

Université de Montréal

Configuration et exploitation d'une machine émotionnelle

par

Amine Trabelsi

Département d'informatique et de recherche opérationnelle

Faculté des arts et des sciences

Mémoire présenté à la Faculté des arts et des sciences
en vue de l'obtention du grade de Maîtrise ès Science (M. Sc.)
en Informatique

Novembre, 2010

© Amine Trabelsi, 2010

Université de Montréal
Faculté des arts et des sciences

Ce mémoire intitulé :

Configuration et exploitation d'une machine émotionnelle

Présenté par :
Amine Trabelsi

a été évalué par un jury composé des personnes suivantes :

Julie Vachon, président-rapporteur
Claude Frasson, directeur de recherche
Esma Aïmeur, membre du jury

Résumé

Dans ce travail, nous explorons la faisabilité de doter les machines de la capacité de prédire, dans un contexte d'interaction homme-machine (IHM), l'émotion d'un utilisateur, ainsi que son intensité, de manière instantanée pour une grande variété de situations. Plus spécifiquement, une application a été développée, appelée *machine émotionnelle*, capable de «comprendre» la signification d'une situation en se basant sur le modèle théorique d'évaluation de l'émotion Ortony, Clore et Collins (OCC). Cette machine est apte, également, à prédire les réactions émotionnelles des utilisateurs, en combinant des versions améliorées des *k plus proches voisins* et des *réseaux de neurones*. Une procédure empirique a été réalisée pour l'acquisition des données. Ces dernières ont fourni une connaissance consistante aux algorithmes d'apprentissage choisis et ont permis de tester la performance de la machine. Les résultats obtenus montrent que la *machine émotionnelle* proposée est capable de produire de bonnes prédictions. Une telle réalisation pourrait encourager son utilisation future dans des domaines exploitant la reconnaissance automatique de l'émotion.

Mots-clés : Reconnaissance automatique de l'émotion ; Apprentissage machine ; Interaction homme-machine ; Informatique affective.

Abstract

This paper explores the feasibility of equipping computers with the ability to predict, in a context of a human computer interaction, the probable user's emotion and its intensity for a wide variety of emotion-eliciting situations. More specifically, an online framework, the *Emotional Machine*, is developed enabling computers to «understand» situations using OCC model of emotion and to predict user's reaction by combining refined versions of Artificial Neural Network and k Nearest Neighbours algorithms. An empirical procedure including a web-based anonymous questionnaire for data acquisition was designed to provide the chosen machine learning algorithms with a consistent knowledge and to test the application's recognition performance. Results from the empirical investigation show that the proposed *Emotional Machine* is capable of producing accurate predictions. Such an achievement may encourage future using of our framework for automated emotion recognition in various application fields.

Keywords: Automated Emotion Recognition; Machine Learning; Human Computer Interaction; Affective Computing

Table des matières

Résumé.....	iii
Abstract.....	iv
Table des matières.....	v
Liste des tableaux.....	vii
Liste des figures.....	viii
Remerciements.....	x
Introduction.....	1
Chapitre 1 Utilisation des émotions en informatique.....	11
1.1 Qu'est-ce qu'une émotion.....	11
1.2 Les théories de l'émotion.....	12
1.2.1 Les théories de l'émotion discrète.....	12
1.2.2 Les théories de l'émotion dimensionnelle.....	13
1.2.3 Les théories d'évaluation cognitive.....	14
1.3 Les émotions dans les interactions homme-machine.....	16
1.3.1 L'informatique affective.....	16
1.3.2 Sources et outils de reconnaissance d'émotion.....	17
1.3.3 Applications de la reconnaissance automatique de l'émotion.....	21
1.3.4 Les domaines inexplorés et les limites.....	22
1.4 L'émotion et les systèmes intelligents.....	24
1.5 Conclusion.....	26
Chapitre 2 Algorithmes d'apprentissage machine.....	28
2.1 Qu'est-ce que l'apprentissage machine ?.....	28
2.2 Apprentissage supervisé.....	30
2.2.1 Présentation.....	30
2.2.2 Phases d'entraînement, validation et test.....	32
2.3 Algorithmes utilisés.....	35
2.4 Choix des algorithmes.....	41
2.5 Conclusion.....	43
Chapitre 3 Procédure empirique: collecte et description numérique des situations émotionnelles.....	44

3.1 Le modèle OCC	45
3.1.1 Les types de situations	45
3.1.2 Les groupes et paires d'émotions dans le modèle OCC.....	47
3.1.3 Relation entre les types de situations et les émotions	48
3.1.4 Les variables d'évaluation du modèle OCC.....	50
3.1.5 La description d'une situation émotionnelle.....	51
3.2 Construction d'une base de situations émotionnelles	53
3.2.1 Collecte de situations	54
3.2.2 Description formelle des situations émotionnelles	58
3.3 Conclusion	64
Chapitre 4 Architecture et exploitation de la machine émotionnelle.....	66
4.1 Construction de l'ensemble d'entraînement.....	66
4.1.1 Importance des informations personnelles.....	67
4.1.2 Collecte des informations personnelles et des réactions émotionnelles.....	69
4.2 Le fonctionnement de la Machine émotionnelle.....	72
4.2.1 Principe général.....	73
4.2.2 Architecture.....	75
4.3 Exploitation et résultats.....	84
4.4 Conclusion	91
Conclusion générale.....	93
Bibliographie.....	97
Annexe I Apprentissage du perceptron multicouche	xi
Annexe II Les situations émotionnelles collectées	xxii
Annexe III Questionnaire de Personnalité EPQR-A.....	xxxvi

Liste des tableaux

Tableau 3.1 Description semi-formelle des situations correspondantes aux 22 émotions...	48
Tableau 3.2 Termes émotionnels équivalents	49
Tableau 3.3 Les variables d'évaluation-intensité du modèle OCC.....	51
Tableau 3.4 Les variables centrales et locales relatives à chaque type de situation	52
Tableau 3.5 La répartition des 250 situations finales selon les émotions et les intensités...	57
Tableau 3.6 Les valeurs prises par les variables d'évaluation-intensité.....	59
Tableau 4.1 Exemple de calcul de vecteur de sortie du k PPV pondéré pour l'émotion.....	78
Tableau 4.2 Exemple de calcul de vecteur de sortie du k PPV pondéré pour l'intensité	78
Tableau 4.3 Exemple de vecteur de sortie du RdN pour l'émotion.....	80
Tableau 4.4 Exemple de vecteur de sortie du RdN pour l'intensité	80
Tableau 4.5 Les pourcentages de prédictions correctes pour les émotions et les intensités	90

Liste des figures

Figure 1.1 Photo d'André Agassi à la fin d'un match	23
Figure 2.1 Schématisation d'un problème de surapprentissage.....	33
Figure 2.2 Schématisation d'un problème de sousapprentissage	34
Figure 2.3 Exemple d'un arbre de décision	35
Figure 2.4 Perceptron multicouche à une seule couche cachée	37
Figure 3.1 Structure du modèle OCC (<i>Ortony, Clore, & Collins, 1988</i>).....	47
Figure 3.2 Saisie des situations émotionnelles.....	55
Figure 3.3 Interface de sélection du type de la situation.....	61
Figure 3.4 Interface de description de la situation avec les variables globales.....	62
Figure 3.5 Interface de description d'une situation avec les variables locales.....	63
Figure 4.1 Saisie des informations personnelles et du questionnaire de personnalité	69
Figure 4.2 Interface de saisie de l'émotion et son intensité.....	71
Figure 4.3 Architecture de la <i>machine émotionnelle</i>	75
Figure 4.4 Réseau de neurones appris pour la prédiction des émotions et leurs intensités.	79
Figure 4.5 Interface d'affichage de la prédiction de la <i>machine émotionnelle</i>	82
Figure 4.6 Distribution des participants selon le sexe.....	85
Figure 4.7 Distribution des participants selon l'âge	86
Figure 4.8 Distribution des participants selon la religion	86
Figure 4.9 Distribution des traits de personnalité majoritaires chez les participants.....	87

à mes parents Abdelwahed et Monia, et mon frère Chiheb

Remerciements

Je tiens, tout d'abord, à remercier mon directeur de recherche, Professeur Claude Frasson, pour le soutien constant qu'il m'a apporté et le suivi régulier qu'il m'a accordé tout au long de ce travail. Ses idées et ses remarques, ainsi que ses encouragements et sa patience m'ont été d'une aide précieuse, sans laquelle ce mémoire n'aurait pas vu le jour.

Un remerciement sincère s'adresse aux membres du jury qui m'ont fait l'honneur d'évaluer ce travail.

J'aimerais également remercier mes amis et collègues étudiants du laboratoire Héron, particulièrement Lotfi, Pierre, Maher, Arezki, Franck, Amine, Soumaya et Ramla, pour leurs conseils et leur disponibilité.

Je voudrais aussi exprimer ma profonde gratitude à mon père pour son soutien continu et sa présence, à ma mère pour ses encouragements et ses conseils précieux, ainsi qu'à mon frère pour le temps qu'il a consacré pour réviser ce mémoire. J'en profite, également, pour remercier mon cousin Aymen qui m'a été d'un grand soutien moral durant ces deux dernières années passées ensemble.

Introduction

L'utilisation des ordinateurs a connu une expansion colossale lors des deux dernières décennies. Désormais, leur manipulation se fait dans le cadre du travail avec l'accroissement du traitement informatique en ce qui concerne les fonctionnalités administratives, les études, ou encore l'enseignement. De plus, avec l'expansion d'internet, l'acquisition de l'information est devenue beaucoup plus facile et rapide. Ainsi, l'utilisation des ordinateurs est devenue très variée, avec la capacité pour tous, à partir de simples clics, d'enrichir son capital de savoir et de culture, de se divertir, de chercher des restaurants, d'acheter des livres, de communiquer avec le monde extérieur à l'aide des réseaux sociaux ou de logiciels de « chats », de donner son avis sur des forums dédiés, de chercher des itinéraires, etc. La présence des ordinateurs est telle qu'elle se fait ressentir de plus en plus tant au niveau professionnel, que personnel et social.

Adapter ces machines (les ordinateurs) et les rendre plus conviviales pour leurs utilisateurs est désormais une nécessité. Il s'agit de rendre leur interaction plus naturelle, se rapprochant davantage du comportement humain. Elles tendraient, par conséquent, à anticiper les besoins des utilisateurs, à reconnaître leurs désirs et à comprendre leur état d'esprit. En fait, le caractère naturel de l'interaction quotidienne ou typique entre les humains réside dans la capacité d'induire ou d'inférer l'émotion de l'autre et d'ajuster, à partir de là, son comportement ou sa réponse, et ce afin d'assurer un meilleur échange (*Fragopanagos & Taylor, 2005*). Plusieurs recherches récentes en interaction homme-machine (IHM) (Human-Computer Interaction en anglais) suggèrent d'ailleurs de doter les ordinateurs de la capacité de reconnaître et de détecter le comportement de l'utilisateur, principalement son comportement affectif ou émotionnel, avant d'initier une interaction basée sur cette information (*Maja Pantic, Pentland, Nijholt, & Huang, 2007; Zeng, Pantic, Roisman, & Huang, 2008*). Cette reconnaissance automatique des émotions pourrait rendre les ordinateurs proactifs avec des interfaces centrées utilisateurs, allant même jusqu'à provoquer l'interaction pour subvenir à un besoin ou à un désir non indiqué préalablement. Ils cesseraient donc d'être passifs en répondant seulement aux commandes. Cela pourrait aussi servir à influencer délibérément l'émotion de l'utilisateur dans une perspective empathique, c'est-à-dire le rendre joyeux, atténuer son mécontentement ou sa frustration, etc. (*Picard, 2000*). Ce genre de procédé, à savoir, détecter automatiquement l'état

émotionnel d'une personne et essayer d'agir en conséquence, constitue l'objectif des recherches en informatique affective ou « *affective computing* », une discipline intégrant l'intelligence artificielle, la psychologie et les sciences cognitives (Picard, 2000; Tao & Tan, 2005). Par conséquent, plusieurs recherches en IHM ont incorporé l'informatique affective et l'ont utilisée dans une variété d'applications (p. ex., Duric et al., 2002; Kapoor, Burlison, & Picard, 2007).

En effet, les interfaces capables de reconnaître l'émotion de la personne sont nombreuses. Elles peuvent être intégrées aux systèmes tuteurs intelligents (p. ex., Conati & Maclaren, 2004; D'Mello, Graesser, & Picard, 2007; Kapoor, Burlison, & Picard, 2007; Ochs & Frasson, 2004) pour reconnaître l'émotion de l'apprenant, ou dans les véhicules (p. ex., Healey, 2000; Lisetti & Nasoz, 2005) afin de repérer un conducteur stressé ou en extrême colère. De même, la reconnaissance automatique de l'émotion dans les centres d'appels peut permettre de détecter la frustration d'un client à partir du ton de sa voix (p. ex., Lee & Narayanan, 2005; Vidrascu & Devillers, 2005). Dans les secteurs des jeux vidéo et des divertissements virtuels, elle peut également instaurer un environnement d'interaction ressemblant à celui des humains (p. ex., Healey, Picard, & Dabek, 1998; Slater, Buckley, & Bechkoum, 2006).

Le recours fréquent à la reconnaissance automatique de l'émotion est dû, entre autres, aux avancées technologiques réalisées dans le traitement des images, des vidéos, de la voix et des signaux physiologiques reliés aux émotions. Ainsi, une machine (un ordinateur, un robot, un agent intelligent, etc.) peut détecter l'émotion de quelqu'un à partir des sources ou canaux habituels utilisés par les humains, soit l'expression du visage, l'intonation de la voix, la posture du corps, les gestes et les mouvements. Elle peut également la détecter en analysant les réponses physiologiques comme par exemple le rythme cardiaque, la tension artérielle, le pouls, la conductance de la peau et la respiration. Ces différentes informations sont recueillies à l'aide d'outils spécifiques, de plus en plus accessibles, comme les caméras détectant avec précision chaque geste, les oculomètres permettant de suivre et d'enregistrer le mouvement des yeux, les souris ou les chaises capteurs de signaux, les capteurs de la peau, etc. Ainsi, les données recueillies à l'aide de ces outils et à partir des différentes sources sont traitées par la machine dans le but d'identifier l'émotion de la personne. Généralement, pour réaliser cet objectif, la plupart

des recherches ont recours aux algorithmes d'apprentissage, particulièrement les algorithmes supervisés. En effet, ces derniers sont utilisés afin de prédire ou classer un phénomène, comme l'émotion d'un individu (joie, colère, peur, etc.), à partir d'informations (l'expression du visage, l'intonation, les signaux, etc.), appelées aussi variables explicatives pouvant contenir des patrons (patterns) qui aideraient à le reconnaître ou le classer. Pour ce faire, ces algorithmes nécessitent un entraînement préalable sur plusieurs cas du phénomène et ses différentes manifestations, décrites par les variables explicatives. Ces différents cas sont rassemblés dans ce qu'on appelle un ensemble d'entraînement ou d'apprentissage. Ainsi, dans des recherches en reconnaissance automatique de l'émotion, plusieurs études empiriques ont pu être réalisées (p. ex., *Gunes & Piccardi, 2006; Lee & Narayanan, 2005*), donnant lieu à plusieurs ensembles ou bases de données d'images, vidéos, conversations, signaux physiologiques, etc.

Parallèlement à l'expansion des techniques de reconnaissance dans les IHM, les émotions ont fait l'objet de recherches dans un cadre plus global, celui des systèmes computationnels intelligents. En effet, ces recherches visent à produire des systèmes qui simulent le raisonnement et le comportement humain au sein, par exemple, de robots ou des agents autonomes intelligents. Dans un premier temps, la plupart de ces travaux tendent à hisser les recherches en intelligence artificielle à un niveau plus élevé, en concevant des machines qui réfléchissent et qui sont capables de percevoir leur environnement. Dans un deuxième temps, leur but est d'aider à mieux comprendre le fonctionnement du cerveau humain et sa manière de raisonner (*Minsky, 2007*). Plusieurs chercheurs affirment que ces « cerveaux artificiels » ne seraient pas complets sans une composante simulant le processus émotionnel humain, ce qui permettrait une synthèse plus fiable du processus de raisonnement (*Franklin, 2002; Minsky, 1988, 2007; Sloman, 2001*).

Ainsi, dans un but d'humaniser la machine, l'incorporation de l'émotion, dans la plupart des travaux portant sur les systèmes intelligents, s'est faite pour simuler le processus humain de génération de l'émotion à l'aide de modèles computationnels inspirés des modèles théoriques établis par des psychologues. Dans la littérature relative à l'émotion, il y a principalement trois courants qui diffèrent selon la spécification des composantes principales intervenant dans le processus de génération (la cognition, le processus physiologique, etc.) et dans les interactions entre ces différentes composantes

(Gratch, Marsella, & Petta, 2009). En outre, les théories de l'*émotion discrète* (p. ex., Ekman, Friesen, & Ellsworth, 1972; Izard, 1971; Oatley & Johnson-Laird, 1987; Tomkins, 1962) soutiennent que les émotions constituent un ensemble discret et qu'elles sont engendrées par un processus ou un programme « capteur-moteur » impliquant une relation ou un lien entre la cognition et les réponses somatiques. Par contre, les théories *dimensionnelles* (p. ex., Carver, 2001; Lang, 1995; Russell, 1980, 2003) réfutent le fait que les émotions soient discrètes et soutiennent que celles qui sont communes (peur, joie, etc.) ne sont que des étiquettes (ou labels) cognitives qu'on utilise pour mieux comprendre une activation physiologique qui ne peut pas être sujette à une discrétisation. Par conséquent, ces théories proposent de mesurer cette activation selon des dimensions bipolaires comme la valence (émotion positive ou négative) et l'excitation (forte ou faible) (Lang, 1995). Néanmoins, les théories les plus appliquées dans les systèmes intelligents et qui représentent le troisième courant sont les théories de l'*évaluation cognitive* (p. ex., Fridja, 1986; Lazarus, 1991; Scherer, 1986). Ces théories focalisent sur la perception cognitive des situations ou des événements engendrant une émotion. L'idée principale repose sur le fait que l'évaluation établie par un individu de l'environnement, des circonstances et des conséquences d'une situation qu'il vit, joue un rôle important dans la génération de sa réponse émotionnelle. Ainsi, ces théories définissent un ensemble de critères ou de dimensions qui sont présumées caractériser le processus d'évaluation cognitive d'une situation engendrant une émotion. Ces dimensions sont appelées variables d'évaluation, comme par exemple, la *surprise* (c.-à-d., le degré auquel une personne s'attend ou ne s'attend pas à l'occurrence de l'évènement), ou le *potentiel d'adaptation* (c. à-d., le point à partir duquel l'individu croit être capable de s'adapter à la situation qui se présente). Par exemple, une situation totalement inattendue, pour laquelle on croit ne pas pouvoir s'adapter peut engendrer une émotion de peur (Gratch, Marsella, & Petta, 2009). Les différentes théories d'évaluation se distinguent les unes des autres, principalement par les variables d'évaluation qu'elles proposent et les détails du processus cognitif les impliquant (le lien entre ces variables et l'émotion engendrée correspondante). Les modèles computationnels émotionnels ont été majoritairement influencés par la théorie d'évaluation d'Ortony, Clore et Collins (1988), plus connue sous le nom du modèle OCC. En la définissant, les auteurs ciblaient la communauté d'intelligence artificielle, ce qui a rendu sa

conversion, dans des programmes informatiques, triviale. De plus, cela explique son utilisation dans les systèmes intelligents (p. ex., *Elliott, 1993; Reilly, 1997*).

Dans ce qui précède, nous avons donné un aperçu quant à l'incorporation des émotions dans les recherches en intelligence artificielle, plus précisément du point de vue de deux sous disciplines : les interfaces homme-machine (avec la reconnaissance automatique de l'émotion) et les systèmes intelligents (à l'aide des modèles computationnels d'émotion). Les travaux en reconnaissance automatique de l'émotion sont autant utiles pour les IHM que pour les systèmes computationnels intelligents. Ils peuvent, par exemple, être intégrés à ces derniers (c.-à-d., les robots, agents intelligents) en tant que composante permettant de percevoir l'environnement, en l'occurrence, l'émotion d'un humain. De plus, ils peuvent aider à comprendre le fonctionnement du modèle émotionnel humain, en appliquant et en testant les modèles émotionnels théoriques pour détecter et découvrir, au moyen des différents algorithmes d'apprentissage, de nouveaux patrons. Ces derniers peuvent mieux expliquer le processus de génération de l'émotion, ses stimuli et sa manifestation, et améliorer, par conséquent, les modèles artificiels et théoriques.

Néanmoins, malgré les progrès réalisés et les différents travaux qui se sont intéressés à la reconnaissance automatique de l'émotion, la performance des machines reste loin d'égaliser celle des humains dans l'induction de l'émotion d'autrui. D'abord, cela est dû au fait que la plupart des recherches se sont concentrées sur la détection « unimodale », à partir d'une seule source (visage-seulement, voix-seulement, etc.), et que les émotions étudiées étaient délibérément ou exagérément affichées (*Zeng, Pantic, Roisman, & Huang, 2008*). Une autre raison importante est le fait de ne pas tenir compte des contextes ou des différentes situations dans lesquelles l'émotion est éprouvée. Par exemple, un système de reconnaissance d'émotions peut détecter un visage « brut » en pleurs, mais ne peut pas affirmer l'émotion derrière cette expression, car cela pourrait être celle d'une tristesse ou bien d'une joie selon la situation à laquelle l'individu est confronté. Par ailleurs, des psychologues (*Carroll & Russell, 1996; Ekman, Friesen, & Ellsworth, 1972*) affirment l'importance d'avoir une information sur la situation ou le contexte avant de déduire l'émotion d'une personne à partir de son comportement, ses gestes, sa voix ou l'expression de son visage. En effet, cette information aide à lever l'ambiguïté en donnant une idée plus claire sur l'émotion probable. À titre d'exemple, si on aperçoit un conducteur roulant à

100 km/h en plein centre-ville, il peut être difficile de cerner l'émotion derrière ce comportement. À priori, on pourrait croire qu'il est content d'avoir une nouvelle voiture, ou bien qu'il est en train de chercher un peu d'excitation, ou encore qu'il est en colère à cause de quelque chose, ou bien tout simplement qu'il n'y a pas d'émotion derrière et que c'est sa façon de conduire. Mais, si on sait qu'un événement indésirable pour ce conducteur vient juste de se produire avant qu'il adopte ce comportement, par exemple qu'il vient d'être démis de ses fonctions, on pourrait alors avoir une idée un peu plus claire de sa probable émotion. En effet, sachant que la plupart des gens virés ont tendance à éprouver une émotion négative, on penchera plutôt vers la colère ou la tristesse et on rejettera les émotions de joie ou d'excitation. Dans la déduction de l'émotion du conducteur, on s'est basé sur deux informations. La première est la situation ou l'événement indésirable qu'a vécu le conducteur. La deuxième est une information implicite (dérivée de la première) sur la réaction probable de la plupart des gens face une situation similaire. Cette dernière est une sorte de connaissance préexistante (antérieure) ou un sens commun des réactions émotionnelles typiques des individus face aux situations. Une information supplémentaire relativement à la réaction émotionnelle de l'individu face à des situations similaires passées aurait été également pertinente (p.ex., sa réaction face à un refus d'admission dans une université). Ce genre d'interprétation du comportement émotionnel des individus, qui semble évident pour les humains, et qui est basé sur la situation génératrice de l'émotion et d'autres connaissances comme l'historique du comportement émotionnelle d'un individu et les réactions typiques ou générales, est souvent négligé dans les travaux sur la reconnaissance de l'émotion (*Speer et al.*, 2009; *Trabelsi & Frasson*, 2010).

Dans le présent mémoire, nous adoptons ce principe d'agrégation des sources d'informations différentes pour améliorer le processus de la reconnaissance automatique de l'émotion. L'objectif principal consiste à concevoir et expérimenter une application qui se base sur cet aspect contextuel ou situationnel. Cette application permet de prédire l'émotion probable d'un individu, ainsi que son intensité face à une situation S_p donnée. Pour générer une telle prédiction, l'application combine deux types d'informations. La première est relative à l'historique des réactions émotionnelles de l'individu face à des situations similaires rencontrées. La deuxième concerne les réactions émotionnelles d'autres individus face à la situation S_p elle-même ou à des situations similaires. Cette dernière information

représente une sorte de sens commun relatif aux réactions émotionnelles typiques (*c'est une sorte de connaissance partagée, ou de patrimoine émotionnel collectif*).

Notre application est utile dans la mesure où elle peut être intégrée dans un système plus complet et global de reconnaissance d'émotions tels que les systèmes tuteurs intelligents (p. ex., *Conati & Maclaren, 2004*) ou les systèmes installés dans des véhicules (p. ex., *Lisetti & Nasoz, 2005*). Pour reconnaître l'émotion, ces systèmes se basent, entre autres, sur les traits du visage, le son, les gestes ou les signaux physiologiques (*Picard, 2000*). Dans ce schéma global, notre application est perçue comme une composante qui fournit une information prédictive complémentaire à ces systèmes et contribue par conséquent, à l'amélioration de leurs performances et leurs fonctionnements. La structure d'un système global de reconnaissance suppose également l'existence d'un historique réel retraçant les réactions émotionnelles spécifiques à un individu. Cet historique serait exploité par notre application pour générer le premier type de prédiction à combiner. Les éléments constitutifs de l'historique émotionnel peuvent être procurés par les méthodes de reconnaissances faciales, sonores ou physiologiques. En cas d'absence d'une telle information, une solution consiste à la créer ou la simuler de manière interactive et collaborative. C'est l'option qui a été retenue dans ce mémoire sachant que l'objectif d'une telle solution (procédure) est d'illustrer les différentes étapes du bon fonctionnement de l'application (voir chapitre 4). Aussi, assembler ces informations individuelles pourrait enrichir l'information sur les réactions typiques ou générales des individus. Quelques exemples simples d'intégration de notre application dans ce genre de système de reconnaissances sont présentés au chapitre 2. Pour la suite de ce travail, le terme « *machine émotionnelle* » sera désormais employé pour désigner l'application dont on vient de présenter les premières prémisses.

Ultimement, notre principal objectif est de rendre cette machine capable de prédire de façon convenable (et non parfaite) l'émotion d'une personne et son intensité, pour une très grande variété de situations émotionnelles de la vie courante, sans se limiter à un contexte ou à un domaine particulier (p.ex., contexte professionnel, contexte familial, contexte social, etc.). Il est souhaitable de revenir sur la notion de prédictibilité qui sera évoquée tout le long du mémoire, et sur le potentiel prédictif de la machine que nous proposons. Il est évident qu'une personne n'est pas capable de prédire ses émotions de

façon parfaite et encore moins celles des autres dans toutes les situations. Goleman (1995) parle de conscience de soi et de conscience sociale. Le premier concept est défini comme étant la capacité à comprendre ses émotions, à reconnaître leur influence et à les utiliser pour guider nos décisions. Le second concept est relatif à la capacité à détecter et à comprendre les émotions d'autrui et à y réagir. Cette capacité varie, donc, d'un individu à un autre. Goleman et al. (2002) poussent les choses un peu plus loin et parlent, même, de compétences émotionnelles. Tout cela c'est pour dire que le potentiel prédictif des personnes et encore plus de notre machine est limité et que toute prédiction, quelle que soit sa nature (humaine ou non) est sujette à l'erreur et à plusieurs types de biais. Mais cela ne doit pas nous empêcher de consentir plus d'effort pour améliorer la capacité prédictive de nos modèles. Notre travail s'insère dans le cadre de cet effort.

Ainsi, tout au long du mémoire, nous tenterons de répondre essentiellement à trois questions :

1. Comment une machine peut-elle « comprendre » ce qu'est une situation émotionnelle ? En d'autres termes, comment pouvons-nous représenter l'information portant sur une situation émotionnelle au sein d'une machine ?
2. Comment pouvons-nous doter la machine d'une connaissance sur les réactions émotionnelles spécifiques (d'une personne) et générales (de la majorité des gens) ?
3. Comment cette machines peut-elle, comme les humains, traiter et utiliser ces informations de manière efficace et produire avec un degré de précision acceptable une prédiction de l'émotion élicitée par une situation S_p ainsi que son intensité.

Afin d'être en mesure de répondre à ces questions, nous avons jugé utile, d'abord, de situer de manière précise et explicite le cadre théorique et conceptuel nécessaire au développement de la machine, ensuite de mener une étude expérimentale pour illustrer les différentes étapes de construction et d'évaluation de la machine. Au travers de cette expérience, nous souhaitons réaliser trois objectifs.

1. Le premier consiste à caractériser et définir formellement (c.-à-d., avec un langage compréhensible par la machine) une situation émotionnelle. Pour accomplir cette tâche, notre choix s'est porté sur un modèle théorique d'évaluation cognitive des émotions que nous avons complété par une spécification appropriée des aspects sociodémographiques

et psychologiques. Cela nous permet, également, de tester empiriquement la pertinence du modèle retenu.

2. Le deuxième consiste à collecter en parallèle un ensemble d'observations (réelles) sur les comportements émotionnels. À cette fin, un système de recueil des réactions de plusieurs personnes est développé. Ces données servent à simuler l'information spécifique, relative à l'historique des réactions d'une personne particulière, et l'information générale, relative au sens commun.
3. Le troisième objectif consiste à doter la machine d'un pouvoir prédictif acceptable en intégrant des algorithmes d'apprentissage appropriés capables d'exploiter, de façon optimale, les données empiriques recueillies. Nous souhaitons également que cette prédiction soit instantanée et facile à mettre à jour de façon rapide. La performance prédictive de la machine est évaluée à l'aide d'une partie des données collectées.

Le plan de ce mémoire est comme suit.

- Le premier chapitre présente une vue d'ensemble sur les recherches menées en intelligence artificielle qui se sont intéressées aux émotions. Principalement, il décrit les différents travaux de reconnaissance automatique de l'émotion dans les IHM ainsi que les recherches sur les modèles computationnels de l'émotion et leur fondement théorique.
- Le deuxième chapitre offre un survol des algorithmes d'apprentissage machine en général et détaille, en particulier, les algorithmes utilisés pour la construction de la *machine émotionnelle*. Pour rendre la lecture plus accessible, un effort de vulgarisation des concepts a été apporté par le biais de nombreux exemples. Une grande partie des détails techniques a été léguée aux annexes.
- Le troisième chapitre détaille toute la procédure empirique mise en œuvre et nécessaire pour la configuration et le fonctionnement de la machine.
- Le quatrième chapitre décrit l'architecture de la machine et l'utilisation des algorithmes d'apprentissage. Puis, il présente l'expérience élaborée pour son évaluation. Enfin, il résume les résultats et les performances de la machine quant à la reconnaissance d'émotions.

- La conclusion éclaire le cheminement général du mémoire, les principaux développements proposés, les résultats obtenus, l'apport de la recherche ainsi que ses limites et perspectives.

Chapitre 1 Utilisation des émotions en informatique

Ce chapitre passe en revue les différents travaux en psychologie et en intelligence artificielle qui ont traité de l'émotion, en mettant l'accent sur les techniques et les théories les plus utilisées dans la reconnaissance automatique et la génération artificielle de l'émotion. Mais, en premier lieu, il convient de clarifier ce qu'est une émotion. Quels sont ses effets sur l'être humain et comment peut-elle l'influencer ?

1.1 Qu'est-ce qu'une émotion

Il n'y a pas de consensus sur la définition du terme émotion, même si dans la littérature il y a eu plusieurs tentatives pour cerner le concept (*Kleinginna & Kleinginna, 1981*). Cependant, une description répandue spécifie que les émotions sont des réactions synchronisées d'une personne, impliquant ses ressources mentales et somatiques, face à un événement stimulus qui est perçu ou évalué comme étant pertinent pour ses besoins, ses objectifs et ses valeurs (*Cohn, 2007; Scherer, 1999, 2004*). En effet, la plupart des recherches en psychologie considèrent que les émotions incluent ou intègrent un processus mental ou cognitif (c.-à-d., évaluation ou perception subjective du stimulus), un processus physiologique ou somatique (accélération des battements de cœur) et des tendances et réponses comportementales (expression corporelle, faciale et vocale) (*Gratch, Marsella, & Petta, 2009*). En outre, plusieurs travaux présentent l'émotion comme un phénomène momentané qui ne dure pas dans le temps, contrairement à l'humeur par exemple, qui elle se prolonge sur une certaine période (*Brave & Nass, 2003*). De plus, plusieurs auteurs affirment que l'émotion est une réaction aux stimuli, souvent associée à une situation, un événement ou une entité. Par conséquent, ils la décrivent comme étant intentionnelle. Quelques recherches désignent l'émotion d'une personne par les termes « état émotionnel » (même si le terme état peut sous-entendre une durée dans le temps et englober par conséquent l'humeur), ainsi que « réaction ou réponse émotionnelle », ou même « état affectif » qui est souvent utilisé dans un contexte clinique. Dans le présent mémoire, nous utilisons ces différents termes pour désigner le même concept à savoir, l'émotion.

Plusieurs psychologues affirment qu'une personne ne peut avoir une pensée ou réaliser une tâche sans engager consciemment ou inconsciemment une émotion (*Picard,*

2000). En effet, ils affirment que les émotions sont impliquées et ont des effets sur le comportement (*LeDoux*, 1998) et plusieurs fonctions cognitives comme l'attention, la mémorisation, ou la prise de décision, etc. (*Brave & Nass*, 2003). Par exemple, un objet ou une situation perçue par un individu comme importante pour ses besoins et ses objectifs peut engendrer une émotion d'excitation qui augmente son niveau de concentration ou d'attention (*Clore & Gasper*, 2000). De même, une émotion de dégoût quant à la situation peut engendrer une distraction. D'autre part, les événements engendrant des émotions sont mieux mémorisés que d'autres événements routiniers (*Thorson & Friestad*, 1989). De plus, concernant la mémorisation, des recherches soutiennent que les personnes se souviennent beaucoup plus des événements avec une très forte émotion négative désagréable que des événements positifs (p. ex., *Newhagen & Reeves*, 1991; *Newhagen & Reeves*, 1992; *Reeves, Newhagen, Maibach, Basil, & Kurz*, 1991). D'un autre côté, les travaux de quelques neurologues et psychologues comme Damasio (1994) ou Isen (2001) ont démontré l'implication et la nécessité des émotions dans la prise de décision et la résolution de problème.

Les émotions sont ainsi très présentes dans la vie de l'homme. Par conséquent, elles ont fait l'objet de plusieurs études psychologiques. Dans la littérature, il existe essentiellement trois théories qui décrivent la manifestation des émotions et spécifient le processus de génération (physiologique ou cognitif), soit les théories de l'émotion discrète, les théories de l'émotion dimensionnelle et les théories de l'évaluation cognitive. Ces théories sont présentées à la section qui suit.

1.2 Les théories de l'émotion

1.2.1 Les théories de l'émotion discrète

Ces théories soutiennent que les émotions représentent un ensemble fini. Elles sont appelées émotions de base et sont universelles, partagées par tous les humains. Les premiers travaux étaient l'oeuvre de Tomkins (1962), lequel a proposé neuf émotions. Ensuite, Ekman et al. (1972) ont présenté leurs six émotions de base, chacune correspondant à une expression faciale : tristesse, joie, colère, peur, dégoût et surprise. Ces émotions ont été très utilisées dans la littérature. La plupart des théories discrètes (*Ekman, Friesen, & Ellsworth*, 1972; *Izard*, 1971; *Oatley & Johnson-Laird*, 1987; *Tomkins*, 1962)

déterminent le nombre d'émotions par rapport aux différents patterns de l'expression faciale. Cela a donné lieu à différentes listes d'émotions de base en type et en nombre selon la théorie discrète (*Fragopanagos & Taylor, 2005*).

Concernant le processus de génération, ces théories supposent que ces émotions de base ne requièrent pas un processus cognitif complexe et qu'elles surviennent principalement suite au déclenchement d'un processus détecteur-moteur ou neuro-moteur inné ou préprogrammé (*Gratch, Marsella, & Petta, 2009*). Par exemple, le fait d'entendre une grande détonation peut déclencher à priori inconsciemment (involontairement) et d'une manière systématique une émotion de peur.

Du fait de la diversité des listes des émotions présentées par chaque théorie, en nombre et en type, d'autres recherches ont proposé des théories qui vont dans un sens totalement opposé aux thèses présentées par les théories discrètes. Ce sont les théories de l'émotion dimensionnelle.

1.2.2 Les théories de l'émotion dimensionnelle

Les théories de l'émotion dimensionnelle stipulent que les émotions ne sont pas discrètes et qu'elles ne peuvent pas être catégorisées. Elles soutiennent également que celles qui sont communément appelées émotions (p. ex., joie, tristesse, etc.) ne sont que des labels ou étiquettes qu'on utilise pour mieux comprendre le phénomène. Selon ces théories, les émotions correspondent plutôt à des phénomènes continus issus d'activations ou d'excitations physiologiques (p. ex., accroissement des battements du cœur, troubles respiratoires, etc.) (*Lang, 1995*) et peuvent être représentées à l'aide de points dans un repère à dimensions bipolaires. La plupart des théories dimensionnelles (p. ex., *Carver, 2001; Lang, 1995; Russell, 1980, 2003*) ont caractérisé les émotions à l'aide deux dimensions. La première est la valence, avec ses deux pôles opposés reflétant le fait que l'évènement déclencheur de l'émotion est perçu comme positif et bénéfique ou bien négatif et nuisible. La deuxième dimension correspond à l'excitation, appelée aussi activation, et traduit l'intensité de l'excitation physiologique qui révèle la prédisposition d'agir ou d'entreprendre une action suite à l'évènement générateur de l'émotion. Cette activation ou excitation varie entre deux valeurs extrêmes « élevée » et « basse » (faible). *Lang (1995)* soutient que ces deux dimensions (valence et excitation) sont presque suffisantes pour

balayer tout l'espace des émotions possibles déjà introduites dans la littérature. Ainsi, une transformation pourrait se faire depuis les émotions de base vers l'espace valence-excitation. Par exemple, quelqu'un qui sent une excitation élevée et qui perçoit sa situation comme étant nuisible peut avoir l'émotion communément connue sous le terme « colère » (*Gratch, Marsella, & Petta, 2009*). En outre, il y a eu d'autres théories qui ont rajouté une troisième dimension appelée dominance (p. ex., *Mehrabian & Russell, 1980*). Cette dernière indique à quel point la personne a un contrôle et un pouvoir d'action sur la situation, ou bien à quel point elle est soumise et impuissante.

Les théories dimensionnelles ont fait l'objet de plusieurs études dans la littérature, encouragées par le cadre théorique simple utilisé pour la caractérisation de l'émotion. Cependant, des psychologues (*Lazarus, 1991; Ortony, Clore, & Collins, 1988*) soutiennent que cette caractérisation, selon seulement deux ou trois dimensions, peut engendrer une perte d'informations du fait de la richesse, la complexité et le caractère multidimensionnel de l'émotion.

1.2.3 Les théories d'évaluation cognitive

Introduites par Arnold (1960) et Lazarus (1968), les théories d'évaluation cognitive ont été, et sont encore, l'objet d'un grand nombre de recherches au sujet de l'émotion. Elles soutiennent que la réaction émotionnelle, avec ce qu'elle implique comme réactions corporelles ou physiologiques, est le résultat d'un processus cognitif correspondant à la perception ou l'évaluation faite par l'individu de l'impact que pourrait avoir son environnement et des situations qu'il vit. En fait, cela correspond à l'influence des comportements d'autrui, des événements qui surviennent et des objets qui l'entourent sur son bien-être, ses objectifs et ses valeurs (*Ellsworth, 1994; Roseman & Smith, 2001; Sander, Grandjean, & Scherer, 2005*).

Les théories d'évaluation (p. ex., *Frijda, 1987; Lazarus, 1991; Ortony, Clore, & Collins, 1988; Roseman, Antonio, & Jose, 1996; Scherer, 2001; Smith & Kirby, 2000*) affirment que le processus d'évaluation cognitive des situations génératrices de l'émotion peut être caractérisé par un ensemble de dimensions ou critères. Ces dimensions expliquent non seulement le type de la réaction émotionnelle (joie, tristesse, colère, etc.) mais aussi son intensité. Elles sont également appelées variables d'évaluation. Des exemples de

dimensions présentées dans la plupart des théories sont le degré d'influence de la situation sur le bien-être de l'individu, la conformité ou la non-conformité avec ses objectifs, le degré d'anticipation de l'occurrence de cette situation, la capacité de s'y adapter, ou encore, l'agent (autrui ou soi-même) qui en est responsable, etc. (*Gratch, Marsella, & Petta, 2009; Siemer, Mauss, & Gross, 2007*). Par exemple, un événement extrêmement bénéfique qui survient et auquel on ne s'attend pas peut engendrer une émotion forte de joie (gagner au loto). De même, un événement possiblement bénéfique (non pas extrêmement) et auquel on s'attend peut également générer une émotion de joie, mais pas aussi intense que la première (recevoir son salaire la fin du mois). Les différences entre les théories d'évaluation résident dans la spécification de quelques variables d'évaluation et dans la description de leur implication dans le processus d'évaluation cognitive et de leur corrélation avec l'émotion générée (*Gratch, Marsella, & Petta, 2009*).

