cart).

The second level at which the augmentation problem operates is represented by (aug).

$$\dfrac{\vdots}{\varphi \supset O\psi \longrightarrow (\varphi \wedge \rho) \supset O\psi} \quad (\mathrm{aug})$$

This is, strictly speaking, the problem of *augmentation* (cf. Jones 1991), also known as the problem of *strengthening the antecedent* of a deontic conditional (cf. Alchourrón 1996). Omitting the steps for symmetry and associativity (which we will do in all the proofs for the remaining of this paper), we can see that this problem also comes from (Cart).

[22] As we said earlier, there is also the problem of *deontic detachment*. See Loewer and Belzer (1983) for a discussion, but also Vorobej (1986) and Jones (1991).

$$
\begin{prooftree}
\AxiomC{}
\LeftLabel{(1)}
\UnaryInfC{$(\varphi \wedge \rho) \wedge (\varphi \supset O\psi) \longrightarrow (\varphi \wedge \rho) \wedge (\varphi \supset O\psi)$}
\LeftLabel{(Cart)}
\UnaryInfC{$(\varphi \wedge \rho) \wedge (\varphi \supset O\psi) \longrightarrow \varphi \wedge (\varphi \supset O\psi)$}
\AxiomC{}
\RightLabel{(1)}
\UnaryInfC{$\varphi \supset O\psi \longrightarrow \varphi \supset O\psi$}
\RightLabel{(cl)}
\UnaryInfC{$\varphi \wedge (\varphi \supset O\psi) \longrightarrow O\psi$}
\RightLabel{(cut)}
\BinaryInfC{$(\varphi \wedge \rho) \wedge (\varphi \supset O\psi) \longrightarrow O\psi$}
\RightLabel{(cl)}
\UnaryInfC{$\varphi \supset O\psi \longrightarrow (\varphi \wedge \rho) \supset O\psi$}
\end{prooftree}
$$

It is noteworthy that in both cases the proof only depends upon the characteristics of the propositional consequence relation.

It is also worth mentioning that unrestricted detachment comes from the fact that the consequence relation satisfies (wk), which is due to (Cart). Since the consequence relation satisfies (wk), it follows that one can (factually) detach $O\psi$ from $\varphi \supset O\psi$ as soon as φ is in the context, even though some extra relevant information ρ might be present and thwart $O\psi$. Similarly, if we assume that ψ is obligatory under condition φ, (Cart) entails that it is also obligatory under conditions $\varphi \wedge \rho$, and thus $O\psi$ can be detached from $\varphi \supset O\psi$ even though the context is $\varphi \wedge \rho$. Hence, it seems that if we were able to block augmentation while keeping detachment, we would be able to block unrestricted factual detachment.

The other arguments in favour of non-monotonic foundations for deontic logic concern the representation of normative conflicts and deontic explosion. The impossibility of normative conflicts in standard deontic logic comes from the axiom schema for normative consistency (or from the principle *ought implies can*, see Goble 2009, p.458). Following Goble (2004, p.78), deontic explosion comes from the principle *ex falso sequitur quodlibet*: from the false, anything follows. In the modal logic K, deontic explosion comes from aggregation of obligations and then *ex falso sequitur quodlibet*. Categorically speaking, this amounts to considering $\perp$ as an initial object, which is represented by the axiom (0). Thus, if we were to reject (0), $O\varphi \wedge O\neg\varphi$ would *not* entail $O\psi$ for any ψ, whether or not we accept an axiom schema for normative consistency or an axiom schema for aggregation.

In short, augmentation and unrestricted detachment come from (Cart), while deontic explosion comes from (0). From a categorical standpoint, it can thus be argued that these problems come from the fact that we use a Cartesian deductive system to model conditional normative reasoning. Identifying deontic explosion with *ex falso sequitur quodlibet* or (factual) detachment and augmentation with (wk) is not something new. What is new however is that when we look at deontic logic from a categorical perspective, we see that these problems are directly related to the two structural properties that characterize a Cartesian deductive system. From this standpoint, we have an argument in favour of non-Cartesian foundations for conditional normative reasoning. The reader who is not familiar with category theory or who is (unfortunately) against it as a foundational framework will in all likelihood not find this argument very convincing. In this eventuality, it might be relevant to look at our proposal this way: from all the discussions we find within the literature concerning (factual) detachment, augmentation, contrary-to-duty reasoning and normative conflicts, it seems impossible to model conditional normative reasoning within a monadic approach without further operators and without taking deontic conditionals as primitive. As we will see in the remainder of this paper, this is actually possible, and in this respect it is a contribution to the literature. Furthermore, our approach provides a new avenue for future research: even though it is possible to solve the problems of deontic logic by keeping its Cartesian structure and modifying it to accommodate these problems, it is possible to work 'from below' in a weaker structure (see figure 15.2 below). In this respect, $\mathcal{CNR}$ is presented as a foundational framework from which one can approach deontic logic from a different perspective.

Non-Cartesian foundations for conditional normative reasoning

Let $\mathcal{CNR}$ be a deductive system for conditional normative reasoning. It is defined from the language $\mathcal{L}_{\mathcal{CNR}} = \{(,), Prop, \otimes, \top, \multimap, \parr, \bot, O\}$, with $Prop$ a collection of atomic propositions p_i. Propositionally, the well-formed formulas are defined recursively by:

$$\varphi := \top \mid \bot \mid p_i \mid \varphi \otimes \psi \mid \varphi \multimap \psi \mid \varphi \parr \psi$$

As we said at the beginning of this section, it is not our aim to argue in favour (or against) the rules that should govern the deontic operator O, nor is it our aim to argue in favour of the type of formula that should be in its scope. Hence, assuming that ψ is suited for O, we assume that $O\psi$ is well-formed. $\mathcal{CNR}$ can thus be seen as a foundational framework, to which one can add the desired rules, axioms and restrictions for O.

Definition 15.1. *$\mathcal{CNR}$ is defined from $\mathcal{L}_{\mathcal{CNR}}$ and its well-formed formulas as a* symmetric closed deductive system with classical negation and co-tensor*.*

Thus defined, the propositional part of $\mathcal{CNR}$ (without O) is comparable to the multiplicative fragment of linear logic **MLL**. It would be wrong, however, to assume that we are working within the framework of **MLL**. Even though from a categorical standpoint the propositional part of $\mathcal{CNR}$ shares a common structure with **MLL**, it remains that $\mathcal{CNR}$ and **MLL** are two different logics.[23] This is actually an epistemological virtue of category theory: it allows us to see two different logical systems as *identical* from an abstract perspective, when we consider only their categorical structure.[24]

Our conjunction $\otimes$ shares the same properties as the multiplicative conjunction and is commutative, associative but not Cartesian. The conditional $\multimap$ is comparable to the linear implication and is the adjoint of $\otimes$. It will be used to model deontic conditionals. A conditional obligation is formalized by $\varphi \multimap O\psi$ and is read 'under the condition (or context) that φ holds, $O\psi$ holds'.

Since our aim is to show that this framework can solve all the arguments that can be objected to KD in favour of non-monotonic foundations, we will require that $\mathcal{CNR}$ satisfies the axiom schema (**D**$\otimes$) so that the comparison with KD may be more direct.

[23] For instance, the consequence relation is defined between lists of propositions in **MLL**, while it is defined only between propositions in $\mathcal{CNR}$.

[24] $\mathcal{CNR}$ and **MLL** share the same structure but are different, such as the ring $(\mathbb{N}, \times, 1, +, 0)$ and a De Morgan algebra $(\mathbb{P}, \wedge, \top, \vee, \bot)$.

$$\frac{}{O\varphi \otimes O\neg\varphi \longrightarrow \bot}\ (\mathbf{D}\otimes)$$

Recall that the (factual) detachment problem is that we want to avoid *unrestricted* detachment but we still want a form of *restricted* detachment. The first consequence of dropping (Cart) is that our consequence relation no longer satisfies (wk). Indeed, one motivation of Girard when introducing linear logic was to be able to restrict the use of the structural rule weakening (and contraction). The multiplicative conjunction in linear logic is meant to model limited resources. As a result, one cannot detach a member of the conjunction: it must be considered as a whole.[25] It is not because one has φ and ψ together that one has both independently. By rejecting (Cart), one is no longer able to detach $\varphi \otimes (\varphi \multimap O\psi)$ from $\rho \otimes (\varphi \otimes (\varphi \multimap O\psi))$, and thus one cannot apply (cut) to obtain $\rho \otimes (\varphi \otimes (\varphi \multimap O\psi)) \longrightarrow O\psi$. Abandoning (Cart) hence allows us to avoid both faces of the problem of augmentation, and thus it also enables us to avoid unrestricted detachment.

Indeed, we still have a form of detachment since we have:

$$\dfrac{\dfrac{}{\varphi \multimap O\psi \longrightarrow \varphi \multimap O\psi}\ (1)}{\varphi \otimes (\varphi \multimap O\psi) \longrightarrow O\psi}\ (\mathrm{cl})$$

However, this does not yield *unrestricted* detachment insofar as augmentation is avoided. The detachment problem for deontic conditionals is not that we *never* want detachment but rather that detachment is only desirable under certain circumstances. While *unrestricted* detachment is to be avoided, a form of *restricted* detachment is desirable when there is nothing else to defeat it or that conflicts with the obligation which is detached.[26] Hence, the problem of detachment arises from (the two faces of) augmentation. If a deontic conditional satisfies (aug) or (wk), then it entails the problem of unrestricted detachment since one will be able to detach the conditional obligation notwithstanding the extra conditions that are assumed. But the detachment of $O\psi$ is only suitable when we have (and only have) φ and we know that $O\psi$ holds under these conditions.

This is exactly what happens with $\mathcal{CNR}$. The detachment of a conditional obligation $O\psi$ from a context φ and a deontic conditional $\varphi \multimap O\psi$ can only be done when the conditions φ, but nothing else, are assumed. As such, we do not have unrestricted detachment considering that there is no deduction arrow from $\rho \otimes (\varphi \otimes (\varphi \multimap O\psi))$ to $O\psi$. The formalization of deontic conditional by $\varphi \multimap O\psi$ thus allows us to solve the (factual) detachment problem by insuring that detachment is suitable, but only when there are no other conditions in addition to the antecedent of the deontic conditional.

For example, although we can conclude from *Paul has a conditional obligation to drive his pregnant wife to the hospital if she is near delivering* and the fact that *his wife is near delivering* that *Paul ought to drive his wife to the hospital* (cf. 15.28), it cannot be

[25] In computer science, it is usually interpreted as a parallel process.

[26] On this, see Decew (1981, p.69), Loewer and Belzer (1983, p.307), Vorobej (1986, p.23), Jones (1991), Alchourrón (1996) and Bonevac (1998, pp.253-4).

deduced if we add the premise that *Paul is drunk* since there is no deduction arrow from $r \otimes (p \otimes (p \multimap Oq))$ to Oq. To conclude that Paul still ought to drive his wife, we would need to assume that the obligation to drive his wife to the hospital still holds under the conditions that Paul is drunk (i.e., $(r \otimes p) \multimap Oq)$, which is unlikely.

$$p \otimes (p \multimap Oq) \longrightarrow Oq \tag{15.28}$$

It should be emphasized that we do not think that the (ideal) obligation to drive his wife to the hospital is not still in force. In this resepct, the obligation is not simply discarded or cancelled given the situation. Rather, in the context where Paul is drunk, the obligation to drive his wife to the hospital will be overshadowed by the interdiction to drive. As such, Paul still has an obligation to drive his wife to the hospital but it is not *actual* in the sense that it should not be an obligation that guides Paul's action at this moment.

Note that from (t) we can derive 15.29.

$$r \otimes (p \otimes (p \multimap Oq)) \longrightarrow r \otimes Oq \tag{15.29}$$

This perfectly makes sense since the context implies that even though Paul is drunk, the ideal obligation to drive his wife to the hospital is still in force.[27] In this context, the obligation for Paul to drive his wife to the hospital is ideal, and he needs to evaluate his actual obligations given the fact that he is drunk. But although Paul is in no way fit to drive, that does not mean that ideally it is false that he should do it. It is not, however, an actual obligation, and as such Oq cannot be derived alone.

The same phenomenon would happen if the extra information r was irrelevant. For instance, if we added as a premise that *Paul is wearing a white t-shirt*, we would not be able to conclude from $r \otimes (p \otimes (p \multimap Oq))$ that Paul ought to drive his wife to the hospital. This might seem counter intuitive at first glance, but it actually is not.

Although we introduced $\mathcal{CNR}$ as a foundational framework for conditional normative reasoning, the application we have in mind for $\mathcal{CNR}$ is to legal reasoning. Consider for example a case at trial in a Court of justice, where (everyone will agree) conditional normative reasoning is at play. First, the judge will need to establish all the admissible relevant facts, and then he will need to establish all the relevant legislation, case law and doctrine. Only after everything is fixed will the judge be able to reason and draw a conclusion. It is this exact phenomenon that is modeled by $\mathcal{CNR}$: one cannot infer that Paul ought to drive his wife to the hospital without having all the relevant information at one's disposal. As such, when new information is added, one must first examine if the conditional obligation still holds under the new context. For instance, one would need to ask oneself whether or not Paul still has an obligation to drive his pregnant wife to the hospital if she is near delivering *and* he is wearing a white t-shirt. If the additional information is relevant and the conditional obligation still holds under the new context,

[27] See Carmo and Jones (2002) for the distinction between ideal and actual obligations.

then one will assume that $(r \otimes p) \multimap Oq$. Otherwise, if it is irrelevant, then it will simply be discarded from the context.

We will return on this point in the next section, but for now just note that the process of determining the (relevant) context and the conditions under which obligations hold is, from a legal point of view, something which cannot be done by logic alone. This process requires a careful analysis, which often appeals to the general principles that govern the interpretation of the law. Interpreting the law, however, is not something which can be done by logic alone (cf. Côté 2006), and hence if logic is to be of help to legal discourse, then it must be able to model this phenomenon.

The conjunction $\otimes$ enables us to reason from inseparable packages, which is exactly what we need when we reason in a court of justice. Verdicts are drawn from the conjunction of the admissible relevant facts and the relevant legislation (case law and doctrine) considered *all together with no other information*. The conjunction $\otimes$ provides us with a way to represent this kind of conjunction, when everything is considered as a whole and where all parts are glued together and are inseparable. From the perspective of legal discourse, this type of conjunction, considered with its adjoint, enables us to model conditional normative reasoning as it is done within a court of justice.

The logical connectives are not interpreted in terms of truth value but are rather interpreted in terms of 'things that hold'. The conjunction $\otimes$ enables us to consider altogether the context that holds, and then from $\multimap$ we can deduce the obligations that hold *exactly* under that context. This kind of conjunction makes sense within conditional normative reasoning seeing that contexts are things that must be interpreted as a whole. For instance, the context:

1. Paul is at the bar;
2. Paul struck Peter.

is not the same context as:

1. Paul is at the bar;
3. Peter attacked Paul;
2. Paul struck Peter.

While reasoning with the context 1–3, one cannot simply discard 3 from (Cart) and reason with only 1–2 on the grounds that the conjunction of 1–3 is true, hence the conjunction of 1–2 is true and thus it implies that the obligations that can be derived from context 1–2 are also true under context 1–3. When one is doing conditional normative reasoning, one is considering a context altogether, and the question one is trying to answer is "which obligation holds under the context 1–3?": not under the context 1–2, nor in the context 1–4 (whatever 4 is), but within the context 1–3. Interpreting the conjunction in terms of 'things that hold' instead of truthfulness enables us to represent that in a court of justice, one is not reasoning with *facts* but is rather reasoning with all the *admissible* facts, considered together. As such, the task is to determine which context *holds*, and what are the obligations that hold in, and exactly in, that context.

Chisholm's paradox

In order to solve Chisholm's puzzle, a system must be able to:

1. preserve the consistency of Chisholm's set;
2. preserve the independence between the sentences;
3. show how the right conclusion can be obtained.

Fortunately, $\mathcal{CNR}$ enables us to do all these things, and as such it is able to adequately model contrary-to-duty reasoning. The translation of Chisholm's set within $\mathcal{CNR}$ which is the closest to the one in KD is:

$$O\neg p \tag{15.30}$$
$$\neg p \multimap O\neg q \tag{15.31}$$
$$p \multimap Oq \tag{15.32}$$
$$p \tag{15.33}$$

From 15.34 and 15.35 we can conclude 15.36 from (t).

$$p \otimes (p \multimap Oq) \longrightarrow Oq \tag{15.34}$$
$$O\neg p \longrightarrow O\neg p \tag{15.35}$$
$$O\neg p \otimes (p \otimes (p \multimap Oq)) \longrightarrow O\neg p \otimes Oq \tag{15.36}$$

By the same reasoning we obtain:

$$(\neg p \multimap O\neg q) \otimes (O\neg p \otimes (p \otimes (p \multimap Oq))) \longrightarrow (\neg p \multimap O\neg q) \otimes (O\neg p \otimes Oq)$$

This, however, is perfectly consistent. From these premises, we obtain that it is forbidden to steal and that it is obligatory to give back the money that was stolen: we have a consistent package where the deontic conditional holds and the ideal obligation to not steal holds together with the actual obligation for Paul to give back the money that was stolen.

Moreover, all these sentences are independent from one another. The only case which is not trivial is the relationship between 15.31 and 15.33. Sentence 15.31, however, is not entailed by 15.33. To see this, consider how the proof would work:

$$
\begin{prooftree}
\AxiomC{}
\LeftLabel{(1)}
\UnaryInfC{$\neg p \longrightarrow p \multimap \bot$}
\LeftLabel{(cl)}
\UnaryInfC{$\neg p \otimes p \longrightarrow \bot$}
\AxiomC{}
\RightLabel{(0)}
\UnaryInfC{$\bot \longrightarrow O\neg q$}
\RightLabel{(cut)}
\BinaryInfC{$\neg p \otimes p \longrightarrow O\neg q$}
\RightLabel{(cl)}
\UnaryInfC{$p \longrightarrow \neg p \multimap O\neg q$}
\end{prooftree}
$$

But since $\bot$ is not understood as an initial object, we do not have (0) and thus the two sentences are independent. From this proof, we can also see that $\varphi \longrightarrow (\neg\varphi \multimap \psi)$ or $\varphi \longrightarrow (\psi \multimap \varphi)$ do not hold in $\mathcal{CNR}$ (this will be important for the discussion of Prior's paradox below).

Chisholm's paradox is a case of factual defeasibility where the obligation to not steal is not cancelled but is only overshadowed, due to the violation. Put differently, it is assumed that Oq holds under the circumstances that $O\neg p$, $\neg p \multimap O\neg q$ and p. Hence, to derive Oq, one simply has to assume 15.37 so that we can derive 15.38 by (cl) and (1).

$$(O\neg p \otimes ((\neg p \multimap O\neg q) \otimes p)) \multimap Oq \tag{15.37}$$

$$(O\neg p \otimes ((\neg p \multimap O\neg q) \otimes p)) \otimes [(O\neg p \otimes ((\neg p \multimap O\neg q) \otimes p)) \multimap Oq] \longrightarrow Oq \tag{15.38}$$

At this point, one might object that this solution of Chisholm's paradox requires *ad hoc* modifications to the premises. This would, however, be incorrect. The solution of Chisholm's puzzle requires, as we saw, that the logic be able to preserve the consistency of Chisholm's set as it is usually presented and the independence between the sentences, which $\mathcal{CNR}$ does. Moreover, it must be able to *model* conditional normative reasoning, and to do so one needs an appropriate translation from the natural language to $\mathcal{CNR}$'s language. As such, the assumption 15.37 is not an *ad hoc* modification to the premises but is rather the proper translation of the third premise instead of 15.32 (and it preserves the independence between the sentences).

But still, one might object that this is unsatisfying given that one must first evaluate the conditions under which the conditional obligation holds, and as such it requires that the reasoner makes previous inferences to encode the desired outcome within the premises of his argument, and then uses the logic to derive said outcome. Formulated this way, however, this objection presupposes a foundational position regarding what logic is and what it aims at. Of course, if one assumes that logic ultimately aims at the automation of reasoning, then one would expect that the formal system be able to do everything by itself, and thus the fact that the reasoner must first provide an appropriate translation of his argument *before* applying the logic would be seen as a shortcoming.

That being said, such a position is hardly defensible when one considers the role logic has to (and can) play within legal discourse. When interpreting the law, it is admitted (at least in Canada) that logic alone is insufficient to draw proper conclusions and that a human factor is required to determine the *intention* of the legislator (cf. Côté 2006). Moreover, one must take into consideration the aim of the judicial system, the values of the society and the direction the judicial system should follow. But this is not a process that can be done by logic alone. Hence, if logic is to be of help to the (Canadian) legal discourse, then it is not as a tool that could allow for the automation of legal reasoning and eventually replace the reasoners themselves. If logic can be useful to the (Canadian) legal discourse, it is by providing criteria that enable us to evaluate the reasoning process and discriminates between good and bad legal inferences. From a foundational standpoint, this amounts to consider logic as a tool that can help (or assist) the reasoner to evaluate the adequacy of the logical form of his inferences. Such an understanding of logic is normative rather than descriptive: it assumes that the task of logic is not to model how reasoning is done but is rather to model how it should be done. On this account, the aforementioned objection is not relevant anymore. This is not to say that one position is better than the other, but it is rather to say that a system of logic must be evaluated in regards to what it aims to do.

When one looks at (Canadian) legal discourse, one sees that conditional normative reasoning requires that we first gather all the relevant information and evaluate under which context conditional obligations hold. This is a task that cannot be done by logic alone and requires a human factor. Thus, if a logic is to be of help to model (Canadian) legal inferences, then it must make some place for this phenomenon. This is exactly what $\mathcal{CNR}$ does: it allows for detachment of conditional obligations but only in contexts where we know that the obligations hold. This is consistent with the (factual) detachment problem: we can only detach an obligation from a deontic conditional and some context when we can insure that no other information will thwart that obligation. Detachment is only suitable when *all things are considered*. Legally speaking, insuring this amounts to interpreting the law and determining the relevant context, and this is a process that cannot be done deductively by a formal system.[28]

Returning to Chisholm's paradox, the inconsistency of Chisholm's set in the standard system arises from deontic detachment since $O\neg q$ can be derived from $O\neg p$ and $O(\neg p \supset \neg q)$. Let us see what happens if instead we translate it with an ideal obligation $O\neg q$.

$$O\neg p \tag{15.39}$$

$$O\neg q \tag{15.40}$$

$$p \multimap Oq \tag{15.41}$$

$$p \tag{15.42}$$

[28] A last objection would be to say that the relations of cancellation and overshadowing are mentioned by the norms, and hence a formal system should be able to model these situations and show how the proper conclusion can be drawn from the correct premises. This would, however, be a mistake from the perspective of (Canadian) legal norms. Indeed, interpreting the law is a necessary phenomenon *because* everything is not explicit. The principles that guide the interpretation of the law are meant to insure that we can (rightfully) conclude that such obligation override another, or that it cancels it. Establishing such premises, however, requires a human factor and cannot be done by logic alone.

Since we cannot detach any member of the multiplicative conjunction and since augmentation is not satisfied, it follows that *none* of the following are derivable:

$$O\neg p \otimes (O\neg q \otimes (p \otimes (p \multimap Oq))) \longrightarrow Oq \tag{15.43}$$
$$O\neg p \otimes (O\neg q \otimes (p \otimes (p \multimap Oq))) \longrightarrow O\neg q \tag{15.44}$$
$$O\neg p \otimes (O\neg q \otimes (p \otimes (p \multimap Oq))) \longrightarrow Oq \otimes O\neg q \tag{15.45}$$

What we *do* have is

$$\dfrac{\dfrac{\dfrac{}{O\neg p \longrightarrow O\neg p} \qquad \dfrac{}{O\neg q \longrightarrow O\neg q}\,(1)}{O\neg p \otimes O\neg q \longrightarrow O\neg p \otimes O\neg q}\,(\mathrm{t}) \qquad \dfrac{\dfrac{}{p \multimap Oq \longrightarrow p \multimap Oq}\,(1)}{p \otimes (p \multimap Oq) \longrightarrow Oq}\,(\mathrm{cl})}{(O\neg p \otimes O\neg q) \otimes (p \otimes (p \multimap Oq)) \longrightarrow (O\neg p \otimes O\neg q) \otimes Oq}\,(\mathrm{t})$$

But this only yields 15.46 from the previous proof, (cut) and:

$$\dfrac{\dfrac{}{O\neg p \longrightarrow O\neg p}\,(1) \qquad \dfrac{}{Oq \otimes O\neg q \longrightarrow \bot}\,(\mathbf{D}\otimes)}{O\neg p \otimes (Oq \otimes O\neg q) \longrightarrow O\neg p \otimes \bot}\,(\mathrm{t})$$

$$(O\neg p \otimes O\neg q) \otimes (p \otimes (p \multimap Oq)) \longrightarrow O\neg p \otimes \bot \tag{15.46}$$
$$O\neg p \otimes \bot \longrightarrow \bot \tag{15.47}$$

Note, however, that 15.47 is not available in $\mathcal{CNR}$ since it follows from (0) or (Cart) and neither are available.

$$\dfrac{\dfrac{}{\bot \longrightarrow O\neg p \multimap \bot}\,(0)}{O\neg p \otimes \bot \longrightarrow \bot}\,(\mathrm{cl})$$

$$\dfrac{\dfrac{}{O\neg p \otimes \bot \longrightarrow O\neg p \otimes \bot}\,(1)}{O\neg p \otimes \bot \longrightarrow \bot}\,(\mathrm{Cart})$$

This is actually a nice property of $\mathcal{CNR}$: a conjunction which contains $\bot$ is not necessarily a contradiction. Indeed, the formula $\varphi \otimes \bot$ is not isomorphic to $\bot$ (while $\varphi \wedge \bot \cong \bot$ in a Cartesian deductive system). As such, $\bot$ can be seen as a 'failure' in $\mathcal{CNR}$ instead of a contradiction (in terms of truth values). This allows for the representation of conflicting obligations in a set of premises. It is a nice property since equation 15.46 does not entail a total failure. Even though $(O\neg p \otimes O\neg q) \otimes (p \otimes (p \multimap Oq))$ leads to a failure in conjunction with $O\neg p$, it does not imply that it fails altogether. The formula

$(O\neg p \otimes O\neg q) \otimes (p \otimes (p \multimap Oq))$ is not a 'total normative failure' or a 'total normative inconsistency'. Thus constructed, $\mathcal{CNR}$ allows us to represent normative conflicts insofar as it enables us to model that even though a part of our assumptions 'fails', it does not imply that all our assumptions fail altogether.[29] Hence, equation 15.46 enables us to represent the fact that $(O\neg p \otimes O\neg q) \otimes (p \otimes (p \multimap Oq))$ leads to 'a failure *and* $O\neg p$'. Put differently, $\mathcal{CNR}$ allows us to show that some obligations are still in force even though some part of a normative system might fail. Hence, although we have:

$$\dfrac{\dfrac{}{\varphi \longrightarrow \varphi}\,(1) \qquad \dfrac{\vdots}{\psi \otimes \neg\psi \longrightarrow \bot}}{\varphi \otimes (\psi \otimes \neg\psi) \longrightarrow \varphi \otimes \bot}\,(\mathrm{t})$$

It does not follow that:

$$\varphi \otimes (\psi \otimes \neg\psi) \longrightarrow \bot \tag{15.48}$$

We saw that Chisholm's paradox is a case of factual defeasibility in which the contrary-to-duty obligation Oq overshadows the violated (ideal) obligation $O\neg p$ and the deontically detached (ideal) obligation $O\neg q$. Thus, Oq holds under the circumstances that $O\neg p$ is violated and that $O\neg q$ is deontically detached but is overshadowed. This is translated by 15.49 and thus one can derive Oq from 15.50.

$$(O\neg p \otimes (O\neg q \otimes p)) \multimap Oq \tag{15.49}$$

$$(O\neg p \otimes (O\neg q \otimes p)) \otimes ((O\neg p \otimes (O\neg q \otimes p)) \multimap Oq) \longrightarrow Oq \tag{15.50}$$

To derive a contradiction from this one would need to assume $O\neg q$ twice in order to be able to use (t) and obtain:

$$\boldsymbol{O\neg q} \otimes [(O\neg p \otimes (O\neg q \otimes p)) \otimes ((O\neg p \otimes (O\neg q \otimes p)) \multimap Oq)] \longrightarrow Oq \otimes \boldsymbol{O\neg q} \longrightarrow \bot \tag{15.51}$$

However, 15.52 is not the context which is assumed, 15.53 is.

$$O\neg q \otimes (O\neg p \otimes (O\neg q \otimes p)) \tag{15.52}$$

$$(O\neg p \otimes (O\neg q \otimes p)) \tag{15.53}$$

Violation and contrary-to-duty reasoning can thus be formalized within $\mathcal{CNR}$ as long as we specify the conditions under which a contrary-to-duty obligation holds. As such, the paradoxical character of Chisholm's set disappear seeing that only Oq can be derived from the premises. One oddity of Chiholm's paradox is that both $O\neg q$ and Oq seem to hold simultaneously. By restricting the detachment of Oq to the proper context, this oddity is avoided given that $O\neg q$ is used to derive Oq, and as a result there is only one obligation in force.

[29] This property might be relevant to artificial intelligence and programing.

Paul and the Nazi

Conflicts between *prima facie* and conditional obligations can also be formalized. Consider the aforementioned example of Paul and the Nazi. We have the three following sentences:

1. Paul ought to not lie;
2. Paul's lies can save the family;
3. If it can save the family, Paul ought to lie.

Following the translation within standard deontic logic, this can be translated by:

$$O\neg p \tag{15.54}$$

$$q \tag{15.55}$$

$$q \multimap Op \tag{15.56}$$

It is noteworthy that this yields a failure within our framework.

$$\dfrac{\dfrac{\dfrac{\dfrac{}{q \multimap Op \longrightarrow q \multimap Op}\,(1)}{q \otimes (q \multimap Op) \longrightarrow Op}\,(\mathrm{cl}) \qquad \dfrac{}{O\neg p \longrightarrow O\neg p}\,(1)}{(q \otimes (q \multimap Op)) \otimes O\neg p \longrightarrow Op \otimes O\neg p}\,(\mathrm{t}) \qquad \dfrac{}{Op \otimes O\neg p \longrightarrow \bot}\,(\mathbf{D}\otimes)}{(q \otimes (q \multimap Op)) \otimes O\neg p \longrightarrow \bot}\,(\mathrm{cut})$$

However, contra the standard systems, this does not entail that anything is to be considered as obligatory. One of Horty's (1997) arguments in favour of non-monotonic foundations for deontic logic is that in the standard system, a conflict of obligations entails that anything can be seen as obligatory in virtue of the principle that from the false anything follows (*ex falso sequitur quodlibet*), which is KD-valid. This is deontic explosion. It is noteworthy, though, that by moving from a Cartesian deductive system to a symmetric deductive system, we not only rejected (Cart) but also left aside the axiom (0).

$$\dfrac{}{0_\varphi : \bot \longrightarrow \varphi}\,(0)$$

As a result, the following inference pattern does not hold within our approach.

$$\dfrac{\dfrac{\vdots}{(\rho \otimes (\rho \multimap O\varphi)) \otimes O\neg\varphi \longrightarrow \bot} \qquad \dfrac{}{\bot \longrightarrow O\psi}\,(0)}{(\rho \otimes (\rho \multimap O\varphi)) \otimes O\neg\varphi \longrightarrow O\psi}\,(\mathrm{cut})$$

Thus, Horty's objection is met. From a categorical point of view, deontic explosion arises when one considers $\bot$ as an initial object within a logic that aims to model normative reasoning. Hence, if one was to assume that 15.56 was an appropriate translation for the conditional obligation, one would obtain a failure but deontic explosion would be blocked given that (0) is not available.

