

Université de Montréal

Jeu de poursuite sur des modèles du Web et généralisation

par
Tantely Ramanampanoharana

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de Maître ès sciences (M.Sc.)
en informatique

Décembre, 2004

© Tantely Ramanampanoharana, 2004.



Direction des bibliothèques

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

Jeu de poursuite sur des modèles du Web et généralisation

présenté par:

Tantely Ramanampanoharana

a été évalué par un jury composé des personnes suivantes:

Stefan Wolf, président-rapporteur
Geña Hahn, directeur de recherche
Alain Tapp, membre du jury

Mémoire accepté le: ...13/04/05...

RÉSUMÉ

Quand il s'agit de la sécurité d'un environnement, un aspect qui lui est inhérent est la poursuite et, idéalement, la capture d'un éventuel intrus. Ce mémoire considère le problème de la recherche d'intrus dans les graphes. Un voleur se déplace le long des arêtes d'un graphe et un certain nombre de policiers le traquent. Quelle stratégie ces derniers doivent-ils alors adopter pour remplir leur mission, soit capturer le voleur ? En empruntant des arguments de la théorie des graphes aléatoires, on prouve que presque tous les graphes donnent avantage à l'intrus. Un graphe qui a justement retenu notre attention et dans lequel l'application de ce problème peut-être nombreuse, est le *Web*. Nous étudions sa structure et son comportement vis-à-vis du jeu de poursuite par le biais de différents modèles.

Mots-clés : Graphe aléatoire, Graphe homogène, Jeu de poursuite, Police-voleur, Webgraph.

ABSTRACT

In this thesis, we consider the problem of pursuit-evasion on a graph G , in which an intruder and a set of pursuers move along the edges of G until capture occurs eventually, i.e. one of the pursuers occupies the same vertex as the intruder. Such problem arises in many situations especially when the security of an environment is of concern. Of particular interest is the *Webgraph*. We study its structure and its relation to the graph-searching problem by the means of various models. Some of our analysis use tools from random graphs theory and model theory.

Keywords : Cops and Robber game, Graph-searching, Homogenous graph, Pursuit-evasion, Random graph, Webgraph.

TABLE DES MATIÈRES

RÉSUMÉ	iv
ABSTRACT	v
TABLE DES MATIÈRES	vi
LISTE DES FIGURES	ix
LISTE DES NOTATIONS ET DES SYMBOLES	x
LISTE DES TABLEAUX	xii
DÉDICACE	xiii
REMERCIEMENTS	xiv
INTRODUCTION	1
CHAPITRE 1 : JEU DE POURSUITE DANS LES GRAPHERS . .	3
1.1 Notions de théorie des graphes	4
1.1.1 Relation	4
1.1.2 Graphes, sous-graphes	6
1.1.3 Adjacence, domination	7
1.1.4 Connexité	8
1.1.5 Homomorphisme, contraction et mineur	10
1.1.6 Quelques graphes	11
1.2 Notions d'analyse	11
1.3 Théorie des jeux	14
1.4 Jeux de poursuite	16
1.4.1 Fouille	16
1.4.2 Police-voleur	26

CHAPITRE 2 : MODÈLES DU WEB	35
2.1 Mesures	36
2.1.1 Distribution des degrés	36
2.1.2 Noyaux bipartites	38
2.1.3 Composantes connexes	39
2.1.4 Diamètre	42
2.2 Modèles	42
2.2.1 Erdős-Rényi (ER)	43
2.2.2 Attachement préférentiel (BA)	45
2.2.3 Copie évolutive (BJ)	49
2.3 Discussions	50
CHAPITRE 3 : LIMITES DES MODÈLES DU WEB	52
3.1 Limites	53
3.1.1 $G(n, p)$	54
3.1.2 $BA(m, t)$	61
3.1.3 $BJ(p, \phi, H)$	65
3.2 Jeu de poursuite	69
CHAPITRE 4 : GRAPHES HOMOGENES	72
4.1 Cas fini	74
4.1.1 Caractérisation	75
4.1.2 Jeu de poursuite	75
4.2 Cas infini	76
4.2.1 Amalgamation	77
4.2.2 Jeu de poursuite	83
CHAPITRE 5 : GRAPHES ALÉATOIRES	86
5.1 Évolution Erdős-Rényi	87
5.1.1 Le vide	87
5.1.2 Les arbres	87