Ainsi, contrairement aux autres théories (discrètes et dimensionnelles), les théories d'évaluation fournissent un cadre théorique plus riche qui permet de mieux cerner le caractère multidimensionnel de l'émotion et proposent des liens et des relations détaillées entre le processus cognitif de génération de l'émotion et la réaction émotionnelle produite (*Sander, Grandjean, & Scherer, 2005*). Du fait de cette richesse, elles ont été très utilisées dans le domaine de l'intelligence artificielle, principalement dans les travaux sur les modèles computationnels d'émotions qui essaient de simuler le processus de génération de l'émotion (p. ex., *Frijda & Swagerman, 1987; Gratch & Marsella, 2004; Scherer, 1993; Staller & Petta, 2001*) (voir section 1.4). La plupart des chercheurs s'intéressant aux modèles computationnels d'émotions s'accordent pour dire que la théorie d'évaluation d'Ortony, Clore et Collins (1988), plus connue sous l'abréviation « modèle OCC », reste l'une des plus utilisées (*André, Klesen, Gebhard, Allen, & Rist, 2000; Bates, Loyall, & Reilly, 1992; Elliott, 1993*). Cette théorie sera détaillée ultérieurement dans le chapitre 3.

Une autre utilisation intéressante des théories d'évaluation est la reconnaissance automatique de l'émotion, dans un contexte d'interaction homme-machine (IHM), où la machine serait capable d'inférer l'évaluation cognitive de l'individu (*Sander, Grandjean, & Scherer, 2005*). L'application des théories d'évaluation dans l'intelligence artificielle ne s'est pas restreinte à la simulation du processus émotionnel. Quelques travaux visant à reconnaître l'émotion d'un individu dans un contexte d'interaction homme-machine ont

également utilisé ces théories (*Elliott & Brzezinski, 1998*). Néanmoins, les recherches existantes en reconnaissance automatique sont davantage basées sur les informations visibles et mesurables, comme les expressions et mouvements du corps ou les signaux physiologiques, que sur l'aspect cognitif de génération de l'émotion.

Dans ce qui suit, nous essaierons de présenter les travaux qui ont tenté d'intégrer les émotions au niveau des interactions homme-machine au moyen de la reconnaissance automatique de l'émotion. Ensuite, nous présenterons quelques modèles computationnels reposant sur les théories d'évaluation.

1.3 Les émotions dans les interactions homme-machine

Dans l'introduction de ce mémoire, nous avons souligné l'omniprésence des ordinateurs et des machines dans la vie quotidienne de l'humain. Nous avons également expliqué la nécessité d'améliorer la qualité des interactions entre l'humain et la machine en rendant cette dernière capable de reconnaître les états émotionnels que peut éprouver l'utilisateur. Cette problématique est abordée par les recherches en informatique affective.

1.3.1 L'informatique affective

L'informatique affective introduite par Picard (2000) vise à doter les ordinateurs de la capacité de reconnaître et d'interpréter les émotions et les comportements affectifs d'une personne afin de fournir des réponses plus intelligentes, d'initier une conversation, d'agir sur l'état de la personne et d'enrichir par conséquent l'interaction en la rendant plus « humaine » (*Tao & Tan, 2005*). Ainsi, l'informatique affective peut être considérée comme une sous-discipline de l'interaction homme-machine.

La reconnaissance de l'émotion se fait à partir de plusieurs sources (p. ex., intonation de la voix, expression du visage, gestes du corps, signaux physiologiques, etc.). Généralement, la reconnaissance à partir de ces sources se fait à l'aide d'algorithmes d'apprentissage machine. Les recherches en informatique affective ont connu une expansion considérable et ont été appliquées à une variété de domaines, comme les jeux vidéo, les systèmes tuteurs intelligents, les centres d'appel, etc. Ces différents aspects sont présentés ci-dessous.

1.3.2 Sources et outils de reconnaissance d'émotion

(A) Reconnaissance faciale

La majorité des travaux en reconnaissance automatique de l'émotion utilisent les expressions du visage comme source d'information. En effet, le travail d'Ekman et al. (1972), en particulier, a contribué au développement de ce domaine en proposant un système de mesure ou de codification des expressions faciales FACS (Face Action Coding System). Ils définissent 27 mouvements élémentaires et indépendants du visage appelés unités d'action (Action Units AUs). La combinaison de ces « AU » permet de décrire une variété d'expressions faciales. Une correspondance entre chaque émotion de base parmi les six (Ekman, Friesen, & Ellsworth, 1972) et les mouvements « AU » qui lui sont pertinents a été alors établie au moyen du système EMFACS¹ (Emotional FACS).

Généralement, l'extraction (le décodage) des traits du visage qui dénotent une expression faciale émotionnelle peut se faire selon deux approches. La première approche se concentre sur les formes géométriques de quelques composantes du visage (p. ex., la bouche, l'œil, etc.) et l'emplacement des points saillants du visage (p. ex., les coins des yeux, sourcils, bouche, etc.) (Zeng, Pantic, Roisman, & Huang, 2008). Par exemple, des sourcils détendus avec une bouche large ayant les coins tirés vers les oreilles peuvent témoigner d'une émotion de joie (Parke & Waters, 2008). Dans la deuxième approche, on se préoccupe plutôt de l'apparence globale du visage en incluant les rides, les sillons du front, etc.

Plusieurs travaux se sont intéressés à la première approche. Nous pouvons citer ceux de Chang et al. (2006) ou Pantic et Rothkrantz (2004) qui ont utilisé plusieurs points de repère autour des yeux, du nez, de la bouche, etc. pour extraire les formes. D'autres chercheurs, comme Bartlett et al. (2006) ou Valstar et al. (2004) se sont plutôt intéressés à la deuxième approche basée sur l'apparence globale. Une combinaison des deux approches a aussi fait l'objet de quelques travaux (p. ex., Tian, Kanade, & Cohn, 2005).

Concernant les outils et les techniques utilisés pour la reconnaissance faciale des émotions, il existe des caméras et des oculomètres (appareils destinés à suivre le mouvement des yeux) utilisés pour enregistrer les différentes expressions émotionnelles du

¹ <http://face-and-emotion.com/dataface/general/homepage.jsp>

visage. La reconnaissance des émotions, à partir des traits extraits, est réalisée à l'aide des algorithmes de classification. Par exemple, Bartlett et al. (2006) ont utilisé une variante de l'algorithme des machines à vecteurs supports (SVM), Valster et al. (2004) se sont basés sur l'algorithme des k plus proches voisins (k NN), et Pantic et al. (2004) ont préféré utiliser un algorithme à base de règles.

(B) Reconnaissance vocale

Dans l'interaction de tous les jours entre les humains, la voix et les paroles représentent l'un des modes ou l'une des sources les plus fiables de reconnaissance de l'émotion. Par exemple, une voix faible et cassée accompagnée d'un discours « peu enthousiaste » peut traduire une émotion de tristesse, de même qu'une voix intense et forte peut traduire une émotion de colère. Dans la reconnaissance automatique des émotions à partir des signaux sonores de la parole, les chercheurs se réfèrent aux caractéristiques prosodiques comme l'intonation, le rythme, le ton, l'accent, l'énergie dégagée, l'élocution, etc. (p. ex., *Cowie et al.*, 2001; *Juslin & Scherer*, 2005; *Vidrascu & Devillers*, 2005). La grande majorité des travaux s'intéresse à l'extraction de ces caractéristiques principalement pour reconnaître les six émotions de base (*Tao & Tan*, 2005). Mais, il existe de nouvelles méthodes, moins répandues, visant à détecter d'autres émotions comme la frustration, l'anxiété ou l'ennui en essayant d'extraire le son des rires, des cris, des soupirs ou du bâillement (*Russell, Bachorowski, & Fernandez-Dols*, 2003). D'autres recherches ont tenté de rajouter les caractéristiques linguistiques et sémantiques de la parole (les mots ou la répétition d'une expression qui peuvent indiquer l'expérience d'une certaine émotion) aux caractéristiques prosodiques afin d'améliorer les performances de la détection de l'émotion (p. ex., *Lee & Narayanan*, 2005; *Litman & Forbes-Riley*, 2004). Cependant, ces tentatives restent peu nombreuses, du fait de la difficulté d'extraire les caractéristiques linguistiques comme les expressions émotionnelles à partir des paroles et du fait de la différence de ces expressions d'un langage à un autre.

Les données ou signaux sonores émotionnels peuvent être recueillis en demandant à des acteurs de parler en simulant une certaine émotion (p. ex., *Kwon, Chan, Hao, & Lee*, 2003), ou bien en collectant des conversations enregistrées dans un cadre susceptible d'engendrer des émotions, comme les sondages effectués par les centres d'appel (p. ex., *Lee & Narayanan*, 2005; *Vidrascu & Devillers*, 2005), les interviews (p. ex., *Hirschberg et*

al., 2005) ou les meetings (p. ex., *Neiberg, Elenius, & Laskowski*, 2006). Sur ces signaux collectés, plusieurs techniques d'extraction et de traitement de signal sont appliquées, accompagnées des techniques d'apprentissage pour la classification des émotions. Ainsi, parmi les travaux que nous avons cités, Litman et Forbes-Reily (2004) ont utilisé l'algorithme des arbres de décision C4.5 (*Quinlan*, 1993), alors que Vidrascu et Devillers (2005) se sont basés sur les SVM.

(C) Reconnaissance à partir des signaux physiologiques

Comme indiqué à la section 1.2.2, les théories dimensionnelles affirment que les émotions reflètent des manifestations physiologiques, comme l'accélération du rythme cardiaque ou les troubles respiratoires, etc. Ainsi, quelques travaux se sont intéressés à mesurer ces manifestations sous forme de signaux, et ensuite à les corrélérer avec les deux dimensions de valence et d'excitation (*Wagner, Kim, & André*, 2005). Les principaux signaux traités sont :

- l'électromyographie (EMG) qui mesure l'activité électrique au niveau des muscles et qui peut être corrélée à la valence (*Lang*, 1995);
- la réponse électromodale de la peau (Galvanic Skin Response) qui est utilisée pour mesurer la conductance de la peau et qui renseigne sur l'excitation;
- le rythme cardiaque qui est selon Lang (1995) corrélé à la valence;
- la fréquence de la respiration qui peut renseigner sur l'excitation (*Haag, Goronzy, Schaich, & Williams*, 2004).

D'autres signaux peuvent être mesurés et combinés, comme la pression artérielle systolique et diastolique, la dilatation pupillaire, la température, la pression du doigt, etc. (*Picard*, 2000).

Les travaux qui se sont intéressés à ce mode physiologique de détection n'ont pas été aussi nombreux que ceux pour la détection faciale ou vocale. Cela est essentiellement dû au fait que les appareils utilisés sont, pour la plupart, intrusifs (p. ex., capteurs conductance de la peau, capteurs pour la pression artérielle, thermomètre, etc.), ce qui rend leur manipulation plus complexe et la mise en place d'un environnement pour une expérimentation plus difficile (*Leon, Clarke, Callaghan, & Sepulveda*, 2007; *Sebe, Cohen, Gevers, & Huang*, 2005). Néanmoins, quelques tentatives ont été effectuées, comme le

travail de Picard et al. (2001) qui consistait à utiliser quelques signaux pour prédire ou reconnaître, à l'aide d'algorithmes statistiques avancés (p. ex., projection de Fischer), l'émotion d'un individu après lui avoir présenté des images pouvant susciter son émotion. Aussi, Nasoz et al. (2003) ont utilisé des films afin de provoquer l'émotion d'un participant avant de la mesurer. Ils se sont basés sur l'analyse factorielle discriminante et l'algorithme des k plus proches voisins pour la reconnaissance. Deux autres travaux sont souvent cités dans la littérature. Le premier est celui de Healey (2000) qui a essayé de reconnaître l'émotion à partir de signaux physiologiques provenant de systèmes de mesures portables (wearable systems). Ces systèmes ont été testés sur des personnes effectuant des tâches de la vie quotidienne. Healey a utilisé, entre autres, des fonctions linéaires discriminantes pour la reconnaissance. Le deuxième est celui de Haag et al. (2004) qui ont utilisé plusieurs signaux. Dans ce travail, les émotions ont été provoquées à l'aide de plusieurs photos présentées aux personnes participantes. Pour la reconnaissance, Haag et ses collaborateurs ont utilisé l'algorithme des réseaux de neurones artificiels.

(D) Autres modes de reconnaissance de l'émotion

La majorité des travaux en reconnaissance automatique de l'émotion utilise comme seule source de reconnaissance l'expression faciale ou la parole, ou dans une moindre mesure, les signaux physiologiques. Récemment, il y a eu de plus en plus d'intérêt pour la reconnaissance à partir d'autres sources, comme les mouvements et gestes du corps, les ondes cérébrales, ou même la reconnaissance multimodale qui combine plusieurs sources (p. ex., faciales, physiologiques et vocales, etc.).

La posture du corps, les mouvements et les gestes des mains peuvent renseigner sur certains états émotionnels. Brandir les mains en haut en sautant peut révéler une émotion de joie intense. La plupart des efforts liés à ce type de reconnaissance se sont concentrés sur le mouvement des mains (*Aggarwal & Cai, 1999; Pavlovic, Sharma, & Huang, 1997*) plutôt que sur le corps dont les postures restent difficiles à interpréter, surtout dans le cadre d'applications en temps réel qui nécessitent une précision extrême dans la capture des mouvements (*Tao & Tan, 2005*).

Une autre nouvelle tendance, dans la détection de l'émotion, consiste à mesurer l'activité cérébrale de l'individu à l'aide des signaux produits par l'électro-encéphalographie (EEG). Quelques travaux s'y sont intéressés, à l'instar d'Isinho et

Hagiwara (2003) qui ont tenté de reconnaître 4 émotions en utilisant un réseau de neurones. De même, récemment, Lin et al. (2009) ont appliqué l'algorithme des machines à vecteurs de support sur les signaux EEG recueillis suite à une stimulation musicale des émotions. Ce mode de reconnaissance est prometteur, mais, comme les signaux physiologiques, la difficulté de l'introduire dans des applications à temps réel et la complexité de sa manipulation restent des inconvénients majeurs.

La reconnaissance multimodale est celle qui correspond le mieux à la reconnaissance humaine de l'émotion en se basant sur plusieurs sources simultanément. Les efforts dans ce domaine ont eu du mal à commencer, à cause de la difficulté de synchronisation des différentes sources (*Russell, Bachorowski, & Fernandez-Dols, 2003; Zeng, Pantic, Roisman, & Huang, 2008*). Les combinaisons des sources prises en compte sont différentes d'un travail à un autre. Ainsi, il y a eu des travaux qui ont combiné l'expression faciale et les mouvements ou la posture du corps (p. ex., *Balomenos et al., 2005; Gunes & Piccardi, 2006*). D'autres, comme Lisetti et Nasoz (2002), ont rajouté la détection des signaux physiologiques et celle des paroles à la détection faciale dans un cadre d'un agent intelligent qui reflète l'émotion de l'utilisateur. De même, Kapoor et al. (2007) ont combiné plusieurs signaux physiologiques et l'expression du visage pour la prédiction de la frustration d'un apprenant dans un cadre éducatif.

1.3.3 Applications de la reconnaissance automatique de l'émotion

Les multiples travaux sur la reconnaissance automatique de l'émotion et la diversité des sources de détection employées (p. ex., paroles, gestes, visage, etc.) ont permis son application au niveau de domaines très variés au sein desquels l'émotion joue un rôle important. Par exemple, elle a été intégrée aux systèmes tuteurs intelligents (p. ex., *Conati & Maclaren, 2004; D'Mello, Graesser, & Picard, 2007; Kapoor, Bursleson, & Picard, 2007; Ochs & Frasson, 2004*). Un tuteur pourrait ajuster le contenu de son cours ou bien la vitesse d'exposition de la matière selon que l'étudiant est excité (intéressé) ou bien ennuyé. Dans ce dernier cas de figure, il pourrait même essayer de changer l'état émotionnel de l'étudiant pour le rendre plus attentif grâce à des actions pédagogiques (p. ex., en l'encourageant, en lui remontant le moral). De même, un système automatique de reconnaissance d'émotions peut être incorporé dans les véhicules (p. ex., *Healey, 2000; Lisetti & Nasoz, 2005*) afin de repérer un conducteur en état de colère extrême ou de stress,

et essayer de le reconforter ou de l'alerter sur les conséquences néfastes que cette émotion pourrait avoir sur sa conduite. Similairement, détecter la frustration d'un client, au niveau des centres d'appel ou des services clients automatisés (p. ex., *Lee & Narayanan, 2005; Vidrascu & Devillers, 2005*), causée par l'inadéquation des options fournies, pourrait aider à transférer directement l'appel à un consultant sans devoir passer par les différentes étapes. De même, une analyse des contenus affectifs des boîtes vocales a été réalisée (p. ex., *Inanoglu & Caneel, 2005*). Mais l'utilisation la plus marquée de la reconnaissance automatique de l'émotion reste au niveau des secteurs des jeux vidéo et du divertissement (p. ex., *Healey, Picard, & Dabek, 1998; Slater, Buckley, & Bechkoum, 2006*). Leur objectif est de produire des jeux interactifs ou bien des agents intelligents sociaux, tenant compte de l'état affectif du joueur ou du propriétaire, et tendant à instaurer un environnement d'interactions ressemblant à celui des humains.

1.3.4 Les domaines inexplorés et les limites

Zeng et al. (2008) ont affirmé que, dans la plupart des travaux de reconnaissance faciale et vocale, les émotions sont délibérément exprimées et exagérément affichées, et que la spontanéité est encore difficile à détecter automatiquement même s'il existe quelques tentatives qui essayent d'y parvenir (p. ex., *Littlewort, Bartlett, & Lee, 2007; Valstar, Pantic, Ambadar, & Cohn, 2006*). De plus, on s'est essentiellement intéressé à la reconnaissance des six émotions de base (*Ekman, Friesen, & Ellsworth, 1972*), du fait de leur caractère universel. Cependant, cette liste d'émotions ne peut être exhaustive et englober les possibles émotions que peut éprouver un individu dans sa vie de tous les jours. D'un autre côté, les efforts produits se sont focalisés sur la reconnaissance uni-modale, soit faciale ou vocale. On assiste à de plus en plus de travaux en reconnaissance multimodale, combinant différentes sources (voir section 1.3.2). Ceux-ci demeurent néanmoins minoritaires. Un autre domaine, très peu exploré, est la reconnaissance en temps réel de l'émotion. La cause est que la plupart des recherches en reconnaissance automatique de l'émotion utilisent des techniques d'apprentissage machine. Ces techniques nécessitent une phase, en amont, appelée phase d'entraînement sur les données collectées (p. ex., images, sons, signaux, etc.) qui permet de créer le modèle de prédiction ou de reconnaissance (voir chapitre 2). Les recherches n'ont pas essayé d'adapter ces techniques pour une

reconnaissance instantanée, à l'exception de quelques unes (*Bartlett et al.*, 2006; *Nasoz, Lisetti, Alvarez, & Finkelstein*, 2003).

Zeng et al. (2008) ont également soutenu que les travaux de recherche ne tiennent pas compte d'une composante intuitive et essentielle dans la reconnaissance humaine de l'émotion d'autrui, qui est la prise en compte du contexte ou de l'environnement dans lequel cette émotion a été provoquée ou stimulée. Des psychologues (*Carroll & Russell*, 1996; *Ekman, Friesen, & Ellsworth*, 1972) ont même affirmé que l'interprétation des gestes, des paroles et de l'expression du visage d'un individu dépend entre autres de la situation ou des antécédents qui ont induit ce comportement. Un exemple concret est la photo présente dans la Figure 1.1 qui expose le joueur de tennis André Agassi tendant ses mains en avant et pleurant. En effet, dans ce cas particulier, les recherches sur la reconnaissance automatique de l'émotion pourraient détecter que Agassi tend ses mains vers l'avant et pleure, mais ne pourraient pas inférer si l'émotion derrière est une émotion de joie ou une émotion de tristesse. Cependant si la reconnaissance automatique avait tenu compte du fait que cette photo correspondait à la fin du dernier match joué par Agassi sur un court de tennis professionnel après une grande carrière de champion, l'inférence de l'émotion aurait pu être plus facile.

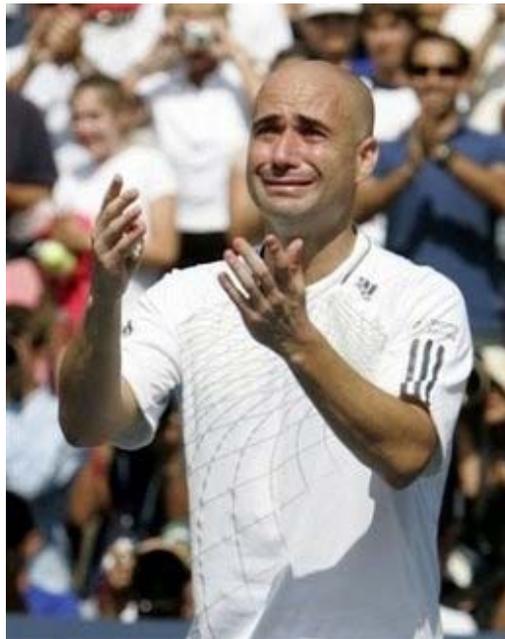


Figure 1.1 Photo d'André Agassi à la fin d'un match

Il existe des travaux qui ont tenu compte du contexte, mais ils se sont limités à un domaine d'application particulier, ou à l'étude d'une émotion particulière au moyen de scénarios d'expérimentation délibérément mis en place et prédéfinis. Par exemple, pour la détection de la fatigue, qui est considérée comme une émotion n'appartenant pas aux émotions de base, Ji et al. (2006) ont modélisé le contexte par quelques variables relatives à la condition du travail, la qualité du sommeil, la condition physique, etc. De même, dans le domaine des systèmes tuteurs intelligents, Chalfoun et al. (2006) ont essayé de prédire l'émotion d'un apprenant selon deux scénarios contextuels, c'est-à-dire selon qu'il ait reçu une note meilleure ou pire que celle à laquelle il s'attendait. Aussi, Kapoor et Picard (2005) ont caractérisé l'environnement, dans le cadre d'un jeu éducatif, par la difficulté de la tâche et le niveau du jeu.

Ainsi, les situations ont toujours été réduites au domaine ou à l'émotion étudiés. Personne ne s'est encore penché sur une modélisation générale qui permettrait à une machine interagissant avec un humain de pouvoir « assimiler », comme le ferait un humain, la situation dans laquelle se trouve son vis-à-vis, et ce, quelque soit le domaine. Un des objectifs de ce mémoire est de trouver une façon de modéliser les différentes situations pouvant engendrer une émotion, et essayer de valider empiriquement cette modélisation.

1.4 L'émotion et les systèmes intelligents

Nous avons expliqué dans l'introduction de ce mémoire que les émotions sont étudiées également dans le cadre des recherches en systèmes computationnels intelligents. Ces systèmes tendent à simuler principalement le raisonnement humain dans des robots ou agents intelligents. Selon plusieurs chercheurs, les émotions jouent un rôle dans ce processus de raisonnement. Goleman (1995), par exemple, soutient que les émotions sont des caractéristiques de l'intelligence humaine. Minsky, aussi, dans son livre « The Emotion Machine » (*Minsky, 2007*), propose une modélisation artificielle du raisonnement humain et soutient que selon son état émotionnel, une personne raisonne d'une certaine manière et que les émotions, par conséquent, représentent des « manières de raisonner » ou de penser. Cependant, même si son livre s'intitule « The Emotion Machine », Minsky se focalise principalement sur le modèle de raisonnement humain, dont les émotions ne constituent qu'un aspect parmi d'autres. D'un autre côté, il existe d'autres travaux qui se sont focalisés

essentiellement sur la simulation de la génération de l'émotion au sein des systèmes computationnels intelligents produisant ce qu'on appelle les systèmes computationnels de l'émotion ou émotionnels. Ces systèmes s'inspirent des théories de l'émotion, principalement des théories d'évaluation cognitive. Slater et al. (2008) affirment qu'il existe plusieurs publications sur les systèmes computationnels de l'émotion, mais que la plupart de ces publications se réfèrent principalement à ou font partie de trois grands projets de recherches dans ce domaine, à savoir, *Cognition and Affect Project (CogAff)* (Sloman & Poli, 1995), *Tok Project* (Bates, Loyall, & Reilly, 1992), et *Émile Project* (Gratch, 2000).

Le CogAff a été réalisé au sein de l'université de Birmingham. Son but global est de modéliser le comportement d'un agent pour les jeux. Une des caractéristiques de l'architecture développée est la capacité de simuler l'émotion. Les auteurs de ce travail ont suggéré une implémentation à l'aide des techniques IA comme la machine à état fini et les réseaux de neurones. Néanmoins, l'aspect émotionnel dans l'architecture n'est pas basé sur une des théories émotionnelles (Slater, Moreton, Buckley, & Bridges, 2008).

Le projet *Tok* a été conçu à l'université de Carnegie Mellon. C'est une architecture agent intégrée à un environnement virtuel, appelé l'environnement *Oz* (Reilly & Bates, 1992). *Tok* contient un module comportemental orienté objectif appelé *HAP* (Loyall & Bates, 1991) et un autre module qui s'occupe des émotions appelé *Em* (Reilly & Bates, 1992). Ce dernier est essentiellement basé sur le modèle cognitif de l'émotion OCC (Ortony, Clore, & Collins, 1988). L'objectif de l'agent est spécifié dans le module HAP. Il peut avoir une émotion selon le niveau de réalisation ou non-réalisation de cet objectif. Dans cette architecture le processus de génération de l'émotion est dynamique, l'émotion change lorsque les désirs ou les objectifs changent au cours du temps. Cependant, cette architecture n'a pas été adaptée à un environnement multi-agent.

Émile est aussi une architecture agent qui se base sur l'architecture du module *Em* (Reilly & Bates, 1992) et sur celle du travail d'Elliot (1992) *The Affective Reasoner*. L'agent est capable non seulement d'évaluer une situation par rapport à ses besoins à l'aide du modèle OCC, mais aussi d'essayer de s'adapter à elle en planifiant et en changeant d'objectifs (Gratch & Marsella, 2003). Ces objectifs peuvent aussi être influencés par la présence et par l'action d'autres agents. Ainsi, *Émile* permet une simulation dans un

contexte où cohabitent plusieurs agents. Cette architecture a été étendue et renommée EMA (Emotion and Adaptation) (*Gratch & Marsella, 2004*).

Il existe d'autres modèles computationnels de l'émotion intégrés à des architectures d'agents intelligents comme *BlueEyes*², *FLAME* (*El-Nasr, Yen, & Ioerger, 2000*), *Poeple and Robot*³, *SAFIRA*⁴, etc.

Les chercheurs s'accordent pour dire que la majorité de ces travaux, comme dans les projets *Tok* et *Émile*, sont basés sur le modèle OCC. Ce dernier est décrit de manière à faciliter sa conversion dans des logiciels agents. Le modèle OCC a été surtout utilisé au sein des modèles computationnels pour simuler la génération de l'émotion de l'agent dans un environnement multi-agent. Un modèle computationnel basé sur OCC pourrait aussi servir à la compréhension du mécanisme derrière la génération de l'émotion d'un individu dans un contexte d'interaction homme-machine (*Trabelsi & Frasson, 2010*). Néanmoins, ce genre d'approche a suscité la curiosité de peu de chercheurs, à part quelques-uns dans des domaines spécifiques. C'est le cas de Conati et Maclaren (2004) qui ont tenté de modéliser l'émotion de l'apprenant en se basant sur OCC, dans un système éducatif particulier. Dans ce mémoire, nous essayons de nous inspirer du modèle OCC dans l'optique de doter la machine d'une capacité de compréhension de l'évaluation cognitive que fait un individu par rapport à sa situation. Cette capacité permettrait alors à la machine de prédire l'émotion. Cette démarche est présentée en détail, accompagnée de la présentation du modèle OCC, au chapitre 3.

1.5 Conclusion

Au cours de ce chapitre, nous avons introduit dans un premier temps la notion d'émotion. Ensuite, nous avons présenté le rôle de l'émotion en intelligence artificielle, plus précisément en reconnaissance automatique de l'émotion et dans les systèmes intelligents.

Pour la reconnaissance automatique, nous avons passé en revue les différents modes (facial, vocal, physiologique, etc.) utilisés et les travaux qui s'y sont intéressés en mentionnant quelques algorithmes d'apprentissage machine utilisés pour la reconnaissance.

² <http://www.almaden.ibm.com/cs/BlueEyes/index.html>

³ <http://www.peopleandrobots.org/>

⁴ <http://www.sics.se/safira>

Puis, nous avons présenté leurs limites qui résident, par exemple, dans le nombre réduit des émotions étudiées (les six émotions de bases), le peu de travaux qui se sont intéressés à la reconnaissance instantanée, et, particulièrement, la négligence de l'environnement et des situations qui provoquent la réaction émotionnelle.

D'un autre côté, nous avons présenté quelques recherches dans le domaine des systèmes intelligents qui avaient pour objectif de créer des systèmes computationnels permettant de simuler la génération de l'émotion, et par conséquent d'interpréter les situations qui provoquaient ces émotions dans un environnement virtuel en se basant sur des théories d'évaluation cognitive, plus précisément le modèle OCC. Cependant, une utilisation de ces modèles, dans un contexte d'interaction homme-machine, pour l'interprétation des situations émotionnelles que peut vivre un individu, n'a pas été étudiée.

Ainsi, dans ce mémoire, nous tentons, dans un contexte d'interaction homme-machine, de trouver un cadre général de modélisation pour une grande variété de situations émotionnelles possibles, en nous basant sur le modèle OCC. Puis, nous visons à utiliser cette modélisation pour la configuration d'une application, la *machine émotionnelle*, capable de prédire l'émotion d'un individu et son intensité pour une situation donnée. Cette machine aura pour but de prédire les 22 émotions du modèle OCC d'une manière instantanée en appliquant les techniques d'apprentissage appropriées. Elle serait utile, une fois intégrée dans des systèmes multimodaux de reconnaissance de l'émotion (voir section 1.3.2), puisqu'elle leur fournirait une information additionnelle relative à l'émotion que pourrait engendrer une situation donnée chez la personne. Ainsi, le chapitre 2 présente les algorithmes d'apprentissage utilisés dans la *machine émotionnelle*. Il présente également quelques exemples simples mettant en valeur l'utilité des prédictions que pourrait réaliser la machine dans des systèmes de reconnaissance d'émotion plus complets.

Chapitre 2 Algorithmes d'apprentissage machine

Le chapitre présente les principes fondamentaux de l'apprentissage machine. Nous nous concentrons sur l'apprentissage supervisé. Nous discutons en détail le *réseau de neurones perceptron multicouche* et l'algorithme des *k plus proches voisins*. Ces deux algorithmes sont utiles pour le développement de notre machine émotionnelle.

2.1 Qu'est-ce que l'apprentissage machine ?

En sciences cognitives, l'apprentissage (humain) est considéré comme étant le processus visant l'amélioration des performances au fur et à mesure de l'exercice d'une activité. Pour apprendre, il faut s'exercer, s'entraîner, acquérir une expérience afin d'être performant dans une activité donnée. L'amélioration de la performance passe aussi par la capacité de l'humain à inférer des enseignements et des règles suite à son vécu. Il peut alors réaliser de nouvelles tâches d'une manière plus efficace.

L'apprentissage machine est né de la volonté des chercheurs en intelligence artificielle de doter la machine avec la capacité d'apprendre et d'améliorer automatiquement sa performance dans l'exécution d'une tâche donnée (*Bishop, 2006; Duda, Hart, & Stork, 2001; Hastie, Tibshirani, & Friedman, 2001; Mitchell, 1997; Vapnik, 2000*). Les algorithmes d'apprentissage machine assurent ce processus. Nous recensons les arbres de décision, les réseaux de neurones, les *k*-moyennes, les machines à vecteurs supports comme des algorithmes parmi d'autres. Cette automatisation de la performance vient en opposition à l'idée d'une machine qui exécute des programmes explicitement définis par l'humain.

Certes, les chercheurs ne sont pas encore arrivés à concevoir des algorithmes permettant aux machines d'avoir une faculté d'apprentissage comparable à celle de l'homme. Cela est dû à la complexité du fonctionnement des structures responsables de l'exécution d'une telle tâche chez l'être humain. Néanmoins, ces algorithmes suscitent de l'intérêt dans plusieurs disciplines et sont appliqués dans des domaines très variés. Ils sont généralement appliqués à de grandes bases de données contenant plusieurs cas ou exemples du phénomène à étudier. Leur but consiste à identifier des patrons récurrents et des

régularités implicites. Ils sont utilisés, entre autres, pour détecter des cas de fraude par cartes de crédit, des cas d'intrusions informatiques, ou pour trouver les meilleures catégories de consommateurs à cibler pour des campagnes de publicité. Ces algorithmes sont également appliqués dans le cadre des interfaces homme-machine pour apprendre à la machine d'une manière autonome et dynamique à s'adapter aux désirs et préférences des utilisateurs. Les sites Web adaptatifs de Google et d'Amazon sont des exemples. Ils permettent aussi de reconnaître automatiquement des formes, des visages, des paroles, l'écriture manuscrite, l'émotion, etc.

Les performances des algorithmes varient selon le domaine d'application et le type d'apprentissage à effectuer. À performances d'apprentissage égales, on préfère l'algorithme le plus simple à implémenter, le plus facile à interpréter, le moins coûteux en termes de complexité spatiale et temporelle, parmi d'autres critères. L'apprentissage est divisé principalement en trois types : l'*apprentissage supervisé*, l'*apprentissage non supervisé* et l'*apprentissage par renforcement*.

- En *apprentissage supervisé*, on dispose d'observations qui prennent la forme de couples d'entrées sorties. L'algorithme consiste alors à chercher une fonction reliant les sorties aux entrées. Lorsque la sortie prend des valeurs discrètes, la tâche de la fonction est dite de *classification*. Lorsque les valeurs sont continues, la tâche est dite de *régression*. L'algorithme des réseaux de neurones (*Rumelhart & McClelland, 1986*) est un exemple d'algorithme d'*apprentissage supervisé*.
- L'*apprentissage non-supervisé*, par opposition au supervisé, s'applique à des observations n'ayant pas de sorties. Sa tâche consiste à décrire l'association entre les entrées et l'organisation des observations. Il existe trois familles dans l'*apprentissage non-supervisé*: la *catégorisation*, la *réduction de dimensionnalité* et l'*estimation de densité de probabilité*. La *catégorisation* (clustering) consiste à diviser un ensemble d'observations disponibles en des groupes similaires. L'algorithme des k-moyennes (*MacQueen, 1965*) en est un exemple. La *réduction de dimensionnalité* consiste à projeter les observations d'un espace multidimensionnel sur une autre dimension plus réduite. Le cas de l'analyse en composantes principales (*Jolliffe, 1986*) est l'un des plus connus. L'*estimation de densité de probabilité* permet de modéliser la vraie distribution

des observations. Le modèle de mélange de densités gaussiennes (*McLachlan & Basford, 1988*) en est un exemple.

- Par opposition, l'apprentissage par renforcement (*Sutton & Barto, 1998*) consiste en l'entraînement dynamique d'un agent intelligent qui maximise la valeur espérée de son utilité dans un environnement incertain. Pour ce faire, l'agent exécute différentes actions qui lui permettent, au moyen d'un processus d'essais erreurs, d'adapter son comportement afin d'optimiser sa fonction objectif. L'algorithme Q-Learning (*Watkins & Dayan, 1992*) en est un exemple.

Dans ce mémoire, nous nous concentrons sur les algorithmes de l'apprentissage *supervisé* utilisés pour résoudre les problèmes de classification et de prédiction. Notre travail porte sur la prédiction de l'émotion d'une personne (sortie) étant donné son profil, sa personnalité et la situation à laquelle il fait face (entrées). Les algorithmes supervisés sont donc bien adaptés pour répondre à notre besoin.

2.2 Apprentissage supervisé

2.2.1 Présentation

Nous introduisons les notions clés de l'apprentissage supervisé à l'aide d'un exemple simple tiré de notre expérience empirique. À l'aide de questionnaires en ligne, nous avons recensé le profil et la personnalité de plusieurs personnes ainsi que leurs réactions émotionnelles face à une grande variété de situations. Cette expérience est détaillée davantage dans les deux chapitres qui suivent.

Nous souhaitons classer l'émotion d'une personne comme étant positive ou négative avec une intensité forte ou faible, advenant des informations sur la situation génératrice de l'émotion et des caractéristiques relatives à la personne, comme sa personnalité. En apprentissage supervisé, la personnalité et les informations sur la situation sont appelées *variables explicatives*, ou encore *variables d'entrée*. Elles sont notées par le vecteur $\mathbf{x} = (x_1, \dots, x_d)$. L'émotion ainsi que l'intensité sont appelées *variables cibles*, ou encore *variables de sortie*. Les *variables cibles* sont notées $\mathbf{y} = (y_1, \dots, y_m)$. Une observation i est définie comme un couple de variables d'entrées et de variables cibles $(\mathbf{x}_i; \mathbf{y}_i)$. Elle est

également appelée *un exemple*. L'ensemble des *exemples* disponibles constitue *l'ensemble d'entraînement*. Il est noté D_n , où n est le nombre d'exemples.

Dans notre cas, un *exemple* correspond au couple $(x_{1i}, x_{2i}, x_{3i}; y_{1i}, y_{2i})$.

1. La valeur x_{1i} caractérise la personnalité d'une personne particulière. Pour simplifier, nous supposons qu'elle indique si la personne est *extravertie* ou non. Une personne extravertie a tendance à extérioriser ses sentiments, à être dynamique et à établir des liens sociaux (Bouvard, 2009).
2. Les valeurs x_{2i} et x_{3i} caractérisent une situation donnée que vit la personne. Elles correspondent respectivement aux aspects de *surprise* et de *désirabilité*. La *surprise* indique si la situation vécue est attendue ou inattendue. La *désirabilité* indique si la situation est désirable ou indésirable. Si cette dernière information n'est pas disponible, on dira que la *désirabilité* est inconnue. Une situation devrait être caractérisée avec plus de variables, mais, pour simplifier, nous nous limitons dans cette section à ces deux dernières. Plus loin dans le mémoire, toutes les variables utilisées pour décrire la personnalité et caractériser une situation sont présentées avec plus de détails.
3. La valeur y_{1i} désigne l'émotion, positive ou négative, ressentie par la personne de personnalité x_{1i} face à la situation décrite par x_{2i} et x_{3i} .
4. La valeur y_{2i} indique une intensité forte ou faible de cette émotion.

En apprentissage supervisé, on suppose que les variables cibles sont reliées aux variables d'entrée par une fonction $f: \mathbf{x} \rightarrow \mathbf{y}$, sauf que f dans la vraie vie est inconnue. L'algorithme d'apprentissage recherche une approximation viable de la fonction f , notée \hat{f} , basée sur l'ensemble d'entraînement D_n . Ce processus de recherche (apprentissage) est amélioré d'une manière automatique au fur et à mesure que de nouveaux exemples se rendent disponibles. Pour \mathbf{x} donné, $\hat{f}(\mathbf{x}) = \hat{\mathbf{y}}$ est dite la *prédiction* de \mathbf{y} par l'algorithme d'apprentissage. L'apprentissage supervisé permet essentiellement deux tâches.

1. La *classification* est utilisée lorsqu'une variable cible y_i admet un nombre fini de valeurs. La fonction approchée \hat{f} est alors appelée *classifieur*. Lorsque la variable cible a deux valeurs possibles, comme dans le cas de l'émotion positive ou négative, la classification est dite *binnaire*. Lorsqu'elle prend plus que deux valeurs, la classification est alors dite *multiple* ou *multiclasses*. Il existe également la classification à plusieurs

cibles, comme le cas de classification de l'émotion et de l'intensité. Cette éventualité est moins traitée dans la littérature.

2. La *régression* est utilisée lorsqu'une variable cible y_i est continue. Dans ce cas, la tâche consiste à estimer ou approximer la cible. La régression est l'une des méthodes les plus utilisées dans tous les domaines scientifiques.

La plupart des algorithmes supervisés peuvent réaliser les deux tâches de classification et de régression.

2.2.2 Phases d'entraînement, validation et test

(A) Phase d'entraînement

La *phase d'entraînement*, dite aussi *d'apprentissage*, consiste à rechercher la fonction \hat{f} qui:

1. relie au mieux les variables d'entrée aux variables cibles pour les couples (\mathbf{x}, \mathbf{y}) de l'ensemble d'entraînement;
2. assure une généralisation adéquate pour les nouveaux couples (\mathbf{x}, \mathbf{y}) de l'univers qui ne sont pas dans l'ensemble d'entraînement.