That said, the example of Paul and the Nazi is not inconsistent within the natural language, and therefore, if we want to avoid a paradox, we need to be able to represent it within $\mathcal{CNR}$. As it happens, the failure comes from a bad translation within our language. Looking at the second premise, one is likely to accept that:

3. Even if Paul ought to not lie, if lying can save Peter's life, then Paul ought to lie.

This is translated by 15.57.

$$(O\neg p \otimes q) \multimap Op \tag{15.57}$$

In other words, when the conditional obligation is properly translated and we assume that the (actual) obligation to lie overshadows the (*prima facie*) obligation to not lie in a context where it can save the family, we obtain 15.58.

$$(O\neg p \otimes q) \otimes [(O\neg p \otimes q) \multimap Op] \longrightarrow Op \tag{15.58}$$

This is perfectly consistent and allows us to conclude that Paul should lie to save the family. To obtain an inconsistency, one would need, as in Chisholm's paradox, to assume $O\neg p$ twice to obtain:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{(O\neg p \otimes q) \multimap Op \longrightarrow (O\neg p \otimes q) \multimap Op}\,(1)}{(O\neg p \otimes q) \otimes [(O\neg p \otimes q) \multimap Op] \longrightarrow Op}\,(\mathrm{cl}) \qquad \dfrac{}{O\neg p \longrightarrow O\neg p}\,(1)}{O\neg p \otimes ((O\neg p \otimes q) \otimes [(O\neg p \otimes q) \multimap Op]) \longrightarrow Op \otimes O\neg p}\,(\mathrm{t}) \qquad \dfrac{}{Op \otimes O\neg p \longrightarrow \bot}\,(\mathbf{D}\otimes)}{O\neg p \otimes ((O\neg p \otimes q) \otimes [(O\neg p \otimes q) \multimap Op]) \longrightarrow \bot}\,(\mathrm{cut})$$

But again, it is 15.59 which is assumed and not 15.60.

$$(O\neg p \otimes q) \otimes [(O\neg p \otimes q) \multimap Op] \tag{15.59}$$

$$O\neg p \otimes ((O\neg p \otimes q) \otimes [(O\neg p \otimes q) \multimap Op]) \tag{15.60}$$

The fence example

The fence example can also be analyzed through our framework:

1. there should be no fence;

2. but if the cottage is by the sea, then there can be a fence.

The closest translation of these sentences to KD's would be:

$$O\neg q \tag{15.61}$$
$$p \multimap \neg O\neg q \tag{15.62}$$
$$p \tag{15.63}$$

These are inconsistent within $\mathcal{CNR}$ since one can derive:

$$\dfrac{\dfrac{\dfrac{}{p \multimap \neg O\neg q \longrightarrow p \multimap \neg O\neg q}\,(1)}{p \otimes (p \multimap \neg O\neg q) \longrightarrow \neg O\neg q}\,(\mathrm{cl}) \qquad \dfrac{}{O\neg q \longrightarrow O\neg q}\,(1)}{O\neg q \otimes (p \otimes (p \multimap \neg O\neg q)) \longrightarrow \neg O\neg q \otimes O\neg q}\,(\mathrm{t})$$

Since one can also derive:

$$\dfrac{\dfrac{\dfrac{}{\neg O\neg q \otimes O\neg q \longrightarrow \neg O\neg q \otimes O\neg q}\,(1)}{\neg O\neg q \otimes O\neg q \longrightarrow O\neg q \otimes \neg O\neg q}\,(\mathrm{b}) \qquad \dfrac{\dfrac{\dfrac{}{\neg O\neg q \longrightarrow \neg O\neg q}\,(1)}{\neg O\neg q \longrightarrow O\neg q \multimap \bot}\,(\mathrm{def}\ \neg)}{O\neg q \otimes \neg O\neg q \longrightarrow \bot}\,(\mathrm{cl})}{\neg O\neg q \otimes O\neg q \longrightarrow \bot}\,(\mathrm{cut})$$

It is possible to apply (cut) to obtain:

$$\dfrac{\dfrac{\vdots}{O\neg q \otimes (p \otimes (p \multimap \neg O\neg q)) \longrightarrow \neg O\neg q \otimes O\neg q} \qquad \dfrac{\vdots}{\neg O\neg q \otimes O\neg q \longrightarrow \bot}}{O\neg q \otimes (p \otimes (p \multimap \neg O\neg q)) \longrightarrow \bot}\,(\mathrm{cut})$$

Contra the two aforementioned paradoxes, the fence example follows from a propositional inconsistency instead of a normative one. That being said, depending on whether we assume that $\neg O\neg q$ overshadows $O\neg q$ or cancels it, the fence example can be solved either by properly translating the conditional permission by assuming that $\neg O\neg q$ holds under the conditions that $O\neg q$ and p, which gives us 15.64, or by discarding the obligation that there be no fence if it is cancelled, hence we get 15.65 and $O\neg q$ is not in the context anymore.

$$(O\neg q \otimes p) \otimes ((O\neg q \otimes p) \multimap \neg O\neg q) \longrightarrow \neg O\neg q \tag{15.64}$$
$$p \otimes (p \multimap \neg O\neg q) \longrightarrow \neg O\neg q \tag{15.65}$$

It is noteworthy that although we have 15.66 in classical logic, 15.67 is *not* derivable in $\mathcal{CNR}$.

$$(\varphi \wedge \psi) \supset \neg\psi \longrightarrow \neg(\varphi \wedge \psi) \tag{15.66}$$
$$(\varphi \otimes \psi) \multimap \neg\psi \longrightarrow \neg(\varphi \otimes \psi) \tag{15.67}$$

Indeed, simply notice that to derive 15.68 one must first derive 15.69 through (Cart) and then apply (Cart) again together with 15.70 to obtain 15.71, which by (cut) gives 15.72 and by (cl) yields 15.73.

$$(\varphi \wedge \psi) \supset \neg\psi \longrightarrow \neg(\varphi \wedge \psi) \tag{15.68}$$
$$(\varphi \wedge \psi) \wedge [(\varphi \wedge \psi) \supset \neg\psi] \longrightarrow \psi \tag{15.69}$$
$$(\varphi \wedge \psi) \wedge [(\varphi \wedge \psi) \supset \neg\psi] \longrightarrow \neg\psi \tag{15.70}$$
$$(\varphi \wedge \psi) \wedge [(\varphi \wedge \psi) \supset \neg\psi] \longrightarrow \psi \wedge \neg\psi \tag{15.71}$$
$$(\varphi \wedge \psi) \wedge [(\varphi \wedge \psi) \supset \neg\psi] \longrightarrow \bot \tag{15.72}$$
$$(\varphi \wedge \psi) \supset \neg\psi \longrightarrow (\varphi \wedge \psi) \supset \bot \tag{15.73}$$

In standard deontic logic, we could not assume that $\neg O \neg q$ holds under the condition $O\neg q \wedge p$ when these conditions are met since it would be inconsistent. In $\mathcal{CNR}$, however, it is possible. This is also a nice property of this system: it is possible to assume that φ does not hold under the circumstances that a package including φ holds without obtaining a failure. Normatively, this is an interesting property: even though there can be a *prima facie* or an ideal obligation that holds, when it is considered to hold together with some package that contains relevant information it might happen that as an actual obligation it does not hold. Similarly, it allows for conflicts of obligations that come from conflicting norms where one is more specific than the other: it is possible to assume that even though a general obligation holds, when it is considered together with some context it might happen that it does not anymore.

The dog example and the pragmatic oddity

The dog example is constructed from the following sentences (cf. Prakken and Sergot 1997):

1. There should be no dog.

2. However, if there is a dog, then there should be a fence.

3. There is a dog.

The usual translation of the dog example is:

$$O\neg p \tag{15.74}$$

$$p \multimap Oq \tag{15.75}$$

$$p \tag{15.76}$$

The dog example leads to what Prakken and Sergot (1996, p.95) call the *pragmatic oddity*, represented by deduction 15.77. The pragmatic oddity follows from the intuitive interpretation of O in terms of ideal worlds, where $O\varphi$ holds for a scenario when φ holds in every accessible ideal scenario. In this respect, equation 15.77 seems to suggest that in all ideal scenarios, there is no dog but there is a fence.

$$O\neg p \otimes (p \otimes (p \multimap Oq)) \longrightarrow O\neg p \otimes Oq \tag{15.77}$$

Carmo and Jones (2002, p.279) argued that a logic that aims to model contrary-to-duty reasoning must be able to avoid the pragmatic oddity. In $\mathcal{CNR}$, the pragmatic oddity can be avoided if one uses the proper translation for the contrary-to-duty obligation and specify adequately that the obligation that there be a fence holds under the circumstances that there is a dog even though there should not be a dog. By using 15.78 instead of 15.75, one avoid the pragmatic oddity and only obtains Oq.

$$(p \otimes O\neg p) \multimap Oq \tag{15.78}$$

Contraposition

It has been pointed out to us that modeling conditional obligations with $\multimap$ might lead to undesirable results when we consider contraposition. Think for instance of Horty's (1994, p.52) asparagus example, which is structurally identical to the aforementioned example of Paul and the Nazi:

1. Don't eat with your fingers.
2. If you are served asparagus, you ought to eat with your fingers.

The usual translation of these sentences is:

$$O\neg p \tag{15.79}$$

$$q \multimap Op \tag{15.80}$$

It is easy to show that in a context where asparagus is served, we obtain $\perp$. In this respect, and assuming that this is a situation where a normative conflict arises from specificity (i.e., from two conflicting norms, one more specific contradicts another more general), this situation could be solved as in the fence example. There is, however, a more disturbing consequence that can be derived from these premises. First, note that contraposition can easily be proven.

\begin{prooftree}
\AxiomC{$\vdots$}
\UnaryInfC{$\varphi \longrightarrow \psi$}
\AxiomC{}
\RightLabel{(1)}
\UnaryInfC{$\psi \multimap \bot \longrightarrow \psi \multimap \bot$}
\RightLabel{(cl)}
\UnaryInfC{$\psi \otimes (\psi \multimap \bot) \longrightarrow \bot$}
\RightLabel{(cl)}
\UnaryInfC{$\psi \longrightarrow (\psi \multimap \bot) \multimap \bot$}
\RightLabel{(cut)}
\BinaryInfC{$\varphi \longrightarrow (\psi \multimap \bot) \multimap \bot$}
\RightLabel{(cl)}
\UnaryInfC{$(\psi \multimap \bot) \otimes \varphi \longrightarrow \bot$}
\RightLabel{(cl)}
\UnaryInfC{$\psi \multimap \bot \longrightarrow \varphi \multimap \bot$}
\UnaryInfC{$\neg\psi \longrightarrow \neg\varphi$}
\end{prooftree}

From the assumption that it is forbidden to eat with your fingers and the conditional obligation to eat with your fingers when asparagus is served, one can conclude that it is false that asparagus is served in virtue of the axiom schema (**D**$\otimes$).

\begin{prooftree}
\AxiomC{}
\RightLabel{($\mathbf{D}\otimes$)}
\UnaryInfC{$O\neg p \otimes Op \longrightarrow \bot$}
\RightLabel{(cl)}
\UnaryInfC{$O\neg p \longrightarrow Op \multimap \bot$}
\UnaryInfC{$O\neg p \longrightarrow \neg Op$}
\AxiomC{}
\RightLabel{(1)}
\UnaryInfC{$q \multimap Op \longrightarrow q \multimap Op$}
\RightLabel{(t)}
\BinaryInfC{$O\neg p \otimes (q \multimap Op) \longrightarrow \neg Op \otimes (q \multimap Op)$}
\AxiomC{$\vdots$}
\UnaryInfC{$q \multimap Op \longrightarrow \neg Op \multimap \neg q$}
\RightLabel{(cl)}
\UnaryInfC{$\neg Op \otimes (q \multimap Op) \longrightarrow \neg q$}
\RightLabel{(cut)}
\BinaryInfC{$O\neg p \otimes (q \multimap Op) \longrightarrow \neg q$}
\end{prooftree}

This result is counterintuitive given that it violates the semantical dichotomy between facts and norms. As David Hume pointed out in his *Treatise on human nature*, one cannot conclude a descriptive proposition from a set of purely normative premises (see also Jørgensen's 1937 dilemma). From the last result, however, it is possible to conclude the fact that $\neg q$ from a conditional obligation and an obligation, and this, strictly speaking, violates the semantical dichotomy. This is not so damaging, however, since by assuming a deontic conditional we assumed that there is a relation between q and Op via $q \multimap Op$. If one assumes that Op holds under the context q, then one also assumes that $\neg q$ holds under the context $\neg Op$ (or that q does not hold when Op does not), and since from (**D**$\otimes$) we can conclude $\neg Op$ from $O\neg p$, it follows that if $O\neg p$ is in force, then Op does not hold, which implies that q does not hold either.

But the situation becomes cumbersome when we consider the following theorem.

\begin{prooftree}
\AxiomC{}
\RightLabel{(1)}
\UnaryInfC{$q \longrightarrow q$}
\AxiomC{$\vdots$}
\UnaryInfC{$O\neg p \otimes (q \multimap Op) \longrightarrow \neg q$}
\RightLabel{(t)}
\BinaryInfC{$q \otimes (O\neg p \otimes (q \multimap Op)) \longrightarrow q \otimes \neg q$}
\end{prooftree}

Hence, it would appear that the list $q \otimes (O\neg p \otimes (q \multimap Op))$ does not only yield a normative conflict, as it was shown by the example of Paul and the Nazi, but that it also entails a *factual* conflict, which is clearly counterintuitive (note that this can also be derived in KD).

We will show in the remainder of this section how this problem can be solved by a proper translation in $\mathcal{CNR}$. But first, let us see how we can make sense of this last result. It is noteworthy that in both cases, the conflict comes from (**D**$\otimes$). Indeed, even in the case of the 'factual' conflict, the derivation of 15.81 depends on the use of the axiom schema for normative consistency.

$$O\neg p \otimes (q \multimap Op) \longrightarrow \neg q \tag{15.81}$$

The 'factual' failure thus comes from (**D**$\otimes$). That said, it is possible to make sense of the last result if we leave aside the terms 'normative' and 'factual' failure, and only speak of a *failure*.

$$\begin{tikzcd}
\varphi \arrow[dr, "n"] \arrow[drr, "m"] \arrow[ddr, "o"'] & & \\
& q \otimes (O\neg p \otimes (q \multimap Op)) \arrow[r, "i"] \arrow[d, "f"'] & Op \otimes O\neg p \arrow[d, "j"] \\
& q \otimes \neg q \arrow[r, "g"'] & \bot
\end{tikzcd}$$

This situation actually forms a pullback square: if there are $\varphi \xrightarrow{o} q \otimes \neg q$ and $\varphi \xrightarrow{m} Op \otimes O\neg p$, then there is one and only one n that makes the above diagram commute. The path from $q \otimes (O\neg p \otimes (q \multimap Op))$ to $\bot$ is not relevant at all since $gf = ji$. The only thing that matters is that $q \otimes (O\neg p \otimes (q \multimap Op))$ implies a failure, and this fundamentally comes from (**D**$\otimes$).

Furthermore, one could argue that the aforementioned counterintuitive result only depends on a bad translation of the asparagus example within $\mathcal{CNR}$. Firstly, it can be argued that Op actually holds under circumstances that generally you should not eat with your fingers but you ought to if you are served asparagus, i.e., $(q \otimes O\neg p) \multimap Op$. Following the lines of the previous proof, we can still derive 15.82, but we do *not* get 15.83 since we only have 15.84 and we dot *not* have 15.85. To derive a contradiction, one would need to assume $O\neg p$ twice.

$$O\neg p \otimes ((O\neg p \otimes q) \multimap Op) \longrightarrow \neg(O\neg p \otimes q) \tag{15.82}$$

$$q \otimes (O\neg p \otimes ((O\neg p \otimes q) \multimap Op)) \longrightarrow \bot \tag{15.83}$$

$$q \otimes (O\neg p \otimes ((O\neg p \otimes q) \multimap Op)) \longrightarrow \neg(O\neg p \otimes q) \otimes q \tag{15.84}$$

$$q \otimes (O\neg p \otimes ((O\neg p \otimes q) \multimap Op)) \longrightarrow \neg(O\neg p \otimes q) \otimes (q \otimes O\neg p) \tag{15.85}$$

Secondly, it is also possible to argue that the asparagus example would actually be better translated by:

$$\neg q \multimap O\neg p \tag{15.86}$$

$$q \multimap Op \tag{15.87}$$

Indeed, $O\neg p$ only holds when no asparagus is served. It is easy to show from (cl) and (cut) that these premises entail both $\neg O\neg p \multimap Op$ or its contraposition $\neg Op \multimap O\neg p$ (the proof is left to the reader). This result perfectly makes sense and only states that one obligation holds when the other does not. Moreover, in order to derive a normative conflict $Op \otimes O\neg p$, one would need to assume a factual inconsistency $q \otimes \neg q$, which is unlikely.

Finally, one could also argue that the proper translation for the asparagus example is rather:

$$O\neg p \tag{15.88}$$
$$(O\neg p \otimes q) \multimap Or \tag{15.89}$$

This would mean that one ought to not eat with one's fingers, but when served asparagus one must eat *asparagus* with one's fingers (since in the context where asparagus is served, the obligation to eat with your fingers only holds for asparagus, otherwise the norm would lead to undesirable circumstances, for instance if asparagus is served as a side dish to soup). Note that even if we assumed that $r \multimap p$, one would still obtain Or by specifying that it holds under the context $(r \multimap p) \otimes (O\neg p \otimes q)$.[30]

In short, the difficulty raised by contraposition can be solved by translating properly the asparagus example within $\mathcal{CNR}$'s language and interpreting $\perp$ as a failure.

Input-output logics

One particularity of $\mathcal{CNR}$ is that, as in linear logic, its inputs can only be used once. For instance, if we were to use φ together with $\varphi \multimap \psi$ to derive ψ, then φ would not be usable anymore, since it would have been 'depleted' while going through $\varphi \multimap \psi$ and providing the result ψ. As it happens, Makinson and van der Torre (2000)[31] provide a framework which enables us to model a comparable (but different) phenomenon. Syntactically, input-output logics are defined from the following rules:

$$\text{(SI)}\ \frac{(a,x)}{(b,x)}\ (\text{with } b \vdash_{LC} a) \qquad \text{(AND)}\ \frac{(a,x)\ (a,y)}{(a, x \wedge y)} \qquad \text{(OR)}\ \frac{(a,x)\ (b,x)}{(a \vee b, x)}$$

$$\text{(WO)}\ \frac{(a,x)}{(a,y)}\ (\text{with } x \vdash_{LC} y) \qquad \text{(CT)}\ \frac{(a,x)\ (a \wedge x, y)}{(a,y)}$$

A pair (a, x) represents that an input a gives an output x. Four different systems

[30] The definition of $\mathcal{CNR}$ does not imply that it satisfies any distribution principle (cf. Goble 2009, p.451), as for example $r \multimap p \longrightarrow Or \multimap Op$. But even if it did, we would not obtain a contradiction since we would only be able to derive Op from $(O\neg p \otimes q) \multimap Or$, $O\neg p \otimes q$ and $r \multimap p$. To derive $Op \otimes O\neg p$, one would need to assume $O\neg p$ twice

[31] See also Makinson and van der Torre (2001; 2003a; 2003b).

are obtained, depending on which set of rules we adopt.

$$R_1 = \{(SI), (AND), (WO)\}$$
$$R_2 = \{(SI), (AND), (WO), (OR)\}$$
$$R_3 = \{(SI), (AND), (WO), (CT)\}$$
$$R_4 = \{(SI), (AND), (WO), (OR), (CT)\}$$

The important points that must be underlined are that in these systems, the inputs are not found within the outputs and contraposition does not hold (Makinson and van der Torre 2000, p.407). There is a fundamental difference between $\mathcal{CNR}$ and input-output logics given that in the former, an input can only be used once, while in the latter it can be used twice as long as it is considered as an input.[32] That said, although the addition of (CT) provides a system where the output is reusable, none of these system satisfy (I).

$$(\mathrm{I})\,\frac{}{(x, x)}$$

As such, inputs cannot be considered as outputs, unless the logics satisfies (I). In this respect, input-output logics can be compared to $\mathcal{CNR}$ given that what comes *in* does not necessarily come *out*.

The addition of the rule (CT) yields that output can be reused as input. Replacing an input-output pair (a, x) by a deduction arrow $a \longrightarrow x$, one can easily see that (CT) and (AND) do not hold in $\mathcal{CNR}$.[33] The rule (AND) enables one to recover a Cartesian structure, which our approach does not satisfy.

From a categorical point of view, there is a fundamental distinction between Cartesian and monoidal structures. Makinson's and van der Torre's approach works very well, as many others do. Nonetheless, these approaches use a Cartesian structure. This is not to say that $\mathcal{CNR}$ is better suited than other Cartesian frameworks to deal with conditional normative reasoning: our aim is to illustrate that even though some Cartesian approaches work perfectly fine, there are other possibilities when one looks at deontic logic from a categorical perspective. Instead of working 'over' the line of figure 15.2 by keeping the Cartesian structure and modifying the logic to accommodate the problems that arise, one can work 'below' the line within another weaker structure.

[32] For instance, we do not have $\varphi \otimes ((\varphi \multimap \psi) \otimes (\varphi \multimap \rho)) \longrightarrow \psi \otimes \rho$ in $\mathcal{CNR}$ but we do have a derivation of $(p, q \wedge r)$ from (p, q) and (p, r). I am indebted to the comments of an anonymous referee for this precision.

[33] Strictly speaking, *constrained* input-output logics, and not simply input-output logics, are meant to model contrary-to-duty reasoning and normative conflicts (see Makinson and van der Torre 2001 for constrained input-output logics). In a nutshell, the idea is to impose a constraint that guarantees that the outputs are consistent with each other and with the inputs. We do not discuss further constrained input-output logics given that they are built from the same syntactical rules than input-output logics, and as such the comparison we make is sufficient.

Cartesian

– –

Monoidal

Figure 15.2: Monoidal and Cartesian categories

Deontic linear logic

We conclude this section by discussing the only approach within the literature which, to the best of our knowledge, is comparable to $\mathcal{CNR}$. Lokhorst (1997) proposed an Andersonian reduction[34] based on the **MALL** fragment of linear logic (i.e., the multiplicative and additive fragment, without exponentials) to solve the paradoxes that come from the idempotence of the conjunction in standard deontic logic. An operation $\circ$ is *idempotent* if φ is isomorphic to $\varphi \circ \varphi$. In the standard system, the idempotence of the conjunction comes from (Cart).[35]

$$\dfrac{(1)\,\dfrac{}{\varphi \longrightarrow \varphi} \qquad \dfrac{}{\varphi \longrightarrow \varphi}\,(1)}{\varphi \longrightarrow \varphi \otimes \varphi}\,(\text{Cart}) \qquad\qquad \dfrac{\dfrac{}{\varphi \otimes \varphi \longrightarrow \varphi \otimes \varphi}\,(1)}{\varphi \otimes \varphi \longrightarrow \varphi}\,(\text{Cart})$$

The idempotence of the conjunction in standard deontic logic becomes problematic when one looks at the two following theorems.

$$\vdash_{KD} O\varphi \equiv (O\varphi \wedge O\varphi) \tag{15.90}$$

$$\vdash_{KD} O\varphi \equiv O(\varphi \wedge \varphi) \tag{15.91}$$

As it happens, it is not because one has an obligation to do φ that one has two obligations to do φ, which is implied by 15.90, and it is not because one has an obligation to do φ once that one also has an obligation do to φ twice, which is implied by theorem 15.91. To solve this, Lokhorst's (1997) proposal is to use the **MALL** fragment of linear logic (see footnote 20 above for the definition) together with a reduction of the deontic operators to a constant δ that represents what is desired. This kind of reduction of deontic logic to alethic logic is inspired by the work of Kanger (1957, 1971) and Anderson (1958a). The definition of $O\varphi$ is $\delta \multimap \varphi$, that is, if φ is obligatory, then if 'what is desired' is true, φ follows. Since the multiplicative conjunction is defined within a symmetric closed deductive system, it follows that it does not satisfy (Cart), and hence that it does not satisfy idempotence, thus avoiding the two aforementioned theorems.

The paradoxes of deontic logic are not analyzed by Lokhorst but a few theorems and non-theorems are listed (see Lokhorst 1997, p.6). Among the theorems of deontic linear

[34] An anonymous referee pointed out to me that Lokhorst does not use an Andersonian reduction, which would rather use a constant σ for *sanctions* (or a *bad thing*), but uses a reduction that was introduced by Kanger (1971). See McNamara (2010) for a comparison.

[35] Similarly, the idempotence of the disjunction comes from (co-Cart).

logic we find:

$$(\varphi \multimap \psi) \multimap (O\varphi \multimap O\psi) \tag{15.92}$$
$$(O\varphi \wedge O\psi) \cong O(\varphi \wedge \psi) \tag{15.93}$$
$$(\varphi \multimap O\psi) \cong O(\varphi \multimap \psi) \tag{15.94}$$

Theorem 15.92 represents deontic consequence and theorem 15.93 implies that O still satisfies aggregation with the additive conjunction, although it does not with the multiplicative one. Theorem 15.94 is a rather weird consequence of deontic linear logic, which blurs the distinction between a conditional obligation and an obligation which is a conditional. Note that in KD neither of the following is a theorem.

$$(\varphi \supset O\psi) \supset O(\varphi \supset \psi) \tag{15.95}$$
$$O(\varphi \supset \psi) \supset (\varphi \supset O\psi) \tag{15.96}$$

It is actually a very unpleasant consequence of deontic linear logic since it entails:

$$\varphi \multimap O\psi \cong O(\varphi \multimap \psi) \cong O(\neg\psi \multimap \neg\varphi) \cong \neg\psi \multimap O\neg\varphi \tag{15.97}$$

This makes deontic linear logic an unlikely candidate to model contrary-to-duty reasoning and conditional obligations. Counter-examples to 15.97 are quite easy to construct. By contraposition, 15.97 yields 15.98.

$$\varphi \multimap O\psi \cong \neg O\neg\varphi \multimap \psi \tag{15.98}$$

Counter-example: if Paul is in Canada, he ought to respect the law, therefore if it is permitted for Paul to be in Canada (which it is, if he is a citizen), it follows that Paul respects the law. Unfortunately for deontic linear logic, many Canadian citizens are in prison.

Theorem 15.97 entails undesirable consequences for many of the examples using conditional obligations. For brevity's sake, we will limit our discussion to Chisholm's case. Chisholm's puzzle within deontic linear logic is translated by

$$O\neg p \tag{15.99}$$
$$\neg p \multimap O\neg q \tag{15.100}$$
$$p \multimap Oq \tag{15.101}$$
$$p \tag{15.102}$$

In virtue of theorem 15.94, it makes no difference whether or not we use 15.103 or 15.104 instead of 15.100 or 15.101.

$$O(\neg p \supset \neg q) \tag{15.103}$$
$$O(p \supset q) \tag{15.104}$$

A first thing to note is that we do *not* have the usual form of deontic detachment in deontic linear logic. To see this, simply note that although in KD we have 15.105, $\varphi \multimap \rho$ is *not* a consequence of 15.106.

$$(\varphi \supset \psi) \wedge (\varphi \supset (\psi \supset \rho)) \vdash_{KD} \varphi \supset \rho \tag{15.105}$$

$$(\varphi \multimap \psi) \otimes (\varphi \multimap (\psi \multimap \rho)) \tag{15.106}$$

This follows from the properties of the multiplicative conjunction: to derive ρ from 15.106, one would need to assume φ twice. Hence, what we have is:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\varphi \multimap \psi \longrightarrow \varphi \multimap \psi}\,(1)}{\varphi \otimes (\varphi \multimap \psi) \longrightarrow \psi}\,(\mathrm{cl}) \qquad \dfrac{\dfrac{}{\varphi \multimap (\psi \multimap \rho) \longrightarrow \varphi \multimap (\psi \multimap \rho)}\,(1)}{\varphi \otimes (\varphi \multimap (\psi \multimap \rho)) \longrightarrow \psi \multimap \rho}\,(\mathrm{cl})}{[\varphi \otimes (\varphi \multimap \psi)] \otimes [\varphi \otimes (\varphi \multimap (\psi \multimap \rho))] \longrightarrow \psi \otimes (\psi \multimap \rho)}\,(\mathrm{t}) \qquad \dfrac{\dfrac{}{\psi \multimap \rho \longrightarrow \psi \multimap \rho}\,(1)}{\psi \otimes (\psi \multimap \rho) \longrightarrow \rho}\,(\mathrm{cl})}{[\varphi \otimes (\varphi \multimap \psi)] \otimes [\varphi \otimes (\varphi \multimap (\psi \multimap \rho))] \longrightarrow \rho}\,(\mathrm{cut})}{[(\varphi \multimap \psi)] \otimes [\varphi \otimes (\varphi \multimap (\psi \multimap \rho))] \longrightarrow \varphi \multimap \rho}\,(\mathrm{cl})$$

Therefore, deontic detachment is only possible when δ is assumed, in which case we get:

$$\delta \otimes [(\delta \multimap \varphi) \otimes (\delta \multimap (\varphi \multimap \psi))] \longrightarrow \delta \multimap \psi \tag{15.107}$$

By definition, this enables us to derive $O\psi$ from 15.108.

$$\delta \otimes (O\varphi \otimes O(\varphi \multimap \psi)) \tag{15.108}$$

Otherwise, the only form of deontic detachment available is:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{O\varphi \longrightarrow O\varphi}\,(1) \qquad \dfrac{\vdots}{O(\varphi \multimap \psi) \longrightarrow \varphi \multimap O\psi}\,(15.97)}{O\varphi \otimes O(\varphi \multimap \psi) \longrightarrow O\varphi \otimes (\varphi \multimap O\psi)}\,(\mathrm{t})}{O\varphi \otimes O(\varphi \multimap \psi) \longrightarrow (\delta \multimap \varphi) \otimes (\varphi \multimap (\delta \multimap \psi))} \qquad \dfrac{\vdots}{(\delta \multimap \varphi) \otimes (\varphi \multimap (\delta \multimap \psi)) \longrightarrow \delta \multimap (\delta \multimap \psi)}}{O\varphi \otimes O(\varphi \multimap \psi) \longrightarrow \delta \multimap (\delta \multimap \psi)}\,(\mathrm{cut})}{O\varphi \otimes O(\varphi \multimap \psi) \longrightarrow OO\psi}$$

It follows from the transitivity of $\multimap$, which follows from the proof below.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\varphi \multimap \psi \longrightarrow \varphi \multimap \psi}\,(1)}{\varphi \otimes (\varphi \multimap \psi) \longrightarrow \psi}\,(\mathrm{cl}) \qquad \dfrac{\dfrac{\dfrac{}{\psi \multimap \rho \longrightarrow \psi \multimap \rho}\,(1)}{\psi \otimes (\psi \multimap \rho) \longrightarrow \rho}\,(\mathrm{cl})}{\psi \longrightarrow (\psi \multimap \rho) \multimap \rho}\,(\mathrm{cl})}{\varphi \otimes (\varphi \multimap \psi) \longrightarrow (\psi \multimap \rho) \multimap \rho}\,(\mathrm{cut})}{(\psi \multimap \rho) \otimes (\varphi \otimes (\varphi \multimap \psi)) \longrightarrow \rho}\,(\mathrm{cl})}{(\psi \multimap \rho) \otimes (\varphi \multimap \psi) \longrightarrow \varphi \multimap \rho}\,(\mathrm{cl})$$

As a result, the inconsistency of Chisholm's set can be avoided in deontic linear logic if one does *not* presuppose δ. If one was to suppose it, however, we would obtain a conflict of obligations by:

$$\dfrac{\dfrac{\vdots}{\delta\otimes(O\neg p\otimes(\neg p\multimap O\neg q))\longrightarrow O\neg q}\,(15.108)\qquad \dfrac{\dfrac{}{p\multimap Oq\longrightarrow p\multimap Oq}\,\text{(1)}}{p\otimes(p\multimap Oq)\longrightarrow Oq}\,\text{(cl)}}{(\delta\otimes(O\neg p\otimes(\neg p\multimap O\neg q)))\otimes(p\otimes(p\multimap Oq))\longrightarrow O\neg q\otimes Oq}\,\text{(t)}$$

This entails by (cut) and contraposition that 'if we do have the desired result, then we do not have the desired result' since:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\delta\multimap q\longrightarrow\delta\multimap q}\,\text{(1)}}{\delta\otimes(\delta\multimap q)\longrightarrow q}\,\text{(cl)}\qquad \dfrac{\dfrac{\dfrac{}{q\multimap\neg\delta\longrightarrow q\multimap\neg\delta}\,\text{(1)}}{q\otimes(q\multimap\neg\delta)\longrightarrow\neg\delta}\,\text{(cl)}}{q\longrightarrow(q\multimap\neg\delta)\multimap\neg\delta}\,\text{(cl)}}{\delta\otimes(\delta\multimap q)\longrightarrow(q\multimap\neg\delta)\multimap\neg\delta}\,\text{(cut)}}{(q\multimap\neg\delta)\otimes(\delta\otimes(\delta\multimap q))\longrightarrow\neg\delta}\,\text{(cl)}}{(q\multimap\neg\delta)\otimes(\delta\multimap q)\longrightarrow\delta\multimap\neg\delta}\,\text{(cl)}$$

This leads to an impasse since it entails that no desired result can come out of Chisholm's paradox, even though contrary-to-duties are meant to provide the desired outcome in sub-ideal situations.