5.1.3	Composante géante	88
5.1.4	Connexité	89
5.2	Logique du premier ordre	89
5.2.1	Langage	90
5.2.2	Jeu de Ehrenfeucht-Fraïssé	92
5.2.3	Expressivité	92
5.3	Loi Zéro-Un	98
5.4	Seuil	101
CONCLUSION		102
BIBLIOGRAPHIE		105

LISTE DES FIGURES

1.1	Contraction le long de e	11
1.2	Quelques graphes	12
1.3	Graphe Y^2	17
1.4	Réduction homéomorphique	24
1.5	Mineurs interdits pour $sw(G) \leq 2$	24
1.6	$sw(G) = 2$	26
1.7	Police-voleur sur un arbre perché	30
1.8	Triangles avec sens d'orientation	34
2.1	Distribution des degrés entrants	37
2.2	Distribution des degrés sortants	38
2.3	Noyau biparti	39
2.4	Connexité du web	41
2.5	$G(7, 1/2)$	43
2.6	Attachement préférentiel	46
2.7	Copie évolutive : le noeud prototype est en triangle et les liens aléatoires sont en pointillés	50
4.1	Un graphe ultrahomogène	73
5.1	Jeu Ehrenfeucht-Fraïssé imitateur gagnant	96

LISTE DES NOTATIONS ET DES SYMBOLES

$A \setminus B$:	la soustraction ensembliste $\{x : x \in A \wedge x \notin B\}$
$ A $:	la cardinalité de l'ensemble A
$\mathcal{P}(A)$:	l'ensemble des parties de A $\{X : X \subseteq A\}$
$\mathcal{P}_k(A)$:	l'ensemble des parties à k éléments de A $\{X : X \subseteq A \wedge X = k\}$
\mathbb{N} :	l'ensemble des entiers naturels $\{0, 1, 2, \dots\}$
\mathbb{Z} :	l'ensemble des entiers relatifs $\{\dots, -2, -1, 0, 1, 2, \dots\}$
\mathbb{Q} :	l'ensemble des rationnels
\mathbb{R} :	l'ensemble des nombre réels
\mathbb{R}_+ :	l'ensemble des réels positifs $\{x \in \mathbb{R} : x \geq 0\}$
\mathbb{R}_- :	l'ensemble des réels négatifs $\{x \in \mathbb{R} : x \leq 0\}$
$[n]$:	l'ensemble des entiers naturels strictement plus petit que n
\aleph_0 :	la cardinalité de \mathbb{N}
2^{\aleph_0} :	la cardinalité de \mathbb{R}
$V(G)$:	l'ensemble des sommets du graphe G
$E(G)$:	l'ensemble des arêtes du graphe G
$\Pr[G \models P]$:	la probabilité qu'un graphe satisfait la propriété P
$d_G(u)$:	degré du sommet u dans G
$d_G^+(u)$:	degré sortant du sommet u dans G
$d_G^-(u)$:	degré entrant du sommet u dans G
$\delta(G)$:	degré minimum du graphe G
$\Delta(G)$:	degré maximum du graphe G
$\gamma(G)$:	nombre de domination de G
$\alpha(G)$:	nombre d'indépendance de G
$N_G(u)$:	voisinage ouvert du sommet u dans G

$N_G[u]$:	voisinage fermé du sommet u dans G
$N_G^+(u)$:	voisinage sortant du sommet u dans G
$N_G^-(u)$:	voisinage entrant du sommet u dans G
$AUT(G)$:	le groupe des automorphismes de G
$G \sqcup H$:	union disjointe de G et H
$t \cdot G$:	t copies disjointes de G
$G \upharpoonright X$:	le sous-graphe de G induit par $X \subseteq V(G)$
$G - u$:	$G \upharpoonright V(G) \setminus \{u\}$
$H \leq G$:	H est un sous-graphe de G
$H \preceq G$:	H est un sous-graphe induit de G
K_n :	le graphe complet sur n éléments
E_n :	le graphe vide sur n éléments
C_n :	le cycle de longueur n
$G(n, p)$:	l'espace de probabilité des graphes aléatoires sur n sommets et avec probabilité d'une arête p
$G_n(m)$:	l'espace de probabilité des graphes aléatoires sur n sommets ayant m arêtes
R :	le graphe aléatoire infini dénombrable et est la limite de Fraïssé des graphes finis
H_n :	le graphe de Henson et est la limite de Fraïssé de graphes finis n'admettant pas K_n
$\pi_i(x)$:	la projection sur la i -ème composante de x où $x \in \prod A_j$ et $\pi_i(x) \in A_i$

LISTE DES TABLEAUX

2.1	Noyaux bipartis	40
2.2	Taille de la plus grosse composante connexe survivant la suppression des sommets de degré entrant au moins k	40
2.3	Taille des régions du web	42

A mes filles Ashley et Casey.