En apprentissage supervisé, on suppose l'existence d'une fonction de perte à minimiser. Elle mesure la qualité de prédiction pour un exemple $(\mathbf{x}_i, \mathbf{y}_i)$ de l'univers. Elle dépend des sorties de la fonction apprise $\hat{\mathbf{y}}_i = \hat{f}(\mathbf{x}_i)$ et les valeurs correctes des *cibles* \mathbf{y}_i . Dans la phase d'entraînement, l'objectif est de minimiser la perte empirique sur les exemples de l'ensemble d'entraînement. Cette perte est également appelée *erreur d'entraînement*. Elle peut prendre la forme d'une erreur quadratique pour une régression ou un pourcentage d'exemples mal classés dans le cas d'une classification. Pour rechercher la fonction \hat{f} adéquate, l'algorithme:

1. se limite à une famille de fonctions (p. ex., les fonctions linéaires);
2. minimise l'erreur d'entraînement.

Par ailleurs, dans l'apprentissage, il arrive que la fonction obtenue soit beaucoup plus complexe que la vraie fonction f à prévoir (génératrice des exemples). Cela peut engendrer une *erreur d'entraînement* presque nulle, mais peut également donner lieu à des

erreurs de prédiction importantes pour *les exemples* hors du domaine d'entraînement. Ce problème est connu comme étant du *surapprentissage*. Il se peut aussi que la fonction apprise soit trop simple par rapport à la vraie fonction inconnue. On parle alors de *sousapprentissage*. L'art de l'apprentissage se présente donc comme un compromis entre le *surapprentissage* et le *sousapprentissage*.

En guise d'illustration, nous considérons les deux situations suivantes.

1. f est une fonction linéaire de la forme $w_0 + w_1x$ et \hat{f} est un polynôme de haut degré de la forme $w_0 + w_1x^1 + w_2x^2 + \dots + w_dx^d$. Cela débouche sur un *surapprentissage* (Figure 2.1).
2. f est un polynôme de haut degré et \hat{f} est une fonction linéaire. Cela débouche sur un *sousapprentissage* (Figure 2.2).

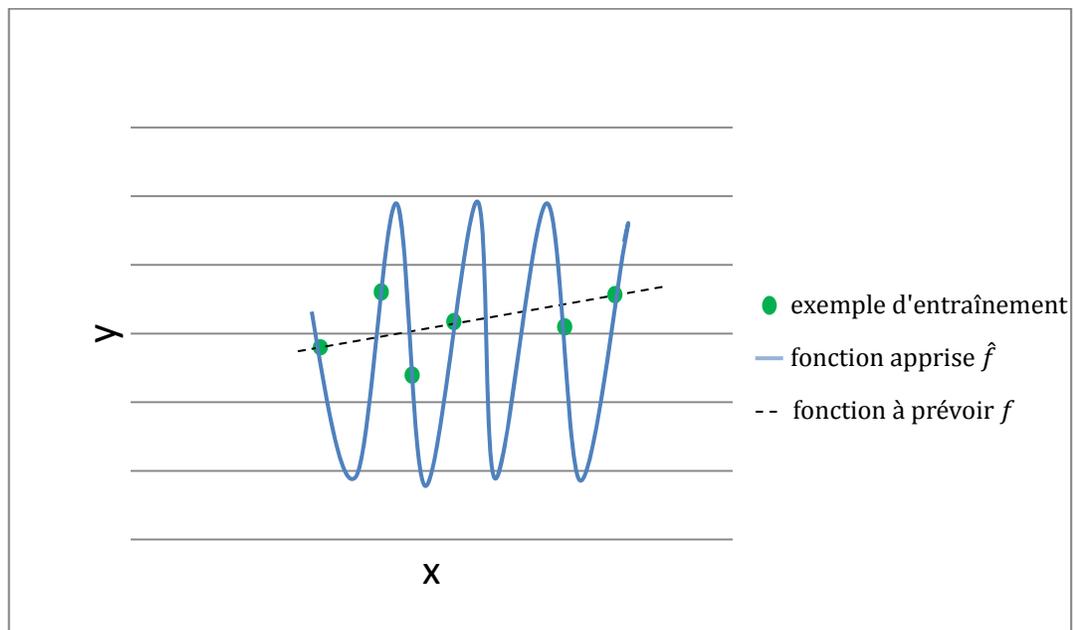


Figure 2.1 Schématisation d'un problème de surapprentissage

(B) Phases de validation et de test

Plusieurs méthodes de validation sont proposées dans la littérature (*Bishop, 2006*). Nous en décrivons une. On suppose qu'on dispose d'un nombre initial d'observations. Dans un premier temps, on réserve une partie de ces observations pour l'*ensemble d'entraînement* et une autre de plus petite taille appelée l'*ensemble test*. Par la suite, on réserve une partie de l'ensemble d'entraînement obtenu, appelée *ensemble de validation*.

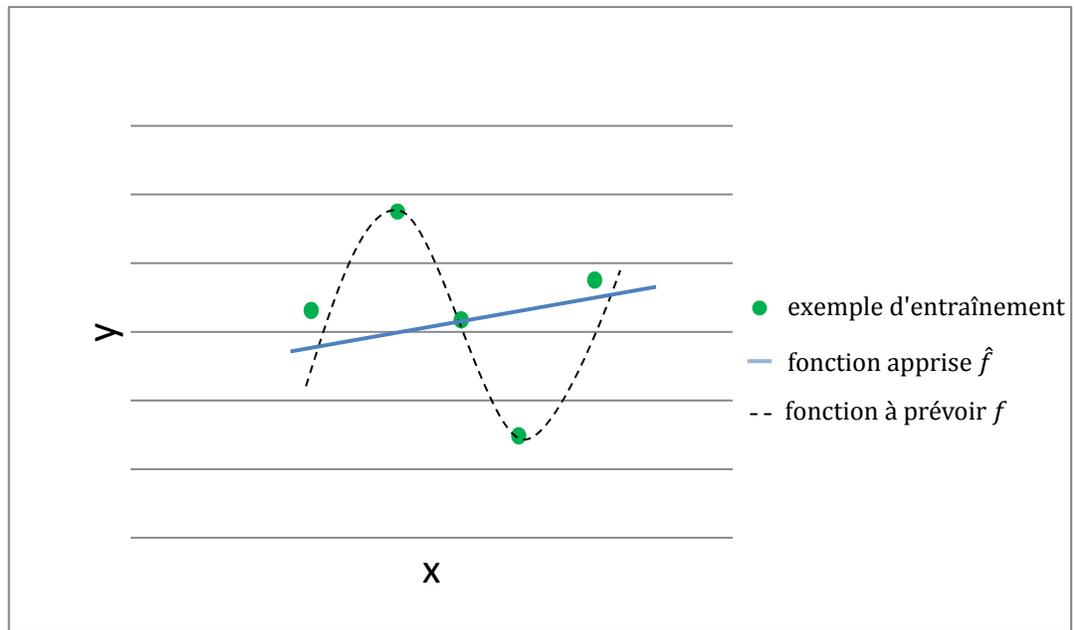


Figure 2.2 Schématisation d'un problème de sous-apprentissage

On suppose également que l'entraînement donne suite à plusieurs alternatives de modèles identifiées par des paramètres additionnels appelés *hyperparamètres* (commentées plus loin dans ce chapitre lors de la présentation des algorithmes). Un exemple d'hyperparamètre est le nombre k de voisins dans l'algorithme des *k plus proches voisins*. Selon la valeur d'un hyperparamètre, la classification ou la régression peut donner des prédictions différentes pour un même *exemple*.

L'algorithme utilise les ensembles d'entraînement, de validation et de tests de la manière suivante.

1. Il apprend un premier modèle avec un premier hyperparamètre.
2. Il mesure la performance de ce modèle appris à l'aide de *l'ensemble de validation*.
3. Il répète les étapes 1-2 avec de nouveaux hyperparamètres jusqu'au moment d'obtenir un modèle final (ayant la meilleure performance).
4. Il réapprend le modèle final à l'aide de l'ensemble d'entraînement complet (y compris l'ensemble de validation).
5. Il estime la performance de généralisation du modèle à l'aide de l'ensemble test.

L'étape 5 constitue la *phase de test*. Elle rend possible, entre autres, la comparaison des performances de plusieurs algorithmes.

2.3 Algorithmes utilisés

Il existe plusieurs algorithmes pour l'apprentissage supervisé qui diffèrent selon la forme du modèle appris, lequel caractérise la relation entre les variables d'*entrées* et les variables *cibles*. Nous citons comme exemple les régressions polynomiales et les règles d'affectation. En guise d'introduction à la classification, nous présentons, dans un premier temps, l'algorithme des arbres de décision. Cet algorithme produit un modèle facile à interpréter sous forme de règles. Dans un deuxième temps, nous nous attardons aux algorithmes des *réseaux de neurones* et des *k plus proches voisins*, utilisés dans notre travail. Nous justifions également le choix de ces algorithmes. La présentation des différents algorithmes est accompagnée d'exemples de classification générés suite à un *entraînement* sur des *exemples* réels collectés lors de l'enquête menée en guise de support à notre travail (Chapitres 3-4). Quelques contextes d'application possibles sont également proposés.

(A) Les arbres de décision

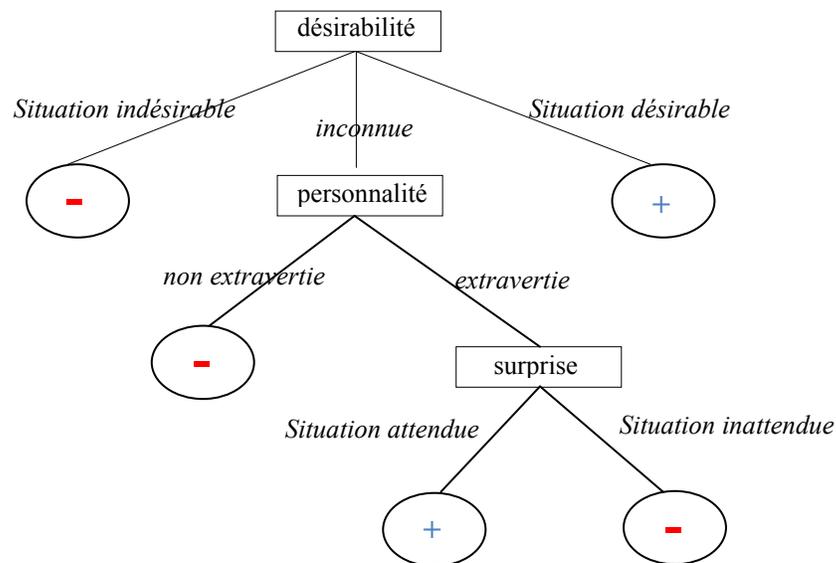


Figure 2.3 Exemple d'un arbre de décision

Nous présentons un arbre de décision à titre introductif. Nous expliquons la manière de l'utiliser à des fins de classification de l'*émotion* (variable cible) advenant la *personnalité*, la *surprise* et la *désirabilité* (variables d'entrées).

La structure d'un arbre de décision (Figure 2.3) est composée de *nœuds* (les rectangles) et de *branches* (Breiman, Friedman, Olshen, & Stone, 1984; Quinlan, 1993). Un nœud correspond à une *variable d'entrée* et chaque branche qui en sort constitue une valeur possible de cette variable (si la variable est discrète) ou un intervalle (si la variable est continue). Les *nœuds feuilles* (les cercles) de l'arbre représentent généralement les valeurs dominantes (plus fréquentes) de la *variable cible* (« + » pour l'émotion positive et « - » pour l'émotion négative). L'arbre est interprété comme un ensemble de règles ou chemins, allant de la racine aux feuilles. Par exemple, le premier chemin de l'arbre à partir de la droite indique que si la situation vécue par une personne est désirable alors l'émotion est à dominance positive. La règle relative au deuxième chemin est comme suit :

Si désirabilité = *inconnue et* personnalité = *extravertie et* surprise = *innatendue,*

Alors émotion = *négative*

Les autres chemins s'interprètent d'une manière similaire. Tous les chemins sont créés de manière à différencier au mieux les *exemples d'entraînement* entre les nœuds feuilles (valeurs de variable cible) (Breiman, Friedman, Olshen, & Stone, 1984; Quinlan, 1993).

Pour terminer, nous citons un exemple réaliste d'un système de reconnaissance d'émotions lorsqu'une personne embarque dans une voiture. Supposons que le système dispose d'une description de la personnalité du conducteur, obtenue suite à un questionnaire de personnalité administré a priori. Supposons que le système puisse reconnaître la désirabilité de la situation selon les gestes du conducteur, le ton de sa voix, ou encore les traits de son visage, perçus à l'aide d'une caméra et d'un microphone intégrés. Supposons également que le système sache reconnaître si une situation est attendue ou non suite aux réactions du conducteur (un freinage brusque, une parole prononcée dénotant un état de surprise). À partir de ces informations, si le système reconnaît le deuxième chemin de notre arbre débouchant sur une émotion négative, on s'attend à ce qu'il réagisse pour calmer ou reconforter le conducteur à temps, avant que la situation ne s'aggrave et n'entraîne un accident.

(B) Les réseaux de neurones artificiels: le perceptron multicouche

Rappelons le problème de base. Nous cherchons la meilleure fonction \hat{f} qui relie les *variables d'entrée* $\mathbf{x} = (x_1, x_2, \dots, x_d)$ aux *variables cibles* $\mathbf{y} = (y_1, y_2, \dots, y_m)$. Un réseau de

neurones commence par transformer les *variables d'entrées* en des *variables intermédiaires* : $z_1 = h_1(\mathbf{x}, \mathbf{w})$, $z_2 = h_2(\mathbf{x}, \mathbf{w})$, ..., $z_q = h_q(\mathbf{x}, \mathbf{w})$, où les \mathbf{w} sont des pondérations (des coefficients) qui indiquent l'importance de chaque *variable d'entrée* dans des combinaisons non linéaires h_1, h_2, \dots, h_q . Ces transformations non linéaires permettent d'approximer des fonctions riches et complexes. Elles peuvent aussi diminuer la dimension du problème sans pour autant perdre de la signification ($q \ll d$). De plus, elles peuvent servir à extraire des caractéristiques plus faciles à traiter mathématiquement (créer des variables plus discriminantes). Par la suite, le réseau de neurones relie les *variables intermédiaires* $\mathbf{z} = z_1, z_2, \dots, z_q$ aux *variables cibles* $y_1 = f_1 = g_1(\mathbf{z}, \mathbf{v})$, ..., $y_m = f_m = g_m(\mathbf{z}, \mathbf{v})$, avec les \mathbf{v} comme pondérations, de manière optimale, afin de minimiser *l'erreur d'entraînement*. Il en résulte des valeurs apprises $\hat{\mathbf{w}}$ et $\hat{\mathbf{v}}$ pour les poids \mathbf{w} et \mathbf{v} . Les fonctions g_1, \dots, g_m représentent des fonctions de transformation qui varient selon que la tâche soit de régression ou de classification. Par exemple, dans le cas de la classification, elles donnent une valeur qui peut être interprétée comme une probabilité de réalisation d'une classe donnée par rapport aux autres classes. Le choix de ces fonctions est détaillé à l'Annexe I. Les paramètres $\hat{\mathbf{w}}$ et $\hat{\mathbf{v}}$ sont appris par le modèle via des principes d'optimisation itérative. Les détails techniques d'optimisation sont exhibés en Annexe I.

Le procédé décrit ci-dessus peut être schématisé comme suit :

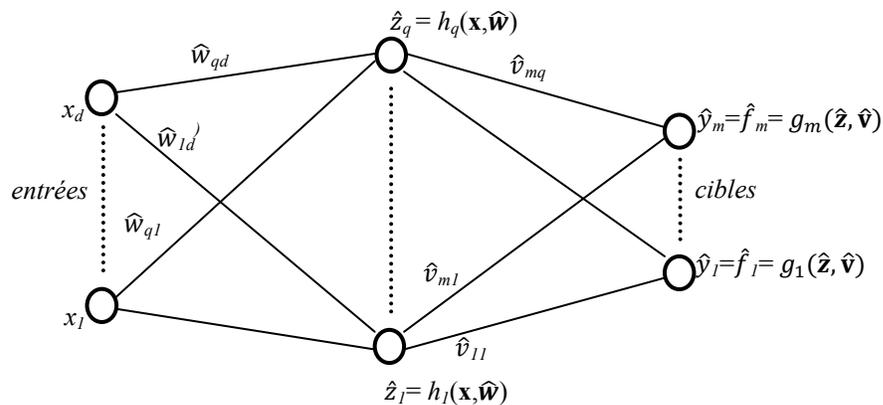


Figure 2.4 Perceptron multicouche à une seule couche cachée

Cette connexion multicouche entre les variables d'entrées et les variables cibles représente un *réseau de neurones artificiel* où les *variables d'entrées*, *intermédiaires* et *cibles* représentent des *neurones* interreliés entre eux comme l'est un réseau de neurones humain. Ce type d'architecture est appelé *perceptron multicouche à une couche cachée*. La

couche cachée est la couche intermédiaire des transformations. Le nombre q de neurones cachés est un hyperparamètre (voir section 2.2.2). Cette architecture est l'une des plus utilisées (*Bishop, 2006*).

Dans cette configuration, il arrive qu'un neurone caché soit relié à toute ou une partie des neurones d'entrée. De même, les neurones de sortie peuvent être reliés à toute ou une partie des neurones cachés.

Un point fort d'un réseau de neurones consiste en sa capacité d'approximer n'importe quelle fonction continue réelle. Cette capacité est connue dans la littérature comme la propriété d'approximation universelle (*Funahashi, 1989*). Aussi, il est capable d'actualiser les poids \mathbf{w} et \mathbf{v} au fur et à mesure que de nouvelles observations se rendent disponibles de manière plus rapide grâce à une variante d'optimisation instantanée appelée *descente de gradient en ligne* ou *stochastique* (*Le Cun et al., 1989*). De plus, l'adaptation de la classification d'une à plusieurs *variables cibles* est simple et elle consiste à modifier la fonction d'erreur au lieu d'apprendre plusieurs réseaux pour chaque *variable cible*.

Nous reprenons l'exemple introduit au début de la section pour la classification de l'*émotion* et de l'intensité (variables cibles), advenant la *personnalité*, la *surprise* et la *désirabilité* (variables d'entrée):

- x_1 = personnalité {extravertie, non extravertie};
- x_2 = désirabilité {situation désirable, situation indésirable, inconnue};
- x_3 = surprise {situation attendue, situation inattendue};
- y_1 = émotion {positive, négative};
- y_2 = intensité {forte, faible}.

Le réseau de neurones appris sur nos données d'enquête donne deux fonctions apprises (\hat{f}_1 , \hat{f}_2) comme fonctions complexes de x_1 , x_2 et x_3 via les transformations h_1 , h_2 , h_3 . Ici, nous avons fixé à trois le nombre de transformations intermédiaires q . Dans le cas général présenté au chapitre 4, le nombre de neurones cachés est un hyperparamètre identifié par l'algorithme d'une manière endogène en recourant à l'*ensemble de validation*. Nous présentons quelques résultats :

1. x_1 = extravertie, x_2 = inconnue, x_3 = inattendue alors y_1 =négative et y_2 =faible

2. $x_1 =$ non extravertie, $x_2 =$ désirable, $x_3 =$ attendue *alors* $y_1 =$ positive et $y_2 =$ forte
3. $x_1 =$ extravertie, $x_2 =$ indésirable, $x_3 =$ inattendue *alors* $y_1 =$ négative et $y_2 =$ forte

Discutons maintenant le cas d'un utilisateur qui apprend une langue étrangère à l'aide d'un logiciel intelligent. Prenons pour acquis la personnalité de l'utilisateur. Le logiciel pourrait reconnaître la surprise à partir des traits de visage, des gestes et des paroles de l'utilisateur. La désirabilité pourrait être perçue au moyen du temps mis par cet apprenant pour répondre aux questions, le niveau de concentration, ou encore l'intensité des frappes sur clavier. Anticiper une émotion négative à partir de ces informations permettrait au logiciel de diminuer la difficulté des exercices proposés et d'encourager l'apprenant à faire mieux.

(C) Les k plus proches voisins pondérés

Les algorithmes précédents représentent des méthodes *paramétriques* où la fonction à apprendre dépend d'un vecteur de paramètres à estimer. Par ailleurs, il existe des algorithmes qui ne contiennent pas de *paramètres* à apprendre, appelés algorithmes *non paramétriques*. La technique des *k plus proches voisins* (*kPPV*) (*Devijver & Kittler, 1982*) en constitue l'une des plus connues. Le *kPPV* est l'un des algorithmes d'apprentissage les plus simples à implémenter. Nous supposons que nous disposons de n observations $(\mathbf{x}_i, \mathbf{y}_i)$, avec $i=1..n$, et d'une nouvelle observation $(\mathbf{x}, .)$ dont la valeur de la cible \mathbf{y} est inconnue. L'algorithme agit comme suit:

1. calcule des distances entre \mathbf{x} et les \mathbf{x}_i pour $i=1..n$;
2. identifie les voisins les plus proches de \mathbf{x} , soit $\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \dots, \mathbf{x}_{(k)}$;
3. considère $\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \dots, \mathbf{y}_{(k)}$ correspondant aux observations $(\mathbf{x}_{(1)}, \mathbf{y}_{(1)}), (\mathbf{x}_{(2)}, \mathbf{y}_{(2)}), \dots, (\mathbf{x}_{(k)}, \mathbf{y}_{(k)})$;
4. classe \mathbf{x} en repérant la classe majoritaire parmi $\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \dots, \mathbf{y}_{(k)}$, soit $\hat{\mathbf{y}}$;

Le nombre k est un *hyperparamètre* qui peut être spécifié ou bien choisi après un test de différentes valeurs sur un *ensemble de validation*. Dans le cas d'une régression, la prédiction est une moyenne $\hat{\mathbf{y}} = \frac{\mathbf{y}_{(1)} + \dots + \mathbf{y}_{(k)}}{k}$. Cet algorithme, basé sur une approche d'approximation locale, peut être un avantage par rapport aux autres algorithmes

paramétriques lorsque la fonction à apprendre se révèle très complexe (Mitchell, 1997). Il y a plusieurs variantes de l'algorithme selon, par exemple, la mesure de distance et la pondération des classes.

La version du *k plus proches voisins* présentée a été ensuite améliorée et raffinée en donnant à chacun des *k* exemples des poids dans la classification selon leur distance. Une mesure communément utilisée est la distance *Euclidienne standard*. Ainsi, plus un exemple est proche de l'exemple à classer, plus il a du poids, et plus il est influent ou déterminant dans la classification. Cette variante de l'algorithme est appelée le *kPPV pondéré* ou *pondéré-distance* (*Distance-Weighted* en anglais). Une manière de faire consiste à donner comme poids pour un exemple l'inverse de sa distance au carré (Mitchell, 1997). Dans ce cas, la classification est comme suit :

- Si la distance entre l'exemple à classer et un exemple d'entraînement est égale à 0, la classe prédite est celle de ce dernier.
- S'il existe plusieurs exemples identiques à l'exemple à classer, on lui affecte la classe majoritaire parmi eux.
- S'il n'existe pas d'exemples identiques, appliquer le *kPPV pondéré*.

Selon Mitchell (1997), le *kPPV pondéré* peut s'avérer très utile pour la résolution de plusieurs problèmes pratiques, car il est robuste contre le bruit des données et efficace lorsqu'il utilise de grandes quantités de données. Cette variante de l'algorithme sera une composante essentielle de notre *machine émotionnelle* (voir Chapitre 4).

Lorsque nous appliquons le principe du *kPPV pondéré* sur nos données pour classer l'émotion comme positive ou négative ou l'intensité comme forte ou faible selon la personnalité, la désirabilité et la surprise, nous obtenons :

- personnalité = extravertie, désirabilité = inconnue et surprise = attendue alors émotion = positive avec une majorité de 68%;
- personnalité = non extravertie, désirabilité = indésirable et surprise = inattendue alors intensité = forte avec une majorité de 77%.

Les agents intelligents intégrés aux jeux vidéo de divertissement peuvent profiter de ce genre de modèle prédictif. Leur objectif est de produire des jeux interactifs tenant

compte de l'état affectif du joueur et d'instaurer un environnement d'interactions ressemblant à celui des humains. Nous supposons qu'un agent dispose de l'information sur la personnalité du joueur. Au moyen d'un oculomètre (eye tracking camera), il est également capable de détecter le mouvement des yeux et le niveau de concentration. Ces informations rajoutées au niveau de difficulté du jeu choisi (amateur, professionnel), le temps restant à jouer et le niveau d'avancement atteint par le joueur peuvent renseigner sur la désirabilité de la situation. Aussi, l'historique des niveaux de difficulté et d'avancement atteint auparavant par le joueur peut indiquer si la situation dans laquelle il se trouve est attendue ou inattendue. Si l'émotion probable est positive, l'agent pourrait initier une interaction en disant par exemple au joueur « Excellent, aujourd'hui vous êtes apparemment de bonne humeur et en forme ! Que diriez-vous d'un autre challenge encore plus difficile ? ». Dans le cas où l'émotion prévue est négative, l'agent peut diminuer le niveau de difficulté sans que le joueur ne s'en aperçoive afin que ce dernier ne quitte pas le jeu. L'agent peut également l'encourager et le rassurer en disant par exemple « Cette étape n'a pas été facile pour la plupart des joueurs ».

2.4 Choix des algorithmes

Notre objectif est de développer une *machine émotionnelle* capable de prédire instantanément l'émotion d'un individu et son intensité étant donné l'information sur une situation qu'il rencontre, ainsi qu'une description de sa personnalité et de quelques caractéristiques sociodémographiques. Les algorithmes supervisés sont les mieux adaptés pour réaliser cet objectif (voir section 2.2). Ainsi, notre choix s'est posé sur les algorithmes supervisés des k plus proches voisins et des réseaux de neurones (perceptron multicouche).

Le choix de ces algorithmes est motivé et dicté principalement par l'aspect de prédiction instantanée de mise à jour rapide d'un modèle. Cet aspect est d'une importance capitale pour les applications de reconnaissance d'émotions dans lesquels notre machine pourrait être intégrée. Généralement, dans ces systèmes d'interaction homme-machine, l'interprétation rapide des informations et la reconnaissance rapide des émotions est nécessaire. Cela permet par exemple de prévenir un risque d'accident à temps dans le cas d'un conducteur en colère ou de garder l'attention d'un étudiant avant qu'il ne s'ennuie, comme indiqué à la section précédente.

Cet aspect nécessite donc un apprentissage continu et rapide au fur et à mesure qu'un nouvel exemple, relatif à un utilisateur et sa situation, est perçu. Or, la plupart des algorithmes paramétriques (voir section précédente) nécessitent une phase d'entraînement qui utilise tout l'ensemble d'entraînement, ce qui introduit une certaine lenteur dans l'exécution et allonge de manière significative la durée d'apprentissage. Par ailleurs, les algorithmes paramétriques ne possèdent pas d'alternatives pour un apprentissage séquentiel. Les algorithmes non paramétriques sont, de ce point de vue, plus adaptés puisqu'ils ne nécessitent pas un entraînement permanent du modèle de classification. Ainsi, le choix s'est porté sur le *k*PPV, surtout qu'il permet de limiter à *k* le nombre d'exemples à prendre en compte. Le choix de l'algorithme des réseaux de neurones est justifié par le fait qu'il est un des rares algorithmes paramétriques qui possède une variante permettant une adaptation continue et instantanée de la fonction apprise au fur et à mesure que de nouveaux exemples sont présentés (*Le Cun et al.*, 1989). Une deuxième raison qui explique le choix de l'algorithme des *réseaux de neurones* consiste en la facilité de l'adapter pour la classification de deux variables cibles, à savoir l'émotion et l'intensité, en modifiant la fonction d'erreur à minimiser au lieu de recourir à l'apprentissage de deux fonctions ou modèles de classifications séparément (voir Annexe I). En outre, le *k*PPV et les réseaux de neurones représentent deux algorithmes d'apprentissage des plus connus et des plus performants et sont très souvent utilisés pour résoudre différents types de problèmes.

Dans ce contexte, notre travail vise à :

1. collecter des observations réalistes ;
2. configurer et exécuter notre machine émotionnelle afin de prévoir l'émotion et l'intensité pour de nouvelles observations ayant des valeurs d'entrée (caractéristiques sociodémographiques, personnalité, situation) non observées et des valeurs de sortie (émotion et intensité) inconnues ;
3. Mesurer la performance de prédiction de la machine

Ce qui nous importe, ce sont les possibilités d'intégration et d'application du système développé, avec l'information qu'il fournit, dans des systèmes réels de reconnaissance d'émotion comme expliquées à la section 2.3.

2.5 Conclusion

Dans de ce chapitre, nous avons tenté d'expliquer la notion d'apprentissage machine, particulièrement celle de l'apprentissage supervisé. Ensuite, nous avons présenté les algorithmes utilisés, les *réseaux de neurones* et les *k plus proches voisins*, et argumenté leur choix. Ce choix tient compte du fait qu'ils contiennent des variantes convenant aux aspects de mise à jour rapide, de prédiction instantanée, et de la classification plusieurs-cibles. Ces deux algorithmes constituent le corps de notre machine dont l'objectif est de classer ou reconnaître correctement l'émotion relative à une situation qui lui est décrite. Comment cette situation est-elle décrite ? Comment pouvons-nous collecter ces situations et construire une base consistante que la machine utilisera ? Le chapitre qui suit présente les réponses à ces questions et la démarche empirique nécessaire à la mise en place de l'application.

Chapitre 3 Procédure empirique: collecte et description numérique des situations émotionnelles

Nous avons indiqué, tout au long des chapitres précédents, que l'objectif de la *machine émotionnelle* est de prédire l'émotion probable ressentie par un individu et son intensité en utilisant *une information* sur la situation et le contexte aux quels ce dernier est confronté. La qualité de cette prédiction est, bien entendu, tributaire de la qualité de l'information recueillie et de la représentation qui permet de la produire. Dans ce chapitre le problème que nous nous posons est de savoir comment représenter ou modéliser cette *information* et comment caractériser une situation à l'aide des variables ou des dimensions ? La réponse à ces questions faciliterait l'établissement d'une transformation ou traduction de la notion abstraite de situation émotionnelle en un formalisme et un langage compréhensible par la machine.

Certains psychologues ont tenté, à l'aide de quelques dimensions, de caractériser l'environnement de l'humain et les situations pouvant générer des réactions émotionnelles. Mehrabian et Russell (1980) ont essayé d'identifier les variables pertinentes qui expliquent les stimuli environnementaux. De même, en compilant les travaux de Kurt Lewin, De Rivera (1976) a décrit les situations qui affectent le comportement de l'individu à la suite d'une émotion ressentie. Dans sa description du phénomène émotionnel l'auteur a proposé un grand nombre de dimensions dans un contexte de psychologie topologique (Lewin, 1936). Néanmoins, ces différentes caractérisations sont fondées sur des notions psychologiques complexes qui rendent difficile leur conversion en un formalisme simple et précis.

Pour capturer le processus émotionnel nous avons alors pensé recourir à un formalisme plus simple et intuitif. Ce dernier permet de décrire une situation en fonction de la perception et de l'évaluation cognitive, que fait l'individu, des conséquences probables qu'elle peut initier ou éliciter. Les théories d'évaluation correspondent à ce type de raisonnement puisqu'elles fournissent un cadre théorique pour l'évaluation cognitive des situations émotionnelles à l'aide des *variables d'évaluation* (voir section 1.2.3). Le modèle d'Ortony, Clore et Collins ou OCC (Ortony, Clore, & Collins, 1988), est particulièrement intéressant. Il a été construit pour être facilement adapté aux travaux d'intelligence

artificielle. Ce modèle est devenu très vite populaire grâce à son succès dans la simplification apportée à la représentation des états émotionnels.

Au cours de ce chapitre nous présentons, dans une première étape, le modèle OCC et *ses variables d'évaluation*. Dans une seconde étape, nous expliquons les détails d'une procédure que nous avons conçue et mise en œuvre pour le recueil d'un grand nombre de situations émotionnelles nécessaires au fonctionnement de notre machine. Finalement nous essayons de décrire, en fonction des variables du modèle OCC, l'information cognitive relative à ces différentes situations. Cette description numérique est rendue possible grâce à une deuxième procédure de collecte de données que nous avons développée. Les détails de cette seconde procédure font aussi l'objet des analyses qui suivent.

3.1 Le modèle OCC

Comme toute théorie d'évaluation cognitive, l'objectif du modèle OCC est la modélisation de l'interprétation cognitive que fait un individu des situations génératrices d'émotions. Le modèle propose une description détaillée de cette interprétation. Plus précisément le modèle propose une taxonomie des types d'émotions différenciées par des critères cognitifs et analysent les principaux facteurs cognitifs qui déterminent le type et l'intensité de l'expérience émotionnelle selon la spécificité de la situation rencontrée.

3.1.1 Les types de situations

Ortony, Clore et Collins (1988) font valoir que les émotions sont provoquées chez les individus lorsque la situation influence leurs objectifs, affecte leurs normes et valeurs, ou fait appel à leur goût. Ainsi, selon l'évaluation faite par la personne, une situation génératrice d'émotions comprend essentiellement trois catégories ou formes:

- Un *évènement* ayant des conséquences sur les besoins et les objectifs.
- Une *action d'agent* qui répond aux valeurs et normes ou qui en dévie.
- Un aspect d'*objet* qui correspond ou non au goût.

Un *évènement* a des conséquences sur l'individu lui-même ou sur autrui et engendre généralement un sentiment de plaisir ou déplaisir. Un évènement peut être classé selon 3 types de conséquences:

- un évènement qui a des conséquences sur autrui (un ami qui perd sa femme),
- un évènement qui a des conséquences sur les projets futurs et l'avenir de la personne elle-même (le refus d'admission d'une université) et
- un évènement qui agit sur le bien-être actuel de soi (avoir des vacances ; jouer avec ses enfants).

Un *agent* peut être aussi bien une personne, une entité abstraite ou une institution (gouvernement, école, entreprise, etc.) qui peut agir en accord ou non avec ses propres valeurs et normes. L'*agent* peut correspondre au « soi », c'est-à-dire le sujet même, ou bien à une personne ou une entité à laquelle on s'identifie (son pays, son entreprise, son équipe, etc.). Cet agent peut également être une personne ou une entité étrangère à laquelle on ne s'identifie pas. L'action d'un agent engendre une réaction d'approbation ou de désapprobation sans qu'elle ait pour autant des conséquences sur soi (on peut, par exemple, toujours, apprécier une personnalité publique et approuver son geste qui consiste à faire un don important à une association d'orphelins, sans, pour autant, être soi même orphelin).

Un *objet* peut engendrer une appréciation ou un dégoût. Il peut être un objet concret comme une voiture, des habits, un bâtiment ou bien il peut être abstrait comme un fait, une idée ou un concept (racisme, paix, guerre, etc.).

Il existe une quatrième catégorie implicite de situations. Elle correspond à une combinaison de deux catégories, l'*action d'un agent* et l'*évènement*. L'*action d'un agent* (approuvée ou désapprouvée) donne lieu à un *évènement* pouvant affecter les objectifs d'un individu (engendrer un plaisir ou déplaisir), par exemple, un professeur qui prend le temps de répondre aux questions d'un étudiant en dehors des heures du cours.

La classification préliminaire précédente a permis à Ortony, Clore et Collins de définir quatre grandes catégories de réactions émotionnelles associées à ces quatre types de situations. La catégorie relative à l'*évènement* est elle-même divisée en trois sous-catégories selon les conséquences de l'évènement (sur autrui, sur les projets, sur le bien-être). Cela donne lieu, selon les types de situations, à six groupes d'émotions qui sont décrits dans le paragraphe suivant. L'arbre de la Figure 3.1 illustre la structure cognitive des réactions émotionnelles selon le type de situation.

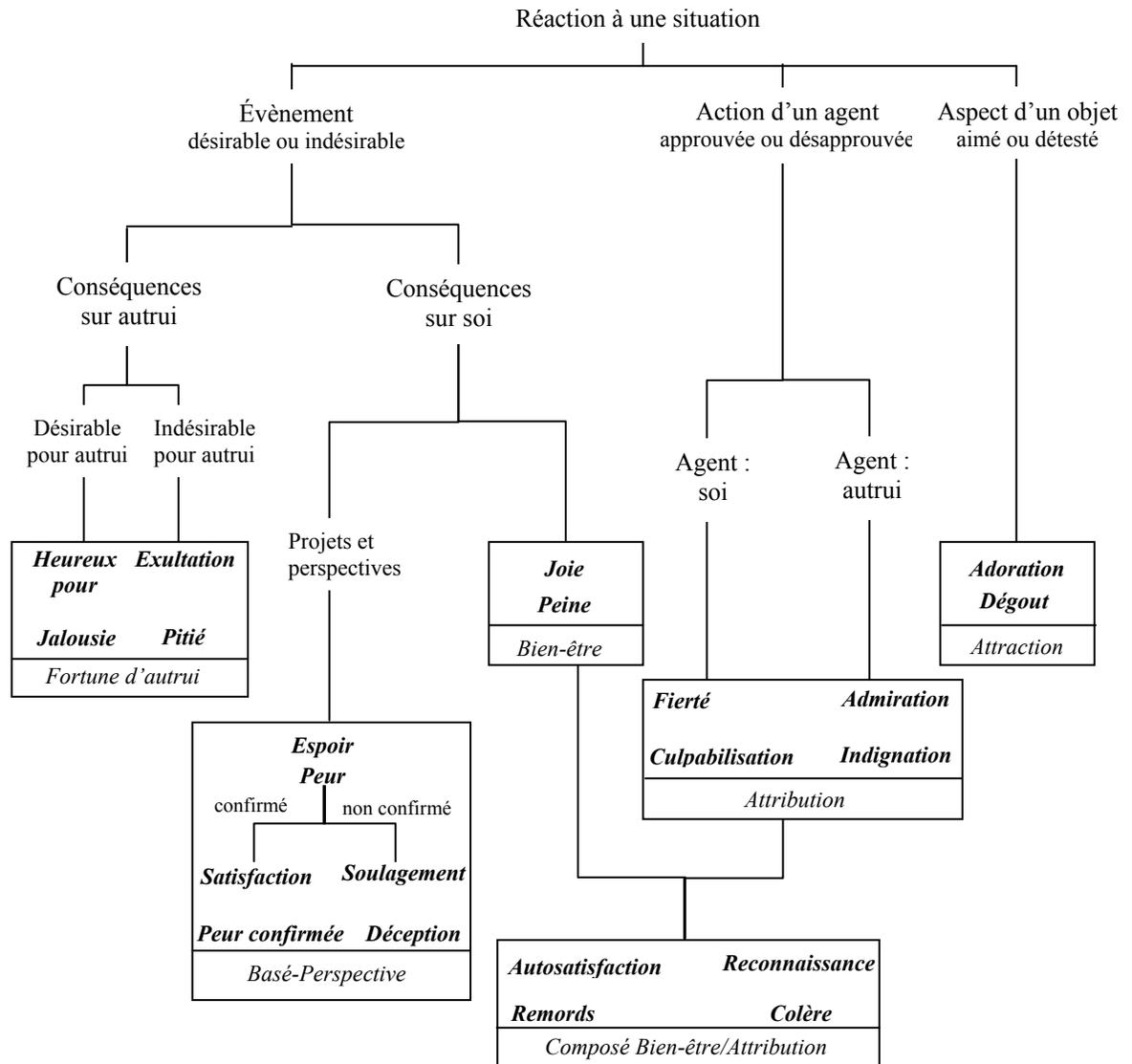


Figure 3.1 Structure du modèle OCC (Ortony, Clore, & Collins, 1988)

3.1.2 Les groupes et paires d'émotions dans le modèle OCC

Les six groupes d'émotions sont :

- le groupe *Fortune d'autrui*, regroupant les événements ayant des conséquences sur autrui;

- le groupe *Basé-perspective*, regroupant les évènements ayant des conséquences sur les projets futurs de soi;
- le groupe *Bien-être*, regroupant les évènements ayant des conséquences sur le bien-être actuel de soi;
- le groupe *Attribution*, désignant les actions d'un agent approuvées ou désapprouvées;
- le groupe composé *Bien-être/Attribution*, désignant les évènements affectant le bien-être de soi, survenant à la suite d'une action d'un agent;
- le groupe *Attraction*, désignant l'appréciation ou le dégoût d'un objet.

Au sein de chaque groupe, il existe des paires d'émotions opposées. Chaque paire contient une émotion positive (de plaisir, d'appréciation : joie, etc.) et une émotion négative (de déplaisir, de dégoût : peine, etc.). Au total, il y a 11 paires, d'où 22 émotions.