Although deontic linear logic has some difficulties to model Chisholm's paradox, it remains that it is able to avoid Prior's. Prior's 1954 paradox of derived obligations aimed to show that von Wright's notion of *commitment*, which was represented by $O(\varphi\supset\psi)$ and meant 'φ commits us to ψ', is inadequate in contexts of derived obligations. Following Åqvist (2002, p.180), Prior's initial formulation of the paradox is represented by theorem 15.111, but more generally the paradox amounts to the following four theorems in the standard system (replacing $\multimap$ by $\supset$ for the translation in KD).

$$\neg\varphi\multimap(\varphi\multimap O\psi) \tag{15.109}$$
$$O\psi\multimap(\varphi\multimap O\psi) \tag{15.110}$$
$$O\neg\varphi\multimap O(\varphi\multimap\psi) \tag{15.111}$$
$$O\psi\multimap O(\varphi\multimap\psi) \tag{15.112}$$

The paradoxical character of 15.111 appears when we interpret it in terms of commitment: if it is forbidden to do φ, then doing φ commits us to anything. This is dubious from a normative point of view: it is forbidden to steal, hence stealing commits us to adultery! As it is pointed out by Lokhorst (1997, p.6), 15.110 is not derivable within deontic linear logic, and given equation 15.97 we can see how 15.111 and 15.112 are not either. Moreover, since the multiplicative connective does not satisfy (0), it follows as we saw earlier that 15.109 is not derivable.

It is noteworthy that since $\mathcal{CNR}$, as deontic linear logic, does not satisfy $\varphi \longrightarrow (\neg\varphi \multimap \psi)$ or $\varphi \longrightarrow (\psi \multimap \varphi)$, it follows that it can avoid Prior's paradox. That said, notwithstanding the objections that can be made against the definition of O in deontic linear logic[36], $\mathcal{CNR}$ is a more likely candidate to model conditional normative reasoning, as we saw with the analysis of Chisholm's puzzle. In virtue of equation 15.97, deontic linear logic fails to model properly deontic conditionals. As such, it is our view that conditional normative reasoning should not be done within deontic linear logic but should rather be done within the foundational framework of $\mathcal{CNR}$, to which one can add the appropriate axioms and rules of inference to model the behaviour of the formulas within the scope of the deontic operator.

Closing remarks

To sum up, we analyzed (factual) detachment, augmentation and normative conflicts from the perspective of category theory. From this standpoint, the usual objections against standard deontic logic and dyadic deontic logic in favour of non-monotonic (or adaptive) foundations can actually be seen as an argument in favour of non-Cartesian foundations for conditional normative reasoning. We showed how the usual objections against monadic and dyadic deontic logics in favour of non-monotonic foundations can be solved within a monadic framework by using a conditional similar to a linear implication instead of a classical (or intuitionistic) implication to model deontic conditionals and contrary-to-duties. In this respect, our approach contributes to the literature insofar as it is usually assumed that it is impossible for monadic frameworks that do not incorporate further operators and that do not consider deontic conditionals as primitive to model properly conditional normative reasoning. Our analysis showed that it is actually possible, and as such it opens an avenue for research that is usually thought to be a deadlock.

We began by showing how Chisholm's paradox undermines monadic frameworks and how dyadic deontic logic can solve this. We then showed that there are still problems with dyadic deontic logic, namely that it cannot allow for conflicting obligations and that it cannot deal properly with factual detachment. We saw how all of these problems can be solved within the framework of non-monotonic logic. In short, the arguments in favour of non-monotonic foundations for deontic logic can be summarized as follows:

1. Deontic logic must allow for conflicting obligations.
2. In case of unsolvable conflicts, it must not satisfy that anything follows.
3. Reasoning with deontic conditional does not satisfy augmentation, i.e., (aug) or (wk).
4. Some form of *restricted* detachment must be allowed, but *unrestricted* detachment must be avoided.

[36] By definition, $O\varphi$ means that $\neg\varphi$ entails what is not desired. However, it is easy to imagine a situation where not respecting an obligation actually does not entail what is not desired.

5. Deontic logic must be able to model situations in which conflicts of obligations are solved.

   (a) It must be able to model violations and contrary-to-duty reasoning.

   (b) It must be able to model conflicts between *prima facie* and actual obligations.

   (c) It must be able to model conflicting norms and specificity.

Using category theory as a foundational framework, we showed how different logical systems can be classified from a proof-theoretical perspective by focusing on the properties of their connectives. The categorical analysis of logic shows that logical connectives do not appear randomly. It begins with the tensor, its adjoint(s), which allows for the definition of negation, and then the co-tensor. It is quite common within the deontic logic literature to see the material conditional as the heart of the problem when trying to model conditional obligations and contrary-to-duties. Interestingly, though, our analysis shows that the problems come from the properties of conjunction.

Our analysis of augmentation, (factual) detachment and normative conflicts showed how (wk) and *ex falso sequitur quodlibet* are the sources of the problem. In itself, this is not new. Nonetheless, a categorical analysis of these problems showed that *ex falso sequitur quodlibet* and (wk) actually come from the two fundamental characteristics of a Cartesian deductive system, namely (0) and (Cart). We thus proposed to reject these two characteristics and do conditional normative reasoning within the closest deductive system at our disposition, namely a symmetric closed deductive system with classical negation and co-tensor.

We only analyzed the problems that come directly from the properties of the consequence relation and left open the issues regarding the rule that should or should not govern the behaviour of the deontic operator.[37] As such, $\mathcal{CNR}$ was presented as a foundational framework to do conditional normative reasoning. The defeasibility of normative inferences comes from the fact that there is a lack of information regarding the context, the norms that are valid in that context and the conditions under which conditional obligations hold. This is actually why detachment is to be avoided, since additional conditions under which the obligation does *not* hold can be added. Incidentally, this is also why *restricted* detachment is allowed. A conditional obligation can be detached under certain circumstances only when we are sure that there is no other information that will defeat the obligation. Hence, restricted detachment is allowed when *all things are considered* and every relevant information regarding the conditional obligation is known.[38] These ideas are captured by our framework. To detach an obligation $O\psi$ from a deontic conditional $\varphi \multimap O\psi$, one must assume that $O\psi$ holds under the circumstances φ, and moreover one must be *exactly* under these circumstances. If additional information regarding the context is provided, then additional information regarding what obligations hold within this

[37] On this point, see Peterson (2014b).

[38] This view is common within the literature. See for instance Alchourrón (1996), Asher and Bonevac (1997) or Morreau (1997).

new context is required. Hence, the conditional obligation $O\psi$ can be detached from a context φ when and only when we know that it holds under these circumstances.[39]

Our analysis was limited to the presentation of $\mathcal{CNR}$ as a foundational framework, and as such the properties of O were not analyzed. In the literature, one can find that there are other criteria that a logic which aims to model conditional normative reasoning should satisfy, as the deontic syllogism $O(\varphi \vee \psi) \wedge \neg O\varphi \longrightarrow O\psi$ (see Horty 1997 and Goble 2009). Analyzing these other criteria was unfortunately beyond the scope of this paper, which is only a first step towards the more general project of defining a categorical deontic logic. It is our hope that the reader who was disappointed that we did not analyze these criteria will be satisfied by our future research, which will treat of these issues. Moreover, we only provided a syntactical analysis of these problems, concentrating only upon the properties of the consequence relation, and hence our approach opens the way to further investigations. For future research, we intend to analyze in details the structure of the propositions that should be in the scope of the deontic operator and we will study the rules and axioms that should be used in conjunction with $\mathcal{CNR}$ to properly model the behaviour of O. A proper semantics for this system will also be provided, and we intend to use category theory to provide a functiorial semantics for $\mathcal{CNR}$.

* * *

Cet article répond au troisième objectif de cette thèse. En analysant les problèmes relatifs aux inférences normatives conditionnelles selon la perspective de la théorie des catégories, nous avons mis en évidence que les problèmes du détachement, de l'augmentation et de l'explosion déontique viennent des propriétés structurelles des systèmes déductifs cartésiens. Ayant maintenant en mains un outils pour modéliser adéquatement les inférences normatives conditionnelles, le prochain chapitre fait la synthèse de ce que nous avons développé jusqu'à présent et propose un système déductif déontique en vue de la modélisation des inférences légales.

[39] Valid deductive normative inferences are thus non-ampliative.

Chapitre 16

The categorical imperative: Category theory as a foundation for deontic logic

Abstract

This article introduces a deontic logic which aims to model Canadian legal discourse. Category theory is assumed as a foundational framework for logic. A deontic deductive system $\mathcal{DDS}$ is defined as two fibrations: the logic for unconditional obligations $\mathcal{OL}$ is defined within a Cartesian closed category on the grounds of an intuitionnistic propositional action logic $\mathcal{PAL}$ and an action logic $\mathcal{AL}$, while a logic for conditional normative reasoning $\mathcal{CNR}$ is defined as a symmetric closed monoidal category. A typed syntax and typed arrows are used to define properly $\mathcal{DDS}$. We show how it can solve the paradoxes of deontic logic and we provide some examples of application to legal reasoning.[1]

Keywords: Legal discourse, Categorical logic, Deductive system, Conditional normative reasoning, Conflicting obligations, Paradoxes

Introduction

Deontic logic was introduced in analogy with modal logic by von Wright (1951) to model normative reasoning.[2] After the developments of possible world semantics with the work of Hintikka, Montague and Kripke (see Wolenski 1990), von Wright's initial approach was redefined within the framework of modal logic. This gave rise to the well-known standard system of deontic logic, the modal logic KD. Many objections were raised against von Wright's initial approach[3], but Chisholm's (1963) paradox was the most damaging to the standard systems. It showed that monadic deontic logics cannot properly model conditional normative reasoning, which is central to the normative discourse.

Chisholm's objection was followed by various proposals. Some authors argued that

[1] A revised version of this chapter has been published in Peterson (2014b).
[2] See Føllesdal and Hilpinen (1970) and McNamara (2010) for an introduction.
[3] See McNamara (2010) for an overview.

contrary-to-duty reasoning should be modeled through a dyadic framework, specifying the conditions under which the obligations hold (see for example van Fraassen 1972; Al-Hibri 1978).[4] Others argued that Chisholm's paradox arises because standard systems do not take into account the temporal dimension implicit to conditional normative reasoning (e.g., Decew 1981), and this lead to the introduction of temporal deontic logics (see for example Thomason 1981; van Eck 1982a; 1982b). But still, other issues remained with these approaches, such as their inability to properly model conflicting obligations and factual detachment. This lead some authors to introduce different solutions to answer the problems of detachment of deontic conditionals and, more importantly, to model conflicting conditional obligations. In addition to Makinson and van der Torre (2000; 2001; 2003a), who introduced input/output logics to model normative conditional and unconditional reasoning, one can find various proposals in non-monotonic (see Nute and Yu 1997 for an introduction) and adaptive logics (see for example Straßer 2011; Beirlaen et al. 2013; Beirlaen and Straßer 2013).

Even though deontic logic was first meant as a formal framework to analyze and evaluate normative reasoning, it has since then been used to serve different purposes. In addition to the analysis of inferences, deontic logic has been used to model normative systems (e.g., Lindahl and Odelstad 2000; Carmo and Jones 2002; Boella and van der Torre 2006c) and multi-agent normative systems (e.g., Pörn 1970; Royakkers 1998; Horty 2001; Pacheco and Carmo 2003; Herzig et al. 2011; Broersen 2011a). It has been used to model contracts (e.g., Boella and van der Torre 2004; Brown 2005; Prisacariu and Schneider 2012) and obligations with deadlines (e.g., Dignum et al. 2005; Broersen et al. 2004; Demolombe 2014). It has been used in computer science (e.g., Wieringa and Meyer 1993; van der Meyden 1996; Carmo and Jones 1996; Carmo and Demolombe 2001; Boella and van der Torre 2003a), artificial intelligence (e.g., Bringsjord et al. 2011; Bringsjord and Taylor 2012) and in law (e.g., Sartor 1994; Walton and Gordon 2005; Prakken 2005; Gordon and Walton 2009; Boella et al. 2010).

This list does not pretend to be exhaustive and is in all likelihood incomplete. There is, however, a lesson that should be learned from this diversity: a system of deontic logic cannot be criticized independently of its purpose. For instance, a deontic logic which aims to model the evolution of a computer program will not require the same characteristics as one that tries to model the structure of the law. Similarly, a deontic logic that aims to model contracts will not need the same properties as one that models inferences. In the present paper, our aim is to introduce a deontic logic adequate to analyze and evaluate the structure of legal inferences. The aim is not to develop a deontic logic that can represent the structure of the Canadian legislation, nor to develop a formal system that can model how legal reasoning is done. Rather, our objective is to develop a formal system that can help in the analysis of the structure of legal reasoning, specifying how it should be done.

The results of the present paper are built upon previous work. Following the seminal work of Lambek (1958) and what was presented in Peterson (2014f), we introduce a typed

[4] Historically, the building blocks of dyadic deontic logic were introduced by von Wright (1956) in answer to Prior's (1954) paradox of derived obligation. Dyadic deontic logic has however been used as a solution to Chisholm's puzzle.

deontic logic within the framework of categorical logic. There are three main theoretical motivations for this paper. Firstly, our aim is to introduce an alternative framework to modal logic to model unconditional obligations. Secondly, we wish to introduce an alternative framework to Boolean algebras to model actions within deontic contexts. Finally, the objective is to introduce a monadic formal system able to model conditional normative reasoning and conflicting obligations without requiring the techniques of non-monotonic or adaptive logics. As such, the deontic logic we propose will operate at three different levels. In a nutshell, we propose to define a deontic deductive system $\mathcal{DDS}$ on the grounds of an action logic $\mathcal{AL}$ and a propositional action logic $\mathcal{PAL}$ (cf. Peterson 2014a), an obligation logic $\mathcal{OL}$ (cf. Peterson 2014e) to model unconditional obligations and a logic $\mathcal{CNR}$ that can model conditional normative reasoning with a monadic O (cf. Peterson 2014d).

The structure of the paper is as follows. In the next section, we present the rationale of our framework and expose the characteristics that a deontic logic which aims to model Canadian legal discourse should satisfy. This will be followed by a brief exposition of the foundational framework we adopt. Then, we present all the relevant material that is required to define $\mathcal{DDS}$ and we provide the categorical definition. A comparison of $\mathcal{DDS}$ with Goble's (2009) analysis is provided, followed by a discussion of some paradoxes. We also provide examples of applications of $\mathcal{DDS}$ to Canadian legal discourse. We conclude in the last section with remarks for future research.

Preliminaries

Some characteristics

The purpose of the present paper is to develop a formal framework relevant to the analysis of Canadian legal discourse. As such, it ought to be founded on some of the characteristics of Canadian legal discourse. A first thing to mention is that *truth*, in Canadian law, is a matter of construction. Following the honorable Jean-Louis Baudoin (2010, p.5), the truth of normative propositions depends upon what is established by the legislator and the case law. In the deontic logic literature, this is consistent with Alchourrón's and Bulygin's (1981, pp.97,102) conception, for whom obligations depend on norms, which are established by legal authorities. Hence, we assume that normative propositions are declarative statements, but which are not of the same type as descriptive propositions. This difference of type between normative and descriptive declarative sentences can be understood in the light of the semantical dichotomy between facts and norms (cf. Hume 1993; Poincaré 1913; Jørgensen 1937). The semantical dichotomy implies that normative and descriptive propositions are not true in the same conditions. As such, an inference from (solely) normative propositions to a descriptive proposition will always be invalid, and similarly for inferences that go from (solely) descriptive propositions to a normative proposition.

From a legal point of view, *actions*, and not *propositions*, are obligatory, permitted or forbidden. Therefore, a deontic logic that wants to adequately model legal discourse must be of the type *ought-to-do* rather than *ought-to-be* (see von Wright 1999a for the distinction). Such a deontic logic must be founded upon a proper action logic able to model *human* actions. Its deontic operators will be of the type that takes an action and transforms it into a normative proposition. Iteration of deontic operators will thus be impossible.

Assuming that the truth of a normative proposition depends upon a norm established by some authority, it follows that tautologous actions are not obligatory (although, as we will see, there will be no such thing as 'tautologous' actions within $\mathcal{AL}$). As a result, the formal system must not satisfy formulas such as $\vdash O\top$ or $\vdash \varphi \supset O\top$, with $\top$ some tautologous action. In the literature, this position is consistent with Chellas (1974), who argued that it must be possible to model a situation where noting is obligatory (not even tautologous actions), and with Jones and Pörn (1985), who argued that it is always possible for someone to act against one's obligations.

A last characteristic worth mentioning is the presumption of consistency of Canadian legal discourse. Of course, there are contradictory laws, but recall that our aim is not to model the structure of the legislation but is rather to model the structure of the inferences we do with this (potentially contradictory) legislation. As it happens, even though there can be *a priori* contradictions within the law, it remains that legal discourse is presupposed to be consistent *after* interpretation. The presumption of consistency is considered as a fundamental value of our judicial system, and as such the preservation of the legal discourse's consistency is a principle that guides the interpretation of the law (Côté 2006, p.387).

That being said, the presumption of consistency is not only admitted within the law literature but is also recognized by the Supreme Court of Canada as both a fundamental principle and a rational criterion for legal discourse. On the one hand, it is a fundamental principle given that it justifies other principles that govern the interpretation of the law, such as the rule *nemo intelligere possit antequam iterum perlegerit*. This rule, which means that no one can understand a part without reading and rereading the whole in full, is meant to insure a consistent interpretation of the law.[5] On the other hand, consistency has been recognized by the Supreme Court as a criterion that enables us to evaluate the legal discourse's rationality. Quoting Sullivan (1994, p.176), the Supreme Court of Canada stated that the presumption of consistency is a "virtually irrebuttable" principle that insures that legal discourse forms a rational framework.[6]

Foundational framework

In this paper, we follow the foundational framework adopted in Peterson (2014a; 2014d; 2014f). We partially follow Lambek (1958; 1968; 1969) in his categorical understanding of logic and assume category theory as a foundational framework to analyze the proof theory of different systems of logic. In order to make this paper self-sufficient, we ex-

[5] See *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, [1996] 3 RCS 919, paragraphe 207.

[6] *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, [1996] 3 RCS 919, paragraphe 209.

pose the relevant material to construct deductive systems and we refer the reader to the aforementioned papers for the technical details and the categorical definitions.

From an epistemological point of view, category theory offers a powerful foundational framework that can enable us to classify different systems of logic through their proof theory. A *deductive system* is understood as a category $\mathcal{C}$ whose objects are formulas and whose arrows are proofs (deductions). The composition of arrows is represented by the rule (cut)[7] and the deduction of a formula from itself is represented by the identity axiom (1)[8].

$$\frac{\varphi \longrightarrow \psi \qquad \psi \longrightarrow \rho}{\varphi \longrightarrow \rho}\,(\mathrm{cut}) \qquad\qquad \frac{}{\varphi \longrightarrow \varphi}\,(1)$$

The interest of the categorical understanding of logic appears when one considers how to introduce different logical connectives, with respect to the categorical structure of the deductive system. It is noteworthy that the connectives do not appear randomly when one defines a logical system from the proof-theoretical perspective of category theory. They appear in a specific order, and the properties of the logical system will actually depend upon the specific properties of some connectives.

From a deductive system, one can construct a *monoidal deductive system* by adding a tensor product $\otimes$ with a unit $\top$, who behave according to the rules (t) and (a) for the associativity of the tensor product, and (l) and (r) to make $\top$ into the unit of $\otimes$.[9]

$$\frac{\varphi \longrightarrow \top \otimes \psi}{\varphi \longrightarrow \psi}\,(\mathrm{l}) \qquad\qquad \frac{\varphi \longrightarrow \psi \otimes \top}{\varphi \longrightarrow \psi}\,(\mathrm{r})$$

$$\frac{\varphi \longrightarrow \psi \qquad \rho \longrightarrow \tau}{\varphi \otimes \rho \longrightarrow \psi \otimes \tau}\,(\mathrm{t}) \qquad\qquad \frac{\tau \longrightarrow (\varphi \otimes \psi) \otimes \rho}{\tau \longrightarrow \varphi \otimes (\psi \otimes \rho)}\,(\mathrm{a})$$

A *closed monoidal deductive system* is obtained on the grounds of a monoidal deductive system by introducing a right adjoint functor to the tensor product via the rule (cl). This rule is an analogue to the deduction theorem.[10] Having this connective at our disposition, it is possible to introduce a special object $\bot$ to define an intuitionistic negation via $\neg\varphi =_{def} \varphi \triangleright \bot$. In such a situation, we speak of a *closed monoidal deductive system with negation*. If one wants a classical negation, then one simply has to add the axiom

[7] Which respects associativity, i.e., $(hg)f = h(gf)$.

[8] Which respects the identity laws, i.e., $1_\psi f = f$ and $g1_\psi = g$.

[9] A double line means that the rule can be applied both ways. The rules (a), (l) and (r) are natural isomorphisms.

[10] Since in a closed monoidal deductive system the tensor product is not commutative, it follows that it can possess two different right adjoints. See the nlab entry http://ncatlab.org/nlab/show/closed+monoidal+category.

$(\neg\neg)$ to obtain a *closed deductive system with classical negation.*[11]

$$\frac{\varphi \otimes \psi \longrightarrow \rho}{\varphi \longrightarrow \psi \rhd \rho} \ (\text{cl}) \qquad\qquad \frac{}{\neg\neg\varphi \longrightarrow \varphi} \ (\neg\neg)$$

A commutative tensor product is obtained by adding a braiding rule (b), which is its own inverse.[12] We thus obtain a *symmetric deductive system.*[13] If the deductive system is symmetric, then (r) can be proven from (l), and vice versa, as for (cl) and (cl').

$$\frac{\varphi \longrightarrow \psi \otimes \tau}{\varphi \longrightarrow \tau \otimes \psi} \ (\text{b})$$

From a symmetric closed deductive system with classical negation, one can define a *compact deductive system* by adding the axioms (cpt1) and (cpt2).[14]

$$\frac{}{\varphi \rhd \psi \longrightarrow \varphi^* \otimes \psi} \ (\text{cpt1}) \qquad\qquad \frac{}{\varphi^* \otimes \psi \longrightarrow \varphi \rhd \psi} \ (\text{cpt2})$$

A *Cartesian deductive system* is a symmetric deductive system where the tensor product is a categorical product and its unit is a terminal object. This is formally represented by the rule (Cart) and the axiom (!).

$$\frac{\varphi \longrightarrow \psi \qquad \varphi \longrightarrow \rho}{\varphi \longrightarrow \psi \otimes \rho} \ (\text{Cart}) \qquad\qquad \frac{}{\varphi \longrightarrow \top} \ (!)$$

The co-tensor $\oslash$ is obtained through the opposite category $\mathcal{C}^{op}$ by reversing the arrows and replacing $\otimes$ and $\top$ respectively by $\oslash$ and $\bot$ to obtain the rules (co-t), (co-a), (co-l), (co-r) and (co-b). We thus speak of a *deductive system with co-tensor* (resp. co-product for Cartesian deductive systems). In the case of a Cartesian deductive system with co-product, the axiom and the rule are (0) and (co-Cart).

$$\frac{\varphi \longrightarrow \rho \qquad \psi \longrightarrow \rho}{\varphi \oslash \psi \longrightarrow \rho} \ (\text{co-Cart}) \qquad\qquad \frac{}{\bot \longrightarrow \varphi} \ (0)$$

The upshot of this mode of presentation is that the logical properties of a deductive system mainly depend upon the properties of the tensor product and the definition of negation. See Peterson (2014a; 2014d) for a proper presentation and a comparison between deductive systems and the literature. For the purpose of this article, simply note that:

[11] With $\varphi \cong \neg\neg\varphi$ a natural isomorphism.

[12] Otherwise the deductive system would only be braided instead of symmetric. The rule is also a natural isomorphism.

[13] If the symmetric deductive system is closed, then there is only one adjoint to the tensor product.

[14] In this case, the unit is the dualizing object. See Barr (1991) for the definition of a dualizing object.

1. the multiplicative fragment of linear logic **MLL** introduced by Girard (1987) corresponds to a closed symmetric deductive system with classical negation and co-tensor;

2. intuitionistic logic is a closed Cartesian deductive system with negation and co-product;

3. classical logic is a closed Cartesian deductive system with classical negation and co-product.

Deontic deductive systems

The deontic logic we propose is called a *deontic deductive system*, $\mathcal{DDS}$ for short. It operates at three different levels, which were previously discussed at length in other papers. It incorporates an action logic, a logic for reasoning with unconditional obligations and a logic to reason with contrary-to-duties and conflicting obligations. In what follows, we only present what is relevant for the construction of $\mathcal{DDS}$ and we refer the reader to these articles for further details, explanations and discussions.

We use a typed syntax and specify the types for the logical connectives, propositions and consequence relations. This will be done to avoid some technical problems when trying to incorporate both a classical negation and an intuitionistic one within $\mathcal{DDS}$. This solution follows the ideas presented in Peterson (2014f) and is inspired by the work of Lambek (1958). Instead of using a unique arrow to model the consequence relation of our deductive system, we will use different types of arrows within different categories.

The introduction of typed arrows is necessary to incorporate both an intuitionistic negation and a classical one within $\mathcal{DDS}$. Otherwise, since initial objects are identical up to isomorphism, the classical negation would absorb the intuitionistic one. Although we could use an intuitionistic logic with strong negation (or equivalently a Nelson algebra, see Brignole 1969; Spinks and Veroff 2008), this would require that we consider the strong negation as a primitive, which would imply that we step outside of our conceptual framework and define the strong negation without using the structural properties of our deductive system. As such, the typed syntax will allow us to stay within the framework of categorical logic.

Action logic

The first step consists in defining a proper action logic to represent the structure of human actions. It was argued in Peterson (2014a) that *choice* and *iteration*, although relevant for programing, are not primitive human actions. On the one hand, even though one may face a choice between two actions, one is not *performing* the action 'choice'. One will *choose*, and then ideally perform one of these actions, but one will not be 'choice-ing'. On the other hand, iteration in programing languages is meant to model recursive programs, and this is not relevant to human actions. Iteration of human actions can be represented by finite sequences, and as such we do not need iteration as a primitive. In addition to atomic actions, the two primitive complex human actions we assume are joint actions $\alpha \bullet \beta$ and sequence of actions $\alpha \curvearrowright \beta$.

Following Peterson (2014a), we distinguish between an *action logic* $\mathcal{AL}$, which represents the structure of actions, and a *propositional action logic* $\mathcal{PAL}$, expressing the structure of the language we use to talk about actions. Assume a category where arrows are proofs and objects are actions. Let *Act* be the collection of all atomic actions a_i. The well-formed formulas $WFF_{\mathcal{AL}}$ of type act are defined recursively from the language $\mathcal{L}_{\mathcal{AL}} = \{Act, \bullet, *, \ominus, \curvearrowright, (,)\}$ by[15]:

$$\alpha := \; a_i \mid \; * \; \mid \alpha \bullet \beta \mid \alpha \ominus \beta \mid \alpha \curvearrowright \beta$$

The connectives of $\mathcal{AL}$ are of the type that takes two actions and transforms them into another action (i.e., of type $\mathsf{act}\backslash\mathsf{act}/\mathsf{act}$).

Definition 16.1. *An* Action Logic $\mathcal{AL}$ *is defined on the grounds of* $\mathcal{L}_{\mathcal{AL}}$*,* $WFF_{\mathcal{AL}}$ *and* $\xrightarrow[\mathsf{AL}]{}$ *arrows by the two following fragments.*

1. *The fragment* $\{Act, \curvearrowright, *, (,)\}$ *is axiomatized by a monoidal deductive system (with* $*$ *the unit of* $\curvearrowright$*).*
2. *The fragment* $\{Act, \bullet, \ominus, *, (,)\}$ *is axiomatized by a compact deductive system, with:*
   (a) $*$ *the unit of* $\bullet$ *(which also dualizes);*
   (b) $\alpha^* =_{def} * \ominus \alpha$*.*

$$\frac{}{\alpha^* \bullet \beta \xrightarrow[\mathsf{AL}]{} \beta \ominus \alpha}\ (cpt1) \qquad \frac{}{\beta \ominus \alpha \xrightarrow[\mathsf{AL}]{} \alpha^* \bullet \beta}\ (cpt2) \qquad \dfrac{\alpha \bullet \beta \xrightarrow[\mathsf{AL}]{} \gamma}{\beta \xrightarrow[\mathsf{AL}]{} \gamma \ominus \alpha}\ (cl)$$

The connective $\bullet$ stands for action conjunction and is understood similarly to the multiplicative conjunction of linear logic (cf. Girard 1987). It is the 'with' connective, $\alpha \bullet \beta$ meaning the action 'α together with β'. In an action logic, actions are *not* understood as declarative sentences. The conjunctive action $\alpha \bullet \beta$ stands for the simultaneous action of α and β, and as such it does not satisfy (Cart).

The connective $\ominus$ is introduced as the right adjoint of $\bullet$. From (cpt1) and (cpt2), it can be defined from $\bullet$ since we have an isomorphism between $\beta \ominus \alpha$ and $\alpha^* \bullet \beta$. It is not, however, interpreted similarly to the linear implication. The formula $\alpha \ominus \beta$ stands for the action 'α without β'. We wrote $\gamma \ominus \alpha$ instead $\alpha \ominus \gamma$ in (cl) to keep the intuitive reading of 'γ without α'. We can think of 'with' and 'without' informally as addition and

[15] Note that in Peterson (2014a) $\otimes$ was used instead of $\bullet$. Here, we reserve the use of $\otimes$ for the analogue to the multiplicative conjunction in $\mathcal{CNR}$.

subtraction: as a complex conjunctive action can be obtained by gluing actions together, an action can also be obtained by removing some parts of a more complex action.[16]

The connective $\curvearrowright$ stands for action sequence, the formula $\alpha \curvearrowright \beta$ meaning 'α and then β'. While action conjunction implies simultaneity, action sequence implies an order. Hence, the tensor product $\bullet$ is both associative and commutative, but sequence $\curvearrowright$ is only associative.

The special action $*$ is the *no change* or *nothing* action. As we will see in the next paragraphs, it is also an action that cannot be done (i.e., it is impossible to perform). Action negation is something which is often introduced without further ado within the literature. Unfortunately, though, we will not propose a throughout analysis of the concept within the present paper since it is beyond its scope (for a discussion, see Peterson 2014a). Action negation is represented by the action α^* and it is understood as some form of complement of α.

An action is minimally understood as something which has causal powers, that is, something that can bring changes in the world. This understanding is uncontroversial within the literature (cf. Goldman 1970; Pörn 1970; Davidson 1967; von Wright 1968; Mossel 2009). As such, *negative* actions are actions too, given that not-acting will have repercussions in the world.[17]

That said, we understand a negative action as something which is either simply not done (omitted) or intentionally not done (forborne). Hence, the complement of an action can be seen as an absence of action (either intentional or not). Action negation is understood as the complement of an action in regards to the *no change* action. From the aforementioned definition, the dual action of α is '*no change* without α'.

An interesting consequence of defining an action logic as a compact deductive system is that we get $\alpha^* \bullet \beta$ isomorphic to $\beta \ominus \alpha$. This result is actually quite plausible: the action 'going to work without wearing a tie' is isomorphic to 'going to work while not wearing a tie'. Another consequence of this definition is that we have $*$ logically equivalent to $\alpha \bullet \alpha^*$. From this we can see why $*$ is an action that cannot be accomplished: it would be impossible to do both an action and its complement at the same time. As such, $*$ can informally be understood as something that does not have any causal power. For instance, if we understand the evolution of a scenario as a discrete process, $*$ can be understood as the vacuous moment between the states.

Having an action logic at our disposition, we can now define a propositional action logic $\mathcal{PAL}$, which uses declarative sentences. We argued in Peterson (2014a) that a propositional action logic should be defined as a closed Cartesian deductive system with negation and co-product, that is, as an intuitionistic deductive system. In short, the argument in favour of an intuitionistic formulation of $\mathcal{PAL}$ revolves around the fact that an action proposition can be considered as true only when we have a 'proof' that the action

[16] Note that the 'without' action is part of our natural language and is often used to refer to actions that are somewhat *less* than other actions.