REMERCIEMENTS

J'aimerais avant tout remercier mon directeur Geña Hahn pour sa gentillesse et sa patience, pour ses encouragements et sa compréhension. J'ai grandement apprécié la qualité de son encadrement, autant sur le point pédagogique que sur le point académique. Je lui en suis très reconnaissant.

Je voudrais ensuite remercier toutes les personnes qui ont contribué de près ou de loin à la réussite de ce mémoire : tout particulièrement l'équipe des *sweepers* de Regina dont Brian Alspach, Denis Hanson, Xiangwen Li, Boting Yang, pour les diverses discussions sur la recherche dans les graphes ; Sif El Harti, pour les exemples et contre-exemples et surtout pour les discussions d'ordre philosophique et métaphysique. Je n'oublie pas de remercier Jeanne MacDonald pour son accueil chaleureux durant mon séjour à Regina.

Je remercie MITACS pour m'avoir soutenu financièrement durant mes années en maîtrise.

Enfin, je voudrais remercier mes parents Dada et Neny, ma femme Bojy et mes enfants Ashley et Casey ainsi que mon frère Tahiry pour leur amour et leur soutien inconditionnel.

INTRODUCTION

Imaginons, un instant, un fugitif en cavale dans les rues de Montréal. Nourri par l'espoir de semer les gardiens de la paix qui se sont lancés à ses trousses, chaque croisement devient le théâtre de dilemme : va-t-on tout droit, à gauche, à droite ou plus raisonnablement se rendre ? Certes, s'il détient un plan de la ville, saura-t-il mieux s'orienter. Néanmoins, la tâche ne relève pas de la trivialité ; pour réussir le fugitif devrait se munir d'une bonne stratégie, et cela vaut aussi pour ceux qui le poursuivent.

Dans un *jeu de poursuite*, un groupe de fureteurs traque un ou plusieurs individus mobiles dans un environnement donné. Si l'on considère ces premiers comme des ressources auxquelles sont associées des coûts, la question de minimisation émerge naturellement. Ce problème a reçu l'attention de bon nombre de chercheurs dans la littérature grâce à la multitude de situations dans lesquelles de tel scénario se manifeste. Dans [54, 63] par exemple, la motivation vient de la robotique où l'on veut coordonner les mouvements de robots à qui on a confié la tâche de surveiller un périmètre. Makedon et Sudborough, dans [75], ont étudié le problème d'optimisation dans la disposition de portes logiques composant un circuit et ont mis à jour un lien avec le nombre minimal de pourchasseurs dans le jeu de poursuite.

Dans ce mémoire, nous abordons le jeu de poursuite dans un univers plus discret : la recherche dans les réseaux. Qu'arrive-t-il si l'information que nous prospectons se déplace ou essaie de nous éviter ? Les techniques de fouille couramment utilisées écartent de telle supposition et, par ce fait même, ignorent la structure sous-jacente. Dans une topologie en anneaux de plus de quatre noeuds par exemple, un code malicieux qui passe d'une machine à l'autre échappe à un agent de recherche s'il prend soin de se déplacer dans le même sens que ce dernier. La structure du réseau revêt un rôle clé dans le jeu de poursuite ; et la compréhension de celle-ci constitue la motivation première du présent mémoire. Il s'inscrit dans un cadre théorique, bien plus que le titre ne le suggère.

La représentation mathématique la plus appropriée des réseaux est le *graphe*. Il sai-

sit en même temps les noeuds et les liens existant entre eux. Il est donc tout à fait naturel d'amorcer ce travail par l'introduction au jeu de poursuite dans les graphes par le biais de formalisations et de caractérisations. C'est l'objet du chapitre 1.

S'il est un énorme graphe, accessible à tout le monde, on peut sûrement retenir le *web*, une collection mondiale de documents pointant les uns vers les autres par des liens hypertextes. Avec ses quelques huit milliards de pages, il constitue un immense terrain où se mêlent bonnes et mauvaises intentions. Penser à ces sites qui proposent des copies illégales de logiciels, qui, malgré les efforts pour les enrayer, subsistent sur la toile mondiale. Le chapitre 2 est consacré à l'étude de la structure du web en tant que graphe. Pour cela, différents modèles sont explorés en rapport avec les propriétés réelles observées sur Internet. Un des constats qui ressortent d'ailleurs de ces observations est la croissance effrénée du web au fil du temps. Il est donc tout à fait légitime de procéder dans le chapitre 3 à l'analyse asymptotique de ces modèles.