3.1.3 Relation entre les types de situations et les émotions

Tableau 3.1 Description semi-formelle des situations correspondantes aux 22 émotions

Émotion	Situation correspondante	Émotion	Situation correspondante
<i>Heureux pour autrui</i>	Un évènement désirable par autrui et souhaité par soi (pour autrui)	<i>Pitié</i>	Un évènement indésirable par autrui et non souhaité par soi (pour autrui)
<i>Exultation</i>	Un évènement indésirable par autrui et souhaité par soi	<i>Jalousie</i>	Un évènement désirable par autrui et non souhaité par soi
<i>Espoir</i>	La perspective d'un évènement désirable	<i>Peur</i>	La perspective d'un évènement indésirable
<i>Satisfaction</i>	La confirmation de la perspective d'un évènement désirable	<i>Peur confirmée</i>	La confirmation de la perspective d'un évènement indésirable
<i>Soulagement</i>	La non-confirmation de la perspective d'un évènement indésirable	<i>Déception</i>	La non-confirmation de la perspective d'un évènement désirable
<i>Joie</i>	Un évènement désirable	<i>Peine</i>	Un évènement indésirable
<i>Fierté</i>	Une action louable de soi n'ayant pas de conséquences sur soi	<i>Culpabilisation</i>	Une action répréhensible de soi n'ayant pas de conséquences sur soi
<i>Admiration</i>	Une action louable d'autrui n'ayant pas de conséquences sur soi	<i>Indignation</i>	Une action répréhensible d'autrui n'ayant pas de conséquences sur soi
<i>Reconnaissance</i>	Une action louable d'autrui ayant de bonnes conséquences sur soi	<i>Colère</i>	Une action répréhensible d'autrui ayant de mauvaises conséquences sur soi
<i>Autosatisfaction</i>	Une action louable de soi ayant de bonnes conséquences sur soi	<i>Remords</i>	Une action répréhensible de soi ayant de mauvaises conséquences sur soi
<i>Adoration</i>	Un objet attractif	<i>Dégoût</i>	Un objet répulsif

Le schéma de la Figure 3.1 représente la structure générale du modèle OCC et résume les liens entre les types de situations (événement, action d'agent, etc.) et les groupes d'émotions engendrées.

Le modèle définit de façon semi-formelle des règles entre chaque paire d'émotions et les types de situations. Par exemple, pour la paire *espoir* et *peur*, il suppose que ces deux émotions peuvent survenir suite à la prise de conscience de la possibilité de réalisation d'un événement ayant des conséquences sur les projets futurs d'un individu et la perspective qu'il a de lui-même. Selon la désirabilité ou la non-désirabilité de cet événement, l'émotion pourrait être, ou bien l'*espoir* (quelqu'un qui entend des rumeurs sur sa possible promotion) ou bien la *peur* (une prévision de tempête imminente).

Tableau 3.2 Termes émotionnels équivalents

Terme émotionnel employé	Termes équivalents	Terme émotionnel employé	Termes équivalents
<i>Heureux pour autrui</i>	Content pour, ravi pour, etc.	<i>Pitié</i>	Compassion, sympathie, triste pour, etc.
<i>Exultation</i>	Jubilation, heureux du malheur d'autrui	<i>Jalousie</i>	Envie
<i>Espoir</i>	Excitation, plaisir anticipé, attendre avec impatience, etc.	<i>Peur</i>	Inquiet, craintif, angoissé, terrifié, appréhensif, effrayé, etc.
<i>Satisfaction</i>	Espoirs réalisés, gratification, etc.	<i>Peur confirmée</i>	Craintes confirmées
<i>Soulagement</i>	Être rassuré	<i>Déception</i>	Frustration, désespoir, désappointement, etc.
<i>Joie</i>	Bonheur, être content, ravi, réjoui, extasié, etc.	<i>Peine</i>	Tristesse, être déprimé, mécontent, chagriné, nostalgique, etc.
<i>Fierté</i>	Être content de soi pour ses actions envers les autres	<i>Culpabilisation</i>	Honte, auto-condamnation, être embarrassé, humilié, etc.
<i>Admiration</i>	Appréciation, estime, respect, être impressionné, etc.	<i>Indignation</i>	Mépriser, reprocher, être scandalisé, choqué, etc.
<i>Reconnaissance</i>	Gratitude, être redevable à quelqu'un, etc.	<i>Colère</i>	Fureur, exaspération, se sentir offensé, outragé, etc.
<i>Autosatisfaction</i>	L'estime de soi, être content de soi suite au succès dans l'accomplissement d'un objectif	<i>Remords</i>	Être en colère contre soi suite à un échec dans l'accomplissement d'un objectif
<i>Adoration</i>	Aimer, aimer bien, adorer, être attiré par, etc.	<i>Dégoût</i>	Détester, aversion, répugnance, etc.

Dans le cas où la réalisation de cet événement potentiel se confirme, l'émotion d'*espoir* peut se transformer en *satisfaction*, si cet événement est désirable, alors que l'émotion de

peur peut *se confirmer* dans le cas contraire. Inversement, la non-confirmation de l'évènement potentiel désirable peut entraîner une *déception*, alors que celle de l'évènement indésirable peut engendrer un *soulagement*. Ainsi, Ortony et al. sont parvenus à caractériser, d'une manière semi-formelle, différentes situations possibles et à associer à chacune une émotion (voir Tableau 3.1).

Selon les auteurs (Ortony, Clore et Collins), les termes qui sont employés pour désigner ces 22 émotions ne sont pas exclusifs. Chaque émotion n'est pas restreinte à un terme ou un mot, mais peut être décrite par d'autres mots ou d'autres expressions équivalentes exprimant le même sentiment. Le Tableau 3.2 présente ces différents substituts.

3.1.4 Les variables d'évaluation du modèle OCC

Ortony, Clore et Collins (1988) plaident en faveur d'une théorie des processus d'évaluation cognitive qui constituent les préalables pour initier les émotions. La plus part des théories d'évaluation stipulent que pour une situation S_p ou un évènement donné le processus émotionnel commence d'abord par les évaluations qui à leur tour initient les changements physiologiques, d'expression et comportementaux, pour aboutir finalement à un état émotionnel donné (Roseman & Smith, 2001). Les évaluations se rapportent à des facteurs ou *variables dites cognitives*. Ces variables sont censées caractériser l'interprétation que fait la personne de la situation S_p dans laquelle il se trouve. Ainsi, la réaction émotionnelle varie, par exemple, selon que la situation est désirable ou non, attendue ou inattendue, facile à gérer ou non, etc. Le modèle OCC présente un certain nombre de *variables d'évaluation* qui décrivent une situation donnée. Ces variables servent à expliquer, interpréter, et prédire la réponse émotionnelle correspondante (p.ex., joie) et particulièrement son intensité (joie extrême, joie moyenne). Le tableau 3.3 présente toutes les variables spécifiées par le modèle OCC et explique leur significations. Ces variables, dites encore *d'évaluation-intensité*, sont divisées en trois catégories :

- les *variables globales*, communes à toutes les situations, incluant la surprise, le *sens de réalité*.

- les *variables centrales*, où chaque variable est associée à un type de situation. Ainsi, la variable *désirabilité* est relative à l'occurrence d'un événement, l'*approbation* est associée à l'action d'un agent, et l'*attirance* concerne un objet.

Tableau 3.3 Les variables d'évaluation-intensité du modèle OCC

	Variable d'évaluation	indique :
Variables globales	<i>Surprise</i>	Si la situation était attendue ou inattendue
	<i>Sens de réalité</i>	Si la situation est réelle et vécue ou bien irréelle et imaginaire
	<i>Dimension temporelle</i> ⁵	Si un événement passé ou futur influe sur la situation présente
Variables centrales	<i>Désirabilité</i>	À quel point l'occurrence de l'évènement est désirable ou indésirable
	<i>Approbation</i>	À quel degré l'action d'un agent est louable ou condamnable
	<i>Attirance</i>	À quel degré l'objet (concept, idée, etc.) est attirant ou répugnant
Variables locales	<i>Désirabilité par autrui</i>	À quel point l'occurrence d'un événement est désirable ou indésirable par autrui
	<i>Estime pour autrui</i>	À quel point autrui est apprécié ou méprisé
	<i>Mérite d'autrui</i>	À quel point on croit qu'autrui mérite ou ne mérite pas ce qui s'est passé
	<i>Vraisemblance</i>	Le degré de croyance que l'évènement anticipé va ou ne va pas se produire
	<i>Réalisation</i>	À quel point l'évènement anticipé s'est ou ne s'est pas finalement réalisé
	<i>Effort</i>	Le degré d'efforts fournis pour que l'évènement anticipé soit évité ou soit réalisé
	<i>Agent</i>	Celui qui réalise l'action, ça peut être « soi » ou autrui
	<i>Puissance du lien</i>	À quel point on s'identifie (on est lié) à la personne ou l'institution qui réalise l'action
	<i>Déviaton</i>	À quel point l'action de l'agent dévie ou sort de la norme (positivement ou négativement)
	<i>Disposition</i>	Si on est disposé ou non disposé à être attiré par l'objet
	<i>Familiarité</i>	À quel point l'objet est familier ou étranger

- les *variables locales* sont également propres à un type particulier de situation. Cependant, elles diffèrent dans le cas d'un événement selon que ce dernier a des conséquences sur autrui ou sur les projets de l'individu.

Les variables centrales et locales impliquées dans la description de chaque type de situation sont indiquées dans le Tableau 3.4.

3.1.5 La description d'une situation émotionnelle

Dans cette section, nous donnons un exemple illustrant l'utilisation des variables d'évaluation-intensité pour décrire une *situation émotionnelle* et représenter (ou restituer)

⁵ Cette variable est définie par Mehrabian et Russell (1980) et n'appartient pas au modèle OCC. Nous avons estimé que son intégration dans notre travail était nécessaire, vu l'information temporelle pertinente qu'elle apporte.

son contenu informationnel. Prenons l'exemple d'une personne qui est restée au chômage pendant plusieurs mois, et qui après avoir passé, sans succès, des dizaines d'entrevues d'embauche, réussit finalement à décrocher un emploi, alors qu'elle commençait à perdre espoir. L'*information* qui caractérise la situation exposée est riche d'enseignements et comporte plusieurs aspects. On pourrait, toujours, avoir des interprétations différentes. Mais, nous estimons que l'interprétation la plus commune (faite par la plupart des gens) est la suivante. Pour la personne concernée, il s'agit d'un évènement *souhaité, inattendu* et totalement *invraisemblable* (perte d'espoir) et que cet évènement aurait des conséquences sur ses plans et ses projets, qu'il *désire* probablement beaucoup. Nous estimons, aussi, que l'*effort* fourni par la personne, pour que l'évènement *se réalise* finalement, est considérable (les différentes entrevues passées).

Tableau 3.4 Les variables centrales et locales relatives à chaque type de situation

Situations Variables	Évènement ayant des conséquences sur			Action d'un agent	Action agent + évènement	Aspect d'un objet
	Autrui	Plans	Bien-être			
<i>Désirabilité</i>	✓	✓	✓		✓	
<i>Approbation</i>				✓	✓	
<i>Attirance</i>						✓
<i>Désirabilité par autrui</i>	✓					
<i>Estime pour autrui</i>	✓					
<i>Mérite d'autrui</i>	✓					
<i>Vraisemblance</i>		✓				
<i>Réalisation</i>		✓				
<i>Effort</i>		✓				
<i>Agent</i>				✓	✓	
<i>Puissance du lien</i>				✓	✓	
<i>Déviation</i>				✓	✓	
<i>Disposition</i>						✓
<i>Familiarité</i>						✓

L'interprétation précédente correspondrait à une *description semi-formelle* de cette situation au moyen de quelques *variables d'évaluation-intensité*, à savoir les variables *Surprise, Vraisemblance, Désirabilité, Effort* et *Réalisation*. Les quatre dernières variables citées font partie des *variables centrales et locales* qui sont utilisées, comme l'indique le tableau 3.4, pour décrire un évènement ayant des conséquences sur les projets futurs. La variable *Surprise* est une variable globale. Elle est donc incluse dans la description de tout type de situation, de même que les variables *Sens de réalité* et *Dimension temporelle* (Tableau 3.3). De ce fait, nous pouvons rajouter dans notre description de la situation énoncée que l'évènement correspond à un fait *réel* (non imaginaire) et que sa réalisation

aidera l'individu à gagner de l'argent dans un futur proche, donc que l'évènement actuel *est lié à un autre évènement futur*. Ainsi, en suivant les schémas de description avec les variables centrales et locales du Tableau 3.4, et également les variables globales, nous pouvons décrire, selon Ortony, Clore et Collins (1988), une grande partie des situations émotionnelles qui peuvent survenir dans la vie courante.

Toutefois, le modèle OCC se limite à l'identification des variables impliquées et ne spécifie pas formellement comment décrire les situations, c'est-à-dire quelles sont les combinaisons de valeurs que ces variables pourraient prendre pour une situation donnée et comment une certaine combinaison de ces valeurs peut donner lieu à une intensité d'une émotion parmi les 22. Il décrit, plutôt, informellement la situation qui peut engendrer une émotion, comme dans le Tableau 3.1, et ne donne pas d'indication sur sa probable intensité. En présentant une formalisation de la description d'une situation (voir section 3.2.2) et en développant notre *machine émotionnelle* (voir chapitre 4), nous proposons des réponses possibles à ces interrogations.

Mais, dans un premier temps, il est essentiel d'avoir beaucoup de situations, comme la situation décrite au début de cette section, pour pouvoir les formaliser et les transformer numériquement.

3.2 Construction d'une base de situations émotionnelles

Généralement, un humain est capable de reconnaître ou de prédire, dans une certaine mesure, l'émotion de quelqu'un d'autre vivant une situation donnée, après avoir expérimenté, vécu ou bien vu plusieurs cas de situations similaires. Fridja (1989) illustre l'importance de ce type d'apprentissage dans l'amélioration de notre capacité de reconnaître nos émotions et de l'enrichissement de notre répertoire émotionnel. De la même manière, notre *machine émotionnelle* doit intégrer dans ses connaissances plusieurs types de situations relatives à plusieurs émotions afin de pouvoir apprendre et prédire l'émotion d'une personne et son intensité de façon acceptable. Ainsi, l'objectif est, ici, de construire une base textuelle qui contient un grand nombre de situations différentes décrites formellement.

3.2.1 Collecte de situations

L'apprentissage mentionné nécessite une base de connaissance riche et variée. La question que nous nous sommes posée, dans un premier temps, est la suivante : Comment trouver un ensemble de situations émotionnelles (comme celle présentée à la section 3.1.5) qui permet d'engendrer les 22 émotions du modèle OCC?

Pour construire un tel ensemble nous nous sommes référés aux travaux de quelques psychologues (*Scherer, 1993; Schwartz & Weinberger, 1980*). La consultation de ces travaux nous a aidé dans la conception de notre base de situations et d'émotions.

Shwartz et Weinberger (1980) ont cherché à identifier l'existence de patrons de réactions émotionnelles en présence de situations différenciées. Pour ce faire, ils ont pu mettre en place une procédure empirique à trois étapes. Au cours de la première étape, ils ont procédé au développement d'un questionnaire de situations émotionnelles. Ils s'intéressaient, particulièrement, à l'étude de trois niveaux d'intensité (faible, modéré, fort) pour chacune des six émotions de base (voir section 1.2.1) (tristesse, joie, colère, peur, dégoût et surprise). Par la suite, ils ont demandé à des étudiants, présents sur le campus de l'Université de Yale, de remplir les formulaires préalablement conçus et de suggérer une phrase simple décrivant une situation correspondant à une émotion donnée avec une intensité particulière. Avec les six émotions et les trois niveaux d'intensité, ils ont pu récupérer 18 groupes de situations émotionnelles décrites textuellement. Les situations de chaque groupe étaient destinées à provoquer une émotion avec une intensité particulière. Après des modifications syntaxiques, ils ont fini par rassembler 360 situations. Ensuite, dans une deuxième étape, ils ont demandé à un autre groupe d'étudiants de valider l'appartenance des situations aux groupes. Finalement, pour l'analyse et la détection de patrons de réactions éventuels, ils ont demandé à un troisième groupe d'étudiants d'imaginer les émotions qu'ils pourraient ressentir face à quelques situations collectées au cours des étapes précédentes.

De la même manière, tout le long des années 90 et dans le cadre de l'étude internationale sur les antécédents et les réactions émotionnelles ISEAR (*International Study of Emotion Antecedents and Reactions*), Scherer et ses collaborateurs (*Scherer, 1993, 1997; Scherer & Wallbott, 1994*), ont demandé à un très grand nombre d'étudiants de différentes cultures et de différents pays de détailler une expérience émotionnelle de la vie courante

qu'ils auraient vécue au cours des dernières semaines ou derniers mois. Les émotions retenues dans l'étude étaient la colère, la peur, la tristesse, la joie, le dégoût, la honte, et la culpabilité. Une très grande base de données de situations relatives à ces émotions a été constituée. Un des objectifs de ce grand sondage était d'étudier les patrons universels de l'expérience émotionnelle et également les différences culturelles de leur symbolisation et régulation (Scherer & Wallbott, 1994).

Université de Montréal

Département d'informatique et de recherche opérationnelle

Veillez détailler dans ces espaces textes, à travers quelques phrases, une première situation que vous avez vécu et qui vous a rendu *moyennement ou quelque part joyeux(se)*, puis une deuxième qui vous a procuré *une joie extrême*.

Moyennement joyeux, content, ravi, réjoui, extasié, etc. :

Extrêmement joyeux, content, ravi, réjoui, extasié, etc. :

Page Suivante Quitter

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

Figure 3.2 Saisie des situations émotionnelles

Pour collecter des situations émotionnelles de la vie quotidienne (vie professionnelle, familiale, etc.), nous nous sommes appuyés sur les travaux précédents. Nous avons pu concevoir un questionnaire approprié au recueil des informations spécifiées par le cadre théorique d'évaluation cognitive proposé plus haut. Nous nous sommes intéressés aux 22 émotions d'OCC, avec deux niveaux d'intensité, à savoir *moyen* et *fort*. A ce stade on peut noter l'étendue et la spécificité de notre travail qui touche à toutes les 22 émotions d'OCC alors que les autres travaux cités se sont uniquement intéressés à six ou sept émotions. La mise en ligne du questionnaire préalablement conçu

est faite à l'aide du langage de programmation *PHP*⁶ (*Hypertext Preprocessor*) utilisé pour les sites dynamiques. Nous avons envoyé ce questionnaire, par courriel, à des étudiants de baccalauréat, maîtrise et doctorat du département d'informatique et de recherche opérationnelle de l'Université de Montréal. Nous leur avons demandé de nous suggérer (de manière anonyme et en quelques phrases) des situations qui selon eux provoqueraient les 22 émotions avec les deux intensités *moyenne* et *forte*. La Figure 3.2 présente un exemple d'une page du questionnaire.

Dans cette page nous demandons à l'utilisateur de saisir une situation vécue engendrant une joie moyenne et une autre pouvant donner lieu à une joie forte ou extrême. Afin d'expliquer davantage la signification des émotions demandées, nous avons présenté les synonymes ou substituts du Tableau 3.2 relatifs à chaque émotion. Ainsi, chaque participant devait remplir 22 pages relatives aux 22 émotions suivant le modèle de la page illustrée dans la Figure 3.2. Pour passer d'une page à une autre, nous n'avons pas établi de restrictions de saisie afin que le participant ne se sente pas obligé de répondre à toutes les questions et puisse décider, par conséquent, de quitter prématurément le questionnaire. Suite à ce procédé, plusieurs participants n'ont pas répondu à la totalité des questions et ont proposé un nombre de situations inférieur aux 44 demandées (22 émotions avec leurs 2 intensités). Toutes les situations saisies ont été enregistrées dans des fichiers.

Il y a eu 57 participants qui ont répondu, partiellement ou entièrement, à notre questionnaire en ligne. La plupart d'entre eux ont proposé des situations relatives aux émotions de *Joie*, *Peine*, *Adoration*, *Dégoût*, *Colère* et *Peur*. Ceci peut s'expliquer du fait que la plupart de ces émotions sont des émotions de base et qu'elles représentent les émotions communes les plus éprouvées dans la vie courante. Les situations relatives aux autres émotions étaient beaucoup moins représentées dans l'échantillon recueilli. Notre collecte de situations émotionnelles diffère de celle de Schwartz et Weinberger (1980). En effet elle offre des situations supposées engendrer d'autres émotions, en plus de celles dites de base, et ce malgré le fait qu'elles soient peu représentées. Ces situations sont d'une grande utilité pour le développement de notre machine émotionnelle (chapitre 4). Nous remarquons que les situations provoquant une émotion donnée étaient presque uniformément distribuées selon les intensités.

⁶ <http://php.net/index.php>

Étant donné que les participants n'ont pas répondu à toutes les questions demandées, nous avons obtenu 1002 situations au lieu de 2508 (57x44). Ainsi, en moyenne, chaque participant a proposé presque 18 situations. Les situations entrées étaient réparties sur 44 groupes (chaque groupe correspond à une émotion donnée avec une intensité particulière). Elles contenaient beaucoup de redondances, c'est-à-dire que beaucoup d'entre elles tournaient au tour du même sujet ou thème. Par exemple, les situations proposées et supposées provoquer une émotion de joie avec une grande intensité tournaient principalement au tour des thèmes de la réussite professionnelle, du bonheur familial, de l'argent et de l'amour. De plus, plusieurs de ces situations étaient brièvement décrites à l'aide de quelques mots. Pour la joie on trouve des mots comme « réussite », « bon emploi » et « bonheur ». Pour le cas de la peine les mots cités sont « décès », « échec » et « infidélité ». Nous avons alors tenté d'enlever ces redondances en regroupant les situations répétées dans une seule qui résume l'idée proposée. Puis, nous avons essayé de développer davantage les situations brièvement décrites en incluant les aspects introduits par le modèle OCC, comme la surprise, l'effort fourni pour l'accomplissement d'une tâche, le degré de réalisation d'un événement souhaité, etc. En moyenne, quatre situations traitant du même thème ont été regroupées dans une seule. Après ce traitement, le nombre de situations est passé de 1002 à 250 situations émotionnelles décrites textuellement, et constituant ainsi un sous-échantillon représentatif de l'échantillon initial.

Tableau 3.5 La répartition des 250 situations finales selon les émotions et les intensités

Émotion	<i>Heureux pour</i>	<i>Exultation</i>	<i>Espoir</i>	<i>Satisfaction</i>	<i>Soulagement</i>	<i>Joie</i>	<i>Fierté</i>	<i>Admiration</i>	<i>Reconnaissance</i>	<i>Auto-satisfaction</i>	<i>Adoration</i>
Intensité moyenne	4	3	5	5	5	10	3	3	3	5	13
Intensité extrême	3	3	5	5	4	11	3	4	4	6	6
Émotion	<i>Pitié</i>	<i>Jalousie</i>	<i>Peur</i>	<i>Peur confir.</i>	<i>Déception</i>	<i>Peine</i>	<i>Culpabilis.</i>	<i>Indignation</i>	<i>Colère</i>	<i>Remords</i>	<i>Dégoût</i>
Intensité moyenne	4	4	10	3	4	7	5	3	7	3	17
Intensité extrême	5	3	7	5	6	8	5	6	7	4	14

Dans notre traitement, nous avons suivi la démarche de Shwartz et Weinberger (1980), présentée au début de cette section. Dans cette démarche, les auteurs ont effectué des

modifications syntaxiques sur les situations sans nuire au sens ou aux idées proposées et ont pu obtenir un ensemble ne contenant pas de redondances. La distribution selon l'émotion et l'intensité des 250 situations est présentée dans le Tableau 3.5. Chaque case du tableau indique le nombre de situations proposées et supposées induire une émotion donnée avec une intensité particulière (en moyenne 5). Les 250 situations sont présentées dans l'Annexe II de ce mémoire.

3.2.2 Description formelle des situations émotionnelles

Les informations décrivant les 250 situations émotionnelles collectées devaient être présentées à la machine. Pour ce faire, dans un premier temps, une formalisation de la description digitale ou numérique de ces situations s'imposait en déterminant les valeurs possibles que peuvent prendre les différentes *variables d'évaluation-intensité*. Dans un deuxième temps, une transformation de l'information de chacune des situations textuelles en un vecteur de valeurs qui lui est relatif était nécessaire. Dans ce qui suit, nous présentons la formalisation utilisée et la procédure mise en place pour la transformation numérique des 250 situations.

(A) Formalisation de la description des situations

Comme indiqué à la section 3.1.5, la description d'une situation se fait à l'aide des trois *variables globales* (Tableau 3.3) et d'un sous-ensemble des *variables centrales* et *locales*. Ainsi, nous avons défini pour chaque variable l'ensemble de valeurs qu'elle peut prendre :

- Les *variables globales*, *Surprise*, *Sens de réalité* et *Dimension temporelle*, prennent des valeurs booléennes, « vrai » ou « faux » (une situation est attendue ou inattendue, réelle ou irréaliste, contenant ou ne contenant pas une dimension temporelle).
- La plupart des *variables centrales* et *locales* prennent des valeurs traduisant leur degré d'intensité, comme le degré de *désirabilité* ou *non-désirabilité* d'un événement, le degré d'*approbation* ou de *condamnation* de l'action d'un agent, le degré d'*attirance* ou de *répugnance* envers un objet, etc. Ainsi, nous avons défini, pour la plupart des variables, cinq valeurs possibles correspondant aux cinq degrés d'intensité (voir Tableau 3.6):

- o La valeur « 1 » correspond à une intensité positive extrême de la variable (une *désirabilité* extrême de l'occurrence d'un évènement, une *approbation* totale de l'action d'un agent, etc.)
- o La valeur « -1 » correspond à une intensité extrêmement négative de la variable (*non-désirabilité* extrême, une *condamnation* ou *non-approbation* totale, etc.)
- o La valeur « 0.5 » correspond à une intensité positive moyenne de la variable (une *désirabilité* moyenne, une *attirance* moyenne, etc.).
- o La valeur « -0.5 » correspond à une intensité négative moyenne de la variable (une *non-désirabilité* moyenne, une *répugnance* ou *non-attirance* moyenne, etc.).
- o La valeur « 0 » correspond à un degré de neutralité, qui traduit le fait que la variable n'est pas incluse dans la description d'une situation donnée.

Il existe des exceptions comme la variable *Puissance du lien* qui prend seulement trois valeurs selon que l'on est extrêmement, moyennement, ou bien pas du tout lié à la personne qui agit.

Tableau 3.6 Les valeurs prises par les variables d'évaluation-intensité

	Variable d'évaluation	Valeurs prises :
Variables globales	<i>Surprise</i>	Vrai (1) ou Faux (0)
	<i>Sens de réalité</i>	Vrai (1) ou Faux (0)
	<i>Dimension temporelle</i>	Vrai (1) ou Faux (0)
Variables centrales	<i>Désirabilité</i>	1, 0.5, 0, -0.5, -1
	<i>Approbation</i>	1, 0.5, 0, -0.5, -1
	<i>Attirance</i>	1, 0.5, 0, -0.5, -1
Variables locales	<i>Désirabilité par autrui</i>	1, 0.5, 0, -0.5, -1
	<i>Estime pour autrui</i>	1, 0.5, 0, -0.5, -1
	<i>Mérite d'autrui</i>	1, 0.5, 0, -0.5, -1
	<i>Vraisemblance</i>	1, 0.5, 0, -0.5, -1
	<i>Réalisation</i>	1, 0.5, 0, -0.5, -1
	<i>Effort</i>	1, 0.5, 0, -0.5, -1
	<i>Agent</i>	Soi (1) ou Autrui (0)
	<i>Puissance du lien</i>	1,0.5,0
	<i>Déviation</i>	1, 0.5, 0, -0.5, -1
	<i>Disposition</i>	1,0.5,0
<i>Familiarité</i>	1, 0.5, 0, -0.5, -1	

Ainsi, une situation peut être décrite numériquement à l'aide d'un vecteur de valeurs correspondant aux intensités des *variables d'évaluation-intensité*. Selon le type de situation,

les variables qui sont impliquées (voir Tableau 3.4) prennent des valeurs différentes. Comme illustration, nous reprenons l'exemple de la situation introduite à la section 3.1.5 décrivant une personne, qui après avoir passé des dizaines d'entrevues et être restée plusieurs mois au chômage, a réussi à obtenir un emploi alors qu'elle commençait à perdre espoir. En utilisant les variables de description relatives à un évènement ayant des conséquences sur les plans futurs, et suivant la description *semi-formelle* de cette situation établie à la section 3.1.5, la situation peut être décrite formellement comme suit :

- la variable *Surprise* prend la valeur « 0 » pour exprimer le fait que la situation était inattendue.
- la variable *Sens de réalité* prend la valeur « 1 » pour exprimer le caractère réel de la situation.
- la variable *Dimension temporelle* prend la valeur « 1 » pour exprimer l'existence d'un évènement futur influant sur la situation présente (décrocher un emploi aidera à gagner de l'argent dans un futur proche).
- la variable *Désirabilité* prend la valeur « 1 » pour exprimer une désirabilité extrême de l'occurrence de l'évènement « décrocher un emploi » par la personne.
- la variable *Vraisemblance* prend la valeur « -1 » pour exprimer l'aspect totalement invraisemblable de la réalisation de l'évènement souhaité aux yeux de la personne vivant la situation.
- la variable *Réalisation* prend la valeur « 1 » afin d'exprimer le fait que l'évènement souhaité est finalement totalement réalisé.
- La variable *Effort* prend également la valeur extrême « 1 » pour les grands efforts déployés afin d'obtenir le boulot (avoir passé des dizaines d'entrevues).

Ainsi, pour cette situation, nous avons présenté une transformation possible de la situation émotionnelle en un vecteur de valeurs possibles en affectant des valeurs aux 7 variables *d'évaluation-intensité* impliquées, et en laissant les autres à leur valeur par défaut, à savoir « 0 ».

La question qui se pose maintenant est la suivante : Comment décrire ou transformer numériquement les 250 situations collectées de cette manière et comment valider ces transformations ? Les éléments de réponse à cette question se trouvent traités à la section suivante.

(B) Transformation numérique des situations émotionnelles

Pour la transformation numérique des 250 situations émotionnelles décrites textuellement, nous avons mis en place une procédure d'évaluation qui permet d'attribuer à chaque situation un vecteur de valeurs. Ce vecteur doit traduire l'évaluation commune que ferait la majorité des individus de la situation décrite, selon les *variables d'évaluation-intensité*. Pour avoir cette information, nous avons choisi de recueillir l'avis de deux ou trois personnes différentes sur chaque situation. Ainsi, dans une étape initiale, une situation est décrite numériquement, un certain nombre de fois, par quelques personnes, avant qu'une description finale ne lui soit assignée. La description finale contient les valeurs majoritaires dans les différentes descriptions initiales. La procédure est détaillée dans les paragraphes qui suivent.

Comment la plupart des gens évalueront-ils cette situation ?

Après avoir passé des dizaines d'entrevues de boulot sans succès et être resté plusieurs mois au chômage, on réussit à avoir un travail alors qu'on commençait à perdre espoir.

Cette situation correspond à:	L'occurrence de l'événement a des conséquences sur :
L'occurrence d'un événement * <input checked="" type="radio"/>	Autrui (non pas sur vous) <input type="radio"/>
L'action d'un agent * <input type="radio"/>	Vos plans et perspectives <input checked="" type="radio"/>
Objet * <input type="radio"/>	Votre bien-être <input type="radio"/>
L'occurrence d'un événement suite à une action d'un agent * <input type="radio"/>	

L'action de l'agent, approuvée ou désapprouvée, engendre un événement qui a des conséquences sur soi ce qui génère un sentiment de plaisir ou déplaisir.

Figure 3.3 Interface de sélection du type de la situation

Afin de permettre à plusieurs personnes de décrire numériquement les situations, nous avons développé des interfaces Web en utilisant la technique *Java Server Pages* (JSP)⁷. Cette technique permet la création des pages Web en se basant sur des balises permettant d'introduire, au sein d'une page *HTML*, du code *Java* pour un traitement donné. Ce code *Java* est compilé côté serveur et génère un contenu *HTML* de manière dynamique côté client. Nous avons utilisé, par conséquent, un serveur Web *Apache Tomcat*⁸, dans sa version 5.0, qui implémente les spécifications de *JSP*. Les descriptions numériques des

⁷ <http://java.sun.com/products/jsp/>

⁸ <http://tomcat.apache.org/>

situations ont été enregistrées dans la base de données MySQL⁹. Nous décrivons dans ce qui suit cette interface.

Dans un premier temps, nous présentons à la personne la situation émotionnelle à décrire (voir Figure 3.3), et nous lui demandons d'indiquer son type (c.-à-d., *évènement*, *action d'un agent* et *aspect d'un objet*) en sélectionnant un bouton radio. Afin que cette personne puisse comprendre la signification d'un *évènement*, d'un *objet* ou de *l'action d'un agent*, etc., nous lui suggérons de placer la souris sur ces termes pour qu'une bulle explicative apparaisse (voir Figure 3.3 la bulle en jaune). Selon le choix fait, de nouvelles interfaces sont présentées contenant les variables globales et les variables centrales et locales impliquées dans la description du type de la situation (Tableau 3.4). Ainsi, une deuxième interface est présentée contenant les variables globales et leurs valeurs possibles (voir Figure 3.4). La personne donne son avis sur la manière avec laquelle la plupart des gens évalueraient la situation selon ces variables (attendue ou inattendue, réelle ou imaginaire, etc.). Des explications sur la signification des variables, comme celles présentées dans le Tableau 3.3, sont affichées lors du placement de la souris sur les termes.

Comment la plupart des gens évalueront-ils cette situation ?

Après avoir passé des dizaines d'entrevues de boulot sans succès et être resté plusieurs mois au chômage, on réussit à avoir un travail alors qu'on commençait à perdre espoir.

Expectation *	Sens de la réalité *	Composante Temporelle *
Inattendue/Imprévisible <input checked="" type="radio"/>	Réelle/Confirmée <input checked="" type="radio"/>	Oui <input checked="" type="radio"/> Est-ce que un événement passé ou futur affecte la situation présente (décrite) ?
Attendue/prévisible <input type="radio"/>	Irréelle/Imaginaire/Non confirmée <input type="radio"/>	Non <input type="radio"/>

Figure 3.4 Interface de description de la situation avec les variables globales

Une fois les valeurs des variables globales attribuées, une troisième interface apparaît (Figure 3.5). Cette interface contient les variables centrales et locales impliquées dans la description selon le type de situation choisi dans la première interface. Si le choix initial avait été, par exemple, *l'action d'un agent*, les variables présentées dans cette troisième

⁹ <http://www.mysql.com/>

interface auraient été *Approbation*, *Agent*, *Puissance du lien* et *Déviation*. Comme pour les interfaces précédentes, les valeurs des variables peuvent être choisies grâce aux boutons radios correspondants aux différentes intensités possibles. La personne rapporte ce qu'elle croit être l'évaluation de la situation selon la majorité des gens.

Comment la plupart des gens évalueront-ils cette situation ?

Après avoir passé des dizaines d'entrevues de boulot sans succès et être resté plusieurs mois au chômage, on réussit à avoir un travail alors qu'on commençait à perdre espoir.

Desirabilité *	Vraisemblance *	Réalisation *	Effort *
À quel point l'occurrence de l'événement anticipé peut être désirée ou non désirée?		Totalement réalisé <input checked="" type="radio"/>	Efforts extrêmes <input checked="" type="radio"/>
Quelque part désirable <input type="radio"/>	Quelque part vraisemblable <input type="radio"/>	Quelque part réalisé <input type="radio"/>	Efforts moyens <input type="radio"/>
Indifférence <input type="radio"/>	Pas de réponse (inconnu) <input type="radio"/>	Pas de réponse (inconnu) <input type="radio"/>	Pas de réponse (inconnu) <input type="radio"/>
Quelque part indésirable <input type="radio"/>	Quelque part invraisemblable <input type="radio"/>	Quelque part irréalisé <input type="radio"/>	Peu d'efforts <input type="radio"/>
Extrêmement indésirable <input type="radio"/>	Extrêmement invraisemblable <input checked="" type="radio"/>	Totalement irréalisé <input type="radio"/>	Aucun effort <input type="radio"/>

Figure 3.5 Interface de description d'une situation avec les variables locales

Une fois la saisie complétée, les choix effectués dans les trois interfaces sont transformés en un vecteur de valeurs numériques décrivant la situation, et ce, en suivant la formalisation présentée à la section précédente. Ce vecteur est enregistré dans la base MySQL. Lorsqu'une deuxième et une troisième personne décrivent la même situation d'une manière différente, ces descriptions sont également enregistrées.

Comme indiqué ci-haut, une description unique et finale est assignée à la situation en attribuant les valeurs majoritaires des variables dans ces différentes descriptions. Par exemple, si deux personnes parmi trois jugent que la situation correspond à un événement affectant les projets, les variables centrales et locales impliquées dans la description finale sont celles relatives à ce type (c.-à-d., Désirabilité, Vraisemblance, Réalisation, Effort). Si les trois types sont différents, alors un type au hasard, parmi les trois, est choisi. Le même processus est répété pour les valeurs des variables : la valeur commune parmi les descriptions est assignée, sinon une d'entre elles est choisie aléatoirement. Par exemple si deux personnes sur trois décrivent la situation comme étant extrêmement désirable, la variable désirabilité prend cette valeur dans la description finale. Au final, une description numérique de la situation traduisant son évaluation générale, par la plupart des gens, est obtenue.

Pour décrire les 250 situations de cette manière, nous avons demandé à quelques étudiants et membres du personnel de l'Université de Montréal de décrire ces situations en ligne, en remplissant ces interfaces. Ainsi, chaque participant décrit 20 situations différentes parmi les 250. Un total de 35 participants, 21 hommes et 14 femmes, a été enregistré. Quatre participants étaient âgés de moins de 18 ans, seize avaient entre 18 et 25 ans, douze entre 26 et 45 ans et trois étaient âgés entre 46 ans et 65 ans. Ainsi, nous avons obtenu 700 (35 x 20) descriptions numériques, où chaque situation est décrite au minimum deux fois. Puis, une description finale a été attribuée à chaque situation, à partir de ces 700 descriptions collectées. Comme il a été expliqué dans le paragraphe précédent, la démarche de description numérique consiste à recenser quelques avis sur les descriptions possibles d'une situation et de retenir les avis majoritaires. Ainsi, une description donnée d'un participant peut être assimilée à un vote. De même, les 35 personnes peuvent être assimilées à une sorte de *comité* qui a pu servir à valider les descriptions finales.

Cette étape a permis de construire une base contenant les descriptions numériques finales caractérisant et identifiant chacune des 250 situations collectées pour servir comme entrées ou ensemble de connaissances qui alimente notre *machine émotionnelle*.

3.3 Conclusion

Dans ce chapitre, nous avons répondu à la première question que nous nous sommes posée dans le chapitre introductif et qui consiste, à savoir, d'abord, comment une machine peut « comprendre » ce qu'est une situation émotionnelle ? Ensuite comment peut-on décrire une situation au sein d'une machine à l'aide d'une combinaison de valeurs numériques ?

Dans un premier temps une formalisation de l'information, relative à une situation, s'est avérée nécessaire. Cette formalisation a été rendue possible par le recours aux théories d'évaluation cognitive. Plus spécifiquement nous nous sommes basés sur les *variables d'évaluation-intensité* du modèle OCC. La représentation de l'information extraite d'une situation à l'aide de ces variables traduit la manière avec laquelle la plupart des gens évaluent cette situation. Par conséquent, nous avons mis en place une procédure empirique qui a permis, d'abord, de collecter selon les différentes émotions, 250 situations textuellement décrites. Ensuite, ces situations ont été évaluées par plusieurs personnes en donnant des valeurs aux variables OCC. Ces valeurs traduisent l'évaluation de la majorité

des gens. Au final, ces descriptions ont été transformées en des vecteurs numériques uniques qui caractérisent ou identifient chacune des situations.

A l'aide de cette procédure, nous avons présenté une réponse possible à une problématique qui n'a pas été traitée dans le modèle OCC, à savoir, la description d'une situation à l'aide d'une combinaison particulière des valeurs des variables. L'inférence de l'émotion et de son intensité pour une combinaison particulière est traitée dans le chapitre 4. La procédure empirique élaborée nous a également permis de créer une base de situations émotionnelles qui est nécessaire pour le fonctionnement de notre *machine émotionnelle*. Cette dernière nécessite un nombre minimum d'expériences ou d'exemples pour s'entraîner et effectuer des inférences de façon convenable. La machine a également besoin d'une connaissance consistante sur les réactions émotionnelles éventuelles de plusieurs personnes face à des situations différenciées. Le chapitre qui suit décrit comment cette connaissance est acquise, et détaille l'architecture de notre machine.