[17] Although all of these authors agree that an action has causal powers, they do not agree, however, on the meaning of action negation. See Peterson (2014a) for a discussion.

was actually done. This is consistent with legal discourse. Not having a proof that the action was not done does not imply that it was done.

Let the collection of action propositions AP be defined by the following condition: if α is a well-formed formula of $\mathcal{AL}$ (of type act), then $\boldsymbol{\alpha}$ is an action proposition of type ap (for action propositions). The bold notation $\boldsymbol{\alpha}$ is used to refer to action propositions as declarative sentences.

The well-formed formulas $WFF_{\mathcal{PAL}}$ of type ap are defined recursively from $\mathcal{L}_{\mathcal{PAL}} = \{\mathsf{AP}, \wedge, \top, \supset, \vee, \bot, (,)\}$ by:

$$\varphi := \ \boldsymbol{\alpha} \mid \top \mid \bot \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \supset \psi$$

Definition 16.2. *A* Propositional Action Logic $\mathcal{PAL}$ *is a closed Cartesian deductive system with negation and co-product built from* $\mathcal{L}_{\mathcal{PAL}}$*,* $WFF_{\mathcal{PAL}}$ *and* $\xrightarrow[\mathsf{PAL}]{}$ *arrows. It has to satisfy the rule* (Ψ)*.*

$$\frac{\alpha \xrightarrow[\mathsf{AL}]{} \beta}{\boldsymbol{\alpha} \xrightarrow[\mathsf{PAL}]{} \boldsymbol{\beta}} \qquad (\Psi) \qquad (\textit{with } \Psi(*) = \bot)$$

The rule (Ψ) enables us to import a part of the structure of $\mathcal{AL}$ into $\mathcal{PAL}$.

Obligation logic

The second level at which $\mathcal{DDS}$ operates regards unconditional normative reasoning. By an unconditional normative reasoning, we mean an inference which deals with unconditional or *all things considered* obligations. These reasonings are modeled through an obligation logic $\mathcal{OL}$ inspired by Peterson (2014e).

Let $\mathcal{L}_{\mathcal{OL}} = \mathcal{L}_{\mathcal{AL}} \cup \mathcal{L}_{\mathcal{PAL}} \cup \{O, P_s\}$. The formulas of $\mathcal{OL}$ are of type np, for normative propositions. They are declarative sentences. The special propositions $\top$ and $\bot$ are either of type ap or np. Similarly, the connectives $\wedge$, $\vee$ and $\supset$ are either of the type that takes two action propositions and transforms them into another action proposition (i.e., of type $\mathsf{ap}\backslash\mathsf{ap}/\mathsf{ap}$), or of the type that takes two normative propositions and transforms them into another normative proposition (i.e., of type $\mathsf{np}\backslash\mathsf{np}/\mathsf{np}$). They do not allow for mixed formulas.

Given α a formula of type act, the well-formed formulas $WFF_{\mathcal{OL}}$ of type np are defined recursively by:

$$\varphi := \ O\alpha \mid P_s\alpha \mid \top \mid \bot \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \supset \psi$$

The operators O and P_s stand respectively for obligation and strong permission. They are of a type that takes an action and transforms it into a normative proposition (i.e., of type np/act). The introduction of strong permission as a primitive operator is necessary to model permissions that are explicitly mentioned by a legal system. From O we define interdiction by (def F) and weak permission by (def P_w).

$$F\alpha =_{def} O\alpha^* \tag{def F}$$

$$P_w\alpha =_{def} \neg F\alpha \tag{def P_w}$$

The definition of weak permission in terms of 'not forbidden' is uncontroversial. If an action is not explicitly forbidden, then it is implicitly permitted. The definition of the F operator in terms of O is, however, sometimes contested.[18] Semantically, interdictions and obligations can be seen as two different things: an interdiction states that an action must *not* be performed, while an obligation states that an action *must* be performed. One says *act* and the other says *do not act*. As such, interdictions dictate the limits we must not trespass, while obligations stipulate the things we are required to do.

We disagree, however, with this semantical distinction. To see this, consider what is required to violate an interdiction and an obligation. In criminal law, there is a distinction between the *actus reus* (the act) and the *mens rea* (the intention). To determine if someone is imputable for a crime (or a fault), it must be established that the accused did indeed commit the crime (or the fault) and that his judgment was not clouded.[19] It does not imply that the accused had the explicit intention to commit the crime (or the fault), but only that his mind was such that he was able to understand the consequences of his actions. Now, consider the two following propositions:

1. It is forbidden to not connect the lamp of a UV disinfection system.[20] ($F\alpha^*$)
2. A sign P-270 indicates that a car driver has the obligation to stop and yield to pedestrians on a pedestrian crossing.[21] ($O\beta$)

When does one violate the interdiction $F\alpha^*$? One violates this interdiction of not connecting the lamp of a UV disinfection system (in the context of sewage treatment) by 'not connecting the lamp of a UV disinfection system'. Hence, one violates $F\alpha^*$ by performing α^*. That said, α^* can be the result of either an omission or a forbearance, and one who does not connect the lamp of a UV disinfection system (in the context of sewage treatment) will be liable in both cases. On the other hand, one violates the obligation to yield to pedestrian at a pedestrian crossing when there is a sign P-270 when 'one does not yield to pedestrians' in that context. As such, one violates $O\beta$ by performing β^*. Hence, it is possible to argue that there is no fundamental semantical distinction between obligations and interdictions insofar as in the end, they mean the same thing: $O\alpha$ means

[18] See for instance Hintikka (1970) and Cox (1978) for a discussion. See also Pacheco and Santos (2004), Pacheco and Carmo (2003) and Broersen (2004).

[19] With some exceptions, as noted in chapter 14 footnote 25.

[20] Règlement sur l'évacuation et le traitement des eaux usées des résidences isolées (*regulation on sewage for isolated residences*), RRQ, c Q-2, r 22.

[21] Règlement sur la signalisation routière (*regulation on road sign*), RRQ, c C-24.2, r 41, A.M. 99-06-15, a. 38.

that α^* should not be performed, as does $F\alpha^*$. Therefore, since both $O\alpha$ and $F\alpha^*$ mean the same thing and are violated in the same conditions, we assume that one can be defined in terms of the other.

Definition 16.3. *$\mathcal{OL}$ is defined by a closed Cartesian deductive system with classical negation, co-product and arrows of type $\xrightarrow[\mathsf{OL}]{}$. $\mathcal{OL}$ has to satisfy the axioms (**D**) and (**P**), and the rules (Δ) and (Π).*

$$\frac{}{O\alpha \xrightarrow[\mathsf{OL}]{} P_s\alpha} \quad (\mathbf{D}) \qquad\qquad \frac{}{P_s\alpha \xrightarrow[\mathsf{OL}]{} \neg O\alpha^*} \quad (\mathbf{P})$$

$$\frac{\boldsymbol{\alpha} \xrightarrow[\mathsf{PAL}]{} \boldsymbol{\beta}}{O\alpha \xrightarrow[\mathsf{OL}]{} O\beta} \quad (\Delta) \qquad (\textit{with } \Delta(\bot) = \bot)$$

$$\frac{\boldsymbol{\alpha} \xrightarrow[\mathsf{PAL}]{} \boldsymbol{\beta}}{P_s\alpha \xrightarrow[\mathsf{OL}]{} P_s\beta} \quad (\Pi) \qquad (\textit{with } \Pi(\bot) = \bot)$$

The rules (Δ) and (Π) allow to represent deontic consequence and can be used to model deontic detachment.[22] Note that $\Delta(\top)$ and $\Pi(\top)$ are undefined. From these rules and (Ψ), one can derive the rules (Σ_O) and (Σ_P) that allow the substitution of actions within the scope of O and P_s.

$$\frac{\alpha \xrightarrow[\mathsf{AL}]{} \beta}{O\alpha \xrightarrow[\mathsf{OL}]{} O\beta} \quad (\Sigma_O) \qquad (\text{with } \Sigma_O(*) = \bot)$$

$$\frac{\alpha \xrightarrow[\mathsf{AL}]{} \beta}{O\alpha \xrightarrow[\mathsf{OL}]{} O\beta} \quad (\Sigma_P) \qquad (\text{with } \Sigma_P(*) = \bot)$$

Axiom (**D**) states that if an action is obligatory, then it is strongly (explicitly) permitted, and axiom (**P**) expresses that if an action is strongly permitted, then it is weakly (implicitly) permitted. Together with (cut), axioms (**D**) and (**P**) allow us to recover the axiom schema (D) of standard deontic logic. From (D), the definition of negation and (cl), we recover the theorem for normative consistency (NC). This theorem is necessary to adequately model legal discourse. It represents the presumption of consistency of legal discourse (after interpretation), which is considered by the Supreme Court of Canada as a criterion of its rationality. As a result, a deontic logic which aims at modeling Canadian legal discourse must validate the presumption of consistency for unconditional obligations.

[22] That is, if $O\alpha$ and $\boldsymbol{\alpha} \longrightarrow \boldsymbol{\beta}$, then $O\beta$.

An action and its complement cannot be obligatory at the same time and under the same conditions.

$$O\alpha \xrightarrow[\mathsf{OL}]{} \neg O\alpha^* \tag{D}$$

$$O\alpha \wedge O\alpha^* \xrightarrow[\mathsf{OL}]{} \bot \tag{NC}$$

Conditional normative reasoning

The last step is to introduce the appropriate material to model conditional normative reasoning and conflicting obligations. In Peterson (2014d), the logic $\mathcal{CNR}$ was presented as a foundational framework that can model conditional normative reasoning through a monadic O without requiring further operators. The properties and the virtues of $\mathcal{CNR}$ as a foundational framework were analyzed thoroughly and we refer the reader to the paper for further details, explanations and discussion.

Let $Prop$ be a collection of atomic descriptive (declarative) formulas p_i of type d. The language for conditional normative reasoning is defined by:

$$\mathcal{L}_{\mathcal{CNR}} = \{(,), \mathsf{AP}, Prop, \otimes, 1, \multimap, \parr, 0, O, P_s\}$$

Assuming that α is a well-formed formula of type act, the well-formed formulas of $WFF_{\mathcal{L}_{\mathcal{CNR}}}$ are defined recursively by:

$$\varphi := 1 \mid 0 \mid \boldsymbol{\alpha} \mid p_i \mid O\alpha \mid P_s\alpha \mid \varphi \otimes \psi \mid \varphi \multimap \psi \mid \varphi \parr \psi$$

As in $\mathcal{OL}$, the operators O and P_s are of type $\mathsf{np/act}$. The main distinction between $\mathcal{L}_{\mathcal{CNR}}$ and $\mathcal{L}_{\mathcal{OL}}$ is that the former allows for mixed formulas. Although the propositions of $\mathcal{CNR}$ are of type np, the connectives $\otimes$, $\multimap$ and $\parr$ are either of the type $\mathsf{np\backslash np/np}$, $\mathsf{ap\backslash np/np}$, $\mathsf{np\backslash np/ap}$, $\mathsf{d\backslash np/np}$ or $\mathsf{np\backslash np/d}$. The special propositions 0 and 1 are of type np or ap. The connective $\otimes$ is similar to the multiplicative conjunction, $\multimap$ is similar to the linear implication and $\parr$ is similar to the multiplicative disjunction of linear logic.[23]

Let $\mathcal{L}_{\mathcal{PAL}^\otimes} = \{(,), \mathsf{AP}, \otimes, 1, \multimap, \parr, 0\}$. The connectives of $\mathcal{PAL}^\otimes$ are of type $\mathsf{ap\backslash ap/ap}$. The well-formed formulas $WFF_{\mathcal{PAL}^\otimes}$ are defined recursively by:

$$\varphi := \boldsymbol{\alpha} \mid 1 \mid 0 \mid \varphi \otimes \psi \mid \varphi \parr \psi \mid \varphi \multimap \psi$$

Definition 16.4. *$\mathcal{PAL}^\otimes$ is defined from $\mathcal{L}_{\mathcal{PAL}^\otimes}$ and $WFF_{\mathcal{PAL}^\otimes}$ by a closed symmetric deductive system with negation, co-tensor and $\xrightarrow[\mathsf{PAL}_\otimes]{}$ arrows. It has to satisfy the rule ($\Psi_\otimes$).*

[23] It would be wrong, however, to assume that we are working within the framework of linear logic. Thus defined, the propositional part of $\mathcal{CNR}$ (without the deontic operators) and the multiplicative fragment **MLL** of linear logic only share the same categorical structure. See Peterson (2014d) for details.

$$\frac{\alpha \xrightarrow[\mathsf{AL}]{} \beta}{\boldsymbol{\alpha} \xrightarrow[\mathsf{PAL}_{\otimes}]{} \boldsymbol{\beta}} \quad (\Psi_{\otimes}) \qquad (with\ \Psi_{\otimes}(*) = 0)$$

Definition 16.5. *$\mathcal{CNR}$ is defined from $\mathcal{L}_{\mathcal{CNR}}$ and $WFF_{\mathcal{L}_{\mathcal{CNR}}}$ by a closed symmetric deductive system with classical negation, co-tensor and $\xrightarrow[\mathsf{CNR}]{}$ arrows. Negation is defined by $\sim \varphi =_{def} \varphi \multimap 0$. It also has to satisfy the axioms ($\mathbf{D}_{\otimes}$) and ($\mathbf{P}_{\otimes}$) together with the rules ($\Delta_{\otimes}$) and ($\Pi_{\otimes}$).*

$$\frac{}{O\alpha \xrightarrow[\mathsf{CNR}]{} P_s\alpha} \ (\mathbf{D}_{\otimes}) \qquad\qquad \frac{}{P_s\alpha \xrightarrow[\mathsf{CNR}]{} \sim O\alpha^*} \ (\mathbf{P}_{\otimes})$$

$$\frac{\boldsymbol{\alpha} \xrightarrow[\mathsf{PAL}_{\otimes}]{} \boldsymbol{\beta}}{O\alpha \xrightarrow[\mathsf{CNR}]{} O\beta} \quad (\Delta_{\otimes}) \qquad (with\ \Delta_{\otimes}(0) = 0)$$

$$\frac{\boldsymbol{\alpha} \xrightarrow[\mathsf{PAL}_{\otimes}]{} \boldsymbol{\beta}}{P_s\alpha \xrightarrow[\mathsf{CNR}]{} P_s\beta} \quad (\Pi_{\otimes}) \qquad (with\ \Pi_{\otimes}(0) = 0)$$

Definition 16.6. *$\mathcal{DDS}$ is defined from $\mathcal{AL}$, $\mathcal{PAL}$, $\mathcal{PAL}^{\otimes}$, $\mathcal{OL}$ and $\mathcal{CNR}$.*

Table 16.2 below summarizes the rules and axioms of the deductive systems presented so far and table 16.1 summarizes the types of each logical connective and specifies the arrow types that use them.

	$\bullet, \ominus, \curvearrowright$	$*$	$\otimes, \multimap, \parr$	$1, 0$	$\wedge, \supset, \vee$	$\top, \bot$	O, P_s
$\mathcal{AL}$	act\act/act	act					
$\mathcal{PAL}$					ap\ap/ap	ap	
$\mathcal{OL}$					np\np/np	np	np/act
$\mathcal{PAL}^{\otimes}$			ap\ap/ap	ap			
$\mathcal{CNR}$			np\np/np ap\np/np d\np/np np\np/ap np\np/d	np			np/act

Table 16.1: Summary of types

	Arrow	Language	Axioms	Rules
$\mathcal{AL}$	$\xrightarrow[\mathsf{AL}]{}$		(1)	(cut)
		$\bullet$		(t),(a),(l),(b)
		$\ominus$		(cl),(cpt1),(cpt2)
		$*$	$(**)$	
		$\curvearrowright$		(t),(a),(l),(r)
$\mathcal{PAL}$	$\xrightarrow[\mathsf{PAL}]{}$		(1)	(cut),(Ψ)
		$\wedge$		(Cart)
		$\supset$		(cl)
		$\vee$		(co-Cart)
		$\bot$	(0)	
		$\top$	(!)	
$\mathcal{PAL}^{\otimes}$	$\xrightarrow[\mathsf{PAL}_{\otimes}]{}$		(1)	(cut),$(\Psi_{\otimes})$
		$\otimes$		(t),(a),(l),(b)
		$\multimap$		(cl)
		$\invamp$		(co-t),(co-a),(co-l),(co-b)
$\mathcal{OL}$	$\xrightarrow[\mathsf{OL}]{}$		(1)	(cut)
		$\wedge$		(Cart)
		$\supset$		(cl)
		$\vee$		(co-Cart)
		$\bot$	(0),$(\neg\neg)$	
		$\top$	(!)	
		O	$(\mathbf{D})$	(Δ)
		P_s	$(\mathbf{P})$	(Π)
$\mathcal{CNR}$	$\xrightarrow[\mathsf{CNR}]{}$		(1)	(cut)
		$\otimes$		(t),(a),(l),(b)
		$\multimap$		(cl)
		$\invamp$		(co-t),(co-a),(co-l),(co-b)
		0	$(\sim\sim)$	
		O	$(\mathbf{D}^{\otimes})$	$(\Delta_{\otimes})$
		P_s	$(\mathbf{P}^{\otimes})$	$(\Pi_{\otimes})$

Table 16.2: Summary of rules and axioms

Categorical definition of $\mathcal{DDS}$

The definition of $\mathcal{DDS}$ will in all likelihood appear to the reader as a bunch of rules and axioms. However, it happens that it has a specific categorical structure. Indeed, the rules that relate the different deductive systems are actually functors that have specific properties: they are fibrations.

Definition 16.7 (Mac Lane 1971, p.13)**.** *A* functor $\mathcal{C} \xrightarrow{\Phi} \mathcal{B}$ *is a morphism between two categories such that:*

1. *there is* $\Phi(c)$ *in* $\mathcal{B}$ *for each* c *in* $\mathcal{C}$*;*
2. *there is* $\Phi(c_1) \xrightarrow{\Phi(f)} \Phi(c_2)$ *in* $\mathcal{B}$ *for each* $c_1 \xrightarrow{f} c_2$ *in* $\mathcal{C}$*;*
3. $\Phi(1_c) = 1_{\Phi(c)}$*;*
4. $\Phi(g \circ f) = \Phi(g) \circ \Phi(f)$*.*

Definition 16.8 (Jacobs 1999, p.27)**.** *Let* $\mathcal{C} \xrightarrow{\Phi} \mathcal{B}$ *be a functor. Then,* $c_1 \xrightarrow{f} c_2$ *in* $\mathcal{C}$ *is* Cartesian over $b_1 \xrightarrow{u} b_2$ *in* $\mathcal{B}$ *when:*

1. $\Phi(f) = u$*;*
2. *for every* $c_3 \xrightarrow{g} c_2$ *in* $\mathcal{C}$ *such that* $\Phi(g) = u \circ w$ *for some* $\Phi(c_3) \xrightarrow{w} b_1$*, there is one and only one* $c_3 \xrightarrow{h} c_1$ *such that* $g = f \circ h$*.*

Definition 16.9 (Jacobs 1999, p.27)**.** $\mathcal{C} \xrightarrow{\Phi} \mathcal{B}$ *is a* fibration *when for every* c_i *in* $\mathcal{C}$ *and* $b_i \xrightarrow{u} \Phi(c_i)$ *in* $\mathcal{B}$*, there is* f *in* $\mathcal{C}$ *Cartesian over* u*.*

Consider the following fragments.[24]

Definition 16.10. $\mathcal{PAL}_\bigstar$ *is the (atomic) fragment of* $\mathcal{PAL}$ *where well-formed formulas are defined by:*

$$\varphi := \boldsymbol{\alpha} \mid \bot$$

Definition 16.11. $\mathcal{OL}_\bigstar$ *is the (atomic) fragment of* $\mathcal{OL}$ *where well-formed formulas are defined by:*

$$\varphi := O\alpha \mid \bot$$

[24] Considering the whole fragment or the atomic fragment together with $\top$ would thwarth the possibility of defining these rules as fibrations. We would have to define arrows Cartesian over $\neg\neg\varphi \longrightarrow \varphi$ and $\varphi \longrightarrow \top$, making the stronger structure absorb the other.

The definition of (Ψ) as a fibration requires a little technical modification to $\mathcal{AL}$. Let us augment $\mathcal{AL}$ to $\mathcal{AL}^\delta$ with a dummy equivalence class of actions $[\delta]$, for which the action δ is the representative. The dummy action has to satisfy the axiom (0), and thus it is understood as a dummy initial object.

$$\frac{}{\delta \longrightarrow \varphi}\ (0)$$

The technical usefulness of δ appears when we consider how to define f Cartesian over 0_α (see the two diagrams below). The first option would be to assign an arbitrary $* \xrightarrow{f} \alpha$ Cartesian over 0_α and define $* \xrightarrow{g} \beta$ as an isomorphism. This option, however, is problematic. Firstly, this would turn $*$ into an initial object in $\mathcal{AL}$, and this would have devastating consequences. Indeed, anything would then be deducible from anything. Recall that $*$ is the unit of $\bullet$. We would have $* \longrightarrow \beta \ominus \alpha$ for any β and any α, hence by (cl) we would have $* \bullet \alpha \longrightarrow \beta$ and by (l) $\alpha \longrightarrow \beta$. The second option would be to define both $* \xrightarrow{f} \alpha$ and $* \xrightarrow{g} \beta$ as isomorphisms. However, this would imply that every action in $\mathcal{AL}$ is isomorphic to $*$, hence impossible to perform.

$$\begin{tikzcd}
* \arrow[d, "1_*"'] \arrow[dr, "f"] & \\
* \arrow[r, "f"'] & \alpha
\end{tikzcd}
\qquad
\begin{tikzcd}
\beta \arrow[d, "g^{-1}"'] \arrow[dr, "f \circ g^{-1}"] & \\
* \arrow[r, "f"'] & \alpha
\end{tikzcd}$$

$$\begin{tikzcd}
\bot \arrow[r, "0_\alpha"] & \boldsymbol{\alpha} \\
\bot \arrow[u, "1_\bot"] \arrow[ur, "0_\alpha"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
\bot \arrow[r, "0_\alpha"] & \boldsymbol{\alpha} \\
\boldsymbol{\beta} \arrow[u, "0_\beta^{-1}"] \arrow[ur, "0_\alpha \circ 0_\beta^{-1}"'] &
\end{tikzcd}$$

The introduction of δ is meant to shadow the Cartesian structure induced by $\bot$ in $\mathcal{PAL}$. Roughly speaking, the idea is to sit the compact closed structure of $\mathcal{AL}$ on some dummy initial object δ. There is an obvious inclusion functor:

$$\mathcal{AL} \hookrightarrow \mathcal{AL}^\delta$$

Definition 16.12. $\mathcal{AL}^\delta \xrightarrow{\Psi} \mathcal{PAL}_\bigstar$ *is a fibration. The functor is defined by:*

$$\Psi(*) = \bot$$
$$\Psi(\delta) = \bot$$
$$\Psi(\alpha) = \boldsymbol{\alpha}$$

1. $\perp \xrightarrow[\mathsf{PAL}]{} \boldsymbol{\alpha}$

$$\begin{tikzcd} \delta \arrow[d, "1_\delta"'] \arrow[dr, "0_\alpha"] & \\ \delta \arrow[r, "0_\alpha"'] & \alpha \end{tikzcd} \qquad \begin{tikzcd} \beta \arrow[d, "0_\beta^{-1}"'] \arrow[dr, "0_\alpha \circ 0_\beta^{-1}"] & \\ \delta \arrow[r, "0_\alpha"'] & \alpha \end{tikzcd}$$

$$\begin{tikzcd} \perp \arrow[r, "0_\alpha"] & \boldsymbol{\alpha} \\ \perp \arrow[u, "1_\perp"] \arrow[ur, "0_\alpha"'] & \end{tikzcd} \qquad \begin{tikzcd} \perp \arrow[r, "0_\alpha"] & \boldsymbol{\alpha} \\ \boldsymbol{\beta} \arrow[u, "0_\beta^{-1}"] \arrow[ur, "0_\alpha \circ 0_\beta^{-1}"'] & \end{tikzcd}$$

2. $\boldsymbol{\alpha} \xrightarrow[\mathsf{PAL}]{} \perp$

$$\begin{tikzcd} * \arrow[d, "f"'] \arrow[dr, "1_*"] & \\ \alpha \arrow[r, "f^{-1}"'] & * \end{tikzcd} \qquad \begin{tikzcd} \beta \arrow[d, "f \circ g^{-1}"'] \arrow[dr, "g^{-1}"] & \\ \alpha \arrow[r, "f^{-1}"'] & * \end{tikzcd}$$

$$\begin{tikzcd} \boldsymbol{\alpha} \arrow[r, "0_\alpha^{-1}"] & \perp \\ \perp \arrow[u, "0_\alpha"] \arrow[ur, "1_\perp"'] & \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{\alpha} \arrow[r, "0_\alpha^{-1}"] & \perp \\ \boldsymbol{\beta} \arrow[u, "0_\alpha \circ 0_\beta^{-1}"] \arrow[ur, "0_\beta^{-1}"'] & \end{tikzcd}$$

The first case deals with the fact that $\perp$ is initial in $\mathcal{PAL}$. It says that when $\boldsymbol{\beta}$ is isomorphic to $\perp$, β is isomorphic to δ.

The second case implies that $\boldsymbol{\alpha}$ is isomorphic to $\perp$. There is a Cartesian f^{-1} which makes α isomorphic to $*$, and if there is $\boldsymbol{\beta}$ isomorphic to $\perp$, then β will also be isomorphic to $*$. We define f^{-1} such that $f^{-1} \circ f = 1_*$ and $f \circ f^{-1} = 1_\alpha$, similarly for g^{-1}.

3. $\boldsymbol{\beta} \xrightarrow[\mathsf{PAL}]{} \boldsymbol{\alpha}$

$$\begin{tikzcd} \delta \arrow[d, "0_\beta"'] \arrow[dr, "0_\alpha"] & \\ \beta \arrow[r, "f"'] & \alpha \end{tikzcd} \qquad \begin{tikzcd} \gamma \arrow[d, "h"'] \arrow[dr, "g"] & \\ \beta \arrow[r, "f"'] & \alpha \end{tikzcd}$$

$$\begin{tikzcd} \boldsymbol{\beta} \arrow[r, "\Psi(f)"] & \boldsymbol{\alpha} \\ \perp \arrow[u, "0_\beta"] \arrow[ur, "0_\alpha"'] & \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{\beta} \arrow[r, "\Psi(f)"] & \boldsymbol{\alpha} \\ \boldsymbol{\gamma} \arrow[u, "\Psi(h)"] \arrow[ur, "\Psi(g)"'] & \end{tikzcd}$$

4. $\perp \xrightarrow[\mathsf{PAL}]{} \perp$

$$\begin{tikzcd} * \arrow[d, "1_*"'] \arrow[dr, "1_*"] & \\ * \arrow[r, "1_*"'] & * \end{tikzcd} \qquad \begin{tikzcd} \alpha \arrow[d, "f^{-1}"'] \arrow[dr, "f^{-1}"] & \\ * \arrow[r, "1_*"'] & * \end{tikzcd}$$

$$\begin{tikzcd} \perp \arrow[r, "1_\perp"] & \perp \\ \perp \arrow[u, "1_\perp"] \arrow[ur, "1_\perp"'] & \end{tikzcd} \qquad \begin{tikzcd} \perp \arrow[r, "1_\perp"] & \perp \\ \boldsymbol{\alpha} \arrow[u, "0_\alpha^{-1}"] \arrow[ur, "0_\alpha^{-1}"'] & \end{tikzcd}$$

In the third case, we assume that if there is an arbitrary $\Psi(f)$, then it is the target of an arbitrary f. We use the dummy δ to define f Cartesian over $\Psi(f)$ for the case where $\perp$ is initial, otherwise we simply use the fact that Ψ is a functor.

In the fourth case, we assume that α is isomorphic to $*$ when $\boldsymbol{\alpha}$ is isomorphic to $\perp$.

Definition 16.13. $\mathcal{PAL}_\bigstar \xrightarrow{\Delta} \mathcal{OL}_\bigstar$ *is a fibration defined by:*

$$\Delta(\perp) = \perp$$
$$\Delta(\boldsymbol{\alpha}) = O\alpha$$

1. $\perp \xrightarrow[\mathsf{OL}]{} O\alpha$

$$
\begin{tikzcd} \perp \arrow[d, "1_\perp"'] \arrow[dr, "0_\alpha"] & \\ \perp \arrow[r, "0_\alpha"'] & \boldsymbol{\alpha} \end{tikzcd}
\qquad
\begin{tikzcd} \boldsymbol{\beta} \arrow[d, "0_\beta^{-1}"'] \arrow[dr, "g"] & \\ \perp \arrow[r, "0_\alpha"'] & \boldsymbol{\alpha} \end{tikzcd}
$$

$$
\begin{tikzcd} \perp \arrow[r, "0_{O\alpha}"] & O\alpha \\ \perp \arrow[u, "1_\perp"] \arrow[ur, "0_{O\alpha}"'] & \end{tikzcd}
\qquad
\begin{tikzcd} \perp \arrow[r, "0_{O\alpha}"] & O\alpha \\ O\beta \arrow[u, "0_{O\beta}^{-1}"] \arrow[ur, "\Delta(g)"'] & \end{tikzcd}
$$

2. $O\beta \xrightarrow[\mathsf{OL}]{} O\alpha$

$$
\begin{tikzcd} \perp \arrow[d, "0_\beta"'] \arrow[dr, "0_\alpha"] & \\ \boldsymbol{\beta} \arrow[r, "f"'] & \boldsymbol{\alpha} \end{tikzcd}
\qquad
\begin{tikzcd} \boldsymbol{\gamma} \arrow[d, "h"'] \arrow[dr, "g"] & \\ \boldsymbol{\beta} \arrow[r, "f"'] & \boldsymbol{\alpha} \end{tikzcd}
$$

$$
\begin{tikzcd} O\beta \arrow[r, "\Delta(f)"] & O\alpha \\ \perp \arrow[u, "0_{O\beta}"] \arrow[ur, "0_{O\alpha}"'] & \end{tikzcd}
\qquad
\begin{tikzcd} O\beta \arrow[r, "\Delta(f)"] & O\alpha \\ O\gamma \arrow[u, "\Delta(h)"] \arrow[ur, "\Delta(g)"'] & \end{tikzcd}
$$

3. $O\alpha \xrightarrow[\mathsf{OL}]{} \perp$

$$
\begin{tikzcd} \perp \arrow[d, "0_\alpha"'] \arrow[dr, "1_\perp"] & \\ \boldsymbol{\alpha} \arrow[r, "0_\alpha^{-1}"'] & \perp \end{tikzcd}
\qquad
\begin{tikzcd} \boldsymbol{\beta} \arrow[d, "f"'] \arrow[dr, "0_\beta^{-1}"] & \\ \boldsymbol{\alpha} \arrow[r, "0_\alpha^{-1}"'] & \perp \end{tikzcd}
$$

$$
\begin{tikzcd} O\alpha \arrow[r, "0_{O\alpha}^{-1}"] & \perp \\ \perp \arrow[u, "0_{O\alpha}"] \arrow[ur, "1_\perp"'] & \end{tikzcd}
\qquad
\begin{tikzcd} O\alpha \arrow[r, "0_{O\alpha}^{-1}"] & \perp \\ O\beta \arrow[u, "\Delta(f)"] \arrow[ur, "0_{O\beta}^{-1}"'] & \end{tikzcd}
$$

4. $\perp \xrightarrow[\mathsf{OL}]{} \perp$

$$
\begin{tikzcd} \perp \arrow[d, "1_\perp"'] \arrow[dr, "1_\perp"] & \\ \perp \arrow[r, "1_\perp"'] & \perp \end{tikzcd}
\qquad
\begin{tikzcd} \boldsymbol{\beta} \arrow[d, "0_\beta^{-1}"'] \arrow[dr, "0_\beta^{-1}"] & \\ \perp \arrow[r, "1_\perp"'] & \perp \end{tikzcd}
$$

$$
\begin{tikzcd} \perp \arrow[r, "1_\perp"] & \perp \\ \perp \arrow[u, "1_\perp"] \arrow[ur, "1_\perp"'] & \end{tikzcd}
\qquad
\begin{tikzcd} \perp \arrow[r, "1_\perp"] & \perp \\ O\beta \arrow[u, "0_{O\beta}^{-1}"] \arrow[ur, "0_{O\beta}^{-1}"'] & \end{tikzcd}
$$

Definition 16.14. *$\mathcal{PAL}_\bigstar \xrightarrow{\Pi} \mathcal{OL}_\bigstar$ is a fibration defined by:*

$$\Pi(\perp) = \perp$$
$$\Pi(\boldsymbol{\alpha}) = P_s\alpha$$

The definition goes along the lines of definition 16.13 (replace O by P_s).