Chapitre 4 Architecture et exploitation de la machine émotionnelle

Tout au long des chapitres précédents, nous avons indiqué que l'objectif du mémoire est d'implémenter une *machine émotionnelle* capable de prédire l'émotion de l'individu et son intensité à partir de l'information relative à la situation qu'il vit. Une tâche de prédiction ou de classification nécessite des données *d'entraînement* (voir Chapitre 2). Pour cette raison, nous avons construit une base contenant des informations sur différentes situations émotionnelles dans un langage compréhensible par la machine (voir Chapitre 3). Néanmoins, ces informations ne sont pas suffisantes pour construire un *ensemble d'entraînement*. Les données sur les réactions possibles de différentes personnes face à ces situations sont également nécessaires pour la prédiction.

Ainsi, ce chapitre présente, dans un premier temps, la collecte des réactions émotionnelles pour la constitution d'un *ensemble d'entraînement* qui servirait d'entrée aux algorithmes d'apprentissage. Dans un deuxième temps, nous présentons l'architecture de notre *machine émotionnelle* et la configuration établie avec les algorithmes d'apprentissage. Nous présenterons en dernier lieu l'exploitation de cette application par un groupe de personnes et les résultats de prédiction obtenus.

4.1 Construction de l'ensemble d'entraînement

Dans le cadre d'un apprentissage supervisé (voir Chapitre 2), un ensemble d'entraînement contient plusieurs exemples du phénomène étudié. Chaque exemple constitue un couple de variables d'entrée et de sortie. Les valeurs des variables d'entrée correspondent à des caractéristiques décrivant un cas particulier du phénomène, et celles de sortie contiennent les réponses associées à ce cas.

Pour la machine émotionnelle proposée dans ce travail, l'ensemble d'entraînement contient, d'un côté, les variables d'entrée caractérisant une personne et la situation émotionnelle à laquelle il fait face, et d'un autre côté, les variables de sortie *émotion* et *intensité* qui leur correspondent. En effet, l'information spécifique à la personne est nécessaire puisque les réactions émotionnelles face une situation ne sont pas les mêmes

pour tous les individus, et deux personnes peuvent avoir deux émotions différentes pour une même situation (Scherer, Wrantik, Sangsue, Tran, & Scherer, 2004).

4.1.1 Importance des informations personnelles

Plusieurs psychologues affirment que les réponses émotionnelles chez les humains sont fonction de plusieurs facteurs distincts. Hess (2001) comme d'autres chercheurs (Bradley, Codispoti, Sabatinelli, & Lang, 2001; Fujita, Diener, & Sandvik, 1991) affirme que la réaction peut varier selon le genre de la personne vivant la situation (homme ou femme). De même, Scherer (2004) indique que l'âge influe sur l'émotion et particulièrement son intensité. Les personnes âgées sont généralement moins émotives que les jeunes personnes. Un autre facteur important est la culture (Diener, Oishi, & Lucas, 2003; Scherer, 1997) de la personne. Ainsi, les facteurs économiques, le degré d'urbanisation, les croyances culturelles, ou encore la religion peuvent jouer un rôle dans la différenciation de l'émotion chez les individus. Mais le facteur le plus étudié comme cause de différenciation de la réponse émotionnelle, reste la personnalité (Scherer, Wrantik, Sangsue, Tran, & Scherer, 2004). Les psychologues (p. ex., Arnold, 1960; Larsen & Ketelaar, 1991; Watson & Clark, 1992) affirment que les individus diffèrent selon leur prédisposition à vivre l'émotion. Quelqu'un de névrotique aurait tendance à s'énerver et à s'inquiéter plus qu'un autre (Rusting & Larsen, 1997).

Ainsi, les informations sociodémographiques et celles relatives à la personnalité de l'individu sont nécessaires pour la reconnaissance de l'émotion et de son intensité. En psychologie, plusieurs travaux stipulent que la personnalité peut être décrite ou caractérisée par des traits ou des dimensions (Bouvard, 2009). Les modèles proposés dans la littérature présentent un nombre varié de dimensions. À titre d'exemple, ceux des *big five* (p. ex., Goldberg, 1993; McCrae & Costa, 1985) présentent cinq dimensions alors que le modèle de Eysenck (Eysenck & Eysenck, 1975) n'en présente que trois. Le point commun qu'on retrouve dans la majorité des modèles est la présence des deux dimensions de *névrosisme* (ou *neuroticisme* selon la traduction de Bouvard (2009) du mot anglais *neuroticism*), et d'*extraversion* (Bouvard, 2009). La mesure de ces dimensions peut se faire à l'aide des questionnaires de personnalité tels que EPQ (Eysenck & Eysenck, 1975). Pour des questions pratiques, des versions abrégées ont été proposées et validées. Un exemple de ce type est le questionnaire EPQR-A (*Abbreviated form of the Revised Eysenck Personality*

Questionnaire en anglais) (Francis, Brown, & Philipchalk, 1992) qui correspond à une version réduite du questionnaire d'Eysenck EPQ (Eysenck, Eysenck, & Barrett, 1985). Cette version a été validée sur des groupes d'étudiants en Angleterre, au Canada, aux États-Unis et en Australie. Le questionnaire contient 24 questions permettant de mesurer quatre dimensions indépendantes :

Le *névrosisme* (*neuroticism* en anglais), symbolisé par la lettre « N », correspond à une hyperactivité émotionnelle en général (Bouvard, 2009), c'est-à-dire une hypersensibilité et une prédisposition à l'anxiété, à l'instabilité, au manque de confiance en soi, au ressenti d'émotions négatives, etc.

L'*extraversion* (*extraversion* en anglais), « E », caractérise particulièrement les personnes hyperactives, dynamiques, qui ont tendance à extérioriser leurs sentiments, souvent à la recherche des sensations fortes, ayant une aisance à établir des liens sociaux, etc. (Bouvard, 2009).

La *dimension psychotique* (*psychoticism* en anglais), « P », selon la traduction de Bouvard (2009), reflète une prédisposition à un comportement antisocial, à une impulsivité, une agressivité, un comportement égoïste et indifférent, une dureté de l'esprit, etc.

L'*échelle de mensonge* (*Lie Scale* en anglais), symbolisée par la lettre « L », est une dimension qui a été rajoutée aux trois précédentes afin de vérifier que les réponses données par le sujet au questionnaire ne sont pas erronées ou biaisées. En d'autres termes, elle mesure le degré de validité de ces informations. Ce questionnaire que nous retenons dans notre recherche comporte des questions qui permettent de l'estimer.

Une fois que le sujet a rempli le questionnaire, quatre notes indépendantes lui sont attribuées à partir de ses réponses aux quatre dimensions. Chaque note peut prendre une valeur entière entre 0 et 6. Ainsi, la personnalité peut être modélisée par un vecteur contenant quatre scores.

Dans le cadre de notre mémoire, nous utilisons ce questionnaire de 24 questions principalement pour recueillir des informations sur la personnalité des individus. Ces informations seront utiles pour améliorer la capacité prédictive de la machine. L'Annexe III présente le questionnaire EPQR-A dans sa version traduite en français (Bouvard, 2009), avec le calcul des scores effectué à partir des réponses.

4.1.2 Collecte des informations personnelles et des réactions émotionnelles

L'objectif ici est de construire l'ensemble d'entraînement avec plusieurs cas. Ces cas comportent d'un côté des informations personnelles (sociodémographiques + personnalité) sur des individus et des situations qu'ils peuvent rencontrer. De l'autre côté, ils contiennent les réactions émotionnelles et leur intensités. Pour recueillir ces cas, nous avons demandé au comité composé des 35 personnes qui ont contribué à la numérisation de l'information des 250 situations collectées (voir la sous-section « transformation numérique des situations émotionnelles » de la section 3.2.2 du chapitre précédent) de nous fournir les données personnelles et de nous faire part de l'émotion ressentie face à chaque situation décrite.

The screenshot shows a web interface with two main sections. The top section, titled 'Informations Personnelles', contains three dropdown menus: 'Age' (set to '18-25 ans'), 'Sexe' (set to 'Masculin'), and 'Religion' (set to 'Bouddhisme'). Below this is a section titled 'Questionnaire de Personnalité' which includes a set of instructions: 'Veillez répondre à chacune des questions en sélectionnant le bouton radio "Oui" ou "Non" qui suit la question. Il n'y a pas de bonne ou mauvaise réponse ni de question piège. Répondez rapidement et ne passez pas trop de temps à réfléchir au sens précis de la question. Veillez répondre à toutes les questions'. Below the instructions is a table of 18 questions, each with two radio buttons for 'Oui' and 'Non'. The 'Oui' button is selected for questions 4, 5, 8, 11, 13, 16, and 17. The 'Non' button is selected for questions 1, 2, 3, 6, 7, 9, 10, 12, 14, 15, and 18.

Question	Oui	Non
1. Votre humeur subit-elle souvent des hauts et des bas?	<input type="radio"/>	<input checked="" type="radio"/>
2. Aimez-vous beaucoup parler?	<input type="radio"/>	<input checked="" type="radio"/>
3. Seriez-vous mal à l'aise si vous aviez des dettes?	<input type="radio"/>	<input checked="" type="radio"/>
4. Êtes-vous plutôt quelqu'un plein d'entrain?	<input checked="" type="radio"/>	<input type="radio"/>
5. Vous est-il arrivé d'avoir si grande envie de quelque chose que vous vous êtes arrangé(e) pour vous servir plus que la part qui vous revenait?	<input checked="" type="radio"/>	<input type="radio"/>
6. Accepteriez-vous de prendre des drogues qui pourraient avoir des effets étranges ou dangereux?	<input type="radio"/>	<input checked="" type="radio"/>
7. Vous est-il arrivé de blâmer quelqu'un pour une faute, sachant que c'est vous qui l'aviez commise?	<input type="radio"/>	<input checked="" type="radio"/>
8. Préférez-vous faire les choses à votre façon plutôt que d'agir en fonction des règles?	<input checked="" type="radio"/>	<input type="radio"/>
9. Vous arrive-t-il souvent d'en "avoir marre"?	<input type="radio"/>	<input checked="" type="radio"/>
10. Vous est-il arrivé de prendre quelque chose appartenant à quelqu'un d'autre (ne serait-ce qu'une épingle ou un bouton)?	<input type="radio"/>	<input checked="" type="radio"/>
11. Vous considérez-vous comme une personne nerveuse?	<input checked="" type="radio"/>	<input type="radio"/>
12. Estimez-vous que le mariage est une institution dépassée qui devrait être abolie?	<input type="radio"/>	<input checked="" type="radio"/>
13. Savez-vous aisément animer une soirée ennuyeuse?	<input checked="" type="radio"/>	<input type="radio"/>
14. Êtes-vous facilement inquiet(e)?	<input type="radio"/>	<input checked="" type="radio"/>
15. Avez-vous tendance à rester en arrière-plan lorsque vous êtes en société?	<input type="radio"/>	<input checked="" type="radio"/>
16. Êtes-vous tourmenté(e) si vous savez qu'il ya des erreurs dans votre travail?	<input checked="" type="radio"/>	<input type="radio"/>
17. Vous est-il arrivé de tricher à un jeu?	<input type="radio"/>	<input checked="" type="radio"/>
18. Souffrez-vous de vos nerfs?	<input type="radio"/>	<input checked="" type="radio"/>

Figure 4.1 Saisie des informations personnelles et du questionnaire de personnalité

A l'aide d'une première interface (Figure 4.1), nous leur avons demandé de choisir la catégorie d'âge à laquelle ils appartiennent. Les catégories retenues sont : moins de 18 ans ; entre 18 et 25 ans ; entre 26 et 45 ans ; entre 46 et 65 ans ; 66 ans et plus. Ils devaient également indiquer leur sexe et sélectionner une religion dans le cas où ils en avaient une. Les religions retenues sont « le Bouddhisme », « le Christianisme », « le Judaïsme », « l'Islam » ou « Autres ». La religion est un indicateur, parmi d'autres, des normes culturelles

chez une personne (Scherer, 1997). Une fois ces informations introduites de manière anonyme, les 35 participants devaient répondre au questionnaire EPQR-A. Le questionnaire demande de répondre rapidement, par un oui ou un non, aux questions posées, sans passer beaucoup de temps à réfléchir au sens précis de la question (voir Annexe III).

Pour la collecte des réactions émotionnelles, nous avons demandé à chaque membre du comité des 35, après avoir décrit une situation particulière suivant la procédure présentée à la section 3.2.2, de divulguer ou rapporter l'intensité de l'émotion qu'il croit pouvoir ressentir s'il vivait cette même situation. Schwartz et al. (1980) affirment que les réponses émotionnelles imaginées sont généralement très proches des émotions réellement ressenties et que l'extraction ou la découverte de patterns est possible en utilisant ce type d'évaluation subjective de l'émotion. Le participant pouvait choisir une des deux intensités, forte ou faible, relatives à une émotion parmi les 22 émotions de OCC. Il lui était également permis d'exprimer l'absence d'émotion face à la situation en ne choisissant aucune des 22 émotions. Ainsi, chacune des 35 personnes a pu fournir ses informations personnelles puis donner ses réactions émotionnelles face à 20 situations. Ces 20 situations sont celles que cette même personne avait auparavant décrites pour les besoins de transformation numérique des situations (voir la sous-section « transformation numérique des situations émotionnelles » de la section 3.2.2 du chapitre précédent). Par conséquent, nous avons obtenu un ensemble d'entraînement de 700 (35 x 20) exemples. Chaque exemple i est formé d'une part d'un vecteur relatif aux variables d'entrée $\mathbf{x}_i (x_{i1}, x_{i2}, \dots, x_{i24})$ contenant 24 valeurs. Ces 24 variables correspondent aux 3 variables sociodémographiques, c'est-à-dire l'âge, le sexe, la religion, aux 4 notes relatives aux quatre traits de personnalité d'un individu (voir section 4.1.1), et aux 17 variables de OCC présentées dans le tableau 3.6, permettant de caractériser une situation particulière S_j parmi les 250 collectées (voir section 3.2.1). D'autre part, les variables cibles de cet exemple sont représentées par un vecteur $\mathbf{y}_i (y_{i1}, y_{i2}, \dots, y_{i25})$ contenant 25 valeurs relatives à l'émotion et l'intensité qu'éprouve l'individu i face à la situation S_j . Ces 25 variables regroupent les 23 états émotionnels possibles (22 émotions + état « pas d'émotion ») et les 2 intensités moyenne et forte.

L'émotion et son intensité ont été saisies au moyen d'une interface (voir Figure 4.2). Cette interface contient 11 échelles où chaque échelle correspond à une paire d'émotions

opposées du modèle OCC (voir section 3.1.2) (p.ex., joie et peine). Chacune des échelles contient cinq seuils possibles.

- Les deux seuils en haut de l'échelle correspondent aux intensités moyenne et forte de l'émotion positive de la paire (p.ex., joie moyenne et joie forte).
- Les deux seuils en bas de l'échelle correspondent aux intensités moyenne et forte de l'émotion négative de la paire (p.ex., peine moyenne et peine forte).
- Le seuil de neutralité, placé au milieu et dans lequel se trouve initialement le curseur de l'échelle, traduit le fait de ne sentir aucune des deux émotions.

Veillez indiquer votre émotion et son intensité lorsque vous faites face à cette situation :

Après avoir passé des dizaines d'entrevues de boulot sans succès et être resté plusieurs mois au chômage, on réussit à avoir un travail alors qu'on commençait à perdre espoir.

Heureux pour	Exultation	Espoir	Satisfaction	Soulagement	Joie	Fierté	Admiration	Reconnaissance	Autosatisfaction	Adoration
Pitié	Jalousie	Peur	Peur confirmée	Déception	Peine	Culpabilisation	Indignation	Colère	Remords	Dégoût

Figure 4.2 Interface de saisie de l'émotion et son intensité

Chaque émotion est expliquée par ses substituts (voir Tableau 3.2) qui apparaissent dans une bulle jaune au passage de la souris sur le terme (voir Figure 4.2). Selon certains psychologues (*Scherer & Tannenbaum, 1986; Schwartz & Weinberger, 1980*), le fait qu'un individu exprime son émotion face à une situation qu'il imagine vivre, ne devrait pas contenir un grand biais suite à une mauvaise interprétation des termes émotionnels.

Une seule émotion avec une seule intensité pouvait être sélectionnée puisque notre travail ne tient pas compte de la cooccurrence des émotions. Ce choix est dû principalement à des raisons de simplification de la tâche ultérieure de reconnaissance d'émotions, puisque dans un cas de cooccurrence, le nombre de variables cibles augmente, de même que les combinaisons de leurs valeurs possibles. Dans ce cas, une bonne classification nécessiterait une masse importante de données comportant plusieurs exemples de combinaisons.

À la fin de cette étape, nous disposons d'un ensemble d'apprentissage qui nous permettait de mettre en place notre *machine émotionnelle*. Nous présentons dans ce qui suit le principe de fonctionnement de la machine, son architecture et son exploitation par quelques personnes, et ce afin de tester sa performance.

4.2 Le fonctionnement de la Machine émotionnelle

L'objectif de la *machine émotionnelle*, comme nous l'avons mentionné tout au long de ce mémoire, est de prédire de manière instantanée l'émotion et l'intensité que peut exprimer un individu face à une nouvelle situation S_p donnée. Comme mentionné à la fin du chapitre introductif, pour générer une telle prédiction, notre application combine deux prédictions distinctes calculées à partir de deux types d'informations différentes.

- La première prédiction se base sur l'historique des réactions émotionnelles de l'individu face à des situations similaires à S_p .
- La deuxième concerne les réactions émotionnelles de d'autres individus face à la situation S_p elle-même ou à des situations similaires. Cette information représente une sorte de sens commun, relatif aux réactions émotionnelles typiques, ou de patrimoine émotionnel collectif.

L'objectif consiste alors à expérimenter cette machine sur des données réelles fournies par des individus. Au cours de cette expérimentation, pour chaque individu, la machine procède de la manière suivante.

- Elle choisit au hasard une situation non vécue S_p parmi les 250 situations (voir section 3.2.2) dont elle dispose.
- Elle saisit les informations sociodémographiques et la personnalité de l'individu à l'aide d'un questionnaire.
- Elle utilise une information sur l'historique des réactions de l'individu face à des situations similaires à S_p pour générer le premier type de prédiction.
- Elle utilise l'ensemble d'entraînement construit (voir section précédente 4.1), regroupant des réactions émotionnelles de plusieurs personnes face à différentes situations, pour générer le deuxième type de prédiction.

Cependant, au cours de l'expérimentation, notre machine ne dispose pas de l'historique des réactions d'un nouvel individu face à des situations similaires à S_p pour pouvoir générer le premier type de prédiction. Une manière de faire consiste à simuler cet historique en demandant à l'individu de nous faire part de ses émotions face à des situations similaires à S_p choisies parmi les 249 situations restantes. Comme indiqué dans l'introduction du mémoire et au chapitre 2, lorsque l'application est intégrée dans un système de reconnaissance d'émotions plus complet, l'historique d'un utilisateur permanent du système devrait être disponible et être procuré par les méthodes de reconnaissances faciales, sonores physiologiques ou autres.

L'approche de prédiction par similarité (ou analyse des similarités), que nous retenons pour produire le premier type de prédiction, est très répandue. Elle est utilisée dans les analyses économiques, en marketing, en physique, en chimie, en biologie, etc. C'est une méthode de prédiction des événements futurs ou inconnus sur la base de vécus similaires ou comparables. En marketing, par exemple, on utilise ce type d'analyse pour la prévision et la reconnaissance des préférences des consommateurs et leurs réactions émotionnelles pour des produits (objets) nouveaux. Pour ces objets inconnus, il n'y a généralement pas de données historiques ou d'habitudes d'achat qu'on pourrait extrapoler. Dans ce cas de figure, même si on sollicite les individus à rapporter les émotions ou préférences, ces derniers se trouveront dans une situation d'incapacité à pouvoir les exprimer. C'est pour cette raison qu'on fait appel à des comportements d'achat d'objets similaires pour prédire les préférences ou émotions ressenties pour ces nouveaux objets. En biologie, par exemple, on utilise des gènes connus pour prédire de nouveaux gènes. Derrière cette méthode, on fait l'hypothèse implicite que des organismes similaires ont probablement des gènes comparables. Donc, il est possible de prédire des gènes inconnus d'une espèce donnée à partir de gènes d'espèces similaires.

4.2.1 Principe général

La machine dispose comme entrée:

- les informations personnelles d'un utilisateur ;
- une nouvelle situation S_p décrite numériquement (voir section 3.2.2) ;
- l'historique des réactions de l'utilisateur dans des situations similaires à S_p ;

- les réactions émotionnelles de plusieurs individus face à différentes situations présentes dans l'ensemble D d'entraînement (voir section 4.1). Celui-ci est intégré dans la machine et constamment mis à jour au fur et à mesure qu'un nouvel individu utilise la machine.

La machine produit comme sorties des prédictions de l'émotion et de l'intensité que ressentirait l'individu face à S_p de la manière suivante.

1. Elle calcule une première prédiction spécifique orientée utilisateur.
 - Elle calcule la distance euclidienne entre la description numérique de S_p et les descriptions des autres situations dont elle dispose.
 - Elle choisit les situations les plus similaires.
 - Elle utilise l'historique des réactions de l'individu face aux situations choisies pour appliquer l'algorithme *des k plus proches voisins pondérés* (voir section 2.3 du chapitre 2).
2. Elle calcule une deuxième prédiction générale traduisant les réactions émotionnelles typiques.
 - Elle utilise l'ensemble d'entraînement D pour entraîner ou bien pour mettre à jour un réseau de neurones artificiel (voir section 2.3 du chapitre 2).
 - Elle fournit les informations personnelles de l'individu et la description numérique de S_p comme des entrées au réseau appris (entraîné), qui fournit en retour sa prédiction.
3. Elle combine les deux prédictions pour synthétiser à la fois la connaissance spécifique sur les réactions habituelles de l'utilisateur et la connaissance préexistante sur les réactions typiques des individus. Les détails de cette combinaison sont expliqués à la section qui suit.

Il est important de signaler qu'au cours de ces différentes étapes, la machine tient compte des informations saisies par un utilisateur, adapte et met à jour l'ensemble d'entraînement et le modèle général des *réseaux de neurones* de manière continue et instantanée. Ces mises à jour sont utilisées ultérieurement pour établir des prédictions pour

d'autres individus. Dans ce qui suit, l'architecture et les interfaces de la machine sont illustrées en détail.

4.2.2 Architecture

L'architecture de l'application est représentée à la Figure 4.3. L'application comporte deux composantes. La première est la *composante de connaissances* contenant toutes les données et les informations nécessaires à la reconnaissance de l'émotion et de son intensité. La deuxième composante est celle de la prédiction, *composante de prédiction*, intégrant les algorithmes des *k plus proches voisins* (*kPPV*) et des *réseaux de neurones* (RdN). Les flèches numérotées et en pointillés dans la Figure 4.3 correspondent à des flux d'informations entre et au sein des deux composantes. Les numéros indiquent un ordre chronologique des étapes. Ci-dessous, nous détaillons ces différentes étapes accompagnées d'un exemple de prédiction.

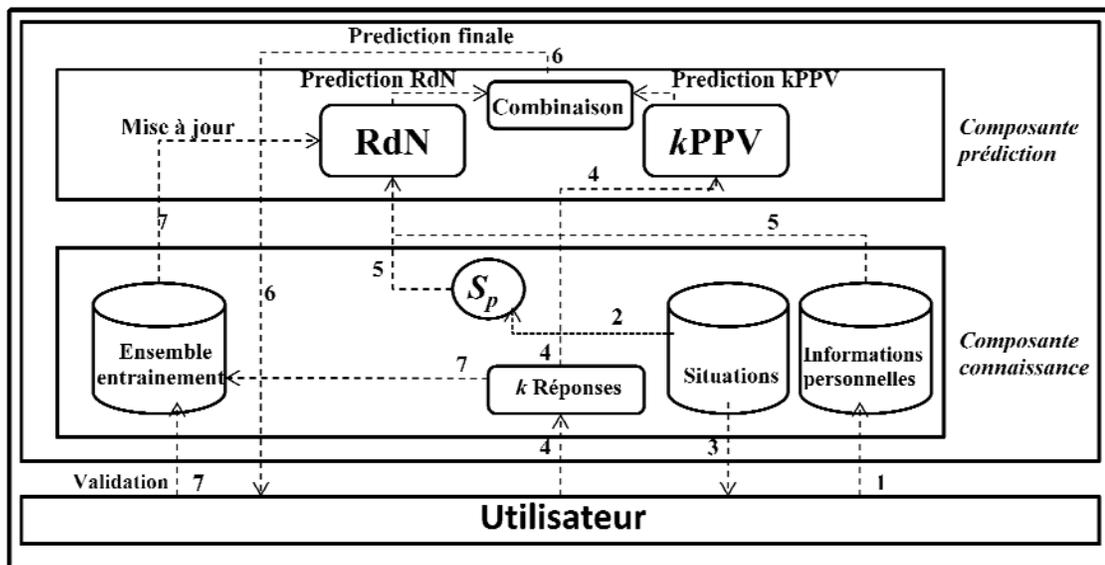


Figure 4.3 Architecture de la *machine émotionnelle*

(A) Étape 1

L'utilisateur de la machine doit entrer ses informations sociodémographiques et répondre au questionnaire de personnalité EPQR-A au moyen d'une interface similaire à celle de la Figure 4.1. Ces données, correspondant à l'âge, le sexe, la religion et les scores des quatre dimensions de personnalités (voir section 4.1.1 et Annexe III), sont enregistrées dans une base de données.

(B) Étape 2

La machine choisit au hasard (selon la loi uniforme), à partir de sa base de situations contenant les 250 situations collectées, une situation S_p . Supposons que la situation S_p choisie corresponde à la situation « *Après dix bonnes années de travail dans votre entreprise, vous êtes, comme prévu, promu au poste de directeur général* ». Les différentes étapes qui suivent viseront à prédire l'émotion et l'intensité de l'individu qui a saisi ses informations personnelles à l'étape 1 face à cette situation.

(C) Étape 3

Cette étape constitue l'étape de départ pour la première prédiction selon l'algorithme k PPV. En effet, la base de situations contient les vecteurs de descriptions numériques des 250 situations selon les 17 *variables d'évaluation-intensité* (voir chapitre 3). Elle contient, également, pour chacune de ces situations une référence sur son émotion-intensité caractéristique (choisie selon sa plus grande pertinence auprès des personnes qui ont décrit la situation, voir section 4.1.2). Cette information est utilisée pour le choix du nombre de situations similaires k , nécessaire lors de la prédiction. Dans un premier temps, un calcul de distances Euclidiennes entre le vecteur de description de la situation S_p et tous les autres vecteurs de la base est effectué. Ensuite, les k situations ayant les plus petites distances sont présentées à l'utilisateur dans une interface comme celle de la Figure 4.2.

Pour le choix de k , s'il existe des situations ayant une description identique à celle de S_p , alors toutes ces situations sont présentées à l'utilisateur. Dans le cas contraire, nous essayons de trouver le nombre de situations proches qui nous permet d'avoir la meilleure prédiction possible. Pour ce faire, nous utilisons les valeurs de l'émotion et l'intensité assignées à chaque situation dans la base de situation. Ces situations et leurs émotions-intensités correspondantes sont assimilées à un ensemble de validation (voir section 2.2.2 du chapitre 2) qui permet de choisir l'hyperparamètre k (voir section 2.3) adéquat en testant la performance selon plusieurs valeurs de ce dernier. Nous avons choisi de varier les valeurs de k entre 1 et 7. Par conséquent, le nombre maximal de situations similaires qui peuvent être présentées à l'utilisateur est de 7. Nous avons choisi ce nombre pour deux raisons. La première est qu'en moyenne le nombre des situations disponibles qui sont supposées provoquer des émotions similaires est à peu près égal à 5, avec une variance assez large, comme indiquée à la section 3.2.1 du chapitre précédent (voir tableau 3.5). La

deuxième est une raison pratique. En effet nous avons estimé qu'un nombre supérieur pourrait dissuader l'utilisateur de continuer l'expérience (puisque l'expérience est en ligne). Ainsi, la valeur retenue pour k est celle qui permet, en appliquant l'algorithme $kPPV$ pondéré, de trouver une émotion et une intensité correspondant à l'émotion et l'intensité de S_p dans la base de situations. Lorsqu'aucune valeur de k ne permet de trouver une émotion et une intensité correspondant à celles de S_p , la valeur maximale que peut prendre k est retenue. Dans ce cas, les 7 situations les plus proches sont présentées.

Considérons la situation S_p , « *Après dix bonnes années de travail dans votre entreprise, vous êtes, comme prévu, promu au poste de directeur général* ». L'application de la procédure, décrite dans le paragraphe précédent, donne $k^*=4$ comme valeur optimale de k . Ainsi, quatre situations sont présentées à l'utilisateur, dans un ordre croissant de distance par rapport à S_p :

1. *Vous réussissez à avoir un boulot mieux rémunéré que l'actuel, mais vous devez quitter vos collègues de travail avec lesquels vous avez développé une grande amitié. (distance=0.5)*
2. *Vous sentez que vous vous approchez d'un objectif dont vous rêviez depuis longtemps. (distance=0.5)*
3. *Vous êtes à un mois d'avoir un nouveau-né. (distance=0.7)*
4. *Vous arrivez à avoir des vacances après une longue année épuisante de travail et vous avez enfin du temps à passer avec vos proches et amis. (distance=0.75)*

Les distances indiquées sont calculées à partir des descriptions numériques établies (voir section 3.2.2).

(D) Étape 4

Au cours de cette étape, l'utilisateur divulgue ses réponses émotionnelles qu'il croit ressentir face à chacune des k situations présentées, au moyen des curseurs sur les échelles d'émotion-intensité (voir Figure 4.2). Ces réponses permettent, comme indiqué au début de la section 4.2, de simuler l'information sur l'historique des réactions émotionnelles de l'utilisateur. Elles sont, dans un premier temps, enregistrées (Figure 4.3) avant d'être

acheminées vers la composante de $kPPV$ pour calculer la première prédiction. Ainsi, en appliquant l'algorithme $kPPV$ pondéré (voir section 2.3), nous obtenons deux vecteurs contenant des scores de prédiction ou de classification respectivement pour les émotions et les intensités. Pour illustrer cela, nous reprenons notre exemple. Supposons que l'utilisateur répond aux quatre situations présentées à l'étape précédente de cette manière :

- Situation 1: émotion *Satisfaction* et intensité *moyenne*
- Situation 2: émotion *Joie* et intensité *forte*
- Situation 3: émotion *Espoir* et intensité *forte*
- Situation 4: émotion *Soulagement* et intensité *forte*

Le $kPPV$ pondéré (voir section 2.3 du chapitre 2) affecte à chaque émotion et à chaque intensité d'une situation donnée un poids pour le calcul d'un score de prédiction. Ce poids est fonction de la distance de la situation par rapport à S_p . Dans notre cas, le poids correspond à l'inverse de la distance au carré (Mitchell, 1997). Ainsi, en utilisant les réponses de l'utilisateur aux quatre situations présentées et les distances de chacune par rapport à S_p (voir étape précédente), l'algorithme $kPPV$ pondéré produit les deux vecteurs des tableaux 4.1 et 4.2.

Tableau 4.1 Exemple de calcul de vecteur de sortie du $kPPV$ pondéré pour l'émotion

Satisfaction	Joie	Espoir	Soulagement
$1/(0.5)^2 = 4$	$1/(0.5)^2 = 4$	$1/(0.7)^2 = 2.04$	$1/(0.75)^2 = 1.77$

Tableau 4.2 Exemple de calcul de vecteur de sortie du $kPPV$ pondéré pour l'intensité

Forte	Moyenne
$1/(0.5)^2 + 1/(0.7)^2 + 1/(0.75)^2 = 7.81$	$1/(0.5)^2 = 4$

L'émotion *Satisfaction* est mentionnée une seule fois par l'utilisateur comme réaction à la situation 1. Son score correspond à l'inverse du carré de la distance de cette situation 1 par rapport à S_p . De même, l'utilisateur a indiqué une intensité *forte* ressentie dans les trois situations 2, 3, et 4. Par conséquent, le score de l'intensité *forte* correspond à la somme de l'inverse des carrés des distances de ces trois situations. Lorsqu'il existe des situations identiques à S_p , la version standard du $kPPV$ (voir section 2.3) est utilisée pour calculer les vecteur de sorties.

Ces sorties sont normalisées pour avoir des valeurs entre 0 et 1 et seront par la suite combinées avec les sorties de l'étape 5.

(E) Étape 5

Cette étape correspond au deuxième type de prédiction basé sur un modèle de *réseau de neurones perceptron multicouche* à une couche cachée (voir section 2.3). Ce modèle (voir Figure 4.4) reçoit comme entrées 24 valeurs de variables contenant les 3 informations sociodémographiques sur l'utilisateur (*âge, sexe, religion*) (voir section 4.1.1), les 4 scores relatifs à sa personnalité (*névrosisme, extraversion, psychoticisme, mensonge*) (voir section 4.1.1), et la description de la situation S_p à laquelle il ferait face suivant les 17 variables d'émotion-intensité (voir tableau 3.6). Comme sortie, le modèle fournit un vecteur de 25 valeurs contenant les probabilités d'occurrence des 23 états émotionnels (22 émotions + état « pas d'émotion ») et des 2 intensités (voir Figure 4.4).

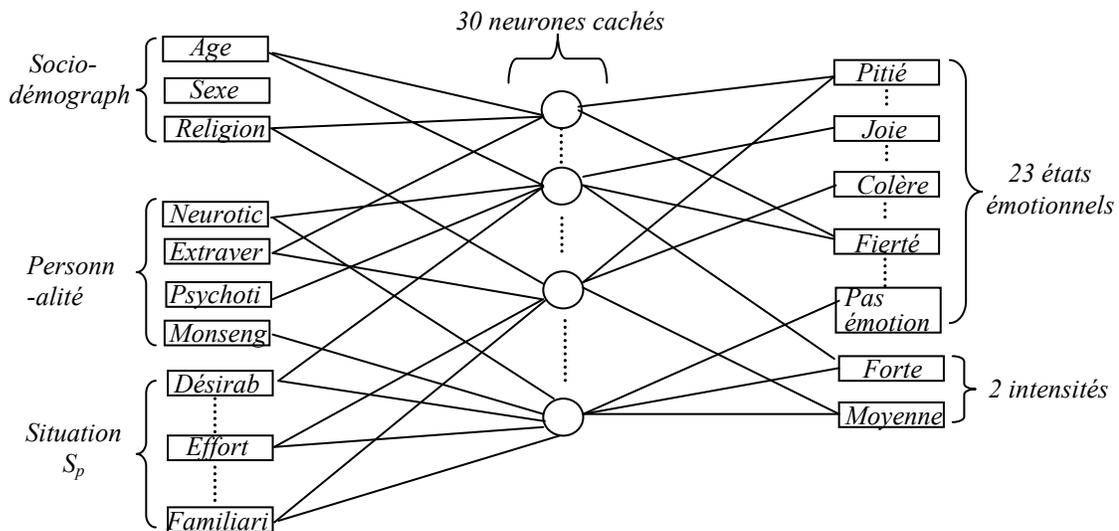


Figure 4.4 Réseau de neurones appris pour la prédiction des émotions et leurs intensités

Ce réseau est construit dans une étape d'initialisation de la *machine émotionnelle* à partir de l'ensemble d'entraînement collecté (voir section 4.1). Cet ensemble contient plusieurs exemples de réactions émotionnelles de différentes personnes face à différentes situations. Ainsi, la prédiction fournie en sortie est basée sur la réaction probable des personnes ayant des caractéristiques sociodémographiques et une personnalité se rapprochant de celles de l'utilisateur face à cette situation S_p ou une situation similaire. Pour la construction du réseau, nous avons, d'abord, cherché à optimiser l'hyperparamètre relatif au nombre de

neurones cachés. L'optimisation est faite sur un ensemble de validation tiré de l'ensemble d'entraînement (voir section 2.2.2). Le nombre optimal trouvé est de 30. Ensuite, les paramètres du réseau ont été appris en utilisant la technique de *descente de gradient* et en appliquant le principe de *rétropropagation* (voir Annexe I) de manière à minimiser une fonction de perte (voir section 2.2.2 et Annexe I).

Reprenons, encore une fois, notre exemple. Supposons que nous présentons à ce modèle entraîné les informations sur l'utilisateur et la description numérique de la situation S_p choisie à l'étape 2. Dans le cas considéré, le modèle donne comme résultat, traduisant la réaction de la plupart des gens ressemblant à l'utilisateur, un vecteur d'émotions ayant des valeurs de probabilités maximales pour la *Joie* et la *Satisfaction* (voir Tableau 4.3) parmi les 23 états possibles et un vecteur d'intensités ayant une valeur maximale pour l'éventualité *forte* (voir Tableau 4.4).

Tableau 4.3 Exemple de vecteur de sortie du RdN pour l'émotion

Espoir	...	Joie	Satisfaction	Soulagement	...	Pas d'émotion
0.01	...	0.5	0.4	0.01	...	0.02

Tableau 4.4 Exemple de vecteur de sortie du RdN pour l'intensité

Forte	Moyenne
0.7	0.3

Ces vecteurs constituent les sorties de la composante RdN qui seront utilisées avec les vecteurs de sortie de la composante $kPPV$ pour la prédiction finale.

(F) Étape 6

Cette étape correspond à la combinaison des résultats de classification ou de prédiction des deux composantes $kPPV$ et RdN. Comme décrit précédemment au début de la section 4.2, l'idée consiste à utiliser, d'une part, une prédiction basée sur des informations spécifiques sur l'utilisateur et son comportement face à des situations similaires à celle à prédire. D'autre part, l'idée consiste à ajuster ou consolider cette information spécifique par une information plus générale ou typique sur les réactions des individus semblables face à ce genre de situations. De ce fait, au cours de la combinaison, un poids plus important est donné à l'information spécifique ou orientée utilisateur, et un autre moins important est assigné à la prédiction générale du RdN. En effet, cette dernière joue un rôle de régulateur ou décideur dans cette combinaison, car elle est déterminante pour la prédiction finale

lorsque le vecteur sortie de la composante $kPPV$ contient quelques valeurs maximales égales ou très proches. L'opération de combinaison s'apparente donc à une prise en considération de deux avis où l'on donne plus de crédibilité à un avis qu'à un autre. Dans le cas où l'avis supposé plus crédible n'est pas tranchant, le deuxième avis permettra, lui, de trancher. Ainsi, cette opération de combinaison des deux prédictions peut être modélisée à l'aide d'une combinaison linéaire. Cette combinaison linéaire attribue à la prédiction spécifique un coefficient (poids) supérieur à celui de la prédiction générale. Une combinaison linéaire des valeurs des vecteurs de sortie des deux algorithmes ($kPPV$ pondéré et RdN) est alors calculée. Pour le choix des coefficients, nous avons testé la performance de quelques combinaisons sur un ensemble de validation: (4/5, 1/5), (3/5, 2/5); (3/4, 1/4) et (2/3, 1/3). La combinaison (2/3, 1/3) a donné les résultats les plus satisfaisants. Ainsi, nous avons affecté à la prédiction du $kPPV$ un poids de 2/3 et à celle du RdN un poids de 1/3.

Pour illustrer la procédure et le schéma de combinaison, nous reprenons notre exemple et les vecteurs de sorties des deux algorithmes dans les tableaux 4.1, 4.2, 4.3 et 4.4. Le vecteur de sortie du $kPPV$, relatif aux émotions, présenté dans le Tableau 4.1 est, d'abord, normalisé pour avoir des valeurs entre 0 et 1. Les valeurs prises par les émotions *Satisfaction* et *Joie* sont toutes les deux égales à 0.34, tandis que les émotions *Espoir* et *Soulagement* prennent respectivement les valeurs 0.17 et 0.15. Pour que le nombre de valeurs soit le même pour les deux vecteurs de sortie des deux algorithmes, nous rajoutons au vecteur de sortie de $kPPV$ des valeurs nulles pour les états émotionnels qui n'ont pas été indiqués par l'utilisateur à l'étape 2. En combinant les valeurs du vecteur $kPPV$ normalisées et celles du RdN présentes dans le Tableau 4.3, nous obtenons les valeurs suivantes pour les émotions les plus probables *Satisfaction* et *Joie*.

$$Satisfaction = \frac{2}{3} * 0.34 + \frac{1}{3} * 0.4 = 0.36$$

$$Joie = \frac{2}{3} * 0.34 + \frac{1}{3} * 0.5 = 0.39$$

Ainsi, la plus grande valeur de la combinaison finale correspond à l'émotion *Joie*. En d'autres termes, l'émotion de l'utilisateur prédite par la *machine émotionnelle* pour la situation S_p , « *Après dix bonnes années de travail dans votre entreprise, vous êtes, comme prévu, promu au poste de directeur général* », est la *Joie*. La prédiction du modèle RdN a

été déterminante et décisive pour le choix de cette prédiction finale, puisque le modèle $kPPV$ donnait des valeurs égales aux deux émotions de *Joie* et de *Satisfaction* et que la prédiction du RdN basée sur la connaissance des réactions typiques penchait légèrement vers la *Joie*.

La même procédure est répétée pour les vecteurs de prédiction de l'intensité. Dans notre exemple, l'intensité *forte* est prédite. Lorsque les vecteurs finaux de l'émotion ou de l'intensité contiennent plusieurs valeurs maximales égales, les prédictions se font de manière aléatoire.

Une fois la combinaison réalisée, une interface est présentée à l'utilisateur contenant la situation S_p (que l'utilisateur ne connaît pas) et la prédiction finale de la machine. La Figure 4.5 en présente un exemple.

(J) Étape 7

Au cours de cette étape, l'utilisateur indique, via l'interface de la Figure 4.5, si la prédiction est correcte ou pas en validant le choix de la machine ou en choisissant une autre alternative au moyen des échelles d'émotion-intensité.

Situation Prédite :

Après dix bonnes années de travail dans votre entreprise, vous êtes, comme prévu, promu au poste de directeur général.

Prédiction de la MACHINE :

Joie extrême

Veuillez modifier le curseur si la prédiction est erronée et/ou appuyer sur valider pour continuer

Heureux pour	Exultation	Espoir	Satisfaction	Soulagement	Joie	Fierté	Admiration	Reconnaissance	Autosatisfaction	Adoration
Pitié	Jalousie	Peur	Peur confirmée	Déception	Peine	Culpabilisation	Indignation	Colère	Remords	Dégoût

Figure 4.5 Interface d'affichage de la prédiction de la *machine émotionnelle*

Cela nous permet de tester la performance de prédiction de la *machine émotionnelle* en calculant le pourcentage de prédictions correctes pour plusieurs personnes. On peut aussi

calculer le pourcentage des mauvaises prédictions. Cela revient au même. Les différentes réponses données par l'utilisateur pour S_p et les k situations qui lui sont les plus proches, recueillies à l'étape 4, sont acheminées vers l'ensemble d'entraînement afin de l'enrichir. Ensuite, une mise à jour instantanée du modèle RdN, en tenant compte de ces nouvelles données, est effectuée. Cette mise à jour est possible grâce à la variante de descente de *gradient stochastique* ou *en ligne* pour l'apprentissage des paramètres (voir chapitre 3 et Annexe I). Cela permet de renforcer le modèle pour des prédictions futures et d'assurer un apprentissage et une amélioration continue au fur et à mesure que des personnes utilisent la machine.

(H) Justification de la démarche de validation de la prédiction avec l'utilisateur

En dépit de son caractère expérimental, la machine devait démontrer son caractère ou sa spécificité prédictif (ve) à ses utilisateurs. En effet ces derniers étaient appelés à tester volontairement et en ligne (sur le Web) une machine qui prédit leurs émotions pour une situation inconnue S_p . Les utilisateurs n'étaient pas rémunérés et l'expérience n'était pas contrôlée ou guidée. Le seul moyen de les attirer, et d'avoir par conséquent un nombre significatif de participants, était de rajouter un peu d'interactivité au moyen de cette composante prédictive. Après un effort louable et une assiduité appliquée dans la saisie d'un grand nombre de pages sur le Web, l'utilisateur s'attend à être récompensé de ses efforts pour voir si la machine est capable de deviner l'émotion qu'il imagine ressentir. C'est la raison pour laquelle nous avons, délibérément, choisi l'option que la machine affiche à notre utilisateur la prédiction qu'elle génère avant de lui demander de fournir la sienne. Maintenant imaginons que cette opération se fasse en sens inverse. Dans ce cas, la crédibilité de la machine serait mise en cause par l'utilisateur. Notre choix a certainement un coût. Celui-ci réside dans le biais résultant de l'influence que pourrait exercer la machine sur une certaine catégorie d'utilisateurs influençables.

Avant de poursuivre notre analyse, il serait utile de rappeler que la mesure usuelle du biais est le taux d'erreur moyen. Dans notre cas, on peut déduire que le biais résultant de notre choix est d'autant plus grand que le nombre d'individus influençables est d'autant plus grand. Autrement dit, si le rapport du nombre de personnes influençables par rapport à celui des non influençables est suffisamment petit, l'ampleur du biais créé par l'annonce de

la prédiction de la machine ne sera pas prononcée. Au fur et à mesure que l'écart en faveur de la seconde catégorie se creuse, le biais moyen sera éliminé de façon graduelle.

L'évaluation de ce type de biais dépend donc des statistiques (répartition) dont on peut disposer sur les utilisateurs influençables. Or, nous ne disposons pas de ce genre de statistiques ni de la méthode adéquate capable de le mesurer. De plus, il est important de noter que plus le nombre N d'utilisateurs augmente plus l'ampleur de ce genre de biais est moindre. Théoriquement on peut imaginer une expérience avec un N suffisamment élevé pour rendre ce biais asymptotiquement négligeable.

Pour finir, notons que le biais relatif au report subjectif de l'émotion (self-report) existe (Picard, 2000; Schwartz & Weinberger, 1980). Cela n'a pas empêché plusieurs travaux en reconnaissance automatique (Chalfoun, Chaffar, & Frasson, 2006; Lisetti & Nasoz, 2002; Sebe et al., 2007) d'utiliser cette méthode pour l'apprentissage de modèles prédictifs et l'évaluation des performances. Vapnik (2000) estime que le biais est un mal nécessaire et affirme qu'un apprentissage utile sans biais est impossible. De plus, le biais dans la reconnaissance d'émotions en général est toujours existant provenant d'abord de la nature floue de l'émotion qui n'est pas un phénomène déterministe c'est à dire parfaitement prédictible ou reconnaissable (comme l'est un phénomène physique). Par ailleurs, le biais pourrait également provenir du choix de l'algorithme d'apprentissage, de la représentativité de l'ensemble d'entraînement et de la pluralité des termes utilisés pour décrire l'émotion (El-Nasr, Yen, & Ioerger, 2000).

4.3 Exploitation et résultats

Nous avons implémenté, en ligne (sur le Web), l'architecture de la machine décrite en détail à la section précédente. Nous nous sommes basés, comme dans la phase de la description des situations (voir section 3.2.2), sur la technique de *JavaServer Pages* (JSP) qui permet d'introduire, dynamiquement, au sein d'une page *HTML* le résultat d'un traitement donné programmé en code *Java*. Ainsi, les deux algorithmes (voir section 2.3, chapitre 2) et leur combinaison ont été implémentés en *Java* et compilés du côté d'un serveur Web *Apache Tomcat*¹⁰ (version 5.0) pour générer la prédiction dans un contenu *HTML* de manière dynamique du côté client. Toutes les données de la *composante de*

¹⁰ <http://tomcat.apache.org/>

connaissances (Figure 4.3), nécessaires à la prédiction, étaient enregistrées dans une base de données MySQL. Lors du déploiement de l'application, nous nous sommes heurtés à des difficultés de dimensionnement et d'affichage des interfaces à cause des différences de standards entre les navigateurs existants (p.ex., Internet Explorer, Mozilla Firefox, Google Chrome, etc.). L'application s'exécute d'une manière satisfaisante selon les standards du navigateur Internet Explorer.

Après avoir configuré la machine, nous avons envoyé un courriel, contenant l'adresse¹¹ de cette application Web, à quelques étudiants et employés de l'Université de Montréal. La machine devait prédire l'émotion d'un participant et son intensité pour 5 situations inconnues S_p . Ainsi, le processus décrit à la section 4.2 était répété 5 fois pour chaque participant. Le but était de récolter un maximum de réactions afin d'avoir une quantité de données significatives permettant d'évaluer la performance de la machine et d'enrichir la base, ce qui permet d'améliorer le modèle de prédiction. La participation était assez faible (environ 40 personnes) dans un premier temps ce qui nous a amenés à renvoyer l'adresse à une liste plus étendue d'employés et étudiants.

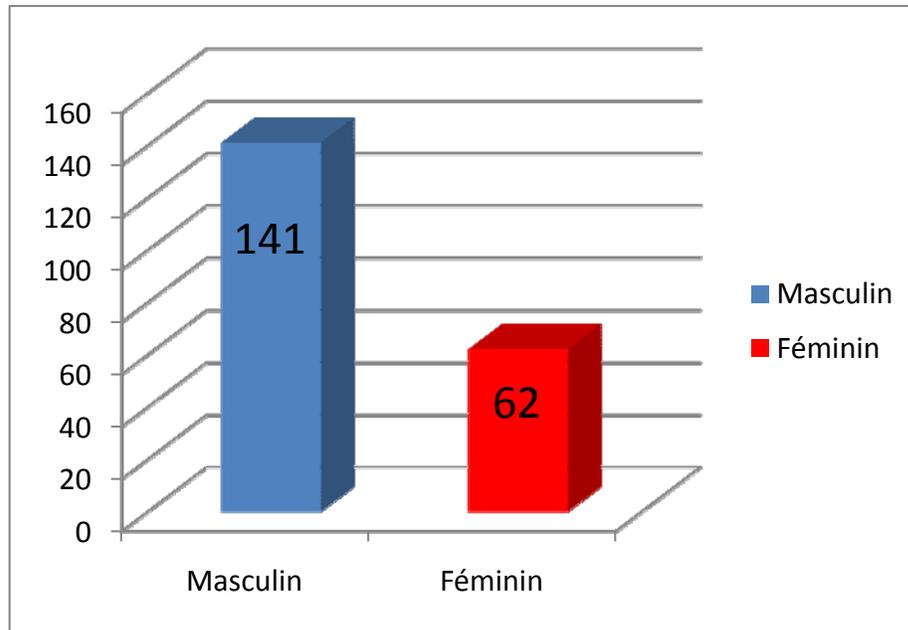


Figure 4.6 Distribution des participants selon le sexe

¹¹ <http://www-etud.iro.umontreal.ca:8080/~trabelam/>

Au final, 203 personnes ont répondu à notre courriel et ont utilisé la machine. La plupart de ces participants étaient de sexe masculin, comme schématisé dans la Figure 4.6. Les catégories d'âge les plus représentées étaient celles des 18 à 25 ans avec 116 personnes et celles des 26 à 45 ans avec 75 personnes (voir Figure 4.7). Cela peut s'expliquer par le fait que le questionnaire était destiné aux étudiants et employés de l'Université de Montréal.

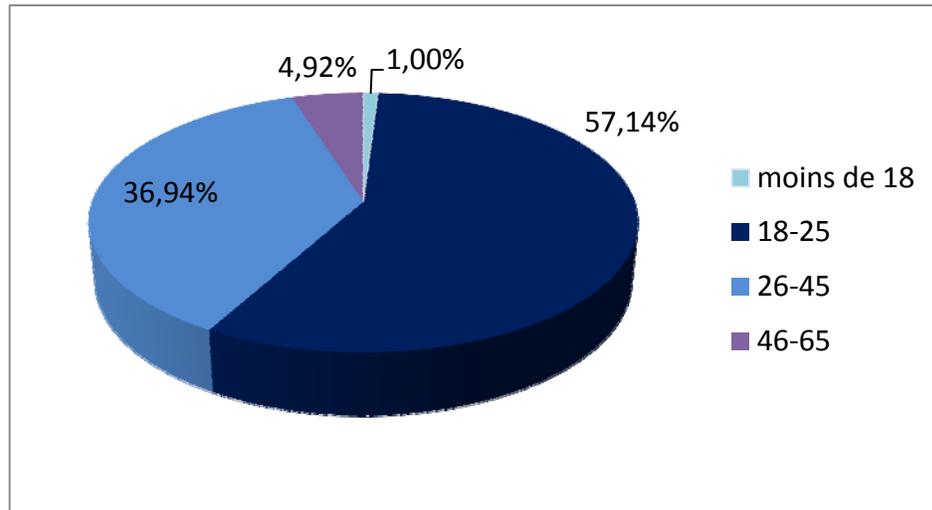


Figure 4.7 Distribution des participants selon l'âge

Les religions majoritairement présentes sont le christianisme (90 personnes) et l'islam (60 personnes) (voir Figure 4.8). Pour ce qui est de la personnalité, 154 participants parmi les 203 ont des scores majoritaires pour un des quatre traits de personnalité.

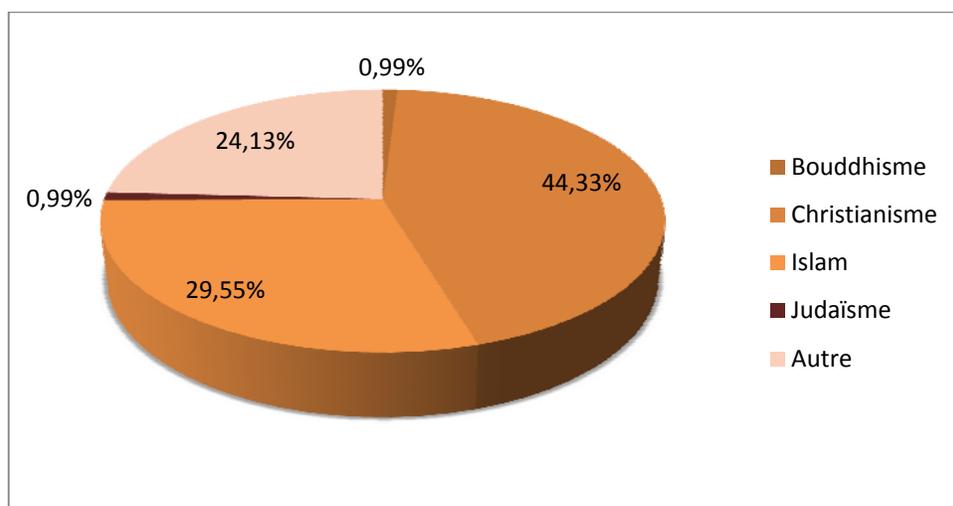


Figure 4.8 Distribution des participants selon la religion

En effet, 55 participants ont un score majoritaire pour *l'extraversion*, 56 pour le *névrosisme*, 13 pour le *psychoticisme* et 30 pour le *mensonge*.

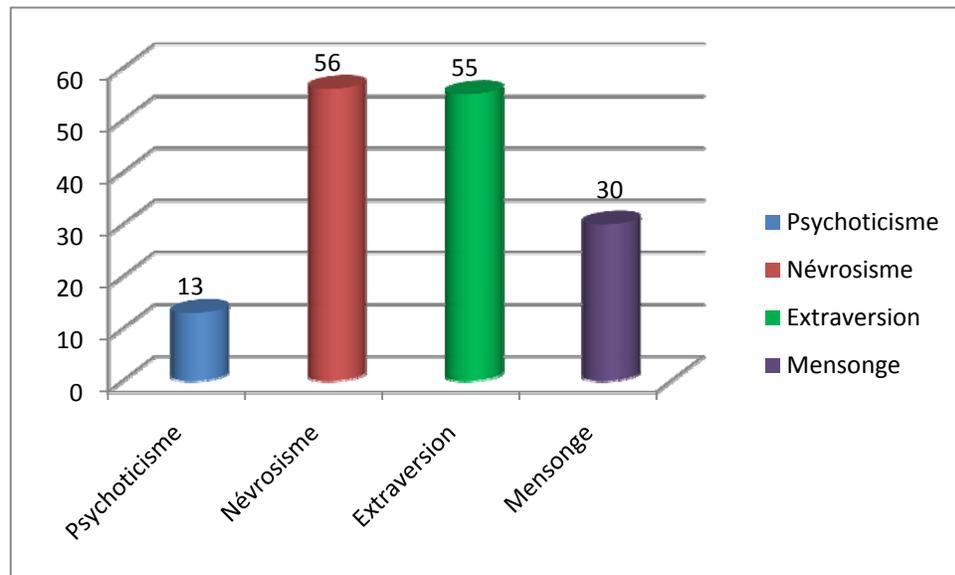


Figure 4.9 Distribution des traits de personnalité majoritaires chez les participants

Puisque l'expérience était effectuée en ligne (sur le Web) et qu'elle n'était pas contrôlée, un participant pouvait quitter notre application lorsqu'il le souhaitait. Cependant, toutes les informations saisies, avant de quitter, étaient enregistrées. Les 203 participants ont rapporté un total de 3544 réactions émotionnelles face aux différentes situations de la base, c'est-à-dire une moyenne d'environ 18 réactions rapportées par participant. Ces 3544 réactions correspondent à la fois aux réactions indiquées par les différents participants face aux situations similaires à une situation S_p à prédire (voir étape 3, section 4.2.2), et également aux réactions face aux situations prédites (voir étape 7, section 4.2.2). La machine est supposée prédire l'émotion d'un participant et son intensité pour 5 situations S_p . Mais, comme nous venons de le mentionner, puisque l'expérience était menée en ligne, plusieurs participants pouvaient donc quitter l'application sans compléter les 5 prédictions. Par conséquent, au lieu de réaliser 1015 (203×5) prédictions, la machine n'a réalisé que 634, soit une moyenne d'un peu plus de 3 prédictions par participant, au lieu de 5. Pour 411 prédictions sur les 634, l'émotion générée par la machine a été validée correctement par le participant, comme étant celle qu'il éprouverait face à la situation proposée (voir Figure 4.5). En d'autres termes, la *machine émotionnelle* peut reconnaître adéquatement l'état émotionnel des personnes, et ce parmi 23 états possibles (22 émotions + pas d'émotion),

avec un pourcentage de 64.83 % pour des situations très variées de la vie réelle. Ces situations constituent un échantillon représentatif d'un contexte élargi de situations de la vie réelle proposées par 57 personnes d'âge et de sexe différents. En effet, elles ont été extraites à partir d'un échantillon initial de 1002 situations variées (voir section 3.2.1). Les situations qu'ils avaient proposées n'étaient pas limitées au domaine professionnel. Elles traduisaient, également, leur vécu quotidien. Ce vécu correspond à plusieurs aspects de la vie quotidienne (familiale, sociale, spirituelle, etc.).

Certes, la performance de prédiction (ou classification) de la machine peut être améliorée en ayant plus de données d'entraînement, car le nombre de classes (23) pour l'émotion est assez élevé. Chaque classe devrait être extrêmement représentée avec plusieurs exemples pour avoir une très bonne prédiction. C'est l'une des principales raisons pour laquelle la plupart des travaux en reconnaissance d'émotions ne se sont intéressés qu'aux 6 émotions de base (Zeng, Pantic, Roisman, & Huang, 2008). D'autre part, ce pourcentage peut ne pas refléter de manière exacte la performance de la machine, car les utilisateurs peuvent confondre les émotions similaires (Steidl, Levit, Batliner, Nöth, & Niemann, 2005) du fait qu'elles sont très détaillées et proches dans leur signification, comme le *Remord* et la *Culpabilisation* ou bien la *Satisfaction* et la *Joie*, etc. Par conséquent, ces utilisateurs peuvent indiquer une émotion différente de celle qui est ressentie en réalité. Toutefois, dans les interfaces implémentées, les termes émotionnels sont expliqués au moyen des bulles qui apparaissaient lors du passage de la souris (voir Figure 4.2, section 4.1.2).

Pour ces différentes raisons, nous avons calculé le pourcentage de prédictions correctes pour un nombre de classes d'émotions moins élevé, et ce en regroupant les émotions similaires parmi les 22, selon les groupes définis par le modèle OCC et présentés au chapitre précédent (voir Figure 3.1). En effet, les nouvelles classes sont constituées comme suit :

Classe 1: regroupe les émotions *Heureux Pour Autrui* et *Exultation*. Ces émotions représentent les émotions positives du groupe *Fortune d'autrui*.

Classe 2: regroupe les émotions négatives du groupe *Fortune d'autrui*, à savoir *Jalousie* et *Pitié*.

- Classe 3: regroupe l'émotion *Joie* et les émotions positives du groupe *Basé Perspective, Espoir, Satisfaction, et Soulagement*.
- Classe 4: regroupe l'émotion *Peine* et les émotions négatives du groupe *Basé Perspective, Peur, Peur confirmée, et Déception*.
- Classe 5: regroupe les émotions positives du groupe *Attribution*, c'est-à-dire *Fierté* et *Admiration*, avec celles du groupe composé *Bien-être/Attribution*, à savoir, *Reconnaissance* et *Autosatisfaction*.
- Classe 6: regroupe les émotions négatives des groupes *Attribution* et *Bien-être/Attribution* (*Culpabilisation, Indignation, Colère* et *Remords*)
- Classe 7: correspond à *Adoration*, l'émotion positive du groupe *Attraction*
- Classe 8: correspond à *Dégoût*, l'émotion négative du groupe *Attraction*
- Classe 9: correspond à l'état *Pas d'émotion*

En recalculant le pourcentage de prédictions correctes selon cette configuration des classes d'émotion, nous trouvons un taux de 72.24 %. Puis, nous avons également essayé de mesurer la capacité de la machine à reconnaître l'émotion de l'utilisateur face à la situation associée à une émotion *positive, négative* ou bien *inexistante*. En d'autres termes, nous avons encore une fois regroupé les émotions, mais cette fois-ci en 3 classes. Le taux de prédictions correctes obtenu est de 86.90 %. D'un autre côté, la reconnaissance correcte des intensités de ces émotions en tant que *forte* ou *moyenne* a atteint un taux de 79.83 %. La comparaison de notre travail avec d'autres travaux en reconnaissance d'émotions reste difficile puisque les contextes des travaux et les émotions visées diffèrent, ainsi que les sources et la nature des données (audio, vidéo, multimodale, physiologique, etc.), et même les bases de données utilisées. Néanmoins, nous allons essayer de comparer notre travail avec des travaux connus dans la littérature en termes de pourcentages de prédiction par rapport au nombre de classes. Par exemple, pour la prédiction ou la classification faciale de 6 classes d'émotions, Kaliouby et Robinson (2005), et Wang et al. (2006) ont réalisé respectivement des taux de classification correcte de 77.4% et 83.6%. Avec un nombre légèrement supérieur de classes à prédire (neuf), notre machine réalise un pourcentage qui n'est pas très loin de ces pourcentages, considérés comme des références en la matière. D'un autre côté, pour la reconnaissance de 3 classes, nous pouvons citer les travaux de Ioannou

(2005) qui ont utilisé les traits du visage, de Forbes-Riley et Litman (2004) et de Liscombe et al. (2005). Ces deux derniers ont eu recours à des signaux sonores. Les pourcentages de classification correcte pour ces trois travaux sont respectivement de 78%, 84.75% et 91.3%. Avec un pourcentage de 86.90% pour la prédiction de 3 classes d'émotions, notre machine ne sort pas de la norme.

Tableau 4.5 Les pourcentages de prédictions correctes pour les émotions et les intensités

	Émotions			Intensités
	23 classes	9 classes	3 classes	
Pourcentage de prédictions correctes	64.83%	72.24%	86.90%	79.83%

Les différents pourcentages réalisés, résumés dans le Tableau 4.5, sont des indicateurs sur la capacité de la *machine émotionnelle* construite pour « assimiler » ou « comprendre » la signification d'une situation émotionnelle de la vie courante indépendamment du domaine (social, professionnel, familial, etc.). En fait, les pourcentages obtenus confirment que les situations similaires recherchées et présentées par la machine (voir étape 3, section 4.2.2) avant de calculer la prédiction sont effectivement similaires et proches puisque les réponses de l'utilisateur face à elles sont utilisées pour la prédiction finale. Cela nous mène à dire également que la modélisation ou la description numérique de situations, mise en place à l'aide des *variables d'évaluation-intensité* de OCC à la section 3.2.2, s'est avérée concluante. Ces performances ont également permis de vérifier la combinaison inspirée du modèle humain de reconnaissance d'émotions basé sur l'information spécifique relative aux réactions de l'utilisateur et l'information générale relative à la connaissance commune des réactions émotionnelles. La combinaison s'est faite de manière instantanée et continue pour chaque utilisateur en se basant sur ses informations personnelles et les informations collectées de tous les utilisateurs passés auparavant. Cette reconnaissance instantanée et continue rajoutée aux bons pourcentages de prédiction des intensités et des émotions pour une situation émotionnelle donnée, indépendamment du domaine, font de la *machine émotionnelle* un outil potentiellement utilisable dans plusieurs applications de la reconnaissance d'émotions. Nous nous attarderons sur ce sujet au cours de la conclusion générale du mémoire.

4.4 Conclusion

Dans ce chapitre, nous avons d'abord tenté d'expliquer la procédure mise en place pour la construction d'un ensemble d'entraînement nécessaire pour le fonctionnement de la *machine émotionnelle*. La procédure consistait à collecter des informations personnelles sur plusieurs personnes et leurs réactions émotionnelles face aux différentes situations de la vie courante collectées auparavant (voir chapitre 3). Ensuite, nous avons détaillé l'architecture de la machine et les différentes étapes entreprises pour la prédiction de l'émotion que ressent une personne face à une situation donnée, ainsi que son intensité. Le corps de la machine est constitué des deux algorithmes de *réseaux de neurones* et des *k plus proches voisins pondérés*. Le modèle appris par les réseaux de neurones utilise l'ensemble d'entraînement construit et traduit par conséquent une connaissance typique et générale des réactions émotionnelles chez les humains. Cela nous a permis de répondre à la deuxième question que nous nous sommes posée lors du chapitre introductif : Comment pouvons-nous équiper les machines d'une connaissance préexistante ou d'un sens commun sur les réactions émotionnelles typiques des humains face aux situations de la vie courante ? Nous avons également essayé, en combinant les deux algorithmes, de répondre à la troisième question (du chapitre introductif) : Comment les machines peuvent-elles, comme les humains, utiliser les informations efficacement de manière à donner de bonnes prédictions sur les émotions des individus et leur intensité ?

En dernier lieu, une exploitation et une expérimentation ont été faites pour vérifier et tester empiriquement la validité du travail effectué et mesurer la performance prédictive de la machine pour de nouveaux utilisateurs. La machine a été implémentée en ligne (sur le Web) et sa prédiction était instantanée. L'apprentissage se faisait de manière continue au fur et à mesure que des nouvelles données étaient entrées. Au final, 203 participants ont testé l'application. Un ensemble de données contenant 3544 exemples a été collecté et des pourcentages satisfaisants de prédictions ont été observés lorsqu'on compare avec les travaux dans la littérature qui ont prédit le même nombre de classes d'émotions. Ces pourcentages peuvent être améliorés avec une plus grande quantité de données. Ils témoignent également de la capacité de notre *machine émotionnelle* à prédire pour un individu donné sa réaction émotionnelle face à une grande variété de situations émotionnelles décrites indépendamment du domaine. Cette caractéristique peut être

intéressante pour de futures recherches dans plusieurs domaines appliquant la reconnaissance automatique de l'émotion.

Conclusion générale

Dans ce mémoire, nous avons passé en revue, dans le premier chapitre, les différents modes utilisés pour la reconnaissance automatique de l'émotion (facial, vocal, physiologique, etc.) et les travaux qui s'y sont intéressés dans un contexte d'interaction homme-machine, en indiquant leurs limites. Ces limites concernent, entre autres, le nombre réduit des émotions étudiées, le peu de travaux englobant la reconnaissance instantanée et la négligence de l'information relative au contexte et aux situations qui engendrent la réaction émotionnelle. Nous avons également présenté les principaux systèmes intelligents visant à simuler le processus de génération d'émotion suite à une évaluation des situations dans des environnements virtuels d'interaction entre agents. La plupart de ces travaux se basent sur la théorie d'évaluation cognitive OCC (*Ortony, Clore, & Collins, 1988*). Néanmoins, l'utilisation de cette théorie, dans un contexte d'interaction homme-machine, pour l'évaluation ou l'interprétation des situations émotionnelles que peut vivre un individu, n'a pas fait l'objet de plusieurs études.

Partant de ce constat, nous avons tenté de développer une application, la *machine émotionnelle*, dans un cadre d'interaction homme-machine, capable d'interpréter de manière instantanée l'information relative à une grande variété de situations de la vie courante et de prédire l'émotion probable parmi 22 possibles, ainsi que son intensité, pour une personne particulière. Pour ce faire, la machine devait d'abord « comprendre » ou « assimiler » la signification d'une situation émotionnelle comme pourrait la percevoir la plupart des gens. Une formalisation de cette information par le biais des variables d'évaluation du modèle OCC a été établie dans le chapitre 3. En effet, une procédure empirique a été mise en place pour collecter, selon les différentes émotions, 250 situations textuellement décrites. Ensuite, ces situations ont été évaluées par plusieurs personnes en donnant des valeurs aux variables d'évaluation-intensité du modèle OCC. Au final, nous avons obtenu une base de situations contenant des descriptions sous forme de vecteurs numériques qui caractérisent ou identifient chacune des 250 situations collectées.

Cette base était nécessaire pour le fonctionnement de notre *machine émotionnelle* qui repose sur des algorithmes d'apprentissage machine nécessitant plusieurs cas d'exemples d'individus, faisant face à un nombre donné de situations, pour pouvoir classer

l'émotion et son intensité. Dans le chapitre 2, nous nous sommes attardés sur la notion d'apprentissage en général et nous avons détaillé, en particulier, les algorithmes utilisés dans le cadre d'une spécification propre à notre machine. Ces derniers sont des versions améliorées des algorithmes des *k plus proches voisins* et des *réseaux de neurones à perceptron multicouche*. Nous avons argumenté leur emploi par leur aspect pratique dans la prédiction instantané réalisée par la machine et dans la classification plusieurs-cibles (émotion + intensité) qu'elle produit.

La mise en place de la machine et l'utilisation de ces algorithmes dans une architecture globale ont été présentées au niveau du chapitre 4. Au préalable, une collecte d'informations sur les caractéristiques sociodémographiques et la personnalité de quelques individus, ainsi que leurs réactions émotionnelles face aux situations prélevées au chapitre 3, a été réalisée. Cette procédure empirique a permis de construire un ensemble d'entraînement à partir duquel la machine pouvait extraire une connaissance sur les réactions générales ou typiques des individus. Puis, pour une prédiction de l'émotion d'une personne, ainsi que son intensité face à une situation donnée, la machine effectue une combinaison entre cette connaissance générale et une autre connaissance spécifique relative aux réactions rapportées de cette personne face à des situations similaires. Afin de vérifier et tester empiriquement la validité du travail, la machine a été déployée en ligne (sur le Web). Un total de 203 étudiants et membres du personnel de l'Université de Montréal l'ont testée. La machine apprenait et s'adaptait au fur et à mesure que de nouveaux participants l'exploitaient. Au final, 3544 exemples d'apprentissage ont été collectés, où chaque exemple contient des informations personnelles sur un individu particulier, une description d'une situation donnée à laquelle cet individu fait face et sa réaction émotionnelle ainsi que son intensité. Les performances de prédiction de la machine ont été satisfaisantes et ont atteint les 64.83 % de prédictions correctes pour un nombre de classes d'émotions ou d'états émotionnels égal à 23 (22 émotions + état pas d'émotion). En calculant la performance pour une granularité de classes moins élevées, de 9 puis de 3, nous avons pu observer des améliorations substantielles avec des pourcentages de 72.24 % et 86.90 % respectivement. De même, les intensités, moyenne et forte, ont été prédites correctement avec un taux de 79.83 %. Cela témoigne de la capacité de notre *machine émotionnelle* à prédire pour un individu donné sa réaction émotionnelle face à une grande variété de situations émotionnelles décrites indépendamment du domaine.

Pour résumer, nous avons pu réaliser nos trois objectifs et répondre aux trois questions que nous nous sommes posées au chapitre introductif. En effet, dans un premier temps, nous avons pu représenter l'information relative à une situation émotionnelle de la vie courante, au sein d'une machine, en la formalisant selon les variables d'évaluation-intensité du modèle OCC. Ainsi, nous avons su appliquer et vérifier la pertinence de ce modèle théorique.

La deuxième réalisation consiste en la construction de l'ensemble d'apprentissage contenant 3544 exemples de personnes faisant face à une situation émotionnelle et leurs réactions correspondantes. Cet ensemble de données pourrait être utilisé dans les recherches en apprentissage machine puisqu'il contient plusieurs exemples et plus d'une variable cible (émotion+intensité). Il pourrait également constituer un cadre de travail pour des psychologues, afin d'étudier les possibles corrélations qui peuvent exister entre les différentes variables, par exemple entre les caractéristiques sociodémographiques et l'intensité de l'émotion ressentie, etc. De plus, les descriptions textuelles des situations collectées et l'information sur les émotions qui leur sont associées pourraient servir dans des recherches en informatique linguistique, visant à trouver un moyen d'extraire l'information émotionnelle à partir d'un texte. Par exemple, on pourrait extraire les mots les plus fréquents dans les situations provoquant une émotion donnée, comme la colère. Ces mots sont ensuite annotés et indexés à cette émotion, selon leur pertinence, puis utilisés dans des travaux relatives à d'autres textes. Cela pourrait être appliqué à l'extraction de l'opinion et l'analyse des sentiments à partir des textes présents dans des forums de discussion, dans des blogs, dans des commentaires rapportés dans des sites de journaux, etc.

La troisième réalisation est le développement de la *machine émotionnelle* qui permet une prédiction instantanée des émotions des utilisateurs, ainsi que leur intensité, pour une grande majorité de situations (indépendamment du domaine) de manière continue et en se basant sur des données réelles rapportées. Ces caractéristiques la rendent intéressante pour plusieurs applications dans les interactions homme-machine. Par exemple, elle peut être utilisée dans les systèmes tuteurs intelligents pour donner à un agent compagnon (compagnon agent) l'information relative à la probable réaction de l'apprenant face à une situation donnée. Elle peut également être intégrée au niveau des agents

intelligents et des robots pour participer à la simulation du comportement émotionnel comme réponse à un stimulus environnemental extérieur. Une autre perspective très intéressante consiste à combiner ce travail avec d'autres travaux sur la reconnaissance d'émotion à partir des expressions faciales et des intonations vocales dans un cadre de prédiction multimodale. Dans ce cas, la machine pourrait fournir une information supplémentaire et pertinente, par exemple, une connaissance à priori sur la probable émotion de l'utilisateur selon la situation dans laquelle il se trouve.

Néanmoins, notre travail contient quelques limites. Les émotions rapportées par les individus peuvent contenir un biais. Comme nous l'avons expliqué au chapitre 4 (voir sections 4.1.2 et 4.2.2), une prédiction sans biais est impossible (*Vapnik, 2000*). De plus, pour une meilleure performance de la prédiction des 23 états émotionnels, une quantité plus importante de données doit être collectée. En outre, nous n'avons pas considéré le cas de cooccurrence de deux ou plusieurs émotions puisque cela nécessitait une quantité énorme d'exemples pour chaque type de combinaison des émotions. Nous n'avons pas considéré, non plus, le cas d'intensités prenant des valeurs continues. Ainsi, plusieurs améliorations peuvent encore être apportées à la *machine émotionnelle*.

Bibliographie

- Aggarwal, J. K., & Cai, Q. (1999). Human motion analysis: a review. *Computer Vision and Image Understanding*, 73, 428-440.
- André, E., Klesen, M., Gebhard, P., Allen, S., & Rist, T. (2000). Integrating Models of Personality and Emotions into Lifelike Characters. Dans A. Paiva (éd.), *Affective interactions* (1814, pp. 150-165): Springer Berlin / Heidelberg.
- Arnold, M. (1960). *Emotion and Personality: Neurological and physiological aspects* (Vol. 2). New York: Columbia University Press.
- Balomenos, T., Raouzaïou, A., Ioannou, S., Drosopoulos, A., Karpouzis, K., & Kollias, S. (2005). Emotion Analysis in Man-Machine Interaction Systems. Dans S. Bengio & H. Bourlard (éds.), *Machine Learning for Multimodal Interaction* (3361, pp. 318-328): Springer Berlin / Heidelberg.
- Bartlett, M. S., Littlewort, G., Frank, M., Lainscsek, C., Fasel, I., & Movellan, J. (2006). Fully Automatic Facial Action Recognition in Spontaneous Behavior. *International Conference on Automatic Face and Gesture Recognition* (pp. 223-230).
- Bates, J., Loyall, A. B., & Reilly, W. S. (1992). An Architecture for Action, Emotion, and Social Behavior. *European Workshop on Modeling and Autonomous Agents in a Multi-Agent World* (pp. 55-68).
- Bishop, C. (2006). *Pattern Recognition and Machine Learning*: Springer New York.
- Bottou, L. (1991). *Une Approche théorique de l'Apprentissage Connexionniste: Applications à la Reconnaissance de la Parole*. Université de Paris XI, Orsay, France.
- Bouvard, M. (2009). *Questionnaires et échelles d'évaluation de la personnalité*: Elsevier Masson.
- Bradley, M., Codispoti, M., Sabatinelli, D., & Lang, P. (2001). Emotion and motivation II: Sex differences in picture processing. *Emotion*, 1(3), 300-319.
- Brave, S., & Nass, C. (2003). Emotion in Human-Computer Interaction. Dans J. Jacko & A. Sears (éds.), *The Human-Computer Interaction Handbook* (pp. 81-96). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Monterey, CA: Wadsworth & Brooks Cole Advanced Books & Software.

- Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology*, 70(2), 205-218.
- Carver, C. S. (2001). Affect and the functional bases of behavior: On the dimensional structure of affective experience. *Personality and Social Psychology Review*, 5(4), 345-356.
- Chalfoun, P., Chaffar, S., & Frasson, C. (2006). Predicting the emotional reaction of the learner with a machine learning technique. *Workshop on Motivational and Affective Issues in International Conference on Intelligent Tutoring Systems* (pp. 13-20).
- Chang, Y., Hu, C., Feris, R., & Turk, M. (2006). Manifold based analysis of facial expression. *Image and Vision Computing*, 24(6), 605-614.
- Clore, G., & Gasper, K. (2000). Feeling is believing: Some affective influences on belief. Dans *Emotions and beliefs: How feelings influence thoughts* (pp. 10-44): Cambridge: Editions de la Maison des Sciences de l'Homme et Cambridge University Press
- Cohn, J. (2007). Foundations of Human Computing: Facial Expression and Emotion. Dans T. Huang, A. Nijholt, M. Pantic & A. Pentland (éds.), *Artificial Intelligence for Human Computing* (4451, pp. 1-16): Springer Berlin / Heidelberg.
- Conati, C., & Maclaren, H. (2004). Evaluating a probabilistic model of student affect. *International Conference on Intelligent Tutoring Systems Maceio, Brazil* (pp. 55-66).
- Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., et al. (2001). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, 18(1), 32-80.
- D'Mello, S., Graesser, A., & Picard, R. W. (2007). Toward an affect-sensitive AutoTutor. *IEEE Intelligent Systems*, 22(4), 53-61.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam Press.
- Devijver, P. A., & Kittler, J. (1982). *Pattern recognition: A statistical approach*: Prentice Hall.
- Diener, E., Oishi, S., & Lucas, R. E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual Review of Psychology*, 54, 403-425.
- Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification*: Wiley-Interscience.

- Duric, Z., Gray, W. D., Heishman, R., Li, F. Y., Rosenfeld, A., Schoelles, M. J., et al. (2002). Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction. *Proceedings of the IEEE*, 90(7), 1272-1289.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the human face: Guide-lines for research and an integration of findings*: New York: Pergamon Press.
- El-Nasr, M. S., Yen, J., & Ioerger, T. R. (2000). FLAME: Fuzzy Logic Adaptive Model of Emotions. *Autonomous Agents and Multi-Agent Systems*, 3(3), 219-257.
- Elliott, C. (1992). *The affective reasoner: a process model of emotions in a multi-agent system*. Ph. D. inédit, Northwestern University
- Elliott, C. (1993). Using the affective reasoner to support social simulations. *International Joint Conference on Artificial Intelligence* (Vol. 13, pp. 194-200) Morgan Kaufmann.
- Elliott, C., & Brzezinski, J. (1998). Autonomous agents as synthetic characters. *AI magazine*, 19(2), 13.
- Ellsworth, P. (1994). Some reasons to expect universal antecedents of emotion. Dans P. Ekman & R. J. Davidson (éds.), *The nature of emotion: Fundamental questions* (pp. 150-154). New York: Oxford University press.
- Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire (Junior and Adult)*. London: Hodder and Stoughton Sevenoaks.
- Eysenck, S. B. G., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences*, 6(1), 21-29.
- Forbes-Riley, K., & Litman, D. (2004). Predicting emotion in spoken dialogue from multiple knowledge sources. *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT/NAACL)* (pp. 201-208).
- Fragopanagos, N., & Taylor, J. G. (2005). Emotion recognition in human-computer interaction. *Neural Networks*, 18(4), 389-405.
- Francis, L. J., Brown, L. B., & Philipchalk, R. (1992). The development of an abbreviated form of the revised Eysenck personality questionnaire (EPQR-A): Its use among students in England, Canada, the U.S.A. and Australia. *Personality and Individual Differences*, 13(4), 443-449.
- Franklin, S. (2002). "Conscious" software: a computational view of mind. Dans *Soft computing agents: new trends for designing autonomous systems* (pp. 1-45): Physica-Verlag GmbH.
- Fridja, N. H. (1986). *The emotions*: Cambridge University Press.