Since the composition of two fibrations yields another fibration (see lemma 1.5.5 of Jacobs 1999), it follows that we also have substitution rules such that $\Sigma_O = \Delta \circ \Psi$ and $\Sigma_P = \Pi \circ \Psi$ that go directly from $\mathcal{AL}$ to $\mathcal{OL}_\bigstar$.

Definition 16.15. *$\mathcal{PAL}_\bigstar^\otimes$ is the (atomic) fragment of $\mathcal{PAL}^\otimes$ defined from the well-formed formulas:*

$$\varphi := \ \boldsymbol{\alpha} \mid 0$$

Definition 16.16. *$\mathcal{CNR}_\bigstar$ is the (atomic) fragment of $\mathcal{CNR}$ where well-formed formulas are defined by:*

$$\varphi := \ O\alpha \mid 0$$

Considering that 0 is not initial in $\mathcal{PAL}^{\otimes}$, we do not need the dummy δ to define $\Psi_{\otimes}$. Actually, the cost of keeping δ would be to induce $\mathcal{AL}$'s dummy structure into $\mathcal{PAL}^{\otimes}$, turning 0 into an initial object (which is undesirable, otherwise we would obtain deontic explosion).

Definition 16.17. $\mathcal{AL} \xrightarrow{\Psi_{\otimes}} \mathcal{PAL}^{\otimes}_{\bigstar}$ *is a fibration defined by:*

$$\Psi_{\otimes}(*) = 0$$
$$\Psi_{\otimes}(\alpha) = \boldsymbol{\alpha}$$

1. $0 \xrightarrow[\mathsf{PAL}^{\otimes}]{} \boldsymbol{\alpha}$

$$\begin{tikzcd} * \arrow[d, "1_*"'] \arrow[dr, "f"] & \\ * \arrow[r, "f"'] & \alpha \end{tikzcd} \qquad \begin{tikzcd} \beta \arrow[d, "g^{-1}"'] \arrow[dr, "f\circ g^{-1}"] & \\ * \arrow[r, "f"'] & \alpha \end{tikzcd}$$

$$\begin{tikzcd} 0 \arrow[r, "\Psi_{\otimes}(f)"] & \boldsymbol{\alpha} \\ 0 \arrow[u, "1_0"] \arrow[ur, "\Psi_{\otimes}(f)"'] & \end{tikzcd} \qquad \begin{tikzcd} 0 \arrow[r, "\Psi_{\otimes}(f)"] & \boldsymbol{\alpha} \\ \boldsymbol{\beta} \arrow[u, "\Psi_{\otimes}(g^{-1})"] \arrow[ur, "\Psi_{\otimes}(f\circ g^{-1})"'] & \end{tikzcd}$$

2. $\boldsymbol{\alpha} \xrightarrow[\mathsf{PAL}^{\otimes}]{} 0$

$$\begin{tikzcd} * \arrow[d, "f"'] \arrow[dr, "1_*"] & \\ \alpha \arrow[r, "f^{-1}"'] & * \end{tikzcd} \qquad \begin{tikzcd} \beta \arrow[d, "f\circ g^{-1}"'] \arrow[dr, "g^{-1}"] & \\ \alpha \arrow[r, "f^{-1}"'] & * \end{tikzcd}$$

$$\begin{tikzcd} \boldsymbol{\alpha} \arrow[r, "\Psi_{\otimes}(f^{-1})"] & 0 \\ 0 \arrow[u, "\Psi_{\otimes}(f)"] \arrow[ur, "1_0"'] & \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{\alpha} \arrow[r, "\Psi_{\otimes}(f^{-1})"] & 0 \\ \boldsymbol{\beta} \arrow[u, "\Psi_{\otimes}(f\circ g^{-1})"] \arrow[ur, "\Psi_{\otimes}(g^{-1})"'] & \end{tikzcd}$$

In the first case, we define f^{-1} such that $f^{-1} \circ f = 1_*$ and $f \circ f^{-1} = 1_{\alpha}$, similarly for g^{-1}, when there is an arrow from 0 to $\boldsymbol{\alpha}$ (resp. to $\boldsymbol{\beta}$). We require that $\Psi_{\otimes}(f^{-1}) = \Psi_{\otimes}(f)^{-1}$, and similarly for $\Psi_{\otimes}(g^{-1})$. As such, if there is an arrow from 0 to $\boldsymbol{\alpha}$, then it is an isomorphism, making α isomorphic to $*$.[25]

The same strategy is applied to the second case.

3. $\boldsymbol{\beta} \xrightarrow[\mathsf{PAL}^{\otimes}]{} \boldsymbol{\alpha}$

$$\begin{tikzcd} * \arrow[d, "g"'] \arrow[dr, "f"] & \\ \beta \arrow[r, "f\circ g^{-1}"'] & \alpha \end{tikzcd} \qquad \begin{tikzcd} \gamma \arrow[d, "g"'] \arrow[dr, "f"] & \\ \beta \arrow[r, "h"'] & \alpha \end{tikzcd}$$

$$\begin{tikzcd} \boldsymbol{\beta} \arrow[r, "\Psi_{\otimes}(h)"] & \boldsymbol{\alpha} \\ 0 \arrow[u, "\Psi_{\otimes}(g)"] \arrow[ur, "\Psi_{\otimes}(f)"'] & \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{\beta} \arrow[r, "\Psi_{\otimes}(h)"] & \boldsymbol{\alpha} \\ \boldsymbol{\gamma} \arrow[u, "\Psi_{\otimes}(g)"] \arrow[ur, "\Psi_{\otimes}(f)"'] & \end{tikzcd}$$

4. $0 \xrightarrow[\mathsf{PAL}^{\otimes}]{} 0$

$$\begin{tikzcd} * \arrow[d, "1_*"'] \arrow[dr, "1_*"] & \\ * \arrow[r, "1_*"'] & * \end{tikzcd} \qquad \begin{tikzcd} \alpha \arrow[d, "f^{-1}"'] \arrow[dr, "f^{-1}"] & \\ * \arrow[r, "1_*"'] & * \end{tikzcd}$$

$$\begin{tikzcd} 0 \arrow[r, "1_0"] & 0 \\ 0 \arrow[u, "1_0"] \arrow[ur, "1_0"'] & \end{tikzcd} \qquad \begin{tikzcd} 0 \arrow[r, "1_0"] & 0 \\ \boldsymbol{\alpha} \arrow[u, "\Psi_{\otimes}(f^{-1})"] \arrow[ur, "\Psi_{\otimes}(f^{-1})"'] & \end{tikzcd}$$

[25] This would not have been possible if 0 was initial.

The third case also uses the strategy applied in the first and the second ones. On the right side, we use the fact that $\Psi_\otimes$ is a functor.

The fourth case requires a proposition isomorphic to 0 and an action isomorphic to $*$, using the same strategy.

Definition 16.18. $\mathcal{PAL}^{\otimes}_{\bigstar} \xrightarrow{\Delta_\otimes} \mathcal{CNR}_{\bigstar}$ *is a fibration defined by:*

$$\Delta_\otimes(0) = 0$$
$$\Delta_\otimes(\boldsymbol{\alpha}) = O\alpha$$

1. $0 \xrightarrow[\mathsf{CNR}]{} O\alpha$

$$\begin{tikzcd} 0 \arrow[d, "1_0"'] \arrow[dr, "f"] & \\ 0 \arrow[r, "f"'] & \boldsymbol{\alpha} \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{\beta} \arrow[d, "g^{-1}"'] \arrow[dr, "f\circ g^{-1}"] & \\ 0 \arrow[r, "f"'] & \boldsymbol{\alpha} \end{tikzcd}$$

$$\begin{tikzcd} 0 \arrow[r, "\Delta_\otimes(f)"] & O\alpha \\ 0 \arrow[u, "1_0"] \arrow[ur, "\Delta_\otimes(f)"'] & \end{tikzcd} \qquad \begin{tikzcd} 0 \arrow[r, "\Delta_\otimes(f)"] & O\alpha \\ O\beta \arrow[u, "\Delta_\otimes(g^{-1})"] \arrow[ur, "\Delta_\otimes(f\circ g^{-1})"'] & \end{tikzcd}$$

2. $O\beta \xrightarrow[\mathsf{CNR}]{} O\alpha$

$$\begin{tikzcd} 0 \arrow[d, "g"'] \arrow[dr, "f"] & \\ \boldsymbol{\beta} \arrow[r, "g\circ f^{-1}"'] & \boldsymbol{\alpha} \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{\gamma} \arrow[d, "h"'] \arrow[dr, "g"] & \\ \boldsymbol{\beta} \arrow[r, "f"'] & \boldsymbol{\alpha} \end{tikzcd}$$

$$\begin{tikzcd} O\beta \arrow[r, "\Delta_\otimes(g\circ f^{-1})"] & O\alpha \\ 0 \arrow[u, "\Delta_\otimes(g)"] \arrow[ur, "\Delta_\otimes(f^{-1})"'] & \end{tikzcd} \qquad \begin{tikzcd} O\beta \arrow[r, "\Delta_\otimes(f)"] & O\alpha \\ O\gamma \arrow[u, "\Delta_\otimes(h)"] \arrow[ur, "\Delta_\otimes(g)"'] & \end{tikzcd}$$

3. $O\alpha \xrightarrow[\mathsf{CNR}]{} 0$

$$\begin{tikzcd} 0 \arrow[d, "f"'] \arrow[dr, "1_0"] & \\ \boldsymbol{\alpha} \arrow[r, "f^{-1}"'] & 0 \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{\beta} \arrow[d, "f\circ g^{-1}"'] \arrow[dr, "g^{-1}"] & \\ \boldsymbol{\alpha} \arrow[r, "f^{-1}"'] & 0 \end{tikzcd}$$

$$\begin{tikzcd} O\alpha \arrow[r, "\Delta_\otimes(f^{-1})"] & 0 \\ 0 \arrow[u, "\Delta_\otimes(f)"] \arrow[ur, "1_0"'] & \end{tikzcd} \qquad \begin{tikzcd} O\alpha \arrow[r, "\Delta_\otimes(f^{-1})"] & 0 \\ O\beta \arrow[u, "\Delta_\otimes(f\circ g^{-1})"] \arrow[ur, "\Delta_\otimes(g^{-1})"'] & \end{tikzcd}$$

4. $0 \xrightarrow[\mathsf{CNR}]{} 0$

$$\begin{tikzcd} 0 \arrow[d, "1_0"'] \arrow[dr, "1_0"] & \\ 0 \arrow[r, "1_0"'] & 0 \end{tikzcd} \qquad \begin{tikzcd} \boldsymbol{\alpha} \arrow[d, "f^{-1}"'] \arrow[dr, "f^{-1}"] & \\ 0 \arrow[r, "1_0"'] & 0 \end{tikzcd}$$

$$\begin{tikzcd} 0 \arrow[r, "1_0"] & 0 \\ 0 \arrow[u, "1_0"] \arrow[ur, "1_0"'] & \end{tikzcd} \qquad \begin{tikzcd} 0 \arrow[r, "1_0"] & 0 \\ O\alpha \arrow[u, "\Delta_\otimes(f^{-1})"] \arrow[ur, "\Delta_\otimes(f^{-1})"'] & \end{tikzcd}$$

As in the previous definition, we put f^{-1} such that $f^{-1} \circ f = 1_*$ and $f \circ f^{-1} = 1_\alpha$, similarly for g^{-1}, when there is an arrow from 0 to $\boldsymbol{\alpha}$ (resp. to $\boldsymbol{\beta}$). We require that $\Delta_\otimes(f^{-1}) = \Psi_\otimes(f)^{-1}$, and similarly for $\Delta_\otimes(g^{-1})$.

Definition 16.19. $\mathcal{PAL}^{\otimes}_{\bigstar} \xrightarrow{\Pi_{\otimes}} \mathcal{CNR}_{\bigstar}$ *is a fibration defined similarly to* $\Delta_{\otimes}$ *(where* P_s *replaces* O*).*

From these two fibrations we can also define two substitution rules such that $\Sigma^{\otimes}_{O} = \Delta_{\otimes} \circ \Psi_{\otimes}$ and $\Sigma^{\otimes}_{P} = \Pi_{\otimes} \circ \Psi_{\otimes}$ are fibrations.

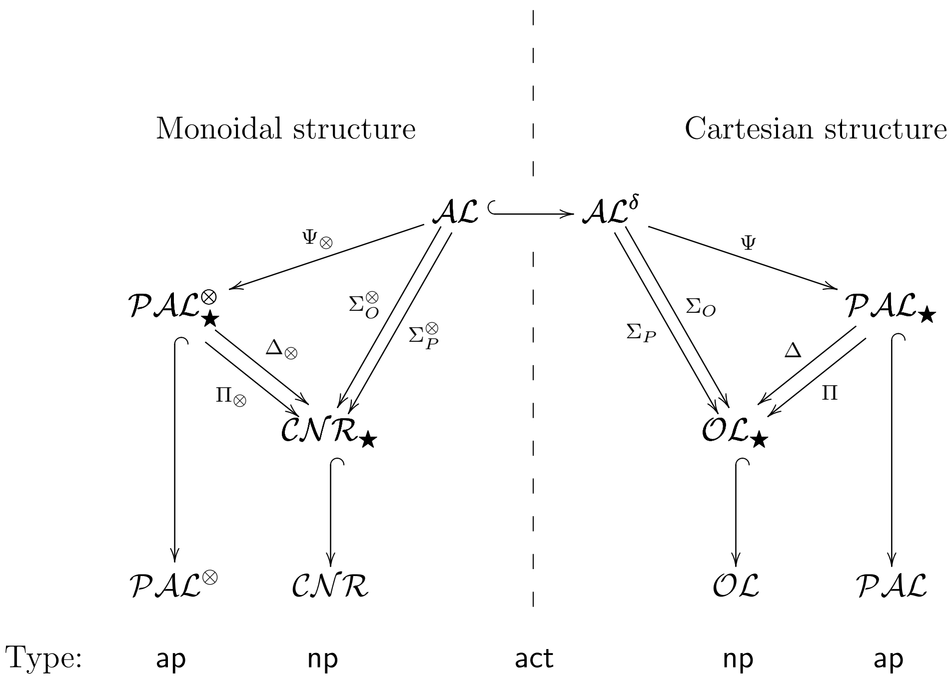


Figure 16.1: $\mathcal{DDS}$

The structure of $\mathcal{DDS}$ is represented in figure 16.1. On the left side of the dashed line, the deductive systems have a monoidal structure, as opposed to the Cartesian structure of the deductive systems on the right side.

Thus defined, $\mathcal{DDS}$ has an interesting property: actions that are isomorphic to $*$ are necessarily *false* in $\mathcal{PAL}$. This is accurate from a philosophical perspective since it would be impossible to accomplish both an action and its complement at the same time.

$$\Psi(\alpha \bullet \alpha^{*} \xrightarrow[\mathsf{AL}]{} *) = \boldsymbol{\alpha} \bullet \boldsymbol{\alpha^{*}} \xrightarrow[\mathsf{PAL}]{} \perp$$

The definition of $\mathcal{DDS}$ as a fibration allows us to preserve the relations between actions and obligations throughout $\mathcal{AL}$, $\mathcal{PAL}$ and $\mathcal{OL}$ (resp. $\mathcal{AL}$, $\mathcal{PAL}^{\otimes}$ and $\mathcal{CNR}$). Hence, if two actions are related (by some arrow) in $\mathcal{AL}$, then they will also be linked in $\mathcal{PAL}$ and $\mathcal{OL}$ (resp. $\mathcal{PAL}^{\otimes}$ and $\mathcal{CNR}$). It also allows us to distinguish between actions that are impossible to perform versus action propositions that are necessarily false. This answers the philosophical intuitions that the notions of *tautologous* and *contradictory* actions do not make any sense. Although a conjunctive action might be impossible to perform, the actions are not *contradictory*: they are only incompatible. That said, if it is impossible to accomplish α and β together, then necessarily $\boldsymbol{\alpha} \bullet \boldsymbol{\beta}$ will be false. This is exactly what we obtain through the definition of (Ψ).

$$\Psi(\alpha \bullet \beta \xrightarrow[\mathsf{AL}]{} *) = \boldsymbol{\alpha} \bullet \boldsymbol{\beta} \xrightarrow[\mathsf{PAL}]{} \bot$$

Combined with (Δ), this entails an interesting property: actions isomorphic to $*$ cannot be obligatory. The same results hold for $\mathcal{PAL}^{\otimes}$ and $\mathcal{CNR}$.

$$\Delta(\boldsymbol{\alpha} \bullet \boldsymbol{\alpha}^{*} \xrightarrow[\mathsf{PAL}]{} \bot) = O(\alpha \bullet \alpha^{*}) \xrightarrow[\mathsf{OL}]{} \bot$$
$$\Delta(\boldsymbol{\alpha} \bullet \boldsymbol{\beta} \xrightarrow[\mathsf{PAL}]{} \bot) = O(\alpha \bullet \beta) \xrightarrow[\mathsf{OL}]{} \bot$$

Looking at (Σ_O), this allows us to formally represent the *ought-implies-can* principle: if it is impossible to accomplish some action α (i.e., $\alpha \cong *$), then it is false that α is obligatory. As such, actions that are isomorphic to the action which is impossible to perform cannot be obligatory. Hence, 'ought' implies 'can'. An interesting feature of $\mathcal{DDS}$ is that it does not require any alethic modality to represent this principle.

To strengthen the ought-implies-can principle, one could add the axiom $(\Diamond)$ to $\mathcal{DDS}$, although it would not add much significance to the ought-implies-can principle.[26]

$$\frac{}{O* \xrightarrow[\mathsf{OL}]{} \bot} \qquad (\Diamond)$$

A comparison with Goble's analysis

Goble (2009) provided a thorough analysis of the conditions that a logic which aims to model normative conflicts should satisfy.[27] His analysis is grounded on two fundamental arguments (Goble 2009, pp.450-1).

Argument 1. Aggregation of obligations together with the ought-implies-can principle lead to the impossibility of normative conflicts.

Argument 2. The distribution principle 'if necessarily α implies β, then $O\alpha$ implies $O\beta$' together with the axiom schema for normative consistency lead to the impossibility of normative conflicts.

[26] Note that $\Delta(*) = O*$.
[27] See also Goble (2004).

The four principles at play within these arguments are:

(Agg) If α is obligatory and β is obligatory, then their conjunction is also obligatory.
(Can) If α is obligatory, then it is possible to accomplish α.
(Dist) If α implies β, then if α is obligatory, then β is obligatory.
(D) If α is obligatory, then not-α is not.

Despite the second argument, Goble (2009, p.469) argues that (Dist) must not be rejected altogether given that it is a fundamental principle for a logic of ought. We agree. It enables us to infer from a set of ought statements other obligations that are not explicitly mentioned themselves. In this respect, a form of distribution is necessary to minimally represent the principle of deontic consequence (i.e., if α implies β, then $O\alpha$ implies $O\beta$).[28]

In his paper, Goble analyzes carefully how all these principles are related and what are the possible modifications one can make to try to answer the two aforementioned arguments. But more importantly, he argues that a logic that wishes to model normative conflicts and answer these arguments must satisfy four adequacy criteria (Goble 2009, pp.458-60, 470)[29]:

1. consistency (a conflict must not entail a contradiction);
2. non-triviality (the logic must avoid deontic explosion);
3. the logic must satisfy the Smith argument;
4. the logic must satisfy the converse of (Agg).

To answer these concerns, Goble (2009, pp.451-2) argues that it is not the ought-implies-can principle which is problematic, nor the aggregation principle for that matter, but that it is rather the unrestricted distribution principle (Dist). Note, however, that in Goble's (2009, p.476) view, the principle (D) should be rejected since it is a 'no-conflict' principle which is not adequate to model obligations within real life. On this point, that (D) should be discarded, we disagree. $\mathcal{DDS}$ aims at modeling the (Canadian) legal discourse, and as such its presupposed consistency (after interpretation) is irrefutable.[30]

The first argument is attributed to Lemmon (1965), who argued that aggregation and the ought-implies-can principle, together with the axiom schema (D) of standard deontic logic, should be rejected since it thwarts the possibility of normative conflicts (see also Lemmon 1962). Strictly speaking, this objection fails to affect $\mathcal{DDS}$ since it does not satisfy (Agg). Nonetheless, it remains that our version of (Can) together with the presumption of consistency (D) do imply the impossibility of normative conflicts.

[28] See Castañeda (1968, p.13). This also enables us to represent the distinction between fixed and derived obligations (cf. Alchourrón and Bulygin 1981, p.102).

[29] Another criterion might be added, namely that the formal system must not allow for the pragmatic oddity (cf. Carmo and Jones 2002). $\mathcal{CNR}$ answers this criterion (see Peterson 2014d).

[30] See *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, [1996] 3 RCS 919, paragraphe 207.

Consider only the case of $\mathcal{CNR}$ for example. Assuming that two actions α and β cannot be performed together (i.e., assuming $\alpha \bullet \beta \xrightarrow[\mathsf{AL}]{} *$), we obtain 16.1 by $(\Sigma^{\otimes}_{O})$. Note that it follows from an instance of (Dist). This allows us to derive that the assumption $O\alpha \otimes O\beta$ fails within $\mathcal{CNR}$ (or that it is false in $\mathcal{OL}$), which is represented by 16.2. From 16.2 we can derive 16.3. The proofs are listed below the theorems (we omit the steps using associativity and symmetry).

$$O\alpha \bullet \beta \xrightarrow[\mathsf{CNR}]{} 0 \tag{16.1}$$

$$O\alpha \otimes O\beta \xrightarrow[\mathsf{CNR}]{} 0 \tag{16.2}$$

$$c \otimes (c \multimap (O\alpha \otimes O\beta)) \xrightarrow[\mathsf{CNR}]{} 0 \tag{16.3}$$

Proof. 16.1, 16.2, 16.3

$$\dfrac{\dfrac{}{\alpha \bullet \beta \xrightarrow[\mathsf{AL}]{} *}\ (\mathrm{H})}{O\alpha \bullet \beta \xrightarrow[\mathsf{CNR}]{} 0}\ (\Sigma^{\otimes}_{O})$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\alpha \bullet \beta \xrightarrow[\mathsf{AL}]{} *}\ (\mathrm{H})}{\alpha \xrightarrow[\mathsf{AL}]{} * \ominus \beta}\ (\mathrm{cl})}{\alpha \xrightarrow[\mathsf{AL}]{} \beta^*}}{O\alpha \xrightarrow[\mathsf{CNR}]{} O\beta^*}\ (\Sigma^{\otimes}_{O}) \quad \dfrac{}{O\beta \xrightarrow[\mathsf{CNR}]{} O\beta}\ (1)}{O\alpha \otimes O\beta \xrightarrow[\mathsf{CNR}]{} O\beta^* \otimes O\beta}\ (\mathrm{t}) \quad \dfrac{\vdots}{O\beta^* \otimes O\beta \xrightarrow[\mathsf{CNR}]{} 0}\ (\mathrm{NC})}{O\alpha \otimes O\beta \xrightarrow[\mathsf{CNR}]{} 0}\ (\mathrm{cut})$$

$$\dfrac{\dfrac{\dfrac{}{c \multimap (O\alpha \otimes O\beta) \xrightarrow[\mathsf{CNR}]{} c \multimap (O\alpha \otimes O\beta)}\ (1)}{c \otimes (c \multimap (O\alpha \otimes O\beta)) \xrightarrow[\mathsf{CNR}]{} O\alpha \otimes O\beta}\ (\mathrm{cl}) \quad \dfrac{\vdots}{O\alpha \otimes O\beta \xrightarrow[\mathsf{CNR}]{} 0}\ (\text{from 16.2})}{c \otimes (c \multimap (O\alpha \otimes O\beta)) \xrightarrow[\mathsf{CNR}]{} 0}\ (\mathrm{cut})$$

□

As such, the following reformulation of argument 1 does affect $\mathcal{DDS}$.

Argument 3. The ought-implies-can principle together with the presumption of normative consistency entail the impossibility of *unconditional* normative conflicts, or of conditional conflicting obligations that hold under the same context.

Thus formulated, however, this argument can be reduced to argument 2 since the (Can) principle is derived from (Dist). That being said, it is important to see within the previous reasonings that the impossibility of normative conflicts happens when we assume that both $O\alpha$ and $O\beta$ hold *together* under the same circumstances. As a result, this objection fails to affect $\mathcal{DDS}$ since it violates the presumption of normative consistency. Legally speaking, conflicting obligations can only arise *a priori* within specific context. Otherwise, if there were unconditional conflicting obligations or (*a posteriori*) conflicting obligations that hold under the same conditions, the discourse would fail to be rational.[31] Recall that $\mathcal{DDS}$ takes place after interpreting the law. As such, the aforementioned normative conflicts are indeed impossible, and so arguments 2 and 3 are actually arguments in favour of $\mathcal{DDS}$.

Now, consider the possibility of conditional conflicting obligations. Suppose that under circumstances c_1, $O\alpha$ holds, that in the context c_2, $O\beta$ holds, and that α and β cannot be accomplished together. There are four possibilities:

1. neither c_1 nor c_2 holds;
2. c_1 holds;
3. c_2 holds;
4. $c_1 \otimes c_2$ holds.

The first case is unproblematic since $(c_1 \multimap O\alpha) \otimes (c_2 \multimap O\beta)$ entails neither $O\alpha$, $O\beta$ nor $O\alpha \otimes O\beta$.

The second and the third cases are analyzed similarly. Under the assumption that $c_1 \otimes ((c_1 \multimap O\alpha) \otimes (c_2 \multimap O\beta))$ holds, we can derive 16.4. Considering that $\alpha \bullet \beta$ is impossible to perform, we also have 16.5.

$$c_1 \otimes ((c_1 \multimap O\alpha) \otimes (c_2 \multimap O\beta)) \xrightarrow[\mathsf{CNR}]{} O\alpha \otimes (c_2 \multimap O\beta) \tag{16.4}$$

$$O\alpha \otimes (c_2 \multimap O\beta) \xrightarrow[\mathsf{CNR}]{} O\beta^{*} \otimes (c_2 \multimap O\beta) \tag{16.5}$$

Proof. 16.4, 16.5.

$$\dfrac{\dfrac{\dfrac{}{c_1 \multimap O\alpha \xrightarrow[\mathsf{CNR}]{} c_1 \multimap O\alpha}\,(1)}{c_1 \otimes (c_1 \multimap O\alpha) \xrightarrow[\mathsf{CNR}]{} O\alpha}\,(\text{cl}) \qquad \dfrac{}{c_2 \multimap O\beta \xrightarrow[\mathsf{CNR}]{} c_2 \multimap O\beta}}{c_1 \otimes ((c_1 \multimap O\alpha) \otimes (c_2 \multimap O\beta)) \xrightarrow[\mathsf{CNR}]{} O\alpha \otimes (c_2 \multimap O\beta)}\,(\text{t})$$

[31] As we said at the beginning of this paper, the Supreme Court of Canada assumes normative consistency as a criterion that enables us to determine whether or not legal discourse is rational (cf. *2747-3174 Québec Inc. c. Québec (Régie des permis d'alcool)*, [1996] 3 RCS 919).

$$\dfrac{\dfrac{\dfrac{\vdots}{\alpha \xrightarrow[\mathsf{AL}]{} \beta^*}\ \text{(from the assumption in the proof of 16.1)}}{O\alpha \xrightarrow[\mathsf{CNR}]{} O\beta^*}\ (\Sigma_O^{\otimes}) \qquad \dfrac{}{c_2 \multimap O\beta \xrightarrow[\mathsf{CNR}]{} c_2 \multimap O\beta}}{O\alpha \otimes (c_2 \multimap O\beta) \xrightarrow[\mathsf{CNR}]{} O\beta^* \otimes (c_2 \multimap O\beta)}\ \text{(t)}$$

□

However, since c_2 is not assumed, $O\beta$ cannot be derived. As a result, under the context c_1, we obtain that α is obligatory although $O\beta$ holds under context c_2, or that β^* is obligatory although $O\beta$ holds under context c_2, which are perfectly consistent. Moreover, their interpretation is meaningful within the natural language: even though β is a conditional obligation that holds under context c_2, if we are under context c_1, then α is an actual obligation.

The fourth case might appear as the most problematic, but it actually is not. Of course, assuming that $c_1 \otimes c_2$ holds would allow us to derive $O\beta^* \otimes O\beta$. However, the presumption of consistency implies that after interpretation, the law must be consistent, and as such if $c_1 \otimes c_2$ holds, then one of the conflicting obligation will have priority over the other, and hence the conflict will be solved.

Prioritizing an obligation in case of a conflict was advocated by Alchourrón and Makinson (1981). More recently, van der Torre and Tan (1997) argued that there are three different types of resolution for normative conflicts, depending upon the nature of the conflict. In a nutshell, they proposed that:

1. If a conflict arises from a contrary-to-duty, then the violated obligation is overshadowed by the contrary-to-duty obligation.

2. If there is a conflict between a *prima facie* and a conditional obligation, the *prima facie* ought is overshadowed by the conditional obligation.

3. If there is a conflict between two norms in a specific context, then one norm will override the other in that context.

Following van der Torre's and Tan's analysis, if $c_1 \otimes c_2$ holds, then one obligation will override the other, either by overshadowing it or by cancelling it. As such, we need to specify which obligation is overridden by stipulating the conditions under which the obligations hold. For instance, if $O\alpha$ overrides $O\beta$ in the context $c_1 \otimes c_2$, then $O\alpha$ holds under the circumstances that c_1, c_2 and $c_2 \multimap O\beta$. Put differently, the assumption that $O\alpha$ overrides $O\beta$ can be translated by:

$$((c_1 \otimes c_2) \otimes (c_2 \multimap O\beta)) \multimap O\alpha$$

From this assumption and the context $(c_1 \otimes c_2) \otimes (c_2 \multimap O\beta)$, one can easily derive that $O\alpha$ is the obligation in force.

All things considered, $\mathcal{DDS}$ can properly model normative conflicts insofar as we properly specify the conditions under which the obligations hold (see Peterson 2014d for details). As such, arguments 1 and 2 fail to affect $\mathcal{DDS}$, although (Can), (Dist) and (D) are satisfied.

Let us now turn our attention to the four criteria advocated by Goble (2009). It was shown in Peterson (2014d) that $\mathcal{CNR}$, as a foundational framework, satisfies the criteria of consistency and non-triviality. This will be exemplified later in the discussion, but for now simply note that consistency is satisfied when we specify the conditions under which the obligations hold, and non-triviality is satisfied given that $\mathcal{CNR}$ does not validate *ex falso sequitur quodlibet*.

The third criterion requires the satisfaction of the following inference pattern.

1.	It is obligatory that either α or β.
2.	It is obligatory that not-α.
$\therefore$	It is obligatory that β.

This inference pattern was initially presented by Horty (1994, 1997) as the Smith argument, which is an argument in favour of the aggregation principle (cf. Goble 2009, p.459). The Smith argument goes along the following lines:

1.	Smith ought to fight in the army or else perform alternative service to his country.
2.	Smith ought to not fight in the army (say, because he needs to stay home to help his mother).
$\therefore$	Smith ought to perform alternative service to his country.