- Frijda, N. (1987). Emotion, cognitive structure, and action tendency. *Cognition & Emotion*, 1(2), 115-143.
- Frijda, N. H., & Swagerman, J. (1987). Can computers feel? theory and design of an emotional system. *Cognition & Emotion*, 1(3), 235-257.
- Fujita, F., Diener, E., & Sandvik, E. (1991). Gender differences in negative affect and well-being: The case for emotional intensity. *Journal of Personality and Social Psychology*, 61(3), 427-434.
- Funahashi, K.-I. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks*, 2(3), 183-192.
- Goldberg, L. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26-34.
- Goleman, D. (1995). *Emotional Intelligence*: Bantam Books.
- Goleman, D., Boyatzis, R., McKee, A., & Borgeaud, E. (2002). *L'intelligence émotionnelle au travail*. Paris, France: Editions Village Mondial / Pearson Education.
- Gratch, J. (2000). Émile: Marshalling passions in training and education. *Proceedings of the fourth international conference on Autonomous agents, Barcelona, Spain* (pp. 325-332) ACM.
- Gratch, J., & Marsella, S. (2003). Fight the way you train: the role and limits of emotions in training for combat. *Brown Journal of World Affect*, 10(1), 63-76.
- Gratch, J., & Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5(4), 269-306.
- Gratch, J., Marsella, S., & Petta, P. (2009). Modeling the cognitive antecedents and consequences of emotion. *Cognitive Systems Research*, 10(1), 1-5.
- Gunes, H., & Piccardi, M. (2006). A bimodal face and body gesture database for automatic analysis of human nonverbal affective behavior. *18th International Conference on Pattern Recognition* (Vol. 1, pp. 1148-1153).
- Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004). Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System. Dans E. André, L. Dybkjaer, W. Minker & P. Heisterkamp (éds.), *Affective Dialogue Systems* (3068, pp. 36-48): Springer Berlin / Heidelberg.
- Hastie, T., Tibshirani, R., & Friedman, J. (2001). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. *NY Springer-Verlag*.
- Healey, J. (2000). *Wearable and Automotive Systems for Affect Recognition from Physiology*. MIT, Cambridge, MA.

- Healey, J., Picard, R. W., & Dabek, F. (1998). A new affect-perceiving interface and its application to personalized music selection. *Workshop on Perceptual User Interfaces, San Francisco, CA* (pp. 4-6).
- Hess, U. (2001). The experience of emotion: situational influences on the elicitation and experience of emotions. *Emotions, Qualia, and Consciousness*, 386-396.
- Hirschberg, J., Benus, S., Brenier, J., Enos, F., Friedman, S., Gilman, S., et al. (2005). Distinguishing deceptive from non-deceptive speech. *European Conference Speech Communication and Technology* (pp. 1833-1836).
- Inanoglu, Z., & Caneel, R. (2005). Emotive alert: HMM-based emotion detection in voicemail messages. *International conference on Intelligent user interfaces, San Diego, California, USA* (pp. 251-253).
- Ioannou, S., Raouzaïou, A., Tzouvaras, V., Mailis, T., Karpouzis, K., & Kollias, S. (2005). Emotion recognition through facial expression analysis based on a neurofuzzy network. *Neural Networks*, 18(4), 423-435.
- Isen, A. M. (2001). An influence of positive affect on decision making in complex situations: Theoretical issues with practical implications. *Journal of Consumer Psychology*, 11(2), 75-85.
- Ishino, K., & Hagiwara, M. (2003). A feeling estimation system using a simple electroencephalograph. *IEEE International Conference on Systems, Man and Cybernetics* (Vol. 5, pp. 4204-4209).
- Izard, C. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- Ji, Q. A., Lan, P. L., & Looney, C. (2006). A probabilistic framework for modeling and real-time monitoring human fatigue. *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans*, 36(5), 862-875.
- Jolliffe, I. (1986). *Principal Component Analysis*: Springer-Verlag.
- Juslin, P. N., & Scherer, K. R. (2005). Vocal expression of affect. Dans J. Harrigan, R. Rosenthal & K. R. Scherer (éds.), *The New Handbook of Methods in Nonverbal Behavior Research* (pp. 65-135): Oxford Univ. Press.
- Kaliouby, R., & Robinson, P. (2005). Real-Time Inference of Complex Mental States from Facial Expressions and Head Gestures. Dans B. Kisačanin, V. Pavlović & T. Huang (éds.), *Real-time vision for human-computer interaction* (pp. 181-200): Springer US.
- Kapoor, A., Bursleson, W., & Picard, R. W. (2007). Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65(8), 724-736.

- Kapoor, A., & Picard, R. W. (2005). Multimodal affect recognition in learning environments. *Annual ACM International Conference on Multimedia, Hilton, Singapore* (pp. 677-682) ACM.
- Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, 5(4), 345.
- Kwon, O., Chan, K., Hao, J., & Lee, T. (2003). Emotion recognition by speech signals. *European Conference on Speech Communication and Technology* (Vol. 1, pp. 125-128).
- Lang, P. J. (1995). The emotion probe: Studies of Motivation and Attention. *American Psychologist*, 50(5), 372-385.
- Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, 61(1), 132-140.
- Lazarus, R. (1968). Emotions and adaptation: Conceptual and empirical relations. Dans W. J. Arnold (éd.), *Nebraska symposium on motivation* (16, pp. 175-270). Lincoln: University of Nebraska Press.
- Lazarus, R. (1991). *Emotion and adaptation*. New York: Oxford University press.
- Le Cun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., et al. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1, 541-551.
- LeDoux, J. E. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. New York: Touchstone books.
- Lee, C. M., & Narayanan, S. S. (2005). Toward detecting emotions in spoken dialogs. *IEEE Transactions on Speech and Audio Processing*, 13(2), 293-303.
- Leon, E., Clarke, G., Callaghan, V., & Sepulveda, F. (2007). A user-independent real-time emotion recognition system for software agents in domestic environments. *Engineering Applications of Artificial Intelligence*, 20(3), 337-345.
- Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill Book Company.
- Lewin, K., & De Rivera, J. (1976). *Field theory as human-science: contributions of Lewin's Berlin group*: Halsted Press.
- Lin, Y. P., Wang, C. H., Wu, T. L., Jeng, S. K., & Chen, J. H. (2009). EEG-Based Emotion Recognition in Music Listening: A Comparison of Schemes for Multiclass Support Vector Machine. *IEEE International Conference on Acoustics, Speech, and Signal Processing, New York* (Vol. 1, pp. 489-492).

- Liscombe, J., Hirschberg, J., & Venditti, J. (2005). Detecting certainness in spoken tutorial dialogues. *European Conference on Speech Communication and Technology* (pp. 1837–1840).
- Lisetti, C., & Nasoz, F. (2002). MAUI: a multimodal affective user interface. *Annual ACM International Conference on Multimedia, Juan-les-Pins, France* (pp. 161-170).
- Lisetti, C., & Nasoz, F. (2005). Affective intelligent car interfaces with emotion recognition. *International Conference on Human Computer Interaction, Las Vegas, USA* (pp. 421-430).
- Litman, D., & Forbes-Riley, K. (2004). Predicting student emotions in computer-human tutoring dialogues. *42nd Annual Meeting of the Association of Computational Linguistics* (pp. 352–359).
- Littlewort, G. C., Bartlett, M. S., & Lee, K. (2007). Faces of pain: automated measurement of spontaneous facial expressions of genuine and posed pain. *International Conference on Multimodal Interfaces, Nagoya, Aichi, Japan* (pp. 15-21).
- Loyall, A., & Bates, J. (1991). *Hap: A reactive, adaptive architecture for agents* (Technical Report CMU-CS-91-147). Pittsburgh, PA: School of Computer Science, Carnegie Mellon University.
- MacQueen, J. (1965). Some methods for classification and analysis of multivariate observations. *Berkley Symposium on Mathematical Statistics and Probability* (pp. 281–297).
- McCrae, R. R., & Costa, P. T. (1985). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and Individual Differences*, 6(5), 587-597.
- McLachlan, G. J., & Basford, K. E. (1988). *Mixture Models*: M. Dekker.
- Mehrabian, A., & Russell, J. A. (1980). *An approach to environmental psychology*: Cambridge, MA: MIT Press.
- Minsky, M. (1988). *The society of mind*. New York: Simon and Schuster.
- Minsky, M. (2007). *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind*. New York: Simon and Schuster.
- Mitchell, T. (1997). *Machine Learning*: McGraw Hill.
- Nasoz, F., Lisetti, C., Alvarez, K., & Finkelstein, N. (2003). Emotion recognition from physiological signals for user modeling of affect. *International Journal of Cognition, Technology and Work Special Issue on Presence*, 6(1), 4-14.

- Neiberg, D., Elenius, K., & Laskowski, K. (2006). Emotion recognition in spontaneous speech using GMMs. *International Conference Spoken Language Processing* (pp. 809-812).
- Newhagen, J., & Reeves, B. (1991). Emotion and memory responses for negative political advertising: A study of television commercials used in the 1988 presidential election. *Television and political advertising: Psychological processes, 1*, 197–220.
- Newhagen, J. E., & Reeves, B. (1992). This evening's bad news: Effects of compelling negative television news images on memory. *Journal of Communication*, 42(2), 25-41.
- Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition & Emotion*, 1(1), 29-50.
- Ochs, M., & Frasson, C. (2004). Emotionally Intelligent Tutoring Systems. *International Florida Artificial Intelligence Research Society Conference* (pp. 251-256).
- Ortony, A., Clore, G., & Collins, A. (1988). *The cognitive structure of emotions*: Cambridge University Press.
- Pantic, M., Pentland, A., Nijholt, A., & Huang, T. (2007). Human Computing and Machine Understanding of Human Behavior: A Survey. Dans T. Huang, A. Nijholt, M. Pantic & A. Pentland (éds.), *Artificial Intelligence for Human Computing* (4451, pp. 47-71): Springer Berlin / Heidelberg.
- Pantic, M., & Rothkrantz, L. J. M. (2004). Facial action recognition for facial expression analysis from static face images. *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics*, 34(3), 1449-1461.
- Parke, F., & Waters, K. (2008). *Computer facial animation*: AK Peters Ltd.
- Pavlovic, V. I., Sharma, R., & Huang, T. S. (1997). Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 677-695.
- Picard, R. W. (2000). *Affective computing*: Cambridge, MA: The MIT Press.
- Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), 1175-1191.
- Quinlan, J. (1993). *C4. 5: programs for machine learning*: Morgan Kaufmann.
- Reeves, B., Newhagen, J., Maibach, E., Basil, M., & Kurz, K. (1991). Negative and positive television messages: Effects of message type and context on attention and memory. *American Behavioral Scientist*, 34(6), 679.

- Reilly, W. S. (1997). A methodology for building believable social agents. *International Conference on Autonomous Agents, Marina del Rey, California, United States* (pp. 114-121).
- Reilly, W. S., & Bates, J. (1992). *Building emotional agents* (Technical Report CMU-CS-92-134). Pittsburgh, PA: School of Computer Science, Carnegie Mellon University.
- Roseman, I., & Smith, C. (2001). Appraisal theory: Overview, assumptions, varieties, controversies. Dans K. R. Scherer, A. Schorr & T. Johnstone (éds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 3-19): Oxford University Press.
- Roseman, I. J., Antonio, A. A., & Jose, P. E. (1996). Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. *Cognition & Emotion, 10*(3), 241-278.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: MIT Press.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145-172.
- Russell, J. A., Bachorowski, J. A., & Fernandez-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54*, 329-349.
- Rusting, C. L., & Larsen, R. J. (1997). Extraversion, neuroticism, and susceptibility to positive and negative affect: A test of two theoretical models. *Personality and Individual Differences, 22*(5), 607-612.
- Sander, D., Grandjean, D., & Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Networks, 18*(4), 317-352.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99*(2), 143-165.
- Scherer, K. R. (1993). Studying the emotion-antecedent appraisal process: An expert system approach. *Cognition & Emotion, 7*(3), 325-355.
- Scherer, K. R. (1997). The role of culture in emotion-antecedent appraisal. *Journal of Personality and Social Psychology, 73*(5), 902-922.
- Scherer, K. R. (1999). Appraisal theory. Dans T. Dalgleish & M. J. Power (éds.), *Handbook of cognition and emotion* (pp. 637-663). Chichester, England: John Wiley & Sons Ltd.

- Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. Dans K. R. Scherer, A. Schorr & T. Johnstone (éds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92-120): Oxford University Press.
- Scherer, K. R. (2004). HUMAINE Deliverable D3c: Preliminary plans for exemplars: theory. . Consulté le 11 June 2009, tiré de.
- Scherer, K. R., & Tannenbaum, P. H. (1986). Emotional experiences in everyday life: A survey approach. *Motivation and emotion*, 10(4), 295-314.
- Scherer, K. R., & Wallbott, H. G. (1994). Evidence for Universality and Cultural Variation of Differential Emotion Response Patterning. *Journal of Personality and Social Psychology*, 66(2), 310-328.
- Scherer, K. R., Wranik, T., Sangsue, J., Tran, V., & Scherer, U. (2004). Emotions in everyday life: probability of occurrence, risk factors, appraisal and reaction patterns. *Social Science Information Sur Les Sciences Sociales*, 43(4), 499-570.
- Schwartz, G., & Weinberger, D. (1980). Patterns of emotional responses to affective situations: Relations among happiness, sadness, anger, fear, depression, and anxiety. *Motivation and emotion*, 4(2), 175-191.
- Sebe, N., Cohen, I., Gevers, T., & Huang, T. (2005). Multimodal approaches for emotion recognition: a survey. *SPIE-IS&T Electronic Imaging* (Vol. 5670, pp. 56-67).
- Sebe, N., Lew, M., Sun, Y., Cohen, I., Gevers, T., & Huang, T. (2007). Authentic facial expression analysis. *Image and Vision Computing*, 25(12), 1856-1863.
- Siemer, M., Mauss, I., & Gross, J. J. (2007). Same situation - Different emotions: How appraisals shape our emotions. *Emotion*, 7(3), 592-600.
- Slater, S., Buckley, K., & Bechkoum, K. (2006). Body Mind and Emotion An Overview of Agent Implementation in Mainstream Computer Games. *ICAPS English Lake District*. (pp. 9-13).
- Slater, S., Moreton, R., Buckley, K., & Bridges, A. (2008). A Review of Agent Emotion Architectures. *Eludamos Journal for Computer Game Culture*, 2(2), 203-214.
- Sloman, A. (2001). Beyond shallow models of emotion. *Cognitive processing*, 2(1), 177.
- Sloman, A., & Poli, R. (1995). SIM_AGENT: A toolkit for exploring agent designs. *Workshop on Agent Theories, Architectures, and Languages, International Joint Conference on Artificial Intelligence, Montréal* (pp. 392-407).
- Smith, C. A., & Kirby, L. D. (2000). Consequences require antecedents: toward a process model of emotion elicitation. Dans J. P. Forgas (éd.), *Feeling and thinking: The role of affect in social cognition* (pp. 83–106). Cambridge: Cambridge University Press.

- Speer, R., Krishnamurthy, J., Havasi, C., Smith, D., Lieberman, H., & Arnold, K. (2009). An interface for targeted collection of common sense knowledge using a mixture model. *International Conference on Intelligent User Interfaces, Sanibel Island, Florida, USA* (pp. 137-146).
- Staller, A., & Petta, P. (2001). Introducing emotions into the computational study of social norms: A first evaluation. *Journal of artificial societies and social simulation*, 4(1), 41.
- Steidl, S., Levit, M., Batliner, A., Nöth, E., & Niemann, H. (2005). "Off All Things the Measure Is Man" Automatic Classification of Emotions and Inter-Labeler Consistency. *International Conference on Acoustic, Speech and Signal Processing* (Vol. 1, pp. 317-320).
- Sutton, R., & Barto, A. (1998). *Reinforcement learning*: MIT Press.
- Tao, J., & Tan, T. (2005). Affective Computing: A Review. Dans J. Tao, T. Tan & R. Picard (éds.), *Affective Computing and Intelligent Interaction* (3784, pp. 981-995): Springer Berlin / Heidelberg.
- Thorson, E., & Friestad, M. (1989). The Effects of Emotion on Episodic Memory for TV Commercials. Dans P. Cafferata & A. Tybout (éds.), *Cognitive and affective responses to advertising* (pp. 305-326). Lexington, MA: Lexington Books.
- Tian, Y. L., Kanade, T., & Cohn, J. F. (2005). Facial Expression Analysis. Dans S. Z. Li & A. K. Jain (éds.), *Handbook of Face Recognition* (pp. 247-275): Springer.
- Tomkins, S. (1962). Affect, imagery, consciousness (Vols. 1&2). *New\ brk: Springer*.
- Trabelsi, A., & Frasson, C. (2010). The Emotional Machine: a Machine Learning Approach to Online Prediction of User's Emotion and Intensity. *IEEE International Conference on Advanced Learning Technologies, Sousse, Tunisia* (pp. 613-617).
- Valstar, M., Pantic, M., & Patras, I. (2004). Motion History for Facial Action Detection from Face Video. *IEEE International Conference on Systems, Man, and Cybernetics* (Vol. 1, pp. 635-640).
- Valstar, M. F., Pantic, M., Ambadar, Z., & Cohn, J. F. (2006). Spontaneous vs. posed facial behavior: automatic analysis of brow actions. *ACM International Conference on Multimodal Interfaces, Banff, Alberta, Canada* (pp. 162-170).
- Vapnik, V. (2000). *The Nature of Statistical Learning Theory*: Springer
- Vidrascu, L., & Devillers, L. (2005). Detection of real-life emotions in call centers. *European Conference on Speech Communication and Technology* (pp. 1841-1844).
- Wagner, J., Kim, J., & André, E. (2005). From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and

- classification. *International Conference on Multimedia and Expo, Amsterdam, Netherlands* (pp. 940-943).
- Wang, J., Yin, L., Wei, X., & Sun, Y. (2006). 3D facial expression recognition based on primitive surface feature distribution. *IEEE International Conference on Computer Vision and Pattern Recognition* (Vol. 2, pp. 1399-1406).
- Watkins, C., & Dayan, P. (1992). Q-learning. *Machine learning*, 8(3), 279-292.
- Watson, D., & Clark, L. A. (1992). On Traits and Temperament: General and Specific Factors of Emotional Experience and Their Relation to the Five-Factor Model. *Journal of Personality*, 60(2), 441-476.
- Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2008). A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31, 39-58.

Annexe I Apprentissage du perceptron multicouche

I.1 Représentation

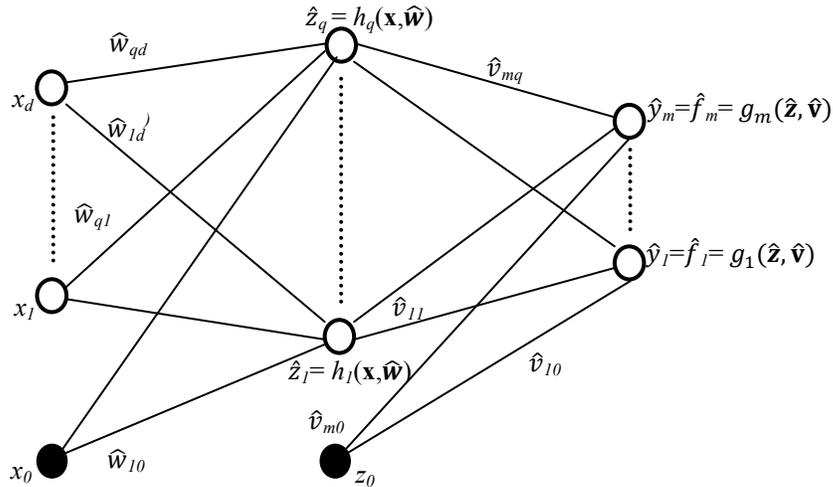


Figure I.1 Perceptron multicouche à une seule couche cachée (Bishop, 2006)

La Figure I.1 illustre le modèle général d'un réseau à une couche cachée présentant une connectivité totale (chaque neurone d'une couche donnée est relié à tous les neurones de la couche précédente). Les neurones de la première couche représentent le vecteur d'entrée et ceux de la dernière couche représentent le vecteur de sortie de la fonction apprise. Dans la couche cachée intermédiaire, chaque neurone ou *unité* j , avec j allant de 1 à q , calcule une combinaison linéaire des entrées (x_1, x_2, \dots, x_d) avec les coefficients ou poids constants qui lui correspondent $(\hat{w}_{j1}, \hat{w}_{j2}, \dots, \hat{w}_{jd})$. Ces poids déterminent la contribution de chaque valeur d'entrée à la sortie du neurone j . La combinaison linéaire est appelée *activation* et prend la forme suivante pour chaque unité cachée j :

$$a_j = \sum_{l=1}^d \hat{w}_{jl} x_l + \hat{w}_{j0} \quad (\text{I.1})$$

Chaque coefficient \hat{w}_{j0} correspond à l'entrée x_0 qui est délibérément initialisée à la valeur 1 (schématisée par le neurone noir dans la première couche de la Figure 1). Ainsi, ce coefficient constitue le *biais* d'une activation donnée. Chaque *activation* est ensuite transformée à l'aide d'une fonction non linéaire dérivable $h(\cdot)$ appelée *fonction d'activation*, pour donner

$$\hat{z}_j = h(a_j) \quad (\text{I.2})$$

Ces quantités \hat{z}_j correspondent aux sorties des neurones cachés. La tangente hyperbolique (\tanh), qui va de R dans $] -1; 1[$, est souvent choisie comme *fonction d'activation* h au niveau de la couche cachée (Bishop, 2006).

$$\tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \quad (\text{I.3})$$

Le même processus est réitéré avec les valeurs \hat{z}_j servant d'entrées pour chacun des m neurones ou *unités* de sortie, donnant lieu aux *activations* suivantes :

$$a_k = \sum_{j=1}^q \hat{v}_{kj} \hat{z}_j + \hat{v}_{k0} \quad (\text{I.4})$$

où k prend les valeurs $1, \dots, m$ et les \hat{v}_{kj} correspondent aux poids entrants du neurone de sortie k . Les \hat{v}_{k0} représentent les biais correspondant à $z_0=1$ (Figure 1). Ces activations a_k sont à leur tour transformées à l'aide d'une fonction non linéaire continue dérivable $g(\cdot)$ donnant lieu à :

$$\hat{f}_k = g(a_k) \quad (\text{I.5})$$

où les \hat{f}_k représentent les sorties finales de la fonction apprise. Le choix de la *fonction d'activation* $g(\cdot)$ reste relatif au type de la tâche d'apprentissage à effectuer. Par exemple, dans le cas de plusieurs classifications binaires la fonction *sigmoïde* est souvent utilisée pour chaque sortie au nombre des classifications à faire (Bishop, 2006). Dans le cas de la régression, la *fonction d'activation* correspond à la *fonction identité* et il n'y a qu'un seul neurone de sortie. Pour la multiclassification (k classes mutuellement exclusives), chaque classe est représentée par un neurone de sortie dont la valeur peut être interprétée comme une probabilité d'appartenance à cette classe plutôt qu'aux autres classes. Cette valeur est obtenue par l'application de la fonction *softmax*, appelée aussi *l'exponentielle normalisée*

$$\text{softmax}(a_k) = \frac{e^{a_k}}{\sum_{p=1}^m e^{a_p}} \quad (\text{I.6})$$

avec $\sum_{p=1}^m e^{a_p}$ représentant la somme, sur toutes les unités de sortie ($1, \dots, m$), de l'exponentielle des activations correspondantes a_p (l'indice p est utilisé afin de ne pas confondre avec l'indice k du numérateur). La valeur finale obtenue d'une unité de sortie k ,

$\hat{f}_k = \text{softmax}(a_k)$, se trouve entre 0 et 1 et la somme sur toutes les sorties est de $\sum_{k=1}^m \hat{f}_k = \sum_{k=1}^m \text{softmax}(a_k) = 1$. Ainsi, nous nous retrouvons avec un vecteur sortie $(\hat{f}_1, \hat{f}_2, \dots, \hat{f}_m)$ dont la composante ayant la plus grande valeur a pour indice le numéro de la classe prédite.

Pour résumer, le *réseau de neurones perceptron multicouche* peut être perçu comme une série de transformations fonctionnelles donnant lieu à une fonction non linéaire générale pouvant avoir comme entrée un vecteur $\mathbf{x} (x_1, x_2, \dots, x_d)$ et comme sortie un vecteur $\hat{\mathbf{f}} (\hat{f}_1, \hat{f}_2, \dots, \hat{f}_m)$, et contrôlée par des vecteurs $\hat{\mathbf{w}}$ et $\hat{\mathbf{v}}$ de paramètres incluant tous les poids et les biais. Pour une tâche de classification multiple avec une fonction *softmax* comme *fonction d'activation* de sortie, et la fonction *tangente hyperbolique* comme *fonction d'activation* au niveau de la couche cachée, cette fonction prend la forme suivante

$$\hat{f}_k(\mathbf{x}, \hat{\mathbf{w}}, \hat{\mathbf{v}}) = \text{softmax} \left(\sum_{j=1}^q \hat{v}_{kj} \tanh \left(\sum_{l=1}^d \hat{w}_{jl} x_l + \hat{w}_{j0} \right) + \hat{v}_{k0} \right) \quad (\text{I.7})$$

Le calcul de cette fonction est généralement considéré comme un processus d'évaluation ou de propagation de l'information à travers le réseau (*Bishop, 2006*), appelé aussi *propagation en avant* (*forward propagation* en anglais).

Lorsqu'on est face à un problème de classification avec deux variables cibles indépendantes, la première ayant m classes exclusives possibles et la deuxième c classes, le vecteur de sortie $\hat{\mathbf{f}}$ prend la forme suivante $(\hat{f}_{11}, \hat{f}_{12}, \dots, \hat{f}_{1m}, \hat{f}_{21}, \hat{f}_{22}, \dots, \hat{f}_{2c})$, avec le premier indice désignant la variable cible et le deuxième indice relatif à la classe de cette cible. Ainsi, la fonction générale $\hat{f}(\mathbf{x}, \hat{\mathbf{w}}, \hat{\mathbf{v}})$ est une composition de deux fonctions.

$$\hat{f}(\mathbf{x}, \hat{\mathbf{w}}, \hat{\mathbf{v}}) = \begin{cases} \hat{f}_{1k}(\mathbf{x}, \hat{\mathbf{w}}, \hat{\mathbf{v}}) = \text{softmax}_1 \left(\sum_{j=1}^q \hat{v}_{kj} \hat{z}_j + \hat{v}_{k0} \right) \\ \text{telque } k = 1, \dots, m \text{ et } \sum_{k=1}^m \text{softmax}_1(a_k) = 1 \\ \hat{f}_{2t}(\mathbf{x}, \hat{\mathbf{w}}, \hat{\mathbf{v}}) = \text{softmax}_2 \left(\sum_{j=1}^q \hat{v}_{tj} \hat{z}_j + \hat{v}_{t0} \right) \\ \text{telque } t = 1, \dots, c \text{ et } \sum_{t=1}^c \text{softmax}_2(a_t) = 1 \end{cases} \quad (\text{I.8})$$

avec

$$a_k = \sum_{j=1}^q \hat{v}_{kj} \hat{z}_j + \hat{v}_{k0} \quad (\text{I. 9})$$

les activations des neurones relatifs à la première cible et

$$a_t = \sum_{j=1}^q \hat{v}_{tj} \hat{z}_j + \hat{v}_{t0} \quad (\text{I. 10})$$

les activations des neurones de sorties correspondants à la deuxième cible.

L'apprentissage de ces fonctions consiste à trouver les paramètres $\hat{\mathbf{w}}$ et $\hat{\mathbf{v}}$ optimaux qui reflètent l'ensemble d'entraînement et dont on peut généraliser l'utilisation pour de nouveaux exemples.

I.2 Apprentissage des paramètres

I.2.1 La fonction d'erreur

Étant donné les exemples $(\mathbf{x}_i, \mathbf{y}_i)$ de l'ensemble d'entraînement D_n avec $i=1, \dots, n$, \mathbf{x}_i le vecteur d'entrée $(x_{i1}, x_{i2}, \dots, x_{id})$ et \mathbf{y}_i le vecteur correspondant aux valeurs vraies ou désirées de sortie $(y_{i1}, y_{i2}, \dots, y_{im})$, l'objectif consiste à minimiser une erreur de classification ou de prédiction de la fonction $\hat{f}(\mathbf{x}, \hat{\mathbf{w}}, \hat{\mathbf{v}})$ sur la totalité des n exemples. Sa forme pour la classification multiple (pour une seule variable cible) de k classes mutuellement exclusives est l'entropie croisée c'est-à-dire l'opposé du logarithme de la *fonction de vraisemblance* (Bishop, 2006)

$$E(\hat{\mathbf{w}}, \hat{\mathbf{v}}) = - \sum_{i=1}^n \sum_{k=1}^m y_{ik} \ln \hat{f}_k(\mathbf{x}_i, \hat{\mathbf{w}}, \hat{\mathbf{v}}) \quad (\text{I. 11})$$

Dans le cas de deux variables cibles indépendantes à plusieurs classes, on rajoute l'erreur par rapport à la deuxième cible

$$E(\hat{\mathbf{w}}, \hat{\mathbf{v}}) = - \sum_{i=1}^n \left(\sum_{k=1}^m y_{i1k} \ln \hat{f}_{1k}(\mathbf{x}_i, \hat{\mathbf{w}}, \hat{\mathbf{v}}) + \sum_{t=1}^c y_{i2t} \ln \hat{f}_{2t}(\mathbf{x}_i, \hat{\mathbf{w}}, \hat{\mathbf{v}}) \right) \quad (\text{I. 12})$$

I.2.2 Minimisation de la fonction d'erreur et la descente du gradient

Le problème consiste à trouver le vecteurs $\widehat{W}=(\widehat{w},\widehat{v})$ qui minimise cette fonction d'erreur. Par conséquent, nous sommes devant un problème d'optimisation qui peut être résolu, du fait que la fonction d'erreur est lisse, continue et non linéaire en W , en trouvant le vecteur W optimal qui annule le gradient de l'erreur $\nabla E(W)$. Le vecteur gradient constitue le vecteur des dérivées partielles de $E(W)$ par rapport à chaque poids et biais du réseau :

$$\nabla E(W) = \left(\frac{\partial E}{\partial w_{11}}, \frac{\partial E}{\partial w_{12}}, \dots, \frac{\partial E}{\partial w_{jl}}, \dots, \frac{\partial E}{\partial v_{kj}}, \dots, \frac{\partial E}{\partial v_{mq}} \right) \quad (\text{I. 13})$$

Cependant, il n'existe pas de solution analytique pour l'équation $\nabla E(W) = 0$. Ainsi, on se réfère aux procédures itératives pour la résolution d'équations continues non linéaires. Les solutions trouvées, dans ce cas, peuvent correspondre à des minima locaux. Le principe général de la plupart de ces techniques consiste à choisir un point de départ W_0 dans l'espace des poids et de bouger en faisant de petits pas successifs et itératifs dans cet espace en cherchant le point ou vecteur poids \widehat{W} qui minimise $E(W)$. Une bonne manière de choisir le point de départ consiste à initialiser les poids et biais uniformément dans l'intervalle $\left[-\frac{K}{nbrentr}, \frac{K}{nbrentr}\right]$ avec K un réel proche de 1 et $nbrentr$ égal à d ou à q . L'entier q désigne le nombre de neurones de la couche liée au poids et biais (Bottou, 1991). La technique d'optimisation la plus utilisée dans les réseaux de neurones est celle de la *descente de gradient* qui utilise l'information sur le gradient à un point donné. En effet le gradient de l'erreur en un point \widehat{W}_t indique la direction de la plus grande pente d'accroissement, dont l'astuce consiste à prendre des pas dans la direction totalement opposée au gradient

$$\widehat{W}_{t+1} = \widehat{W}_t - \eta \nabla E(\widehat{W}_t) \quad (\text{I. 14})$$

Le réel $\eta > 0$, qui représente le pas et qui prend généralement des valeurs assez faibles, est appelé *taux d'apprentissage*. Le gradient est évalué à chaque itération t en \widehat{W}_t avant de déterminer le point ou vecteur suivant \widehat{W}_{t+1} . Le nombre total d'itérations T est un *hyperparamètre* à spécifier. Il en est de même pour η . L'évaluation ou le calcul du gradient de l'erreur $\nabla E(\widehat{W}_t)$ nécessite la prise en compte de la totalité de l'ensemble d'entraînement puisque l'erreur $E(\widehat{W})$ est définie en fonction des n exemples d'entraînement (équation

I.11). Ce cas, où la mise à jour du poids nécessite l'évaluation du gradient en utilisant tous les exemples, correspond au cas ordinaire de la descente de gradient (appelé *batch* en anglais).

Il existe, cependant, une autre variante de *descente de gradient* qui s'avère plus utile, qui est dite *en ligne* (*Le Cun et al.*, 1989), appelée aussi *descente du gradient séquentielle* ou *stochastique*. Cette variante, par opposition à la version ordinaire « batch », permet de faire la mise à jour des poids en utilisant un seul exemple à la fois pour l'évaluation du gradient de l'erreur. En effet, d'après l'équation I.11, l'erreur totale peut être interprétée comme une somme d'erreurs individuelles sur les exemples:

$$E(\widehat{W}) = \sum_{i=1}^n E_i(\widehat{W}) \quad (\text{I.15})$$

Pour la multiclassification à une variable cible

$$E_i(\widehat{W}) = \sum_{k=1}^m y_{ik} \ln \hat{f}_k(\mathbf{x}_i, \widehat{W}) \quad (\text{I.16})$$

D'où la mise à jour des poids en utilisant un exemple à la fois pour le calcul du gradient :

$$\widehat{W}_{t+1} = \widehat{W}_t - \boldsymbol{\eta} \nabla E_i(\widehat{W}_t) \quad (\text{I.17})$$

Cette mise à jour est répétée en itérant sur chaque exemple disponible un certain nombre de fois.

Pour les deux variantes de descente de gradient, la convergence vers un minimum est garantie. Néanmoins, l'avantage majeur de la méthode stochastique est sa capacité à gérer la redondance des données ce qui la rend plus rapide (*Le Cun et al.*, 1989). D'un autre côté, cette méthode permet d'échapper aux minima locaux « aigus » ou « pointus » qui peuvent nuire à la *généralisation*, ce qui la rend plus robuste. En outre, elle peut s'avérer très utile pour le cas de traitement d'une masse importante de données et également dans le cas où les exemples sont entrés en temps réel ou en ligne et l'apprentissage se fait au fur et à mesure.

Nous avons indiqué que la mise à jour des poids dans la descente de gradient nécessite une évaluation du gradient de l'erreur en un point \widehat{W}_t . Cette évaluation revient

alors à calculer les différentes dérivées partielles de l'erreur par rapport au poids et biais (I.13) d'une manière efficace.

1.2.3 La rétropropagation (cas de deux variables cibles)

La *rétropropagation* (*back propagation* en anglais) est la technique qui permet une évaluation efficace, en matière de calcul, des dérivées de l'erreur par rapport aux poids et biais (*Bishop*, 2006). La notion de propagation vient du fait qu'une information (sur l'erreur dans le cas de la rétropropagation) est propagée dans le réseau. En d'autres termes, des calculs intermédiaires sont effectués et par la suite utilisés pour d'autres calculs, comme pour la *propagation en avant* où l'on calcule par exemple des quantités intermédiaires \hat{z}_j (équation I.2) pour calculer les sorties \hat{f}_k (équation I.7). Cependant pour la *rétropropagation*, les calculs se font à l'inverse de la *propagation en avant* en calculant des dérivées à partir des activations de la dernière couche en revenant jusqu'aux activations de la couche d'entrée. Ce calcul est effectué d'une manière récursive en évaluant à chaque fois les dérivées de l'erreur sur les poids et biais entre deux couches données. D'autre part, les quantités évaluées dans la *propagation en avant* sont utilisées dans la *rétropropagation* pour le calcul des dérivées de l'erreur.

La *rétropropagation* repose sur la règle de la dérivation en chaîne. Dans un cas général, lorsque cette règle est appliquée pour dériver une fonction f par rapport à x , où f est fonction de plusieurs autres fonctions $g_1(x)$, $g_2(x)$, ..., $g_n(x)$, elle prend la forme suivante :

$$\frac{\partial f(g_1(x), g_2(x), \dots, g_n(x))}{\partial x} = \sum_{i=1}^n \frac{\partial f}{\partial g_i} \frac{\partial g_i}{\partial x} \quad (\text{I.18})$$

L'équation (I.18) est la somme des produits croisés des dérivées de f par rapport aux fonctions g_i et celles des g_i par rapport à x .

Dans ce qui suit, nous présentons les différentes étapes de calcul des dérivées de l'erreur individuelle $E_i(\hat{W})$ (pour la descente de gradient stochastique), dans le cas du réseau de neurones perceptron multicouche à une couche cachée, ayant comme tâche la multiclassification à deux variables cibles (équation I.8). La fonction d'activation de la

couche cachée est la *tangente hyperbolique* (équation I.3), celle de la sortie est la *softmax*. D'après l'équation I.12, l'erreur individuelle prendrait la forme suivante :

$$E_i(\widehat{W}) = \sum_{k=1}^m y_{i1k} \ln \hat{f}_{1k}(\mathbf{x}_i, \widehat{W}) + \sum_{t=1}^c y_{i2t} \ln \hat{f}_{2t}(\mathbf{x}_i, \widehat{W}) \quad (\text{I.19})$$

Les fonctions $\hat{f}_{1k}(\mathbf{x}_i, \widehat{W})$ et $\hat{f}_{2v}(\mathbf{x}_i, \widehat{W})$ sont définies par l'équation I.8. Dans un premier temps, on calcule les dérivées de l'erreur par rapport aux différentes activations de sortie a_k (pour $k=1, \dots, m$) et a_t (pour $t=1, \dots, c$), relatives aux deux variables cibles. On peut démontrer que ces dérivées donnent les résultats suivants (*Bishop*, 2006):

$$\frac{\partial E_i(\widehat{W})}{\partial a_k} = \hat{f}_{i1k} - y_{i1k} \quad (\text{I.20})$$

$$\frac{\partial E_i(\widehat{W})}{\partial a_t} = \hat{f}_{i2t} - y_{i2t} \quad (\text{I.21})$$

avec \hat{f}_{i1k} et \hat{f}_{i2t} dénotant respectivement $\hat{f}_{1k}(\mathbf{x}_i, \widehat{W})$ et $\hat{f}_{2t}(\mathbf{x}_i, \widehat{W})$. Ces dérivées calculées vont servir à évaluer les dérivées de l'erreur par rapport aux différents poids et biais entre la couche de sortie et la couche cachée : v_{kj} , v_{k0} , v_{tj} et v_{t0} . Elles vont aussi être utiles pour le calcul des dérivées de l'erreur par rapport aux activations des couches cachées a_j . Ainsi, pour calculer la dérivée par rapport aux biais en appliquant la règle de dérivation en chaîne sur les équations I.8, on obtient pour $k=1, \dots, m$

$$\begin{aligned} \frac{\partial E_i(\widehat{W})}{\partial v_{k0}} &= \frac{\partial E_i(\widehat{W})}{\partial a_k} \frac{\partial a_k}{\partial v_{k0}} \\ &= \frac{\partial E_i(\widehat{W})}{\partial a_k} \end{aligned} \quad (\text{I.22})$$

puisque $\frac{\partial a_k}{\partial v_{k0}} = 1$ (équation I.9). De même, pour $t=1, \dots, c$, sachant que $\frac{\partial a_t}{\partial v_{t0}} = 1$ (équation I.10)

$$\begin{aligned} \frac{\partial E_i(\widehat{W})}{\partial v_{t0}} &= \frac{\partial E_i(\widehat{W})}{\partial a_t} \frac{\partial a_t}{\partial v_{t0}} \\ &= \frac{\partial E_i(\widehat{W})}{\partial a_t} \end{aligned} \quad (\text{I.23})$$

Les dérivées de l'erreur par rapport aux poids v_{kj} , qui se trouvent entre les m neurones de sorties relatifs à la première cible et les q neurones cachés, prennent la forme suivante :

$$\frac{\partial E_i(\widehat{W})}{\partial v_{kj}} = \frac{\partial E_i(\widehat{W})}{\partial a_k} \frac{\partial a_k}{\partial v_{kj}} \quad (I.24)$$

En se basant sur l'équation I.9, $\frac{\partial a_k}{\partial v_{kj}} = \hat{z}_j$, d'où :

$$\frac{\partial E_i(\widehat{W})}{\partial v_{kj}} = \frac{\partial E_i(\widehat{W})}{\partial a_k} \hat{z}_j \quad (I.25)$$

De même, pour les dérivées sur les poids v_{tj} relatifs à la deuxième cible

$$\frac{\partial E_i(\widehat{W})}{\partial v_{tj}} = \frac{\partial E_i(\widehat{W})}{\partial a_t} \hat{z}_j \quad (I.26)$$

Pour le calcul des poids et biais entre la couche cachée et la couche d'entrée, la même procédure que nous venons de présenter pour les poids et biais entre la couche de sortie et la couche cachée est réitérée. Dans un premier temps, on calcule les dérivées de l'erreur par rapport aux activations a_j des neurones cachés (équation I.1) selon la règle de dérivation en chaîne :

$$\frac{\partial E_i(\widehat{W})}{\partial a_j} = \sum_{k=1}^m \frac{\partial E_i(\widehat{W})}{\partial a_k} \frac{\partial a_k}{\partial a_j} + \sum_{t=1}^c \frac{\partial E_i(\widehat{W})}{\partial a_t} \frac{\partial a_t}{\partial a_j} \quad (I.27)$$

Les $\frac{\partial E_i(\widehat{W})}{\partial a_k}$ et $\frac{\partial E_i(\widehat{W})}{\partial a_t}$ sont déjà calculés aux étapes précédentes. Pour $\frac{\partial a_k}{\partial a_j}$ et $\frac{\partial a_t}{\partial a_j}$ on applique également la dérivée en chaîne en tenant compte des équations I.2, I.9 et I.10 :

$$\frac{\partial a_k}{\partial a_j} = \frac{\partial a_k}{\partial z_j} \frac{\partial z_j}{\partial a_j} \text{ et } \frac{\partial a_t}{\partial a_j} = \frac{\partial a_t}{\partial z_j} \frac{\partial z_j}{\partial a_j} \quad (I.28)$$

$\frac{\partial a_k}{\partial z_j} = \hat{v}_{kj}$ (d'après l'équation I.9), $\frac{\partial a_t}{\partial z_j} = \hat{v}_{tj}$ (d'après l'équation I.10) et $\frac{\partial z_j}{\partial a_j} = (1 - z_j^2)$ dans le cas où la fonction d'activation de la couche cachée est une tangente hyperbolique car $\frac{\partial \tanh(x)}{\partial x} = 1 - \tanh^2(x)$. Ainsi,

$$\frac{\partial E_i(\widehat{W})}{\partial a_j} = (1 - \hat{z}_j^2) \left(\sum_{k=1}^m \frac{\partial E_i(\widehat{W})}{\partial a_k} \hat{v}_{kj} + \sum_{t=1}^c \frac{\partial E_i(\widehat{W})}{\partial a_t} \hat{v}_{tj} \right) \quad (I.29)$$

Comme lors de l'étape précédente, à partir de cette dérivée on calcule les dérivées de l'erreur par rapport aux poids w_{jl} et biais w_{j0} entre les q neurones de la couche cachée et les d neurones d'entrée. Par conséquent, pour $j=1, \dots, q$, on a

$$\begin{aligned} \frac{\partial E_i(\widehat{W})}{\partial w_{j0}} &= \frac{\partial E_i(\widehat{W})}{\partial a_j} \frac{\partial a_j}{\partial w_{j0}} \\ &= \frac{\partial E_i(\widehat{W})}{\partial a_j} \end{aligned} \quad (\text{I.30})$$

puisque $\frac{\partial a_j}{\partial w_{j0}} = 1$ (voir équation I.1). Pour $j=1, \dots, q$ et $l=1, \dots, d$

$$\frac{\partial E_i(\widehat{W})}{\partial w_{jl}} = \frac{\partial E_i(\widehat{W})}{\partial a_j} \frac{\partial a_j}{\partial w_{jl}} = \frac{\partial E_i(\widehat{W})}{\partial a_j} x_l \quad (\text{I.31})$$

Ainsi, le calcul des différentes dérivées partielles du gradient de l'erreur individuelle est effectué en appliquant la *rétropropagation*. Dans le cas, où la méthode d'optimisation est le gradient ordinaire « batch », on effectue la somme de ces dérivées sur les n exemples d'entraînement.

I.3 Astuces dans l'apprentissage des perceptrons multicouches

I.3.1 Régularisation

Nous avons introduit à la section 2.2.2 du mémoire, la notion de *surapprentissage*. Pour remédier au problème du surapprentissage, on peut rajouter un terme de *régularisation* (*weight decay* en anglais) à la fonction d'erreur. Ce terme consiste, dans sa forme la plus basique, en une pénalité constante qui accroît avec l'augmentation des valeurs des poids et biais lors de l'apprentissage. Ceci mènera vers des vecteurs de poids composés par de petites valeurs, réduisant ainsi le risque de surapprentissage (*Bishop, 2006*). Le terme rajouté peut être sous forme d'une constante λ prédéterminée multipliée par la norme du vecteur \widehat{W} au carré, donnant lieu à l'erreur régularisée suivante (pour une multiclassification avec deux cibles) :

$$E(\widehat{W}) = - \sum_{i=1}^n \left(\sum_{k=1}^m y_{i1k} \ln \hat{f}_{1k}(\mathbf{x}_i, \widehat{W}) + \sum_{v=1}^c y_{i2v} \ln \hat{f}_{2v}(\mathbf{x}_i, \widehat{W}) \right) + \frac{\lambda}{2} \|\widehat{W}\|^2 \quad (\text{I.32})$$

Cette régularisation quadratique est la plus simple (appelée aussi régularisation L2), et le λ constitue un hyperparamètre à spécifier ou à rechercher. Le calcul des dérivées de cette erreur se fait comme nous l'avons indiqué à la section précédente sauf qu'on tiendra également compte du terme de régularisation lors des dérivées par rapport aux poids et biais.

I.3.2 Choix des hyperparamètres

Nous avons introduit les notions d'ensemble de validation et d'hyperparamètre à la section 2.2.2 du mémoire, et nous avons expliqué que cet ensemble sert principalement à tester la performance de l'apprentissage avec plusieurs valeurs d'hyperparamètre afin de choisir celle qui donne la meilleure performance.

Dans un réseau de neurones on peut considérer le taux de *régularisation* λ , le *taux d'apprentissage* η , le nombre T d'itérations lors de la mise à jour des poids, et le nombre q des neurones cachés comme des hyperparamètres. On peut tester les valeurs séparément pour chacun *des* hyperparamètres, comme on peut tester plusieurs combinaisons possibles entre eux, mais cela prolongera beaucoup le traitement.

Annexe II Les situations émotionnelles collectées

Les 250 situations émotionnelles collectées sont regroupées dans 44 groupes relatifs aux 22 émotions du modèle OCC (*Ortony, Clore, & Collins, 1988*) et les deux intensités *forte* et *moyenne*.

II.1 Groupe des situations relatives à l'émotion *Heureux pour*

II.1.1 Intensité moyenne

- Quelqu'un que vous connaissez a réussi à réaliser ses objectifs malgré les obstacles qu'il a rencontrés.
- Un proche à vous arrive à fonder une famille (se marier et avoir des enfants).
- Vous apprenez qu'un ancien camarade de classe était atteint d'un cancer, mais qu'il s'en est sorti après quelques mois de thérapie.
- Votre meilleur ami arrête de fumer.

II.1.2 Intensité forte

- Un proche guérit après avoir contracté une grave maladie, alors que vous n'y croyiez plus.
- Votre meilleur ami arrive enfin à avoir son Ph.D à l'âge de 38 ans après avoir galéré pendant des années.
- Un inculpé accusé à tort est libéré après quelques semaines de prison.

II.2 Groupe des situations relatives à l'émotion *Pitié*

II.2.1 Intensité moyenne

- L'enfant de l'un de vos proches est gravement malade.
- Vous recevez une mauvaise nouvelle : un ami approchant la quarantaine, avec qui vous avez eu beaucoup de problèmes et avez coupé le contact depuis des années, est décédé.
- En prenant le métro, vous apercevez des sans-abris allongés sur les bancs à l'intérieur de la station.
- Un ami d'enfance vient vous demander une aide financière, car il n'a plus de quoi vivre, depuis qu'il a perdu ses jambes dans un accident.

II.2.2 Intensité forte

- Un de vos proches essaye tout pour atteindre ses objectifs, mais n'y arrive pas.
- Le soir, en rentrant du boulot, vous déposez chez lui un collègue avec qui vous venez de faire connaissance. Le lendemain, vous apprenez qu'il est décédé.

- Vous apercevez un enfant errant qui demande à manger.
- Vous entendez parler d'un pays qui vient d'être touché par une catastrophe naturelle qui a fait des milliers de morts.
- Vous apprenez qu'un ancien ami à vous a perdu sa femme et ses enfants dans un accident de voiture.

II.3 Groupe des situations relatives à l'émotion *Exultation*

II.3.1 Intensité moyenne

- Votre équipe sportive préférée bat sa plus grande rivale, après avoir dominé tout au long du match.
- Vous entrez en compétition avec une personne (dans n'importe quel domaine), et vous arrivez à gagner sans fournir d'efforts.
- Vous êtes témoin de la souffrance d'une personne qui vous a causé du mal et qui vous a fait souffrir.

II.3.2 Intensité forte

- Une personne essaye de vous mettre des bâtons dans les roues, mais vous réussissez malgré tout, dans la souffrance, à atteindre vos objectifs.
- Contre toute attente, le candidat (politique) pour lequel vous avez voté a été élu.
- Vous dominez tout le monde au jeu d'échec.

II.4 Groupe des situations relatives à l'émotion *Jalousie*

II.4.1 Intensité moyenne

- Vous voyez quelqu'un réussir, sans trop d'effort, ce qui vous a pris des années de sacrifices à le faire.
- Vous apercevez votre copain (copine) en train de sympathiser avec une belle fille (beau garçon).
- L'équipe sportive rivale à celle que vous supportez, raffle tous les titres et trophées possibles.
- Un très bon ami à vous, qui mérite amplement d'être embauché, est employé après un entretien. Quant à vous, en passant le même entretien, vous ne décrochez pas d'emploi.

II.4.2 Intensité forte

- Votre ex-femme (ex-époux), avec laquelle (lequel) vous avez eu des enfants, se remarie avec un(e) ancien(ne) ami(e) à vous.
- Vous apprenez qu'un jeune père de famille (que vous ne connaissez pas), ayant des enfants âgés de 2 et 4 ans, décède à 35 ans.
- Un collègue pour lequel vous n'éprouvez pas de sympathie et que vous jugez incompetent est promu et devient votre supérieur par la même occasion.

II.5 Groupe des situations relatives à l'émotion *Espoir*

II.5.1 Intensité moyenne

- Prévoir faire la fête avec vos amis après la fin des examens.
- Vous avez passé une entrevue pour avoir un emploi. Vous attendez la réponse depuis deux jours. Le téléphone sonne : vous voyez s'afficher le numéro de l'employeur.
- Vous êtes issu d'une famille pauvre. Vous entendez parler de beaucoup de « stars » (vedettes) issues de familles modestes qui ont fait fortune.
- Vous lisez dans les journaux les gros chiffres d'affaires que réalisent les grandes compagnies, et vous vous imaginez alors dans un futur où vous êtes le patron d'une grande entreprise.
- Vous avez une offre de travail à l'étranger loin de votre famille et de vos proches, vous êtes indécis, mais vos parents vous encouragent à y aller.

II.5.2 Intensité forte

- Vous trouvez de nouvelles idées, de nouveaux projets pour votre avenir et vous réfléchissez à la manière de les concrétiser.
- Vous entendez des collègues à vous dire qu'il y aura peut-être des remaniements au sein de la société et qu'il y a de fortes chances que vous soyez promu.
- Après deux années de travail sans arrêt, votre employeur vous attribue une dernière tâche et vous promet des vacances, si jamais vous l'accomplissez.
- Vous êtes à un mois d'avoir un nouveau-né.
- Vous sentez que vous vous approchez d'un objectif dont vous rêvez depuis longtemps.

II.6 Groupe des situations relatives à l'émotion *Peur*

II.6.1 Intensité moyenne

- Vous êtes à 5 minutes de passer l'examen le plus important de votre cursus.
- Vous attendez le résultat d'un examen médical.
- Votre supérieur vous annonce que vous allez être muté vers une nouvelle filiale dans un autre pays, alors que vous n'avez jamais quitté l'actuel (pays).
- Vous avez enfreint des règles, mais vous n'êtes pas sûre que personne ne s'en est aperçu.
- Votre téléphone sonne à 4 heures du matin. Vous n'avez aucune idée sur l'identité de l'appelant ; vous vous dépêchez pour décrocher.
- Vous attendez le résultat d'un examen, sachant que vous ne l'avez pas bien passé.
- Vous croyez entendre qu'un proche a eu un accident.
- Vous allez devoir faire une présentation orale, pour la première fois, avec une langue que vous ne maîtrisez pas très bien.
- Vous vous préparez pour votre première journée de travail.

- Vous êtes à la cuisine et vous entendez votre bébé pleurer.

II.6.2 Intensité forte

- Vous avez l'impression que quelqu'un vous suit dans la rue.
- Vous ne savez pas ce que vous allez faire à l'avenir (ça semble mal parti pour quelque temps).
- En passant prendre votre enfant de 4 ans, de la garderie, vers la fin de la journée, vous l'apercevez de loin en train de pleurer.
- Vous n'avez pas de nouvelles de votre famille depuis quelques semaines. Ce qui n'est pas du tout habituel.
- Pendant toute une semaine, vous sentez des douleurs inhabituelles au niveau du cœur.
- Vous pensez à la mort.
- Dans la météo, on prévoit qu'il y aurait une grande tempête tropicale qui pourrait s'abattre sur votre ville.

II.7 Groupe des situations relatives à l'émotion *Satisfaction*

II.7.1 Intensité moyenne

- Vous réussissez à avoir un boulot mieux rémunéré que l'actuel, mais vous devez quitter vos collègues de travail avec lesquels vous avez développé une grande amitié.
- Quelqu'un s'attache de plus en plus à vous.
- Vous êtes motivé financièrement pour réaliser un travail.
- Vous assistez à un cours où tout vous semble compréhensible et facile.
- Vous assistez pendant quelques instants à quelque chose d'extraordinaire, d'historique, à un chef-d'œuvre artistique ou un exploit sportif, etc.

II.7.2 Intensité forte

- Avec quelques amis à vous, à partir d'un petit budget, vous réussissez à mettre en place un petit projet.
- Vous arrivez à avoir des vacances après une longue année fatigante de travail et vous avez enfin du temps à passer avec vos proches et amis.
- Après un entretien de travail, vous n'êtes pas sûre d'avoir bien répondu aux questions posées, mais un coup de téléphone vient vous confirmer que vous êtes embauché.
- Vous arrivez, enfin, après des années de travail acharné et de sacrifices, à avoir le succès dont vous avez toujours rêvé.
- Après dix bonnes années de travail dans votre entreprise, vous êtes, comme prévu, promu au poste de directeur général.

II.8 Groupe des situations relatives à l'émotion *Peur confirmée*

II.8.1 Intensité moyenne

- Vous n'avez pas eu le temps de bien préparer un examen. Vous n'êtes pas sûre de pouvoir bien passer. Lorsque vous êtes devant la feuille d'examen, vous n'arrivez pas à répondre aux questions.
- Vous ne réussissez pas un examen d'études très important pour votre avenir, même si au fond vous saviez que les chances que vous le réussissiez étaient minimales.
- Vous obtenez une mauvaise note dans un examen (ordinaire, pas très important) que vous avez très mal passé.

II.8.2 Intensité forte

- On découvre que vous avez triché pour avoir de bonnes notes à l'examen.
- Vous ne vous sentez pas bien. Vous allez voir un médecin. Ce dernier vous confirme que vous êtes malade.
- Vous êtes un passager dans un avion. Ce dernier traverse une zone de turbulences qui dure quelques minutes.
- Vous entendez des rumeurs sur des limogeages de personnels au sein de votre entreprise. Quelques jours plus tard, vous apprenez que vous êtes démis de vos fonctions.
- Vous entamez un projet. En cours de route, vous trouvez beaucoup d'obstacles. Vous avez le pressentiment que vous n'allez pas atteindre vos objectifs. Vous n'abdiquez pas, vous continuez à travailler dur, mais en final vous n'y arrivez pas.

II.9 Groupe des situations relatives à l'émotion *Soulagement*

II.9.1 Intensité moyenne

- Alors que son nom avait presque été oublié, un criminel de guerre est arrêté après des dizaines d'années de cavale.
- Vous n'avez pas de nouvelles de votre famille depuis quelques semaines. Vous ne pouvez pas les joindre. Mais, vous parvenez finalement à avoir leurs nouvelles à travers un proche.
- Dans la météo, on prévoyait un hiver très dur plein de neiges et avec des températures qui avoisineraient les -30°C. Mais finalement, l'hiver a été beaucoup moins dur.
- Votre supérieur vous annonce que vous allez être muté vers une nouvelle filiale dans une autre ville, alors que vous n'avez jamais quitté l'actuelle (ville). Mais finalement, la décision est annulée et vous restez dans la même filiale.

II.9.2 Intensité forte

- Vous faites un accident de voiture, mais vous vous en sortez indemne.
- Vous passez un examen dans le cadre de vos études. Vous pensez que vous l'avez totalement raté. Mais, lors de la remise des résultats, vous obtenez une note bien meilleure que vous ne l'avez cru.

- Vous êtes limité par un délai très court pour remettre un travail. Vous croyez que vous n'allez pas pouvoir le terminer à temps. Après quelques nuits sans sommeil, vous arrivez finalement à le remettre à la dernière minute.
- Vous entendez des rumeurs sur des limogeages de personnels au sein de votre entreprise. En fin de compte, vous n'êtes pas parmi les partants.
- Votre équipe sportive préférée joue un match décisif pour remporter un titre. Alors, qu'elle est malmenée et dominée par l'équipe adverse pendant toute la rencontre, au moment où vous l'attendiez le moins, elle resurgit et gagne en toute fin du match.

II.10 Groupe des situations relatives à l'émotion *Déception*

II.10.1 Intensité moyenne

- Au boulot, les choses ne marchent pas comme vous le souhaitez.
- Vous faites un régime alimentaire pour perdre du poids, vous faites un footing chaque matinée. Mais vous n'arrivez toujours pas à perdre le poids voulu.
- Vous échouez à réaliser une tâche que vous avez qualifiée de facile.
- Vous avez quelque chose d'important à faire, mais vous tombez malade.

II.10.2 Intensité forte

- Vous souhaitez profondément avoir un boulot, car vous êtes au chômage depuis un an. Vous passez un entretien de job qui se déroule bien, mais à la fin, vous n'êtes pas accepté.
- Vous investissez une grande partie de votre argent dans un projet qui, finalement, n'aboutit pas.
- Vous arrivez à avoir une admission dans une université prestigieuse, mais vous n'avez pas les moyens pour payer les frais d'inscription.
- Vous êtes la plupart du temps en manque d'argent, vous n'arrivez pas à faire des projets à avoir des perspectives, à réaliser vos objectifs.
- Vous voulez avoir des études supérieures, mais, à chaque fois, des obstacles vous en empêchent.
- Vous faites des efforts pour gagner la confiance de vos supérieurs, mais vous n'y arrivez pas.

II.11 Groupe des situations relatives à l'émotion *Joie*

II.11.1 Intensité moyenne

- Dans le passé, vous avez été contraint d'abandonner vos études supérieures, mais il y a des moments où vous vous imaginez l'espace de quelques secondes en train de passer une brillante soutenance de thèse pour obtenir le titre de docteur.
- Vous allez vous marier après une longue histoire d'amour avec la femme (l'homme) que vous aimez, mais vous êtes tous les deux contraints de vivre très loin de vos familles.

- Vous vous sentez seul. Vous venez de vivre une expérience amoureuse qui a mal tournée. Vous écoutez de la musique, votre esprit vous emmène loin. Vous imaginez, alors, avoir trouvé l'amour de votre vie.
- Vous passez une bonne journée avec vos meilleurs amis. En rentrant, vous vous allongez sur le canapé et vous revoyez les beaux moments de cette journée.
- Vous regardez la télé, et vous tombez sur ancienne série que vous adoriez lorsque vous étiez beaucoup moins jeune.
- Vous êtes recruté par une très grande firme multinationale, vous allez devoir quitter pendant quelques années votre pays de résidence. Votre conjoint (e) ne pourra vous rejoindre que dans six mois.
- Vous retrouvez après des années sans aucune nouvelle un(e) ami(e) d'enfance avec qui vous entreteniez une très bonne relation.
- Vous rencontrez quelqu'un pour la première fois et vous sentez le courant passé, un feeling partagé.
- Être entouré de vos amis : souper, discuter, s'amuser, faire la fête.
- Très bien habillé(e), vous allez à une fête. Une fois là -bas, une jolie fille (un beau mec) vous drague, ce qui n'est pas habituel pour vous.

II.11.2 Intensité forte

- Après des années de mariage, le médecin vous annonce que vous allez avoir un enfant, juste avant que vous n'abdiquiez et perdiez espoir.
- Vous êtes au chômage depuis des mois. Vous jouez au loto comme chaque semaine, depuis longtemps ! Vous gagnez !
- Vos parents vous affirment qu'ils sont très fiers de vous, alors qu'ils ne vous l'ont jamais dit auparavant.
- Rien ne semble difficile ; pas de crise financière, pas de soucis sur votre avenir et l'avenir de vos proches ; tout ce que vous entreprenez vous le réussissez ; la vie semble simple et facile.
- La découverte d'un nouveau vaccin ou d'un remède contre une maladie incurable.
- Vous mettez fin à votre relation avec la femme (ou l'homme) que vous aimez. Quelques années plus tard, toujours amoureux, vous essayez de renouer contact avec elle (ou lui) et vous parvenez à regagner son cœur.
- Vous êtes très riche.
- Une personne qui vous est chère vous affirme que vous la rendez heureuse.
- Ca fait quelques mois que vous passez les entretiens de job en acceptant d'être sous-payé, mais toujours sans succès. Subitement, vous apprenez que vous venez d'hériter une fortune suite au décès d'un parent lointain dont vous ne connaissiez même pas l'existence.
- Votre pays gagne une médaille olympique.
- Vous êtes réunis avec toute votre famille autour d'un grand barbecue.

II.12 Groupe des situations relatives à l'émotion *Peine*

II.12.1 Intensité moyenne

- Vous marchez dans la rue et vous apercevez, un inconnu en pleure.
- Vous avez des soupçons que votre femme (marie) vous trompe depuis un bon bout de temps, mais vous êtes loin d'en être sûr(e).
- Vous vivez très loin des gens qui vous sont chers, même si ce n'est pas un choix que vous avez fait, mais une nécessité pour une raison particulière.
- Votre père subit une opération chirurgicale à l'hôpital. Vous imaginez le pire se produire.
- Vous vous êtes disputés avec votre meilleur(e) ami(e).
- Vous mettez fin à votre relation avec votre conjoint(e), après de longues années de vie commune parce que vous ne vous entendez plus, même si vous vous aimez encore.
- Voir un proche à vous qui s'engouffre dans la consommation de drogues et de l'alcool, sans rien pouvoir faire.

II.12.2 Intensité forte

- L'homme (femme) que vous aimez vous trompe.
- La perte d'un parent.
- Vos proches ne vous soutiennent pas quand vous en avez besoin
- Vous êtes en conflit avec des personnes qui vous sont très chères.
- Vous vivez seul dans un petit appartement dans un pays étranger, le téléphone n'a pas sonné depuis une semaine, personne ne vous a contacté.
- Votre grand-père, avec qui vous entretenez une relation affective particulière, est décédé à un âge avancé.
- Vous apprenez que quelqu'un de vos proches souffre d'une maladie grave.
- Vous voyez votre mère pleurer sans que vous sachiez la raison.

II.13 Groupe des situations relatives à l'émotion *Fierté*

II.13.1 Intensité moyenne

- Un de vos proches réalise un exploit (scolaire, travail, science, sport, etc).
- Vous réconciliez deux de vos amis qui ne se parlaient plus depuis des années.
- Vous apprenez que vous êtes aimé par beaucoup de gens, alors que vous ne l'avait jamais remarqué.

II.13.2 Intensité forte

- Un ami vous est reconnaissant, parce que vous étiez à ses côtés lorsqu'il en avait besoin.

- Vous arrivez à sauver quelqu'un d'une mort certaine alors que vous n'êtes pas censé le faire.
- Vous lisez un article de journal qui parle de vous (positivement) alors que vous n'étiez pas au courant.

II.14 Groupe des situations relatives à l'émotion *Culpabilisation*

II.14.1 Intensité moyenne

- Un de vos proches a besoin d'aide, mais vous ne vous présentez pas pour montrer votre soutien, parce que vous êtes trop prié par le boulot.
- Vous n'arrivez pas répondre aux attentes d'une personne.
- Alors que vous êtes en train de donner une présentation orale devant un grand public, avec une langue que vous ne maîtrisez pas totalement, la prononciation d'un mot vous pose un problème.
- Un collègue à vous, qui habite loin du lieu de travail, est toujours ponctuel, contrairement à vous qui habitez tout près.
- Vous êtes en classe, votre professeur vous fait une remarque désobligeante parce que vous ne suivez pas attentivement son cours.

II.14.2 Intensité forte

- Vous êtes en retard pour une réunion de travail. Vous traversez les rues en courant, mais vous tombez sur une personne handicapée qui vous demande de l'aider à atteindre l'autre bout de la rue. Vous vous excusez gentiment et vous vous en allez sans l'aider.
- Vous mentez à un de vos proches qui vous fait confiance sans que vous ayez une raison valable. Mais ce dernier finit par le découvrir.
- Vous avez blessé intentionnellement, dans un moment de colère, un proche à vous.
- Vous arrivez toujours en retard à la classe, et, chaque fois, le professeur vous le fait remarquer.
- Vos parents ne sont pas fiers de vous.

II.15 Groupe des situations relatives à l'émotion *Admiration*

II.15.1 Intensité moyenne

- Vous marchez tranquillement dans la rue, et une personne que vous ne connaissez pas vous sourit.
- Vous apprenez que deux pays en guerre se sont réconciliés.
- Vous remarquez qu'il existe quelques médias qui essaient de relater les faits tels qu'ils sont, objectivement, et n'essaient pas de dissimuler la vérité.

II.15.2 Intensité forte

- Vous entendez parler d'une star qui aide, moralement et financièrement, des personnes handicapées et des sans-abris.
- Un proche se souvient de quelque chose qui vous tient à cœur, au moment où vous l'imaginiez le moins.
- Vous recevez des cadeaux, de la part d'une personne qui compte pour vous, sans qu'il y ait une raison particulière
- Vous observez la naissance d'un enfant.

II.16 Groupe des situations relatives à l'émotion *Indignation*

II.16.1 Intensité moyenne

- Vous vous promenez dans votre quartier. De loin, vous apercevez un conducteur qui ne respecte pas les priorités piétonnières.
- Vous constatez l'hypocrisie d'un collègue à vous.
- Vous constatez qu'il existe des médias qui dissimulent la vérité.

II.16.2 Intensité forte

- Vous entendez parler des groupes criminels de trafic d'organes.
- Vous voyez des reportages sur des parents qui maltraitent et violent leurs enfants.
- Vous voyez des actes de vandalisme et des incitations au racisme.
- Vous voyez un reportage sur les tortures que peuvent subir des prisonniers partout dans le monde.
- Vous apprenez qu'il existe des gens vivant dans des pays développés qui n'ont pas d'assurance médicale.
- Voir des gens corrompus qui accèdent à des postes clés dans de grandes entreprises.

II.17 Groupe des situations relatives à l'émotion *Reconnaissance*

II.17.1 Intensité moyenne

- Votre professeur vous aide à mieux comprendre son cours, en étant disponible pour vous, même en dehors des séances de cours.
- Quelqu'un qui compte pour vous et qui vous a fait du mal admet ses erreurs.
- Votre ami(e) est sincère avec vous : il (elle) est toujours là pour vous corriger s'il (elle) remarque que vous faites des erreurs.

II.17.2 Intensité forte

- Vous allez dans une famille d'accueil dans un pays étranger; on vous réserve un accueil chaleureux.
- À un moment important de votre vie, vous recevez de l'aide de la part d'une personne avec laquelle vous n'avez aucun lien affectif.

- Vous êtes nouveau dans une entreprise. Un de vos collègues essaye, sans qu'il en soit obligé, de faire votre connaissance et de vous présenter à tout le staff.
- Un ami est à la hauteur de vos attentes et vous soutient toujours lors des moments difficiles.

II.18 Groupe des situations relatives à l'émotion *Colère*

II.18.1 Intensité moyenne

- Quelqu'un vous vole votre portefeuille.
- Quelqu'un qui compte énormément pour vous a tenu des propos qui vous ont blessé.
- Votre femme fouille dans vos affaires personnelles.
- Vous travaillez sur un projet avec un collègue, mais ce dernier n'effectue pas sa part de la tâche, et vous êtes obligé de la faire à sa place.
- Un proche vous promet quelque chose, mais n'arrive pas tenir sa promesse, ce qui ne s'est jamais produit auparavant.
- Vous faites tomber (et casser) votre bouteille d'eau de toilette, que vous venez d'acheter, à 70\$.
- Quelqu'un pour qui vous avez de l'estime trahit votre confiance.

II.18.2 Intensité forte

- Après dix-huit années de bons et loyaux services au sein de votre entreprise, les responsables décident, subitement, de vous remercier et de vous limoger dans le cadre de réduction d'effectif ou l'intégration de jeunes compétences.
- Vous travaillez durement pour atteindre un objectif, mais vous êtes victime d'une sorte de discrimination (sociale, raciale, etc.) et vous n'y arrivez pas, bien que vous le méritiez.
- On vous prive de la liberté de parole ou de pensée.
- Des secrets à vous dont vous n'êtes pas forcément fier sont dévoilés en public par d'anciens amis.
- Quelqu'un que vous n'avez jamais rencontré prétend qu'il vous connaît et vous accuse à tort dans une affaire de corruption.
- Vous êtes un enseignant : un élève ou un étudiant vous manque de respect.
- Votre enfant est agressé par plusieurs enfants à l'école.

II.19 Groupe des situations relatives à l'émotion *Autosatisfaction*

II.19.1 Intensité moyenne

- Vous gagnez une compétition sans avoir à fournir beaucoup d'effort.
- Vous arrêtez de fumer après dix ans de consommation de tabac.

- Vous arrivez à conquérir le cœur d'une femme (un homme), qui ne semblait pas vous apprécier au début.
- Vous pratiquez une religion.
- Vous réalisez une tâche complexe avec succès.

II.19.2 Intensité forte

- Vous arrêtez temporairement les études universitaires pendant trois ans afin de travailler et pouvoir vous les payer. Après ces trois ans, vous arrivez à reprendre et à obtenir votre diplôme.
- Vous faites tout ce qui est en votre disposition pour que vos enfants reçoivent une bonne éducation. Finalement votre souhait est réalisé.
- Vous réussissez à impressionner vos collègues de travail et votre patron avec la qualité du travail que vous faites.
- Vous faites quelque chose que vous avez toujours voulu faire, après avoir vraiment galéré.
- Vous relevez un défi très difficile, après un effort considérable, alors que tous ceux qui vous entourent prévoyaient le contraire.
- Vous êtes incontestablement le numéro 1 dans un domaine.

II.20 Groupe des situations relatives à l'émotion *Remords*

II.20.1 Intensité moyenne

- Vous faites une bêtise au boulot, mais sans grandes conséquences.
- Vous devez rendre un travail, mais vous n'arrivez pas à respecter les délais, et vous risquez d'être sanctionné.
- Vous grillez un feu rouge, mais vous êtes incertain quant à la présence d'un policier dans les parages. Finalement, quelques mètres plus loin une patrouille de police vous demande de vous arrêter.

II.20.2 Intensité forte

- Vous avez laissé trainer les choses : plusieurs travaux s'accumulaient peu à peu. Maintenant, vous avez un délai très court pour tout remettre.
- Vous ratez l'occasion d'exprimer votre amour à quelqu'un que vous ne reverrez plus jamais.
- Vous ne faites pas assez d'efforts au boulot (toujours en retard, non-respect des délais, etc), ce qui a obligé vos supérieurs à vous sanctionner.
- Vous faites des choses en contradiction avec vos principes.

II.21 Groupe des situations relatives à l'émotion *Adoration*

II.21.1 Intensité moyenne

- Le sourire d'un enfant.
- Les rayons du soleil vous réveillent le matin : une bonne journée ensoleillée s'annonce en perspective.
- Vous vous chamaillez souvent, pour une raison ou une autre, avec des personnes qui sont proches de vous. Mais ces derniers temps, le courant passe beaucoup mieux entre vous.
- Vous contemplez la nature, les magnifiques paysages, les merveilleux endroits partout dans le monde.
- Vous apprenez plein de nouvelles connaissances, en assistant à un cours, en lisant un livre, en surfant sur le net, etc.
- Vous retrouvez après des années sans aucune nouvelle un(e) ami(e) d'enfance avec qui vous entreteniez une très bonne relation.
- Danser comme un fou (une folle) dans tous les sens dans une soirée.
- Vous assistez à un match de votre sport préféré dans un stade ou une salle pleine à craquer, une ambiance très chaude.
- Vous assistez, tout seul sur la plage, à un coucher du soleil.
- Vous allez à un pays étranger, vous visitez les musées.
- Vous vous sentez protégés.
- Être devant sa télé, un samedi soir, alors qu'il pleut dehors.
- Vous apercevez une maman enlaçant très fort dans ses bras son enfant.

II.21.2 Intensité forte

- Être en bonne santé, vous et les personnes qui vous sont chères.
- Voir le sourire innocent de votre enfant.
- Écouter de la musique quand vous vous sentez seul.
- Vous voyagez pour la première fois à un pays étranger, vous faites connaissance avec les gens qui y résident, vous découvrez leur civilisation, leur culture, etc.
- Vous vivez une relation amoureuse très intense.
- Vous vous sentez compris alors que vous ne l'avez jamais été auparavant.

II.22 Groupe des situations relatives à l'émotion *Dégoût*

II.22.1 Intensité moyenne

- Voir sur les écrans de télé des gens qui s'entretuent, des guerres, des génocides, des crimes de guerre contre l'humanité.
- Vous êtes obligé de lire un livre dans un domaine qui ne vous intéresse pas.
- Les jours se suivent et sont tous pareils pour vous (la routine).
- Vous êtes devant un problème délicat à résoudre.
- Vous jouez aux cartes ou à un autre jeu et vous sentez que vous êtes en train de perdre.
- Vous regardez des reportages sur la maltraitance des animaux.
- Vous êtes obligé de travailler en même temps que vous étudiez.

- Vous avez l'impression que le temps passe vite et que vous n'avancez pas dans votre travail.
- Vous constatez que les gens de nos jours ne pensent qu'à l'argent.
- Votre équipe sportive préférée perd un match sans enjeu.
- Votre équipe sportive préférée perd un match, malgré qu'elle a bien joué et qu'elle restait invaincue depuis un an.
- Votre équipe sportive préférée joue très bien, mais n'arrive pas à gagner.
- Un appel téléphonique avec votre fiancé(e) qui dure longtemps comme d'habitude, car elle (il) ne veut pas raccrocher même si vous n'avez plus rien à dire.
- Un professeur qui donne un cours avec un ton monotone, en étant assis sur sa chaise sans bouger.
- Une personne qui compte pour vous n'a pas envie de vous voir, sans que vous sachiez la raison exacte.
- Vous attendez un résultat important (examen ou autre), mais vous ne savez pas quand vous l'aurez.
- Vous passez un temps considérable à faire des recherches sur le net sans trouver des résultats.

II.22.2 Intensité forte

- La malchance s'acharne sur vous, elle vous suit dans chaque projet que vous entamez.
- Vous passez toute une semaine à la maison sans sortir.
- Vous faites la connaissance d'une personne narcissique.
- Vous essayez de suivre un cours, mais vous n'arrivez pas à comprendre ce que le professeur présente.
- Vous oubliez une chose que vous n'aurez pas dû oublier.
- Être obligé de faire un travail qui vous irrite.
- Au cours d'une présentation orale, avec une langue que vous ne maîtrisez pas bien, et devant un grand public, vous n'arrivez pas à trouver vos mots.
- Vous êtes dans un endroit où vous ne connaissez personne, vous ne comprenez pas la langue et vous ne savez pas comment communiquer.
- Voir des gens que vous ne connaissez pas souffrir.
- Vous êtes célibataire depuis un bon bout de temps (quelque chose qui n'est pas habituel pour vous).
- Avoir à attendre quelqu'un qui est chaque fois en retard, alors que vous êtes toujours pile à l'heure.
- Voir des innocents tués dans des guerres, des femmes abusées, des enfants en pleur car ils ont perdu leurs proches.
- Vous vous sentez sous-estimé par vos supérieurs dans le travail, alors que vous bossez dur.
- Vous vous sentez obligé de justifier des actes condamnables d'un proche à vous.

Annexe III Questionnaire de Personnalité EPQR-A

Ce questionnaire constitue la forme abrégée du questionnaire de personnalité révisé de Eysenck (*Eysenck, Eysenck, & Barrett, 1985*).

Instructions : Répondez à chacune des questions en choisissant la réponse « oui » ou « non ». Il n'y a pas de bonnes ou mauvaises réponses ni de questions pièges. Répondez rapidement et ne passez pas trop de temps à réfléchir au sens précis de la question.

Pensez bien répondre à toutes les questions.

N°	Questions	Réponses	
1	Votre humeur subit-elle souvent des hauts et des bas?	Oui	Non
2	Aimez-vous beaucoup parler?	Oui	Non
3	Seriez-vous mal à l'aise si vous aviez des dettes?	Oui	Non
4	Êtes-vous plutôt quelqu'un plein d'entrain?	Oui	Non
5	Vous est-il arrivé d'avoir si grande envie de quelque chose que vous vous êtes arrangé(e) pour vous servir plus que la part qui vous revenait?	Oui	Non
6	Accepteriez-vous de prendre des drogues qui pourraient avoir des effets étranges ou dangereux?	Oui	Non
7	Vous est-il arrivé de blâmer quelqu'un pour une faute, sachant que c'est vous qui l'aviez commise?	Oui	Non
8	Préférez-vous faire les choses à votre façon plutôt que d'agir en fonction des règles?	Oui	Non
9	Vous arrive-t-il souvent d'en « avoir marre »?	Oui	Non
10	Vous est-il arrivé de prendre quelque chose appartenant à quelqu'un d'autre (ne serait-ce qu'une épingle ou un bouton)?	Oui	Non
11	Vous considérez-vous comme une personne nerveuse?	Oui	Non
12	Estimez-vous que le mariage est une institution dépassée qui devrait être abolie?	Oui	Non
13	Savez-vous aisément animer une soirée ennuyeuse?	Oui	Non
14	Êtes-vous facilement inquiet(e)?	Oui	Non
15	Avez-vous tendance à rester en arrière-plan lorsque vous êtes en société?	Oui	Non
16	Êtes-vous tourmenté(e) si vous savez qu'il ya des erreurs dans votre travail?	Oui	Non
17	Vous est-il arrivé de tricher à un jeu?	Oui	Non
18	Souffrez-vous de vos nerfs?	Oui	Non
19	Vous est-il arrivé de profiter de quelqu'un?	Oui	Non
20	Êtes-vous le plus souvent silencieux(se) lorsque vous êtes avec d'autres personnes?	Oui	Non
21	Vous sentez-vous souvent seul(e)?	Oui	Non
22	Est-il mieux de suivre les règles sociales plutôt que de faire les choses à sa façon?	Oui	Non
23	Les autres vous estiment-ils plein d'entrain?	Oui	Non
24	Mettez-vous toujours en pratique ce que vous prêchez?	Oui	Non

Dans ce questionnaire, quatre scores selon quatre échelles relatives aux traits de *névrosisme*, *psychoticisme*, *extraversion* et au *mensonge* sont possibles et calculés de manière indépendante (*Francis, Brown, & Philipchalk, 1992*). Les réponses aux

questions par oui ou non déterminent ces scores entre 0 et 6 pour chaque échelle. L'échelle de mensonge permet de vérifier la validité des réponses du sujet (*Bouvard, 2009*).

L'échelle de *névrosisme* comprend (ou est en relation avec) les questions 1,9,11,14,18, et 21. Une réponse correspondante à « oui » pour une de ces questions rajouterait une valeur de 1 au score du névrosisme.

L'échelle d'extraversion comporte quatre questions, les questions 2, 4, 13 et 23, pour lesquelles une réponse de « oui » rajouterait une valeur de 1 au score. Cette échelle comprend, également, deux autres questions, les questions 15 et 20, pour lesquelles une réponse de « non » augmenterait le score de 1.

De même, l'échelle de psychoticisme comprend les questions 6, 8 et 12 cotées « oui » et les questions 3,16 et 22 cotées « non ».

Similairement, l'échelle de mensonge comprend cinq questions cotées « non », 5,7,10,17 et 19. La question 24 est la seule cotée « oui » pour cette échelle.