This can be seen as an argument in favour of the aggregation principle insofar as the reasoning is usually translated by:

$$\dfrac{\dfrac{O(p \vee q) \qquad O\neg p}{O((p \vee q) \wedge \neg p)}}{Oq}$$

Considering that $((p \vee q) \wedge \neg p) \supset q$, Oq follows from $O((p \vee q) \wedge \neg p)$. As we saw earlier, there is no such thing, in our view, as a *disjunctive* action. As such, assuming that only *actions* can be in the scope of a deontic operator, it follows that the aforementioned translation of the Smith argument is not accurate. Even though this translation is not available within $\mathcal{DDS}$, it happens that it is nonetheless possible to model the validity of the Smith argument. This, however, can be done *without* requiring any aggregation principle. The Smith argument can be translated within $\mathcal{DDS}$'s language by:

$$O\alpha \vee O\beta \tag{16.6}$$

$$O\alpha^{*} \tag{16.7}$$

$$O\beta \tag{16.8}$$

By $(\mathbf{D})$ and $(\mathbf{P})$ we have $O\alpha^* \xrightarrow[\mathsf{OL}]{} \neg O\alpha$, and thus it is easy to show 16.9 using (Cart). By (cut), 16.9 and 16.10, we can easily derive 16.11.

$$O\alpha^* \wedge (O\alpha \vee O\beta) \xrightarrow[\mathsf{OL}]{} \neg O\alpha \wedge (O\alpha \vee O\beta) \tag{16.9}$$

$$\neg O\alpha \wedge (O\alpha \vee O\beta) \xrightarrow[\mathsf{OL}]{} O\beta \tag{16.10}$$

$$O\alpha^* \wedge (O\alpha \vee O\beta) \xrightarrow[\mathsf{OL}]{} O\beta \tag{16.11}$$

Proof. 16.10

$$
\begin{prooftree}
\AxiomC{}
\LeftLabel{(1)}
\UnaryInfC{$\neg O\alpha \xrightarrow[\mathsf{OL}]{} O\alpha \supset \bot$}
\LeftLabel{(cl)}
\UnaryInfC{$\neg O\alpha \wedge O\alpha \xrightarrow[\mathsf{OL}]{} \bot$}
\AxiomC{}
\LeftLabel{(0)}
\UnaryInfC{$\bot \xrightarrow[\mathsf{OL}]{} O\beta$}
\LeftLabel{(cut)}
\BinaryInfC{$\neg O\alpha \wedge O\alpha \xrightarrow[\mathsf{OL}]{} O\beta$}
\LeftLabel{(cl)}
\UnaryInfC{$O\alpha \xrightarrow[\mathsf{OL}]{} \neg O\alpha \supset O\beta$}
\AxiomC{}
\LeftLabel{(!)}
\UnaryInfC{$\neg O\alpha \xrightarrow[\mathsf{OL}]{} \top$}
\AxiomC{}
\RightLabel{(1)}
\UnaryInfC{$O\beta \wedge \top \xrightarrow[\mathsf{OL}]{} O\beta \wedge \top$}
\RightLabel{(Cart)}
\UnaryInfC{$O\beta \wedge \top \xrightarrow[\mathsf{OL}]{} O\beta$}
\RightLabel{(cl)}
\UnaryInfC{$\top \xrightarrow[\mathsf{OL}]{} O\beta \supset O\beta$}
\RightLabel{(cut)}
\BinaryInfC{$\neg O\alpha \xrightarrow[\mathsf{OL}]{} O\beta \supset O\beta$}
\RightLabel{(cl)}
\UnaryInfC{$\neg O\alpha \wedge O\beta \xrightarrow[\mathsf{OL}]{} O\beta$}
\RightLabel{(cl)}
\UnaryInfC{$O\beta \xrightarrow[\mathsf{OL}]{} \neg O\alpha \supset O\beta$}
\LeftLabel{(coCart)}
\BinaryInfC{$(O\alpha \vee O\beta) \xrightarrow[\mathsf{OL}]{} \neg O\alpha \supset O\beta$}
\LeftLabel{(cl)}
\UnaryInfC{$\neg O\alpha \wedge (O\alpha \vee O\beta) \xrightarrow[\mathsf{OL}]{} O\beta$}
\end{prooftree}
$$

□

As a result, $\mathcal{DDS}$ satisfies Goble's third criterion. It is noteworthy that Goble (2004, p.78) argued that although deontic explosion comes from *ex falso sequitur quodlibet*, the principle should not be discarded since otherwise the Smith argument cannot be validated. As it happens, $\mathcal{DDS}$ does satisfy the Smith argument while rejecting *ex falso sequitur quodlibet*.

The fourth criterion regards the following inference pattern.

1.	It is obligatory that α and β.
$\therefore$	It s obligatory that α.

This is the converse of the aggregation principle, translated by:

$$O(p \wedge q) \supset (Op \wedge Oq) \tag{16.12}$$

A first thing to point out is that Goble (2009, p.470) mentions that any logic satisfying (Dist) immediately satisfies the fourth criterion. In light of $\mathcal{DDS}$'s construction, this assumption is inaccurate: $\mathcal{DDS}$ satisfies (Dist) but does not satisfy (Agg) or its converse.

Secondly, Goble (2009, p.469) mentions that this inference pattern holds for any type of ought statement, including legal oughts. On this point we disagree. Contra Goble, we think that the fourth criterion does not represent a valid inference pattern for legal reasoning. This is mainly motivated by the type of action conjunction we use to model human actions. As we saw, an action conjunction $\alpha \bullet \beta$ represents the complex action 'α together with β', and this implies simultaneity. As such, it should be obvious to the reader that $O(\alpha \bullet \beta) \xrightarrow[\mathsf{OL}]{} O\alpha$ does not hold. From a legal point of view, if the (simultaneous) conjunction of two actions is obligatory, it does not imply that the actions are obligatory individually (not simultaneously). Similarly, it is not because two actions α and β are obligatory independently that the conjunctive (simultaneous) action $\alpha \bullet \beta$ is obligatory. For instance, even though it is true that Paul has an obligation to help John and Jane who are drowning, this does not necessarily mean that he ought to help them *at the same time*.

Although neither (Agg) nor its converse are satisfied, it is noteworthy that $\mathcal{DDS}$ can still answer Goble's fourth criterion. Indeed, the converse of (Agg) can be obtained depending on the premises one adopts. For instance, if one was to assume 16.13, then by ($\Delta_{\otimes}$) or (Δ) one would be able to derive 16.14 or 16.15.

$$\boldsymbol{\alpha} \bullet \boldsymbol{\beta} \xrightarrow[\mathsf{PAL}^{\otimes}]{} \boldsymbol{\beta} \tag{16.13}$$

$$O(\alpha \bullet \beta) \xrightarrow[\mathsf{CNR}]{} O\beta \tag{16.14}$$

$$O(\alpha \bullet \beta) \xrightarrow[\mathsf{OL}]{} O\beta \tag{16.15}$$

In addition to the satisfaction of Goble's criteria, it is noteworthy that incidentally, $\mathcal{DDS}$'s construction enables us to answer Schotch's and Jennings's (1981) objection against the modal system KD. Schotch and Jennings argued that the standard system fails to distinguish between the ought-implies-can principle and the presumption of normative consistency given the equivalence between the axiom schema (D) and $\neg O\bot$, which follows from the unrestricted use of the aggregation principle $O\varphi \wedge O\psi \vdash_{KD} O(\varphi \wedge \psi)$.

In our framework, these two principle are independent: while the presumption of normative consistency follows from ($\mathbf{D}$) and ($\mathbf{P}$), the ought-implies-can principle follows from the definition of (Σ_O) as a fibration. As such, whether or not we adopt ($\Diamond$), $\mathcal{DDS}$ answers the philosophical intuition that there is a distinction between the presumption of consistency of obligations and the ought-implies-can principle.

Various aggregation principles could be added to $\mathcal{DDS}$, but it is beyond the scope of this paper to analyze the consequences of adding one or another. Instead, simply notice that adding ($\mathbf{A}$) as an axiom to $\mathcal{OL}$ (resp. $\mathcal{CNR}$) would blur the distinction between the presumption of normative consistency and the ought-implies-can principle. Indeed, $O\alpha \wedge O\alpha^* \xrightarrow[\mathsf{OL}]{} \bot$ would be derivable from ($\mathbf{A}$) and (Σ_O). Thus, adding ($\mathbf{A}$) to $\mathcal{DDS}$ would reopen the door to Schotch's and Jennings's objection.

$$\frac{}{O\alpha \wedge O\beta \xrightarrow[\mathsf{OL}]{} O(\alpha \bullet \beta)} \quad (\mathbf{A})$$

Discussion

The capacity of $\mathcal{CNR}$ to model contrary-to-duty reasoning and normative conflicts has already been discussed at length in Peterson (2014d). For the purpose of this paper, we only exemplify how $\mathcal{DDS}$ satisfies Goble's first and second criteria of consistency and non-triviality. We also provide a discussion of the good Samaritan paradox and conclude with some example of analysis of legal reasoning.

Contrary-to-duty

Chisholm's (1963) paradox shows that von Wright's (1951) initial approach and, more generally, that monadic standard systems cannot model properly contrary-to-duty reasoning. Chisholm's puzzle amounts to the fact that the following set of propositions is inconsistent within a standard system, while it seems perfectly consistent in the natural language.

1. It is forbidden to exceed the speed limit on the highway (100 km/h).
2. If Paul exceeds the speed limit, then he ought to slowdown to 100 km/h.
3. Paul does not ought to slowdown to 100 km/h if he does not exceed the speed limit.
4. Paul exceeds the speed limit.

To solve Chisholm's paradox, a deontic logic must be able to provide an independent translation for each of these sentences (cf. Åqvist 1967; Decew 1981; Tomberlin 1981, 1983; Hansen 1999), and these translations must be consistent with each other. Moreover, one must be able to determine which obligation holds. In $\mathcal{DDS}$, Chisholm's puzzle is translated by:

$$O\alpha^* \qquad \mathsf{np/act} \tag{16.16}$$
$$\boldsymbol{\alpha} \multimap O\beta \qquad \mathsf{ap\backslash np/np} \tag{16.17}$$
$$\boldsymbol{\alpha}^* \multimap\!\sim O\beta \qquad \mathsf{ap\backslash np/np} \tag{16.18}$$
$$\boldsymbol{\alpha} \qquad \mathsf{ap} \tag{16.19}$$

These premises are non-redundant and give us the following result, which is perfectly consistent.

$$(\boldsymbol{\alpha}^* \multimap\!\sim O\beta) \otimes (O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta))) \xrightarrow[\mathsf{CNR}]{} (\boldsymbol{\alpha}^* \multimap\!\sim O\beta) \otimes (O\alpha^* \otimes O\beta) \tag{16.20}$$

Proof. 16.20

$$\dfrac{\dfrac{}{\boldsymbol{\alpha^*} \multimap\!\sim O\beta \xrightarrow[\mathsf{CNR}]{} \boldsymbol{\alpha^*} \multimap\!\sim O\beta}\,(1) \qquad \dfrac{\dfrac{}{O\alpha^* \xrightarrow[\mathsf{CNR}]{} O\alpha^*}\,(1) \qquad \dfrac{\dfrac{}{\boldsymbol{\alpha} \multimap O\beta \xrightarrow[\mathsf{CNR}]{} \boldsymbol{\alpha} \multimap O\beta}\,(1)}{\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta) \xrightarrow[\mathsf{CNR}]{} O\beta}\,(\mathrm{cl})}{O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta)) \xrightarrow[\mathsf{CNR}]{} O\alpha^* \otimes O\beta}\,(\mathrm{t})}{(\boldsymbol{\alpha^*} \multimap\!\sim O\beta) \otimes (O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta))) \xrightarrow[\mathsf{CNR}]{} (\boldsymbol{\alpha^*} \multimap\!\sim O\beta) \otimes (O\alpha^* \otimes O\beta)}\,(\mathrm{t})$$

$\square$

From these premises, we can also derive $O\alpha^* \otimes \sim \boldsymbol{\alpha^*}$ (the proof is left to the reader), which also is a desirable consequence: it is forbidden to exceed the speed limit on the highway but it is false that Paul is not exceeding the speed limit. As such, $\mathcal{DDS}$ is able to represent the fact that an obligation has been violated.

Nute and Yu (1997, p.6) argued that Chisholm's paradox follows when a logic allows for both *factual* and *deontic* detachment.[32] In standard deontic logic, 16.16 and 16.18 are translated respectively as $O\neg p$ and $O(\neg p \supset \neg q)$, which by deontic detachment leads to $O\neg q$. This contradicts Oq, which can be factually detached from the other premises. Let us see what happens if we assumed a form of deontic detachment and instead translated 16.18 by 16.21, expressing that ideally it is false that Paul ought to slow down. From these premises, we would obtain 16.22.

$$\sim O\beta \qquad \mathsf{np/np} \tag{16.21}$$

$$O\alpha^* \otimes (\sim O\beta \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta))) \xrightarrow[\mathsf{CNR}]{} O\alpha^* \otimes (\sim O\beta \otimes O\beta) \tag{16.22}$$

Proof. 16.22

$$\dfrac{\dfrac{}{O\alpha^* \xrightarrow[\mathsf{CNR}]{} O\alpha^*}\,(1) \qquad \dfrac{\dfrac{}{\sim O\beta \xrightarrow[\mathsf{CNR}]{} \sim O\beta}\,(1) \qquad \dfrac{\dfrac{}{\boldsymbol{\alpha} \multimap O\beta \xrightarrow[\mathsf{CNR}]{} \boldsymbol{\alpha} \multimap O\beta}\,(1)}{\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta) \xrightarrow[\mathsf{CNR}]{} O\beta}\,(\mathrm{cl})}{\sim O\beta \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta)) \xrightarrow[\mathsf{CNR}]{} \sim O\beta \otimes O\beta}\,(\mathrm{t})}{O\alpha^* \otimes (\sim O\beta \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta))) \xrightarrow[\mathsf{CNR}]{} O\alpha^* \otimes (\sim O\beta \otimes O\beta)}\,(\mathrm{t})$$

$\square$

[32] See Loewer and Belzer (1983) and Vorobej (1986) for a discussion on detachment.

Note, however, that since $\mathcal{CNR}$ does not satisfy (Cart), it follows that the members of the conjunction cannot be detached. As such, neither $O\alpha^*$, $O\beta$, $\sim O\beta$ or $\sim O\beta \otimes O\beta$ is derivable. What we *do* have, from 16.22 and (cut), is 16.23. This result does *not* yield 0 since neither (Cart) nor (0) are available (see Peterson 2014d). Hence, interpreting 0 as a failure, we obtain that even though some part of a package of normative premises leads to a failure, it does not imply that the premises fail altogether. From a legal point of view, this is an interesting property: it is not because some norms are inconsistent with each other that the whole legal system fails altogether.[33]

$$O\alpha^* \otimes (\sim O\beta \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta))) \xrightarrow[\textsf{CNR}]{} O\alpha^* \otimes 0 \tag{16.23}$$

Proof. 16.23

$$\dfrac{\dfrac{\vdots}{16.22} \qquad \dfrac{(1)\ \dfrac{}{O\alpha^* \xrightarrow[\textsf{CNR}]{} O\alpha^*} \qquad \dfrac{\dfrac{}{\sim O\beta \xrightarrow[\textsf{CNR}]{} O\beta \multimap 0}\ (1)}{\sim O\beta \otimes O\beta \xrightarrow[\textsf{CNR}]{} 0}\ (\text{cl})}{O\alpha^* \otimes (\sim O\beta \otimes O\beta) \xrightarrow[\textsf{CNR}]{} O\alpha^* \otimes 0}\ (\text{t})}{O\alpha^* \otimes (\sim O\beta \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha} \multimap O\beta))) \xrightarrow[\textsf{CNR}]{} O\alpha^* \otimes 0}\ (\text{cut})$$

$\square$

In both cases, the premises do not lead to 0, and as such $\mathcal{DDS}$ satisfies Goble's criterion of consistency. With both translation, $\mathcal{DDS}$ is able to represent the intuitive consistency of Chisholm's set. However, these premises do not allow us to conclude which obligation is in force. To model that $O\beta$ should follow from this set, we need to appropriately translate the second premise. Van der Torre and Tan (1997) argued that Chisholm's paradox is a case of factual defeasibility, where the contrary-to-duty obligation $O\beta$ overshadows the violated obligation $O\alpha^*$. As such, it is assumed that $O\beta$ holds under the circumstances that $\boldsymbol{\alpha}$, $O\alpha^*$ and $\boldsymbol{\alpha^*} \multimap\sim O\beta$ hold. Hence, 16.17 would rather be translated in $\mathcal{DDS}$ by 16.24, and we would thus get the desired result via 16.25.[34]

$$(O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha^*} \multimap\sim O\beta))) \multimap O\beta \tag{16.24}$$

$$(O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha^*} \multimap\sim O\beta))) \otimes (O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha^*} \multimap\sim O\beta))) \multimap O\beta \xrightarrow[\textsf{CNR}]{} O\beta \tag{16.25}$$

[33] This is also an interesting property if $\mathcal{DDS}$ is to have any repercussion on the field of artificial intelligence: some instructions may still be in force even though others fail.

[34] It can be objected that to obtain the desired result, one needs to make *ad hoc* modifications to the premises and do some previous inferences to encode the desired result within the premises. This objection has been dealt with in Peterson (2014d).

Proof. 16.25

$$\dfrac{\overline{(O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha^*} \multimap\!\sim O\beta))) \multimap O\beta \xrightarrow[\mathsf{CNR}]{} (O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha^*} \multimap\!\sim O\beta))) \multimap O\beta}\ (1)}{(O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha^*} \multimap\!\sim O\beta))) \otimes (O\alpha^* \otimes (\boldsymbol{\alpha} \otimes (\boldsymbol{\alpha^*} \multimap\!\sim O\beta))) \multimap O\beta \xrightarrow[\mathsf{CNR}]{} O\beta}\ (\mathrm{cl})$$

□

This last theorem exemplify the fact that $\mathcal{CNR}$ allows for *restricted* (factual) detachment, but only under specific circumstances. It was shown in Peterson (2014d) that $\mathcal{CNR}$ avoids the problem of augmentation (cf. Jones 1991), also known as the problem of strengthening the antecedent of a deontic conditional (cf. Alchourrón 1996). As a result, it does not allow for *unrestricted* detachment. With $\mathcal{CNR}$, (factual) detachment is only possible when we are under the specific context of the conditional obligation. If more information is added to the context, we need to evaluate whether or not the obligation still holds under that context. From a legal point of view, this is a desirable property: to assess whether or not an obligation holds under some conditions, one must interpret the law, and this is a task which is done by lawyers and judges. But more importantly, it is a task that cannot be done by logic alone (cf. Côté 2006).

Normative conflicts

Goble (2004, p.78) identified *ex falso sequitur quodlibet* as the "real culprit" for deontic explosion. Categorically speaking, *ex falso sequitur quodlibet* is represented by the axiom (0) and amounts to consider $\perp$ or 0 as initial objects. Despite this recognition, Goble (2004, 2009) proposed to keep *ex falso sequitur quodlibet*, arguing that it is required to validate the Smith argument.

In $\mathcal{DDS}$, there is no need to keep *ex falso sequitur quodlibet* given that the Smith argument can be modeled through $\mathcal{OL}$. By constructing $\mathcal{CNR}$ as a symmetric closed deductive system, we reject (0) and therefore avoid deontic explosion, satisfying Goble's criterion of non-triviality.

Another interesting property of $\mathcal{DDS}$ is that it enables us to model that even though a norm is in place, it might happen in some situation that the norm does not hold anymore. Formally, this property is represented by the fact that the following does *not* hold for $\mathcal{CNR}$.

$$(\varphi \otimes \psi) \otimes [(\varphi \otimes \psi) \multimap\!\sim \psi] \longrightarrow 0$$

In a Cartesian deductive system, we would be able to obtain both ψ and $\sim \psi$ from $(\varphi \otimes \psi) \otimes [(\varphi \otimes \psi) \multimap\!\sim \psi]$. However, in $\mathcal{CNR}$, $\sim \psi$ is obtained from $(\varphi \otimes \psi)$, which can only be used once. To obtain $\psi \otimes \sim \psi$ from the given premise, one would need to assume φ twice. From a legal perspective, this allows us to model conflicts of norms that arise from specificity (cf. van der Torre and Tan 1997). Even though some general norm holds, it might happen that in some context it is cancelled and does not hold anymore.

The good Samaritan

The good Samaritan paradox was first presented by Prior (1958), but different formulations were offered by Nozick and Routley (1962), Åqvist (1967) and later Forrester (1984). The paradox goes along the following line of reasoning:

1. the good Samaritan ought to help the traveler who has been robbed;
2. if the good Samaritan helps the traveler who has been robbed, then the traveler has been robbed.

By deontic consequence, this imply that it ought to be that the traveler is robbed, which is a peculiar conclusion to be drawn. In $\mathcal{DDS}$, there are three plausible translations of the good Samaritan paradox. The first one is:

$$O(\alpha \bullet \beta) \qquad \textsf{(np/act)} \tag{16.26}$$

$$\boldsymbol{\alpha} \bullet \boldsymbol{\beta} \xrightarrow[\textsf{PAL}^{\otimes}]{} \boldsymbol{\beta} \tag{16.27}$$

From this translation and $(\Delta_{\otimes})$, we obtain the undesired result that $O\beta$ follows from $O(\alpha \bullet \beta)$. This translation is, however, unlikely. The obligation for the good Samaritan is not to help the traveler *while* he is being robbed, but it is rather to help him afterwards. Hence, $O(\alpha \bullet \beta)$ is not an appropriate translation of the first premise given the properties of action conjunction.

The second translation available to us is:

$$O(\alpha \bullet \beta) \qquad \textsf{(np/act)} \tag{16.28}$$

$$\boldsymbol{\alpha} \otimes \boldsymbol{\beta} \xrightarrow[\textsf{PAL}^{\otimes}]{} \boldsymbol{\beta} \qquad \textsf{ap} \tag{16.29}$$

Even though $O(\alpha \bullet \beta)$ is still not an appropriate translation of the first premise, it is noteworthy that with this translation, $(\Delta_{\otimes})$ cannot be applied, and as such we do not obtain a deduction arrow from $O(\alpha \bullet \beta)$ to $O\beta$.

This second translation could be a solution to a modified version of the good Samaritan paradox, which we call *the Hero paradox* (or, as we prefer, *Batman's paradox*):

1. Batman ought to help Robin, who is under attack by the Joker.
2. If Batman helps Robin, who is under attack by the Joker, then Robin is under attack by the Joker.

∴ The Joker ought to attack Robin.

The main difference between Batman's paradox and the good Samaritan paradox is that in the former the conjunctive action is simultaneous while in the latter it is not. As such, $O(\alpha \bullet \beta)$ could be a translation of the first premise in Batman's paradox. If it is, then the translation of the premises by 16.28 and 16.29 enables us to avoid the unlikely conclusion that the Joker has an obligation to attack Robin. That being said, one could easily argue that Batman's obligation is not to 'help Robin while he is under attack', but that it rather is to 'help Robin' under the circumstances that he is under attack. This is a plausible interpretation seeing that if a conjunctive action is obligatory, then it is expected that both parts of the conjunctive action can be accomplished by the same person. In this case, β should not be a part of Batman's obligations since it cannot be accomplished by Batman. The first premise could thus be translated by $\boldsymbol{\beta} \multimap O\alpha$, in which case, assuming that Robin is under attack, one could then conclude that Batman ought to help him.

The third translation of the good Samaritan paradox follows Åqvist (1967) and is more faithful to the meaning of the first premise in the natural language.

$$O\alpha \otimes \boldsymbol{\beta} \qquad \mathsf{(np/act)\backslash np/ap} \tag{16.30}$$

$$\boldsymbol{\alpha} \otimes \boldsymbol{\beta} \xrightarrow[\mathsf{PAL}^{\otimes}]{} \boldsymbol{\beta} \qquad \mathsf{ap} \tag{16.31}$$

Considering that $(\Delta_{\otimes})$ cannot be applied to 16.31, there is no arrow which can link the premises to $O\beta$.

Some useful deductions in $\mathcal{AL}$

We now list some other useful deductions. First, it is quite easy to show from (cpt1) and (cpt2) that $\alpha \bullet \beta^* \cong \alpha \ominus \beta$.

$$\dfrac{(b)\ \dfrac{(1)\ \overline{\alpha \bullet \beta^* \xrightarrow[\mathsf{AL}]{} \alpha \bullet \beta^*}}{\alpha \bullet \beta^* \xrightarrow[\mathsf{AL}]{} \beta^* \bullet \alpha} \qquad \overline{\beta^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} \alpha \ominus \beta}\ (\text{cpt1})}{\alpha \bullet \beta^* \xrightarrow[\mathsf{AL}]{} \alpha \ominus \beta}\ (\text{cut})$$

$$\dfrac{(\text{cpt2})\ \overline{\alpha \ominus \beta \xrightarrow[\mathsf{AL}]{} \beta^* \bullet \alpha} \qquad \dfrac{\overline{\beta^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} \beta^* \bullet \alpha}\ (1)}{\beta^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} \alpha \bullet \beta^*}\ (b)}{\alpha \ominus \beta \xrightarrow[\mathsf{AL}]{} \alpha \bullet \beta^*}\ (\text{cut})$$

Moreover, it is possible to show that $\alpha \cong \alpha^{**}$. The first part of the proof follows from the definition of the dual object by $\alpha^* =_{def} * \ominus \alpha$.

$$\dfrac{\text{(b)}\;\dfrac{\text{(1)}\;\dfrac{}{\alpha \bullet \alpha^* \xrightarrow[\mathsf{AL}]{} \alpha \bullet \alpha^*}}{\alpha \bullet \alpha^* \xrightarrow[\mathsf{AL}]{} \alpha^* \bullet \alpha} \qquad \dfrac{\dfrac{\dfrac{}{\alpha^* \xrightarrow[\mathsf{AL}]{} \alpha^*}\;\text{(1)}}{\alpha^* \xrightarrow[\mathsf{AL}]{} * \ominus \alpha}}{\alpha^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} *}\;\text{(cl)}}{\dfrac{\dfrac{\alpha \bullet \alpha^* \xrightarrow[\mathsf{AL}]{} *}{\alpha \xrightarrow[\mathsf{AL}]{} * \ominus \alpha^*}\;\text{(cl)}}{\alpha \xrightarrow[\mathsf{AL}]{} \alpha^{**}}}\;\text{(cut)}$$

From this proof, one can also see that $\alpha^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} *$. The other part is a bit more tricky.

$$\text{(cut)}\;\dfrac{\text{(t)}\;\dfrac{\dfrac{\text{(from previous proof)}}{\alpha^* \xrightarrow[\mathsf{AL}]{} \alpha^{***}} \qquad \text{(1)}\;\dfrac{}{\alpha \xrightarrow[\mathsf{AL}]{} \alpha}}{\alpha^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} \alpha^{***} \bullet \alpha} \qquad \dfrac{}{\alpha^{***} \bullet \alpha \xrightarrow[\mathsf{AL}]{} \alpha \ominus \alpha^{**}}\;\text{(cpt2)}}{\alpha^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} \alpha \ominus \alpha^{**}}$$

$$\text{(cut)}\;\dfrac{\text{(b)}\;\dfrac{\text{(1)}\;\dfrac{}{\alpha \bullet \alpha^* \xrightarrow[\mathsf{AL}]{} \alpha \bullet \alpha^*}}{\alpha \bullet \alpha^* \xrightarrow[\mathsf{AL}]{} \alpha^* \bullet \alpha} \qquad \dfrac{\vdots}{\alpha^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} \alpha \ominus \alpha^{**}}}{\alpha \bullet \alpha^* \xrightarrow[\mathsf{AL}]{} \alpha \ominus \alpha^{**}}$$

$$\dfrac{\text{(cut)}\;\dfrac{\text{(cl)}\;\dfrac{\text{(r)}\;\dfrac{\text{(1)}\;\dfrac{}{\alpha \bullet * \xrightarrow[\mathsf{AL}]{} \alpha \bullet *}}{\alpha \bullet * \xrightarrow[\mathsf{AL}]{} \alpha}}{* \xrightarrow[\mathsf{AL}]{} \alpha \ominus \alpha} \qquad \text{(cpt2)}\;\dfrac{}{\alpha \ominus \alpha \xrightarrow[\mathsf{AL}]{} \alpha^* \bullet \alpha}}{* \xrightarrow[\mathsf{AL}]{} \alpha^* \bullet \alpha} \qquad \dfrac{\dfrac{}{\alpha^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} \alpha^* \bullet \alpha}\;\text{(1)}}{\alpha^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} \alpha \bullet \alpha^*}\;\text{(b)}}{* \xrightarrow[\mathsf{AL}]{} \alpha \bullet \alpha^*}\;\text{(cut)}$$

$$\dfrac{\text{(r)}\;\dfrac{\text{(1)}\;\dfrac{}{\alpha^{**} \xrightarrow[\mathsf{AL}]{} \alpha^{**}}}{\alpha^{**} \xrightarrow[\mathsf{AL}]{} \alpha^{**} \bullet *} \qquad \dfrac{\dfrac{\dfrac{\vdots}{* \xrightarrow[\mathsf{AL}]{} \alpha \bullet \alpha^*} \qquad \dfrac{\vdots}{\alpha \bullet \alpha^* \xrightarrow[\mathsf{AL}]{} \alpha \ominus \alpha^{**}}}{* \xrightarrow[\mathsf{AL}]{} \alpha \ominus \alpha^{**}}\;\text{(cut)}}{\alpha^{**} \bullet * \xrightarrow[\mathsf{AL}]{} \alpha}\;\text{(cl)}}{\alpha^{**} \xrightarrow[\mathsf{AL}]{} \alpha}\;\text{(cut)}$$

From this proof, we also get $* \cong \alpha \bullet \alpha^*$. Although it might seem curious at first glance, it happens in $\mathcal{AL}$ that we have $* \cong *^*$. This might seem peculiar since in terms of truth value, this would mean that $\top \cong \neg\top$. This is actually an interesting property of compact deductive systems.

$$
\dfrac{\dfrac{\dfrac{}{*^* \xrightarrow[\mathsf{AL}]{} *^*}\,(1)}{*^* \xrightarrow[\mathsf{AL}]{} *^* \bullet *}\,(\mathrm{r}) \qquad \dfrac{\dfrac{\dfrac{}{*^* \bullet * \xrightarrow[\mathsf{AL}]{} *^* \bullet *}\,(1)}{*^* \bullet * \xrightarrow[\mathsf{AL}]{} * \bullet *^*}\,(\mathrm{b}) \qquad \dfrac{\dfrac{\dfrac{}{*^* \xrightarrow[\mathsf{AL}]{} *^*}\,(1)}{*^* \xrightarrow[\mathsf{AL}]{} * \ominus *}}{* \bullet *^* \xrightarrow[\mathsf{AL}]{} *}\,(\mathrm{cl})}{*^* \bullet * \xrightarrow[\mathsf{AL}]{} *}\,(\mathrm{cut})}{*^* \xrightarrow[\mathsf{AL}]{} *}\,(\mathrm{cut})
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{}{* \bullet * \xrightarrow[\mathsf{AL}]{} * \bullet *}\,(1)}{* \bullet * \xrightarrow[\mathsf{AL}]{} *}\,(\mathrm{r})}{* \xrightarrow[\mathsf{AL}]{} * \ominus *}\,(\mathrm{cl})}{* \xrightarrow[\mathsf{AL}]{} *^*}
$$

Here, one can easily see why formulas of $\mathcal{AL}$ should not be interpreted in terms of declarative statements. The formula $\alpha \bullet \alpha^*$ is neither a contradiction nor a tautology, and $*$ is neither *true* nor *false*.

Finally, one can prove that $\alpha \cong \beta$ if and only if $\alpha^* \cong \beta^*$. This will help when we reason with interdictions. The proof is quite straight forward. From left to right, assume $\alpha \cong \beta$, hence $\alpha \cong \beta^{**}$, and thus:

$$
\dfrac{\dfrac{\dfrac{\dfrac{\alpha \xrightarrow[\mathsf{AL}]{} \beta^{**}}{\alpha \xrightarrow[\mathsf{AL}]{} * \ominus \beta^*}}{\beta^* \bullet \alpha \xrightarrow[\mathsf{AL}]{} *}}{\beta^* \xrightarrow[\mathsf{AL}]{} * \ominus \alpha}}{\beta^* \xrightarrow[\mathsf{AL}]{} \alpha^*}
$$

By the same reasoning, we obtain $\alpha^* \xrightarrow[\mathsf{AL}]{} \beta^*$ from $\alpha^{**} \cong \beta$. From right to left, assume $\alpha^* \cong \beta^*$, hence we have (omitting symmetry):

$$
\dfrac{\dfrac{\dfrac{\dfrac{\alpha^* \xrightarrow[\mathsf{AL}]{} \beta^*}{\alpha^* \xrightarrow[\mathsf{AL}]{} * \ominus \beta}}{\beta \bullet \alpha^* \xrightarrow[\mathsf{AL}]{} *}}{\beta \xrightarrow[\mathsf{AL}]{} * \ominus \alpha^*}}{\beta \xrightarrow[\mathsf{AL}]{} \alpha^{**}}
$$

And since we know that $\alpha \cong \alpha^{**}$, it follows from (cut) that $\beta \xrightarrow[\mathsf{AL}]{} \alpha$. We obtain $\alpha \xrightarrow[\mathsf{AL}]{} \beta$ by the same reasoning.

Examples of analysis

We now show how $\mathcal{DDS}$ can help to model legal reasoning with the analysis of three examples.

Example 16.1. Considering that it is forbidden to remove an elm tree without treating the elm stump in a manner satisfactory to an inspector, it follows that it is forbidden to remove an elm tree while not treating the elm stump in a manner satisfactory to an inspector.[35]

This example is presented in an unconditional form. As such, it can be modeled through $\mathcal{OL}$ and the premise can be translated by 16.32. The inference is thus obtained by 16.33.

$$F(\alpha \ominus \beta) \qquad \mathsf{np/act} \tag{16.32}$$
$$F(\alpha \ominus \beta) \xrightarrow[\mathsf{OL}]{} F(\alpha \bullet \beta^*) \tag{16.33}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\text{(from the proofs above)}}{\alpha \ominus \beta \cong \alpha \bullet \beta^*}}{(\alpha \ominus \beta)^* \cong (\alpha \bullet \beta^*)^*}}{O(\alpha \ominus \beta)^* \xrightarrow[\mathsf{OL}]{} O(\alpha \bullet \beta^*)^*}\ (\Delta)}{F(\alpha \ominus \beta) \xrightarrow[\mathsf{OL}]{} F(\alpha \bullet \beta^*)}\ (\text{def } F)$$

Example 16.2. It is forbidden to make tributyltetradecylphosphonium chloride ($F\alpha$), unless the following conditions are satisfied[36]:

1. the manufacture is for export only ($\boldsymbol{\beta}$);
2. the manufacturer has notified the Minister ($\boldsymbol{\gamma}$);
3. the manufacturer uses a fully contained process ($\boldsymbol{\eta}$).

This regulation could be translated either by 16.34 or 16.35.

$$F(\alpha \ominus (\beta \otimes (\gamma \otimes \eta))) \tag{16.34}$$
$$F\alpha \otimes ((F\alpha \otimes (\boldsymbol{\beta} \otimes (\boldsymbol{\gamma} \otimes \boldsymbol{\eta}))) \multimap P_s\alpha) \tag{16.35}$$

However, equation 16.34 does not imply that it is permitted to make tributyltetradecylphosphonium chloride when the conditions are satisfied. Hence, 16.35 would be more accurate to the signification of the regulation within the natural language. It is a case of specificity, where $F\alpha$ is more general and is overruled in some context.

[35] Dutch Elm Disease Regulation, Manitoba Regulation 213/98.
[36] Tributyltetradecylphosphonium Chloride Regulations, SOR/2000-66.

The regulation specifies that the permission holds under the circumstances that $F\alpha$ and $\boldsymbol{\beta} \otimes (\boldsymbol{\gamma} \otimes \boldsymbol{\eta})$ hold, hence the conditional permission can be formulated by:

$$(F\alpha \otimes (\boldsymbol{\beta} \otimes (\boldsymbol{\gamma} \otimes \boldsymbol{\eta}))) \multimap P_s\alpha \tag{16.36}$$

Assuming a context where it is forbidden to make tributyltetradecylphosphonium chloride, but that it is permitted under the aforementioned circumstances, and that these circumstances are met, one can conclude that it is permitted to make tributyltetradecylphosphonium chloride by (1) and (cl). In such a context, the interdiction would not be in force anymore.

$$(F\alpha \otimes (\boldsymbol{\beta} \otimes (\boldsymbol{\gamma} \otimes \boldsymbol{\eta}))) \otimes [(F\alpha \otimes (\boldsymbol{\beta} \otimes (\boldsymbol{\gamma} \otimes \boldsymbol{\eta}))) \multimap P_s\alpha] \xrightarrow[\text{CNR}]{} P_s\alpha \tag{16.37}$$

Example 16.3. The sign P-110-5 indicates that it is forbidden to make a U-turn, unless the vehicle is authorized.[37]

Assume that Paul, who did not drive an authorized vehicle, made a U-turn. In this case, one might want to be able to conclude from the previous regulation that it was forbidden for Paul to do so, and as such that he is liable. The first obvious translation for the regulation is 16.38, which says that it is usually forbidden to make a U-turn but it is permitted if the vehicle is authorized.

$$F\alpha \otimes (\boldsymbol{\beta} \multimap P_s\alpha) \tag{16.38}$$

This, however, would not allow us to conclude that it was forbidden for Paul to make a U-turn. In addition to this regulation, we would have in the context that $F\alpha$ but $\sim \boldsymbol{\beta}$, and since (Cart) is not satisfied, we would not be able to derive $F\alpha$. Another possible translation would be 16.39, which states that the interdiction is conditional to a context where there is a sign P-110-5 (c is of type d), and the U-turn is permitted in a context where there is a sign but the vehicle is authorized.

$$(c \multimap F\alpha) \otimes ((c \ \otimes \boldsymbol{\beta}) \multimap P_s\alpha) \tag{16.39}$$

But again, this would not allow us to conclude that Paul should not have made a U-turn (i.e., that it was forbidden for him to do so), since the context would be $c \otimes \sim \boldsymbol{\beta}$. In order to be able to conclude that it was forbidden for Paul to make a U-turn, we need to assume that this interdiction holds under the circumstances that:

1. there is a sign;
2. the vehicle is unauthorized;
3. there is a conditional permission of making a U-turn when the vehicle is authorized.

[37] Quebec Regulations on Road Signs, RRQ, c C-24.2, r 41.

$$((c \otimes \sim \boldsymbol{\beta}) \otimes (\boldsymbol{\beta} \multimap P_s\alpha)) \multimap F\alpha) \otimes (\boldsymbol{\beta} \multimap P_s\alpha) \tag{16.40}$$

Assuming this together with $c \otimes \sim \boldsymbol{\beta}$, we would then be able to conclude $F\alpha$ from (1) and (cl).

Now assume that a police officer, who saw Paul make a U-turn, also made a U-turn to pursue Paul. In this case, we want to be able to conclude that the police officer was allowed to make the U-turn. The solution is, as in the case of the interdiction, to specify adequately the conditions under which the permission holds. The conditional permission is meant to hold when:

1. there is a sign;
2. the vehicle is authorized;
3. there is a conditional interdiction of making a U-turn when the vehicle is unauthorized.[38]

$$(((c \otimes \boldsymbol{\beta}) \otimes (\sim \boldsymbol{\beta} \multimap F\alpha)) \multimap P_s\alpha) \otimes (\sim \boldsymbol{\beta} \multimap F\alpha) \tag{16.41}$$

Together with $c \otimes \boldsymbol{\beta}$, this yields $P_s\alpha$. This third example shows that interpreting the law is a crucial part of legal reasoning. To do a proper legal inference, one must first determine the context and the normative propositions that hold under that context. As such, there is a human component to legal reasoning, which consists of assessing the appropriate premises of the argument.

Conclusion

Summing up, we introduced a deontic deductive system $\mathcal{DDS}$ built upon an action logic $\mathcal{AL}$, a propositional action logic $\mathcal{PAL}$, an obligation logic $\mathcal{OL}$ and a logic for conditional normative reasoning $\mathcal{CNR}$. An interesting fact about $\mathcal{DDS}$ is that all its components have a different structure. From a philosophical point of view, this is not only a desirable property, but it is foremost a necessary one. To model normative reasoning, one must be able to reason with facts, with actions, with facts about actions, with norms and with conditional norms. The complexity of our natural language relies namely upon the fact that some of its parts do not share the same structure. Reasoning with conditional obligations does not require the same characteristics as reasoning with unconditional obligations. Similarly, reasoning about what actions are is not the same thing as reasoning about which actions are true. By incorporating different structures, $\mathcal{DDS}$ allows for the representation of the complexity of our natural normative language.

[38] We could also use a general interdiction instead of a conditional one, or an interdiction conditional to $c \otimes \sim \boldsymbol{\beta}$.

Another upshot of $\mathcal{DDS}$ is that it possesses a familiar structure. Indeed, fibrations are common structures in mathematics and in type theory. From an epistemological point of view, defining $\mathcal{DDS}$ as a fibration hints at the possibility that the structure of our natural language is more familiar than we think to some mathematical structures.

The main motivation of this article was to introduce a deontic logic that can be used to model Canadian legal discourse. As such, it was constructed upon some of its fundamental characteristics, including the presumption of normative consistency, which insures its rationality. We showed how $\mathcal{DDS}$ can be used to solve some paradoxes of deontic logic and provided some examples of analysis of legal reasoning. Building on previous work, we used category theory as a proof-theoretical foundational framework for logic, and we showed how some problems for deontic logic can be solved by using different types of deductive systems. The present paper was, however, mainly a syntactical contribution to the proof-theory of deontic logic, and as such the task of building a proper semantics for this system remains. For future research, we intend to develop a functorial semantics for $\mathcal{DDS}$ and its components.

All things considered, we hope to have convinced the reader that a categorical analysis of logic is relevant to philosophical issues, hence the categorical imperative.

Chapitre 17

Discussion

Contribution théorique

Tout compte fait, il s'avère que les objectifs fixés au chapitre 10 ont été atteints. Le premier but était de développer un cadre conceptuel alternatif à la logique modale K en vue de l'analyse de la validité des inférences légales inconditionnelles. Cela a été fait au chapitre 11, où une logique de l'obligation a été développée et adaptée aux fins de l'analyse des inférences normatives au chapitre 12. Cela dit, même si OL offre une alternative intéressante au système modal K, il n'en demeure pas moins que ce dernier possède ses limites. Notamment, OL n'est pas basé sur une logique de l'action adéquate à la représentation des actions humaines et échoue à rendre compte de manière acceptable de la distinction entre *action* et *proposition*. Cette limite se comprend dans la mesure où l'objectif était de fournir un cadre de travail différent mais comparable à la logique modale K pour l'analyse des inférences normatives. Néanmoins, OL a été reconstruit au chapitre 16 afin de répondre à ces lacunes.

Le second objectif de la thèse était le développement d'une logique de l'action non-booléenne adéquate à la représentation des actions humaines. Cela a été atteint au chapitre 14, où une logique d'action et une logique propositionnelle d'action ont été proposées.

Finalement, cette thèse visait la modélisation des raisonnements normatifs conditionnels et des conflits d'obligations dans un cadre de travail monadique sans opérateurs supplémentaires. Alors que le chapitre 15 montre comment il est possible de modéliser les inférences normatives conditionnelles dans un cadre de travail monadique, le chapitre 16 utilise la logique typée présentée au chapitre 13 et les résultats obtenus aux chapitre 14 et 15 afin de développer un système déductif déontique capable de modéliser les inférences légales, autant inconditionnelles que conditionnelles.

Le tableau 17.1 indique comment les systèmes proposés permettent de répondre aux objections avancées au chapitre 9. Comme on peut le voir, même si OL succombe aux arguments 2 et 18, sa reconstruction à l'aide d'une logique d'action appropriée permet à $\mathcal{OL}$ d'y répondre. Puisque OL et $\mathcal{OL}$ visent la modélisation des obligations inconditionnelles, les arguments 13-17 ne s'y appliquent pas.

On observe aussi que $\mathcal{CNR}$, qui vise la modélisation des inférences normatives conditionnelles et des conflits d'obligations, répond à la majorité des objections avancées au chapitre 9. Les deux seuls arguments auxquels $\mathcal{CNR}$ ne parvient pas à répondre sont les arguments 3 et 4, qui spécifient que la sémantique d'une logique déontique doit rendre compte de la dichotomie sémantique entre les énoncés descriptifs et normatifs tout en modélisant le fait que la valeur de vérité des propositions normatives dépend de certaines normes, établies pas des autorités.

Ce défaut se comprend notamment à la lumière de notre contribution quant à la théorie de la preuve de la logique déontique. Notre analyse des problèmes dus à la modélisation des inférences normatives conditionnelles et des conflits d'obligations au chapitre 15 a mis en évidence le lien entre ces problèmes et certaines propriétés structurelles des logiques utilisées. Ainsi, outre le système présenté au chapitre 11, la contribution théorique de cette thèse est principalement syntaxique. En assumant la théorie des catégories comme cadre de travail, nous avons développé une théorie de la preuve adéquate à la modélisation des inférences légales, laissant de côté les considérations sémantiques du langage. De fait, à strictement parler, $\mathcal{DDS}$ ne permet pas de répondre aux arguments 3 et 4, qui s'adressent à la sémantique du langage formel. Comme nous le verrons à la section suivante, cela ne constitue cependant pas une impasse pour notre approche.

Arguments	OL	$\mathcal{OL}$	$\mathcal{CNR}$
1	$\checkmark$	$\checkmark$	$\checkmark$
2	$\times$	$\checkmark$	$\checkmark$
3	$\checkmark$	$\times$	$\times$
4	$\checkmark$	$\times$	$\times$
5	$\checkmark$	$\checkmark$	$\checkmark$
6	$\checkmark$	$\checkmark$	$\checkmark$
7	$\checkmark$	$\checkmark$	$\checkmark$
8	$\checkmark$	$\checkmark$	$\checkmark$
9	$\checkmark$	$\checkmark$	$\checkmark$
10	$\checkmark$	$\checkmark$	$\checkmark$
11	$\checkmark$	$\checkmark$	$\checkmark$
12	$\checkmark$	$\checkmark$	$\checkmark$
13	$-$	$-$	$\checkmark$
14	$-$	$-$	$\checkmark$
15	$-$	$-$	$\checkmark$
16	$-$	$-$	$\checkmark$
17	$-$	$-$	$\checkmark$
18	$\times$	$\checkmark$	$\checkmark$
19	$\checkmark$	$\checkmark$	$\checkmark$

Tableau 17.1: Réponse aux arguments

Avenues de recherche future

La logique déontique est un domaine ayant des répercussions dans plusieurs autres disciplines. À la lumière de la contribution théorique de cette thèse, plusieurs avenues de recherche s'offrent à nous. D'abord, considérant que cette thèse offrait principalement une contribution syntaxique quant à l'analyse des inférences légales, un de nos objectifs est de développer une sémantique et de montrer la complétude de $\mathcal{DDS}$.

D'un côté, il s'agira de proposer une sémantique algébrique dans le cadre des treillis résidués (cf. Galatos et al. 2007) et des monoïdes commutatifs résidués involutifs partiellement ordonnés. Brièvement, un monoïde commutatif partiellement ordonné est un ensemble partiellement ordonné $\boldsymbol{M} = \langle M, \leq \rangle$ possédant une opération $\cdot$ associative et commutative ainsi qu'une unité 1. Un monoïde est résidué lorsque celui-ci satisfait ce que l'on nomme la loi de résiduation (R), et il est involutif lorsqu'il satisfait $(p\backslash 0)\backslash 0 = p$.

$$p \cdot q \leq r \text{ si et seulement si } p \leq q\backslash r \tag{R}$$

Or, on peut aisément montrer qu'une telle structure est en fait une instance d'une catégorie $*$-autonome, et par conséquent cela fournit un modèle algébrique pour $\mathcal{CNR}$. Nous avons déjà montré dans Peterson (2014b) que le système déductif construit à partir de la fibration de $\mathcal{AL}$ avec $\mathcal{CNR}$ peut s'interpréter selon cette structure, qui fournit une sémantique adéquate et complète. Ainsi, cela fait un premier pas en vue d'une sémantique pour $\mathcal{DDS}$ et une avenue de recherche consiste à appliquer ces résultats afin d'obtenir une sémantique adéquate et complète pour chaque fragment du système. La sémantique pour $\mathcal{AL}$ requerra probablement l'introduction de la notion de monoïde *compact*, instance d'une catégorie compacte.

Par ailleurs, Baez et Stay (2011) ont montré que les diagrammes à cordes offrent un langage approprié pour les catégories monoïdales. Ainsi, il est possible de construire une sémantique pour $\mathcal{DDS}$ dans à l'aide de tels diagrammes. Nous avons déjà montré dans Peterson (2014c) que les diagrammes à cordes offrent une sémantique adéquate et complète pour $\mathcal{CNR}$ en plus de fournir une procédure de décidabilité. L'intérêt d'une telle approche est que cela permet la représentation visuelle des inférences, et par le fait même cela permet de définir une méthode pour tester la validité des inférences normatives conditionnelles. De fait, une fois que cette méthode sera appliquée à $\mathcal{DDS}$ dans son ensemble, les diagrammes à cordes permettront d'utiliser le système déductif déontique dans une perspective de pensée critique afin d'analyser les inférences normatives.

En plus des pistes qui touchent à la sémantique de $\mathcal{DDS}$, plusieurs applications restent à être explorées. Outre les applications de $\mathcal{DDS}$ quant à l'analyse des arguments légaux, une avenue de recheche intéressante sera de voir dans quelle mesure ce dernier peut être utilisé en intelligence artificielle. Pavlovic (2013) a récemment développé la notion d'*ordinateur monoïdal* sur la base des diagrammes à cordes. En un mot, un ordinateur monoïdal est défini comme une catégorie monoïdale symétrique avec une structure additionnelle. De fait, une avenue de recherche sera de voir dans quelle mesure $\mathcal{CNR}$ peut s'intégrer aux ordinateurs monoïdaux en vue de programmer des agents intelligents dont le comportement est guidé par certaines contraintes.

Conclusion

Somme toute, cette thèse offre un nouveau regard sur la logique déontique. En analysant la théorie de la preuve des systèmes standards selon la théorie des catégories, il est possible de mettre en évidence que certains problèmes importants auxquels la logique déontique fait face sont reliés aux propriétés structurelles des logiques utilisées. Ainsi, l'analyse de la logique déontique sur la base de la théorie des catégories permet d'identifier précisément les aspects problématiques dus à la modélisation des inférences. Par le fait même, cela offre des pistes de solutions en identifiant les structures où ces difficultés peuvent faire surface. Toutes choses considérées, nous espérons sincèrement avoir convaincu le lecteur de la pertinence philosophique d'une analyse de la logique déontique à partir de la théorie des catégories et érigeons comme maxime universelle que tout philosophe devrait aborder la logique selon cette perspective.

Bibliographie

V. Michele Abrusci. A comparison between Lambek syntactic calculus and intuitionistic linear propositional logic. *Mathematical Logic Quarterly*, 36(1):11–15, 1990a.

V. Michele Abrusci. Non-commutative intuitionistic linear logic. *Mathematical Logic Quarterly*, 36(4):297–318, 1990b.

V. Michele Abrusci. Phase semantics and sequent calculus for pure non-commutative classical linear propositional logic. *The Journal of Symbolic Logic*, 56(4):1403–1451, 1991.

V. Michele Abrusci et Paul Ruet. Non-commutative logic I: The multiplicative fragment. *Annals of Pure and Applied Logic*, 101:29–64, 2000.

Kazimierz Ajdukiewicz. Die syntaktische Konnezität. *Studia Philosophica*, 1:1–27, 1935.

Azizah Al-Hibri. *Deontic logic: A comprehensive appraisal and a new proposal*. University Press of America, 1978.

Carlos Alchourrón. Logic without truth. *Ratio Juris*, 3(1):46–67, 1990.

Carlos Alchourrón. Detachment and defeasibility in deontic logic. *Studia Logica*, 57(1): 5–18, 1996.

Carlos Alchourrón et Eugenio Bulygin. *Normative systems*. Springer, 1971.

Carlos Alchourrón et Eugenio Bulygin. The expressive conception of norms. Dans R. Hilpinen, éditeur, *New Studies in Deontic Logic*, pages 95–124. D. Reidel Publishing Company, 1981.

Carlos Alchourrón et David Makinson. Hierarchies of regulations and their logic. Dans R. Hilpinen, éditeur, *New Studies in Deontic Logic*, pages 125–148. D. Reidel Publishing Company, 1981.

Alan Ross Anderson. A reduction of deontic logic to alethic modal logic. *Mind*, 67(265): 100–103, 1958a.

Alan Ross Anderson. The logic of norms. *Logique et Analyse*, 1(2):84–91, 1958b.

Alan Ross Anderson. On the logic of commitment. *Philosophical Studies*, 10(2):23–27, 1959.

Alan Ross Anderson. Reply to Mr Rescher's conditional permission in deontic logic. *Philosophical Studies*, 13(1):6–8, 1962.

Alan Ross Anderson. Some nasty problems in the formal logic of ethics. *Noûs*, 1(4): 345–360, 1967.

Albert J. J. Anglberger. Dynamic deontic logic and its paradoxes. *Studia Logica*, 89(3): 427–435, 2008.

Albert J. J. Anglberger. Alternative reductions for dynamic deontic logics. Dans A. Hieke et H. Leitgeb, éditeurs, *Reduction-Abstraction-Analysis*, volume 11, pages 179–191. Ontos Verlag, 2009.

G. Aldo Antonelli. Non-monotonic logic. Dans E. N. Zalta, éditeur, *The Stanford Encyclopedia of Philosophy*. 2012.

Lennart Åqvist. Good Samaritans, contrary-to-duty imperatives, and epistemic obligations. *Noûs*, 1(4):361–379, 1967.

Lennart Åqvist. Deontic logic. Dans D. M. Gabbay et F. Guenthner, éditeurs, *Handbook of Philosophical Logic*, volume 8, pages 147–264. Kluwer Academic Publishers, 2e édition, 2002.

Lennart Åqvist et Jaap Hoepelman. Some theorems about a "tree" system of deontic tense logic. Dans R. Hilpinen, éditeur, *New Studies in Deontic Logic*, pages 187–221. D. Reidel Publishing Company, 1981.

Nicholas Asher et Daniel Bonevac. Common sense obligation. Dans D. Nute, éditeur, *Defeasible Deontic Logic*, pages 159–203. Kluwer Academic Publishers, 1997.

Steve Awodey. *Category theory*. Oxford University Press, 2e édition, 2006.

Steve Awodey et Kohei Kishida. Topology and modality: The topological interpretation of first-order modal logic. 2007.

Alfred Jules Ayer. *Language, truth and logic*. Victor Gollancz LTD, 1946.

Emmon Bach. Introduction. Dans R. T. Oehrle, E. Bach et D. Wheeler, éditeurs, *Categorial Grammars and Natural Language Structures*, pages 1–16. D. Reidel Publishing Company, 1988.

John C. Baez et Mike Stay. Physics, topology, logic and computation: A Rosetta stone. Dans B. Coecke, éditeur, *New Structures for Physics*, volume 813 de *Lecture Notes in Physics*, pages 95–174. Springer, 2011.

Patrice Bailhache. *Essaie de logique déontique*. Vrin, 1991.

Patrice Bailhache. The deontic branching time: Two related conceptions. *Logique et Analyse*, 36(141-142):159–175, 1993.

Philippe Balbiani. Propositional dynamic logic. Dans E. N. Zalta, éditeur, *The Stanford Encyclopedia of Philosophy*. 2008.

Yehoshua Bar-Hillel. A quasi-arithmetical notation for syntactic description. *Language*, 29(1):47–58, 1953.

Micheal Barr. *$*$-autonomous categories*, volume 752 de *Lecture Notes in Mathematics*. Springer, 1979.

Micheal Barr. $*$-autonomous categories and linear logic. *Mathematical Structures in Computer Science*, 1(2):159–178, 1991.

Jean-Louis Baudoin. *Droit et vérité*. Éditions Thémis, 2010.

Mathieu Beirlaen et Christian Straßer. Two adaptive logics of norm-propositions. *Journal of Applied Logic*, 11(2):147–168, 2013.

Mathieu Beirlaen, Christian Straßer et Joke Meheus. An inconsistency-adaptive deontic logic for normative conflicts. *Journal of Philosophical Logic*, 42(2):285–315, 2013.

Nuel Belnap et Michael Perloff. Seeing to it that: A canonical form for agentives. *Theoria*, 54(3):175–199, 1988.

Katalin Bimbó et J. Michael Dunn. Relational semantics for Kleene logic and action logic. *Notre Dame Journal of Formal Logic*, 46(4):461–490, 2005.

Richard Blute et Philip Scott. Category theory for linear logicians. Dans T. Ehrhard, J.-Y. Girard, P. Ruet et P. Scott, éditeurs, *Linear Logic in Computer Science*, volume 316, pages 3–64. Cambridge University Press, 2004.

Guido Boella, Guido Governatori, Antonino Rotolo et Leendert van der Torre. A formal study on legal compliance and interpretation. Dans *13 International Workshop on Non-Monotonic Reasoning (NMR 2010)*. CEUR Workshop Proceedings, 2010.

Guido Boella et Leendert van der Torre. Local policies for the control of virtual communities. Dans *In proceedings of IEEE/WIC web intelligence conference*, pages 161–167. IEEE Press, 2003a.

Guido Boella et Leendert van der Torre. Permissions and obligations in hierarchical normative systems. Dans *Proceedings of the 9th International Conference on Artificial Intelligence and Law*, ICAIL '03, pages 109–118. ACM, 2003b.

Guido Boella et Leendert van der Torre. Contracts as legal institutions in organizations of autonomous agents. Dans *3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*, pages 948–955. IEEE Computer Society, 2004.

Guido Boella et Leendert van der Torre. An architecture of a normative system: Counts-as conditionals, obligations and permissions. Dans H. Nakashima, M. P. Wellman, G. Weiss et P. Stone, éditeurs, *5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006)*, pages 229–231. ACM, 2006a.

Guido Boella et Leendert van der Torre. Delegation of power in normative multiagent systems. Dans L. Goble et J.-J. Ch. Meyer, éditeurs, *DEON 2006*, volume 4048 de *Lecture Notes in Computer Science*, pages 36–52. Springer, 2006b.

Guido Boella et Leendert van der Torre. A logical architecture of a normative system. Dans L. Goble et J.-J. Ch. Meyer, éditeurs, *DEON 2006*, volume 4048 de *Lecture Notes in Computer Science*, pages 24–35. Springer, 2006c.

Guido Boella, Leendert van der Torre et Harko Verhagen. Introduction to normative multiagent systems. *Computational & Mathematical Organization Theory*, 12(2-3):71–79, 2006.

Daniel Bonevac. Against conditional obligation. *Noûs*, 32(1):37–53, 1998.

Diana Brignole. Equational characterization of Nelson algebra. *Notre Dame Journal of Formal Logic*, 10(3):225–336, 1969.

Selmer Bringsjord et Joshua Taylor. The divine-command approach to robot ethics. Dans P. Lin, G. Abney et K. Abney, éditeurs, *Robot Ethics: The Ethical and Social Implications of Robotics*, pages 85–108. MIT Press, 2012.

Selmer Bringsjord, Joshua Taylor, Trevor Housten, Bram van Heuveln, Konstantine Arkoudas et Micah Clark. Robotethics via category theory: Moving beyond mere formal operations to engineer robots whose decisions are garanteed to be ethically correct. Dans M. Anderson et S. Anderson, éditeurs, *Machine Ethics*, pages 361–374. Cambridge University Press, 2011.

Jan Broersen. *Modal action logics for reasoning about reactive systems.* Thèse de doctorat, Vrije Universiteit Amsterdam, 2003.

Jan Broersen. Action negation and alternative reductions for dynamic deontic logics. *Journal of Applied Logic*, 2(1):153–168, 2004.

Jan Broersen. A logical analysis of the interaction between 'obligation-to-do' and 'knowingly doing'. Dans R. van der Meyden et L. van der Torre, éditeurs, *DEON 2008*, volume 5076 de *Lecture Notes in Computer Science*, pages 140–154. Springer, 2008.

Jan Broersen. A complete stit logic for knowledge and action, and some of its applications. Dans M. Baldoni, T. C. Son, M. van Riemsdijk et M. Winikoff, éditeurs, *Declarative Agent Languages and Technologies VI*, volume 5397 de *Lecture Notes in Computer Science*, pages 47–59. 2009.

Jan Broersen. Deontic epistemic stit logic distinguishing modes of mens rea. *Journal of Applied Logic*, 9(2):137–152, 2011a.

Jan Broersen. Making a start with the stit logic analysis of intentional action. *Journal of Philosophical Logic*, 40(4):499–530, 2011b.

Jan Broersen, Frank Dignum, Virginia Dignum et John-Jules Ch. Meyer. Desinging a deontic logic of deadlines. Dans A. Lomuscio et D. Nute, éditeurs, *DEON 2004*, volume 3065 de *Lecture Notes in Computer Science*, pages 43–56. Springer, 2004.

Mark Brown. Obligation, contracts and negociation: Outlining an approach. *Journal of Applied Logic*, 3(3-4):371–395, 2005.

José Carmo et Andrew J. I. Jones. Deontic logic and contrary-to-duties. Dans D. M. Gabbay et F. Guenthner, éditeurs, *Handbook of Philosophical Logic*, volume 8, pages 265–343. Kluwer Academic Publishers, 2e édition, 2002.

José Carmo et Andrew J. I. Jones. Completeness and decidability results for a logic of contrary-to-duty conditionals. *Journal of Logic and Computation*, 23(3):585–626, 2013.

José Carmo et Olga Pacheco. Deontic and action logics for collective agency and roles. Dans R. Demolombe et R. Hilpinen, éditeurs, *Proceedings of the Fifth International Workshop on Deontic Logic in Computer Science (DEON'00)*, pages 93–124. 2000.

José Carmo et Olga Pacheco. Deontic and action logics for organized collective agency, modeled through institutionalized agents and roles. *Fundamenta Informaticae*, 48(2-3): 129–163, 2001.

José Carmo, Robert Demolombe et Andrew J. I. Jones. An application of deontic logic to information system constraints. *Fundamenta Informaticae*, 48(2-3):165–181, 2001.

José Carmo et Andrew J. I. Jones. Deontic database constraints, violation and recovery. *Studia Logica*, 57(1):139–165, 1996.

Claudia Casadio. Non-commutative linear logic in linguistics. *Grammars*, 4(3):167–185, 2001.

Claudia Casadio et Joachim Lambek. A tale of four grammars. *Studia Logica*, 71(3): 315–329, 2002.

Claudia Casadio, Philip J. Scott et Robert A. G. Seely. Introduction: The Lambek program. Dans C. Casadio, P. J. Scott et R. A. G. Seely, éditeurs, *Language and Grammar*, Studies in Mathematical Linguistics and Natural Language, pages xi–xxxi. Center for the Study of Language and Information, 2004.

Hector-Neri Castañeda. The logic of obligation. *Philosophical Studies*, 10(2):17–23, 1959.

Hector-Neri Castañeda. Acts, the logic of obligation, and deontic calculi. *Philosophical Studies*, 19(1):13–26, 1968.

Hector-Neri Castañeda. On the semantics of the ought-to-do. *Synthese*, 21(3):449–468, 1970.

Hector-Neri Castañeda. Ought, time and the deontic paradoxes. *The Journal of Philosophy*, 74(12):775–791, 1977.

Hector-Neri Castañeda. The paradoxes of deontic logic. Dans R. Hilpinen, éditeur, *New Studies in Deontic Logic*, pages 37–85. D. Reidel Publishing Company, 1981.

Hector-Neri Castañeda. Aspectual actions and the deepest "paradox" of deontic logic. Dans A. A. Martino et F. S. Natali, éditeurs, *Automated Analysis of Legal Texts*, pages 31–50. Elsevier, 1986.

Pablo F. Castro et Tom S. E. Maibaum. Deontic action logic, atomic Boolean algebras and fault-tolerance. *Journal of Applied Logic*, 7(4):441–466, 2009.

Pablo F. Castro et Tom S. E. Maibaum. Towards a first-order deontic action logic. Dans T. Mossakowski et H.-J. Kreowski, éditeurs, *Recent Trends in Algebraic Development Techniques*, volume 7137 de *Lecture Notes in Computer Science*, pages 61–75. Springer, 2012.

Brian F. Chellas. *The logical form of imperatives.* Perry Lane Press, 1969.

Brian F. Chellas. Conditional obligation. Dans S. Stenlund, éditeur, *Logical Theory and Semantic Analysis*, pages 23–33. D. Reidel Publishing Company, 1974.

Brian F. Chellas. *Modal logic: An introduction.* Cambridge University Press, 1980.

Roderick Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, 24(2): 33–36, 1963.

Azizah Cox. Hintikka and the interdefinability of obligation and forbiddance. *The Southwestern Journal of Philosophy*, 9(1):7–10, 1978.

Laszlo Csirmaz. Multi-level permission. Rapport technique 90-25, Rutgers University, 1990.

Pierre-André Côté. *Interprétation des lois.* Éditions Thémis, 3e édition, 2006.

Donald Davidson. The logical form of action sentences. Dans *The Essential Davidson (2006)*. Oxford University Press, 1967.

Judith Decew. Conditional obligation and counterfactuals. *Journal of Philosophical Logic*, 10(1):55–72, 1981.

Pilar Dellunde. On the multimodal logic of normative systems. Dans J. S. Sichman, J. A. Padget, S. Ossowski et P. Noriega, éditeurs, *Coordination, Organizations, Institutions, and Norms in Agent Systems III*, pages 261–274. Springer, 2008.

Robert Demolombe. Obligations with deadlines: A formalization in dynamic deontic logic. *Journal of Logic and Computation*, 24(1):1–17, 2014.

Frank Dignum, Jan Broersen, Virginia Dignum et John-Jules Ch. Meyer. Meeting the deadline: Why, when and how. Dans M. G. Hinchey, J. L. Rash, W. F. Truszkowski et C. A. Rouff, éditeurs, *Formal Approaches to Agent-Based Systems*, volume 3228 de *Lecture Notes in Computer Science*, pages 30–40. Springer, 2005.

Frank Dignum et Ruurd Kuiper. Combining dynamic deontic logic and temporal logic for the specification of deadlines. Dans *System Sciences, 1997, Proceedings of the Thirtieth Hawaii International Conference*, volume 5, pages 336–346. IEEE, 1997.

Frank Dignum et Ruurd Kuiper. Obligations and dense time for specifying deadlines. Dans *System Sciences, 1998, Proceedings of the Thirty-First Hawaii International Conference*, volume 5, pages 186–195. IEEE, 1998.

Frank Dignum, John-Jules Ch. Meyer, Roel Wieringa et Ruurd Kuiper. A modal approach to intentions, commitments and obligations: Intention plus commitment yields obligation. Dans M. Brown et J. Carmo, éditeurs, *Deontic Logic, Agency and Normative Systems*, pages 80–97. Springer, 1996.

Ross Duncan. *Types for Quantum Computing*. Thèse de doctorat, Oxford University, 2006.

Samuel Eilenberg et Saunders Mac Lane. General theory of natural equivalences. *Transactions of the American Mathematical Society*, 58:231–294, 1945.

Dagfinn Føllesdal et Risto Hilpinen. Deontic logic: An introduction. Dans R. Hilpinen, éditeur, *Deontic Logic: Introductory and Systematic Readings*, pages 1–35. D. Reidel Publishing Company, 1970.

James Forrester. Gentle murder, or the adverbial Samaritan. *The Journal of Philosophy*, 81(4):193–197, 1984.

D. M. Gabbay, R. H. Johnson, H. J. Ohlbach et J. Woods, éditeurs. *Handbook of the Logic of Argument and Inference*. Elsevier, 2002.

Dov M. Gabbay. *What is a Logical System?* Clarendon Press, 1994.

Dov M. Gabbay et Karl Schlechta. Critical analysis of the Carmo-Jones system of contrary-to-duty obligations. *ArXiv e-prints*, 1002.3021, 2010.

N. Galatos, P. Jipsen, T. Kowalski et H. Ono, éditeurs. *Residuated lattices: An algebraic glimpse at substructural logics*, volume 151 de *Studies in Logic and the Foundations of Mathematics*. Elsevier, 2007.

James Garson. *Modal logic for philosophers*. Cambridge University Press, 2006.

Gerhard Gentzen. Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift*, 39:176–210;405–431, 1934.

Jean-Yves Girard. Linear Logic. *Theoretical Computer Science*, 50(1):1–102, 1987.

Jean-Yves Girard. Linear logic: Its syntax and its semantic. Dans J.-Y. Girard, Y. Lafont et L. Regnier, éditeurs, *Advances in Linear Logic*, volume 222, pages 1–42. Cambridge University Press, 1995.

Lou Goble. A proposal for dealing with deontic dilemmas. Dans A. Lomuscio et D. Nute, éditeurs, *DEON 2004*, volume 3065 de *Lecture Notes in Computer Science*, pages 74–113. Springer, 2004.

Lou Goble. Normative conflicts and the logic of 'ought'. *Noûs*, 43(3):450–489, 2009.

Robert Goldblatt. *Topoi: The categorical analysis of logic.* Dover Publications, 2006.

Alvin I. Goldman. *A theory of human action.* Princeton University Press, 1970.

Thomas F. Gordon et Douglas Walton. Legal reasoning with argumentation schemes. Dans C. D. Hafner, éditeur, *12th Conference on Artificial Intelligence and Law*, pages 137–146. ACM, 2009.

Guido Governatori et Antonio Rotolo. Defeasible logic: Agency, intention and obligation. Dans A. Lomuscio et D. Nute, éditeurs, *DEON 2004*, volume 3065 de *Lecture Notes in Computer Science*, pages 114–128. Springer, 2004.

Jörg Hansen. Paradoxes of commitment. Dans G. Meggle, éditeur, *Actions, Norms, Values*, pages 255–263. Walter de Gruyter, 1999.

Bengt Hansson. An analysis of some deontic logics. *Noûs*, 3(4):373–398, 1969.

Sven Hansson. Preference-based deontic logic (PDL). *Journal of Philosophical Logic*, 19 (1):75–93, 1990.

David Harel, Dexter Kozen et Jerzy Tiuryn. *Dynamic Logic.* MIT Press, 2000.

Carl Hempel. *Éléments d'épistémologie.* Armand Colin, 1972.

Vincent Hendricks. *Mainstream and formal epistemology.* Cambridge University Press, 2006.

Andreas Herzig, Emiliano Lorini, Frédéric Moisan et Nicolas Troquard. A dynamic logic of normative systems. Dans T. Walsh, éditeur, *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, pages 228–233. AAAI Press, 2011.

Risto Hilpinen et Paul McNamara. Deontic logic: A historical survey and introduction. Dans D. M. Gabbay, J. Horty, X. Parent, R. van der Meyden et L. van der Torre, éditeurs, *Handbook of Deontic Logic and Normative Systems*, pages 3–136. College Publications, 2013.

Jaakko Hintikka. Some main problems of deontic logic. Dans R. Hilpinen, éditeur, *Deontic Logic: Introductory and Systematic Readings*, pages 59–104. D. Reidel Publishing Company, 1970.

Wesley N. Hohfeld. Fundamental legal conceptions as applied to judicial reasoning. *The Yale Law Journal*, 26(8):710–770, 1917.

John Horty. Moral dilemmas and non-monotonic logic. *Journal of Philosophical Logic*, 23 (1):35–65, 1994.

John Horty. Non-monotonic foundations for deontic logic. Dans D. Nute, éditeur, *Defeasible Deontic Logic*, pages 17–44. Kluwer Academic Publishers, 1997.

John Horty. *Agency and deontic logic*. Oxford University Press, 2001.

John Horty et Nuel Belnap. The deliberative stit: A study of action, omission, ability and obligation. *Journal of Philosophical Logic*, 24(6):583–644, 1995.

Jesse Hughes et Lambèr Royakkers. Don't ever do that!: Long-term duties in PD_eL. *Studia Logica*, 89(1):59–79, 2008.

David Hume. *Traité de la nature humaine [1740]*, volume 3. Flammarion, 1993.

Bart Jacobs. *Categorical logic and type theory*, volume 141 de *Studies in Logic and the Foundations of Mathematics*. Elsevier, 1999.

Dale Jacquette. Forrester's paradox. *Dialogue: Canadian Philosophical Review/Revue Canadienne de Philosophie*, 25(4):761–763, 1986.

Arnold Johanson. The logic of normative systems. Dans M. Brown et J. Carmo, éditeurs, *Deontic Logic, Agency and Normative Systems*, pages 123–133. Springer, 1996.

Andrew J. I. Jones. On the logic of deontic conditionals. *Ratio Juris*, 4(3):355–366, 1991.

Andrew J. I. Jones et Ingmar Pörn. Ideality, sub-ideality and deontic logic. *Synthese*, 65 (2):275–290, 1985.

Andrew J. I. Jones et Ingmar Pörn. 'Ought' and 'must'. *Synthese*, 66(1):89–93, 1986.

Andrew J. I. Jones et Marek Sergot. On the characterisation of law and computer systems: The normative systems perspective. Dans *Deontic Logic in Computer Science: Normative System Specification*, pages 275–307. John Wiley & Sons, 1993.

Andrew J. I. Jones et Marek Sergot. A formal characterisation of institutionalised power. *Logic Journal of the IGPL*, 4(3):427–443, 1996.

Jørgen Jørgensen. Imperatives and logic. *Erkenntnis*, 7(1):288–296, 1937.

Stig Kanger. *New foundations for ethical theory*. Stockholm, 1957.

Stig Kanger. New foundations for ethical theory. Dans R. Hilpinen, éditeur, *Deontic Logic: Introductory and Systematic Readings*, pages 36–58. D. Reidel Publishing Company, 1971.

Stig Kanger. Law and logic. *Theoria*, 38(3):105–132, 1972.

Stig Kanger et Helle Kanger. Rights and parliamentarism. *Theoria*, 32(2):85–115, 1966.

Gregory M. Kelly et Miguel L. Laplaza. Coherence for compact closed categories. *Journal of Pure and Applied Algebra*, 19:193–213, 1980.

Stephen Kleene. *Introduction to metamathematics*. ISHI Press International, 1952.

Simo Knuuttila. The emergence of deontic logic in the fourteenth century. Dans R. Hilpinen, éditeur, *New Studies in Deontic Logic*, pages 225–248. D. Reidel Publishing Company, 1981.

Barteld P. Kooi et Allard M. Tamminga. Conflicting obligations in multi-agent deontic logic. Dans L. Goble et J.-J. Ch. Meyer, éditeurs, *DEON 2006*, volume 4048 de *Lecture Notes in Computer Science*, pages 175–186. Springer, 2006.

Dexter Kozen. Kleene algebra with tests. *ACM Transactions on Programming Languages and Systems*, 19(3):427–443, 1997.

Saul Kripke. Semantical analysis of modal logic I: Normal propositional calculi. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 9:67–96, 1963.

Marie La Palme Reyes, John Macnamara et Gonzalo E. Reyes. Reference, kinds and predicates. Dans J. Macnamara et G. E. Reyes, éditeurs, *The Logical Foundations of Cognition*, pages 91–143. Oxford University Press, 1994.

Joachim Lambek. The mathematics of sentence structure. *The American Mathematical Monthly*, 65(3):154–170, 1958.

Joachim Lambek. Contributions to a mathematical analysis of the English verb phrase. *Journal of the Canadian Linguistic Association*, 5:83–89, 1959.

Joachim Lambek. Deductive systems and categories I. *Mathematical Systems Theory*, 2 (4):287–318, 1968.

Joachim Lambek. Deductive systems and categories II. Standard constructions and closed categories. Dans *Category Theory, Homology Theory and their Applications I*, volume 86 de *Lecture Notes in Mathematics*, pages 76–122. Springer, 1969.

Joachim Lambek. Categorial and categorical grammars. Dans R. T. Oehrle, E. Bach et D. Wheeler, éditeurs, *Categorial Grammars and Natural Language Structures*, pages 297–317. D. Reidel Publishing Company, 1988.

Joachim Lambek. Deductive systems and categories in linguistics. Dans D. M. Gabbay, H. J. Ohlback et U. Reyle, éditeurs, *Logic, Language and Reasonings*, pages 279–294. Kluwer Academic Publishers, 1999.

Edward J. Lemmon. Moral dilemmas. *The Philosophical Review*, 71(2):139–158, 1962.

Edward J. Lemmon. Deontic logic and the logic of imperatives. *Logique et Analyse*, 8: 39–71, 1965.

Hans Lenk. Complements and different lattice structures in a logic of action. *Erkenntnis*, 11(2):251–268, 1977.

Lars Lindahl et Jan Odelstad. An algebraic analysis of normative systems. *Ratio Juris*, 13(3):261–278, 2000.

Lars Lindahl et Jan Odelstad. The role of connections as minimal norms in normative systems. Dans T. Bench-Capon, A. Daskalopulu et R. G. F. Winkels, éditeurs, *Legal Knowledge and Information Systems. Jurix 2002: The Fifteenth Annual Conference*, pages 31–40. IOS Press, 2002.

Lars Lindahl et Jan Odelstad. Normative systems and their revision: An algebraic approach. *Artificial Intelligence and Law*, 11(2-3):81–104, 2003.

Lars Lindahl et Jan Odelstad. Normative positions within an algebraic approach to normative systems. *Journal of Applied Logic*, 2(1):63–91, 2004.

Lars Lindahl et Jan Odelstad. Intermediate concepts in normative systems. Dans L. Goble et J.-J. Ch. Meyer, éditeurs, *DEON 2006*, volume 4048 de *Lecture Notes in Computer Science*, pages 187–200. Springer, 2006.

Lars Lindahl et Jan Odelstad. Intermediaries and intervenients in normative systems. *Journal of Applied Logic*, 6(2):229–250, 2008a.

Lars Lindahl et Jan Odelstad. Strata of intervenient concepts in normative systems. Dans R. van der Meyden et L. van der Torre, éditeurs, *DEON 2008*, volume 5076 de *Lecture Notes in Computer Science*, pages 203–217. Springer, 2008b.

Lars Lindahl et Jan Odelstad. Stratification of normative systems with intermediaries. *Journal of Applied Logic*, 9(2):113–136, 2011.

Barry Loewer et Marvin Belzer. Dyadic deontic detachment. *Synthese*, 54(2):295–318, 1983.

Gert-Jan Lokhorst. Deontic linear logic with Petri net semantics. Rapport technique, FICT (Center for the Philosophy of Information and Communication Technology), 1997.

Thierry Lucas. Von Wright's action revisited: Actions as morphisms. *Logique et Analyse*, 49(193):85–115, 2006.

Thierry Lucas. Axioms for action. *Logique et Analyse*, 50(200):103–123, 2007.

Thierry Lucas. Deontic algebras of actions. *Logique et Analyse*, 51(202):367–389, 2008.

Saunders Mac Lane. *Categories for the working mathematician*. Springer, 2e édition, 1971.

Fabrizio Macagno et Douglas Walton. Argument from analogy in law, the classical tradition, and recent theories. *Philosophy and Rhetoric*, 42(2):154–182, 2009.

Fabrizio Macagno et Douglas Walton. Dichotomies and oppositions in legal argumentation. *Ratio Juris*, 23(2):229–257, 2010.

Fabrizio Macagno et Douglas Walton. Presumptions in legal argumentation. *Ratio Juris*, 25(3):271–300, 2012.

David Makinson et Leendert van der Torre. Input/output logics. *Journal of Philosophical Logic*, 29(4):383–408, 2000.

David Makinson et Leendert van der Torre. Constraints for input/output logics. *Journal of Philosophical Logic*, 30(2):155–185, 2001.

David Makinson et Leendert van der Torre. Permission from an input/output perspective. *Journal of Philosophical Logic*, 32(4):391–416, 2003a.

David Makinson et Leendert van der Torre. What is input/output logic? Dans *Foundations of the Formal Sciences II: Applications of Mathematical Logic in Philosophy and Linguistics*, volume 17 de *Trends in Logic Series*, pages 163–174. Kluwer Academic Publishers, 2003b.

Michael Makkai et Gonzalo E. Reyes. Completeness results for intuitionistic and modal logic in a categorical setting. *Annals of Pure and Applied Logic*, 72(1):25–101, 1995.

Ernst Mally. Grundgesetze des Sollens. Elemente der Logik des Willens. Dans K. Wolf et P. Weingartner, éditeurs, *Logische Schriften: Großes Logikfragment, Grundgesetze des Sollens*, page 227–324. D. Reidel Publishing Company, 1926.

Jean-Pierre Marquis. *From a geometrical point of view: A study of the history and philosophy of category theory*. Springer, 2009.

Jean-Pierre Marquis. Category theory. Dans E. N. Zalta, éditeur, *The Stanford Encyclopedia of Philosophy*. 2013.

Audrey M. McKinney. *Conditional obligation and temporally dependent necessity*. Thèse de doctorat, University of Pennsylvania, 1977.

Paul McNamara. Deontic logic. Dans E. N. Zalta, éditeur, *The Stanford Encyclopedia of Philosophy*. 2010.

Elliott Mendelson. *Introduction to mathematical logic*. CRC Press, 5e édition, 2009.

John-Jules Ch. Meyer. A simple solution to the "deepest" paradox in deontic logic. *Logique et Analyse*, 30(117-118):81–90, 1987.

John-Jules Ch. Meyer. A different approach to deontic logic: Deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, 29(1):109–136, 1988.

Michael Moortgat. Symmetric categorial grammar. *Journal of Philosophical Logic*, 38(6): 681–710, 2009.

Michael Moortgat. Typelogical grammar. Dans E. N. Zalta, éditeur, *The Stanford Encyclopedia of Philosophy*. 2012.

Richard Moot et Christian Retoré. *The Logic of Categorial Grammar*, volume 6850 de *Lecture Notes in Computer Science*. Springer, 2012.

Micheal Morreau. Reason to think and act. Dans D. Nute, éditeur, *Defeasible Deontic Logic*, pages 139–158. Kluwer Academic Publishers, 1997.

Benjamin Mossel. Negative actions. *Philosophia*, 37(2):307–333, 2009.

Peter Mott. On Chisholm's paradox. *Journal of Philosophical Logic*, 2(2):197–211, 1973.

Ian Niles. Rescuing the counterfactual solution to Chisholm's paradox. *Philosophia*, 25 (1-4):351–371, 1997.

Robert Nozick et Richard Routley. Escaping the good Samaritan paradox. *Mind*, 71(283): 377–382, 1962.

D. Nute, éditeur. *Defeasible deontic logic*. Kluwer Academic Publishers, 1997.

Donald Nute et Xiaochang Yu. Introduction. Dans D. Nute, éditeur, *Defeasible Deontic Logic*, pages 1–16. Kluwer Academic Publishers, 1997.

Jan Odelstad et Lars Lindahl. Normative systems represented by Boolean quasi-orderings. *Nordic Journal of Philosophical Logic*, 5(2):161–174, 2000.

Kazimierz Opałek et Jan Woleński. Normative systems, permission and deontic logic. *Ratio Juris*, 4(3):334–348, 1991.

Olga Pacheco et José Carmo. A role based model for the normative specification of organized collective agency and agents interaction. *Autonomous Agents and Multi-Agent Systems*, 6(2):145–184, 2003.

Olga Pacheco et Filipe Santos. Delegation in a role-based organization. Dans A. Lomuscio et D. Nute, éditeurs, *DEON 2004*, volume 3065 de *Lecture Notes in Computer Science*, pages 209–227. Springer, 2004.

Hugues Parent. *Traité de droit criminel, Tome II - La culpabilité (actus reus et mens rea).* Éditions Thémis, 2007.

Hugues Parent. *Traité de droit criminel, Tome I - L'imputabilité*. Éditions Thémis, 2008.

Marc Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166, 2002.

Dusko Pavlovic. Monoidal computer I: Basic computability by string diagrams. *Information and Computation*, 226:94–116, 2013. URL `http://arxiv.org/abs/1208.5205`.

Clayton Peterson. *La logique déontique: Une application de la logique à l'éthique et au discours juridique*. Mémoire de maîtrise, Université de Montréal, 2011.

Clayton Peterson. Normative inferences and validity: A heuristic. *Cogency*, 5(1):85–105, 2013a.

Clayton Peterson. *Pensée rationnelle et argumentation*. Presses de l'Université de Montréal, 2013b.

Clayton Peterson. A categorical analysis of action logics. Manuscrit soumis pour publication, 2014a.

Clayton Peterson. The categorical imperative: Category theory as a foundation for deontic logic. *Journal of Applied Logic*, 2014b. http://dx.doi.org/10.1016/j.jal.2014.07.001.

Clayton Peterson. Conditional reasoning with string diagrams. Manuscrit soumis pour publication, 2014c.

Clayton Peterson. Contrary-to-duty reasoning: A categorical approach. Manuscrit soumis pour publication, 2014d.

Clayton Peterson. Formal philosophy and legal reasoning: The validity of legal inferences. Manuscrit soumis pour publication, 2014e.

Clayton Peterson. From linguistics to deontic logic via category theory. Manuscrit soumis pour publication, 2014f.

Clayton Peterson et Jean-Pierre Marquis. A note on Forrester's paradox. *Polish Journal of Philosophy*, VI(2):53–70, 2012.

Guido Pincione. Critical thinking and legal culture. *Rationality, Markets and Morals*, (3): 373–391, 2009.

Henri Poincaré. *Dernières pensées*. Kessinger Publishing, 1913.

Henry Prakken. AI & law, logic and argument schemes. *Argumentation*, 19(3):303–320, 2005.

Henry Prakken et Marek Sergot. Contrary-to-duty obligations. *Studia Logica*, 57(1):91–115, 1996.

Henry Prakken et Marek Sergot. Dyadic deontic logic and contrary-to-duty obligations. Dans D. Nute, éditeur, *Defeasible Deontic Logic*, pages 223–262. Kluwer Academic Publishers, 1997.

Vaughan Pratt. Semantical considerations of Floyd-Hoare Logic. Rapport technique MIT/LCS/TR-168, 1976.

Vaughan Pratt. Application of modal logic to programming. *Studia Logica*, 39(2-3):257–274, 1980.

Vaughan Pratt. Action logic and pure induction. Dans J. Eijck, éditeur, *Logics in AI*, volume 478 de *Lecture Notes in Computer Science*, pages 97–120. Springer, 1991.

Arthur Prior. The paradoxes of derived obligation. *Mind*, 63(249):64–65, 1954.

Arthur Prior. Escapism: The logical basis of ethics. Dans A. I. Melden, éditeur, *Essays in Moral Philosophy*, pages 135–146. University of Washington Press, 1958.

Arthur Prior. *Past, present and future*. Oxford University Press, 1967.

Cristian Prisacariu et Gerardo Schneider. A dynamic deontic logic for complex contracts. *Journal of Logic and Algebraic Programming*, 81(4):458–490, 2012.

Ingmar Pörn. *The logic of power*. Basil Blackwell, 1970.

Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

Nicholas Rescher. An axiom system for deontic logic. *Philosophical Studies*, 9(1):24–30, 1958.

Nicholas Rescher. Conditional permission in deontic logic. *Philosophical Studies*, 13(1): 1–6, 1962.

Gonzalo E. Reyes et Houman Zolfaghari. Topos-theoretic approaches to modality. Dans A. Carboni, M. Pedicchio et G. Rosolini, éditeurs, *Category Theory*, volume 1488 de *Lecture Notes in Mathematics*, pages 359–378. Springer, 1991.

Gonzalo E. Reyes et Houman Zolfaghari. Bi-Heyting algebras, toposes and modalities. *Journal of Philosophical Logic*, 25(1):25–43, 1996.

John Robison. Further difficulties for conditional permission in deontic logic. *Philosophical Studies*, 18(1):27–30, 1967.

Alf Ross. Imperatives and logic. *Theoria*, 7(1):53–71, 1941.

Lambèr Royakkers. *Extending deontic logic for the formalisation of legal rules*. Kluwer Academic Publishers, 1998.

Lambèr Royakkers. Combining deontic and action logics for collective agency. Dans J. Breuker, R. Leenes et R. Winkels, éditeurs, *Legal Knowledge and Information Systems. Jurix 2000: The Thirteenth Annual Conference*, pages 135–146. IOS Press, 2000.

Lambèr Royakkers et Frank Dignum. Defeasible reasoning with legal rules. Dans D. Nute, éditeur, *Defeasible Deontic Logic*, pages 263–286. Kluwer Academic Publishers, 1997.

Young U. Ryu et Ronald M. Lee. Deontic logic viewed as defeasible reasoning. Dans D. Nute, éditeur, *Defeasible Deontic Logic*, pages 123–137. Kluwer Academic Publishers, 1997.

Henrik Sahlqvist. Completeness and correspondence in first and second order semantics for modal logic. Dans S. Kanger, éditeur, *Proceedings of the Thrid Scandanavian Logic Symposium*, pages 110–143. 1975.

Giovanni Sartor. A formal model of legal argumentation. *Ratio Juris*, 7(2):177–211, 1994.

Andrea Schalk. What is a categorical model for linear logic? URL `http://www.cs.man.ac.uk/~schalk/notes/llmodel.pdf`. 2004.

Peter Schotch et Raymond E. Jennings. Non-Kripkean deontic logic. Dans R. Hilpinen, éditeur, *New Studies in Deontic Logic*, pages 149–162. D. Reidel Publishing Company, 1981.

Robert A. G. Seely. Linear logic, $*$-autonomous categories and cofree coalgebras. Dans *In Categories in Computer Science and Logic*, volume 92 de *Contemporary Mathematics*, pages 371–382. 1989.

Krister Segerberg. A completeness theorem in the modal logic of programs. Dans T. Traczyk, éditeur, *Universal Algebra and Applications*, volume 9, pages 31–46. Polish Scientific Publishers, 1982a.

Krister Segerberg. A deontic logic of action. *Studia Logica*, 41(2):269–282, 1982b.

Krister Segerberg. Getting started: Beginnings in the logic of action. *Studia Logica*, 51 (3):347–378, 1992.

Krister Segerberg. A blueprint for deontic logic in three (not necessarily easy) steps. Dans G. Bonanno, J. P. Delgrande, J. Lang et H. Rott, éditeurs, *Formal Models of Belief Change in Rational Agents*. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), 2007.

Krister Segerberg. Blueprint for a dynamic deontic logic. *Journal of Applied Logic*, 7(4): 388–402, 2009.

Krister Segerberg. DΔL: A dynamic deontic logic. *Synthese*, 185(1):1–17, 2012.

Krister Segerberg, John-Jules Meyer et Marcus Kracht. The logic of action. Dans E. N. Zalta, éditeur, *The Stanford Encyclopedia of Philosophy*. 2009.

Kim Solin. Dual choice and iteration in an abstract algebra of action. *Studia Logica*, 100 (3):607–630, 2012.

Kornel Solt. Deontic Alternative Worlds and the Truth-Value of 'OA'. *Logique et Analyse*, 27(107):349–351, 1984.

Matthew Spinks et Robert Veroff. Constructive logic with strong negation in a substructural logic (I). *Studia Logica*, 88(3):325–348, 2008.

Christian Straßer. A deontic logic framework allowing for factual detachment. *Journal of Applied Logic*, 9(1):61–80, 2011.

Ruth Sullivan. *Driedger on the construction of statutes*. Butterworths, 1994.

Alfred Tarski. The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3):341–376, 1944.

Richmond Thomason. Indeterminist time and truth-value gaps. *Theoria*, 36(3):264–281, 1970.

Richmond Thomason. Deontic logic as founded on tense logic. Dans R. Hilpinen, éditeur, *New Studies in Deontic Logic*, pages 165–176. D. Reidel Publishing Company, 1981.

Richmond Thomason. Combinations of tense and modality. Dans D. M. Gabbay et F. Guenthner, éditeurs, *Handbook of Philosophical Logic*, volume 7, pages 205–234. Kluwer Academic Publishers, 2e édition, 2002.

Paul Tomassi. *Logic*. Routledge, 1999.

James Tomberlin. Contrary-to-duty imperatives and conditional obligations. *Noûs*, 15(3): 357–375, 1981.

James Tomberlin. Contrary-to-duty imperatives and Castañeda's system of deontic logic. Dans J. Tomberlin, éditeur, *Agent, Language and the Structure of the World*, pages 231–249. Hackett, 1983.

Robert Trypuz et Piotr Kulicki. A systematics of deontic action logics based on Boolean algebra. *Logic and Logical Philosophy*, 18:253–270, 2009.

Robert Trypuz et Piotr Kulicki. Towards metalogical systematisation of deontic action logics based on Boolean algebra. Dans G. Governatori et G. Sartor, éditeurs, *DEON 2010*, volume 6181 de *Lecture Notes in Computer Science*, pages 132–147. Springer, 2010.

Johan van Benthem. *Language in action: Categories, lambdas and dynamic Logic*. Elsevier, 1991.

Ron van der Meyden. A clausal logic for deontic action specification. Dans V. Saraswat et K. Ueda, éditeurs, *Logic Programming, Proceedings of the 1991 International Symposium*, pages 221–238. MIT Press, 1991.

Ron van der Meyden. The dynamic logic of permission. *Journal of Logic and Computation*, 6(3):465–479, 1996.

Leendert van der Torre et Yao-Hua Tan. The many faces of defeasibility in defeasible deontic logic. Dans D. Nute, éditeur, *Defeasible Deontic Logic*, pages 79–121. Kluwer Academic Publishers, 1997.

Leendert van der Torre et Yao-Hua Tan. Contrary-to-duty reasoning with preference-based dyadic obligations. *Annals of Mathematics and Artificial Intelligence*, 27(1-4): 49–78, 1999.

Job van Eck. A system of temporally relative modal and deontic predicate logic and it's philosophical applications. *Logique et Analyse*, 25(99):249–290, 1982a.

Job van Eck. A system of temporally relative modal and deontic predicate logic and it's philosophical applications (2). *Logique et Analyse*, 25(100):339–381, 1982b.

Bas C. van Fraassen. The logic of conditional obligation. *Journal of Philosophical Logic*, 1(3):417–438, 1972.

Bas C. van Fraassen. Values and the heart's command. *The Journal of Philosophy*, 70(1): 5–19, 1973.

Mark van Roojen. Moral cognitivism vs. non-cognitivism. Dans E. N. Zalta, éditeur, *The Stanford Encyclopedia of Philosophy*. 2011.

Georg H. von Wright. Deontic logic. *Mind*, 60(237):1–15, 1951.

Georg H. von Wright. A note on deontic logic and derived obligation. *Mind*, 65(260): 507–509, 1956.

Georg H. von Wright. *Norm and Action*. Routledge & Kegan Paul, 1963.

Georg H. von Wright. Deontic logics. *American Philosophical Quarterly*, 4(2):136–143, 1967.

Georg H. von Wright. *An essay in deontic logic and the general theory of action*, volume XXI de *Acta Philosophica Fennica*. North-Holland Publishing Company, 1968.

Georg H. von Wright. Norms of higher order. *Studia Logica*, 42(2):119–127, 1983.

Georg H. von Wright. Deontic logic: A personal view. *Ratio Juris*, 12(1):26–38, 1999a.

Georg H. von Wright. Ought to be–ought to do. Dans G. Meggle, éditeur, *Actions, Norms, Values*, pages 3–10. Walter de Gruyter, 1999b.

Mark Vorobej. Conditional obligation and detachement. *Canadian Journal of Philosophy*, 16(1):11–26, 1986.

Douglas Walton. Omitting, refraining and letting happen. *American Philosophical Quarterly*, 17(4):319–326, 1980.

Douglas Walton. New methods for evaluating arguments. *Inquiry: Critical Thinking Across the Disciplines*, 15(4):44–65, 1995.

Douglas Walton. How can logic best be applied to arguments? *Logic Journal of the IGPL*, 5(4):603–614, 1997.

Douglas Walton. Similarity, precedent and argument from analogy. *Artificial Intelligence and Law*, 18(3):217–246, 2010.

Douglas Walton, Katie Atkinson, Trevor Bench-Capon, Adam Wyner et Dan Cartwright. Argumentation in the framework of deliberation dialogue. Dans C. Bjola et M. Kornprobst, éditeurs, *Arguing Global Governance*, pages 201–230. Routledge, 2010.

Douglas Walton et David M. Godden. Informal logic and the dialectical approach to argument. Dans H. V. Hansen et R. C. Pinto, éditeurs, *Reason Reclaimed*, pages 3–17. Vale Press, 2007.

Douglas Walton et Thomas F. Gordon. Critical questions in computational models of legal argument. Dans P. E. Dunne et T. Bench-Capon, éditeurs, *IAAIL Workshop Series, International Workshop on Argumentation in Artificial Intelligence and Law*, pages 103–111. Wolf Legal Publishers, 2005.

Ota Weinberger. A philosophical approach to norm logic. *Ratio Juris*, 14(1):130–141, 2001.

Roel Wieringa et John-Jules Ch. Meyer. Applications of deontic logic in computer science: A concise overview. Dans *Deontic Logic in Computer Science: Normative System Specification*, pages 17–40. John Wiley & Sons, 1993.

Charles Wiseman. The theory of modal groups. *The Journal of Philosophy*, 67(11):367–376, 1970.

Jan Wolenski. Deontic logic and possible world semantics: A historical sketch. *Studia Logica*, 49(2):273–282, 1990.

Ming Xu. On the basic logic of stit with a single agent. *The Journal of Symbolic Logic*, 60(2):459–483, 1995.