Un de ces modèles est justement *le graphe aléatoire*, qui perçoit les liens hypertextes comme étant le fruit du hasard et complètement indépendants les uns des autres. Découverte par Erdős et Rényi [42], sa limite possède des propriétés à la fois belles et surprenantes, qui lui confèrent un rôle central en théorie des graphes. Parmi celles-ci, une qui va retenir notre attention dans le chapitre 4 est l'homogénéité. C'est une propriété qui a été très étudiée en théorie des modèles et qui renforce considérablement la symétrie dans une structure. Malgré cela, presque tous les graphes sont homogènes et c'est le genre de surprise que la théorie des graphes aléatoires abordée dans le chapitre 5 nous réserve.

CHAPITRE 1

JEU DE POURSUITE DANS LES GRAPHEs

Le jeu de poursuite dans les graphes a été introduit par Breisch [26] et Parsons [81] : un homme est perdu dans un réseau de tunnels, probablement dans le complexe minier de Tau Tona en Afrique du Sud pour ne pas dire dans les grottes d'Afghanistan. Un groupe de personnes, ayant une parfaite connaissance de l'endroit, est envoyé à sa recherche. Aux yeux de ces dernières, le déplacement de l'homme est totalement imprévisible et inconnu si bien qu'on ne considère l'homme retrouvé que lorsqu'un des chercheurs réussisse à l'immobiliser, c'est-à-dire les deux occupent le même point de l'espace. Quel est alors le nombre minimum de sauveteurs à dépêcher sur les lieux pour assurer la réussite du secours ? Les tunnels et leur interconnexion sont représentés par un graphe et la question nous mène à caractériser les graphes selon ce nombre.

Par la suite, plusieurs variations du scénario ont été considérées [47]. Elles diffèrent les unes des autres par la nature des déplacements permis et par les informations accessibles aux acteurs. Deux types d'acteurs, ayant par moment des objectifs opposés, sont identifiés ici : celui qui vagabonde dans le noir et ceux qui sont à sa recherche. Les termes *évadé*, *intrus* ou *voleur* seront utilisés pour désigner le premier type, tandis que *chercheurs*, *fouilleurs*, *gendarmes*, *policiers* ou *sauveteurs* seront attribués au second. Notons que le cas où les gendarmes sont à la poursuite de plusieurs évadés se réduit à capturer ces derniers un par un si on n'exige pas qu'ils soient tous arrêté au même moment. En effet, soit k le nombre minimum de chercheurs pour capturer un intrus dans le graphe. La présence de plusieurs intrus ne peut pas augmenter le nombre minimum de chercheurs puisqu'en ayant arrêté le premier intrus, par hypothèse ce même nombre suffit pour capturer le second quelque soit son mouvement et l'endroit où il se trouve à ce moment précis. Et ainsi de suite. D'autre part, il faut au moins k gendarmes pour les capturer tous car les voleurs n'ont qu'à se tenir ensemble sur un même sommet à chaque fois, ce

qui se réduit au cas d'un voleur. Cet argument justifie donc l'emploi du singulier pour le premier type d'acteur. Dans ce mémoire, nous allons nous concentrer sur les deux cas extrêmes :

modèle 1 l'intrus se déplace de manière continue et à une vitesse illimitée dans le graphe, face à une formation de chercheurs aveugles ;

modèle 2 alternativement au tic d'un horloge, les deux opposants se déplacent d'un bout de tunnel à un autre, en ayant une totale connaissance de l'endroit où son adversaire se trouve.

La première déclinaison est la version originale de Parsons. Le terme *fouille* décrit bien la situation, c'est d'ailleurs celui utilisé dans la littérature pour la référencer. La seconde relève de la théorie des jeux : *jeu de police-voleur* [78], un peu à la manière du jeu d'échec vis-à-vis des batailles médiévales.

1.1 Notions de théorie des graphes

1.1.1 Relation

Soient A et B deux ensembles. Le *produit cartésien* de A par B et noté $A \times B$ est l'ensemble de tous les couples ordonnés (a, b) où a est pris dans A et b , dans B . Plus formellement,

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

Plus généralement, le produit cartésien de n ensembles A_1, A_2, \dots, A_n , noté

$$\prod_{i=1}^n A_i = A_1 \times A_2 \times \dots \times A_n,$$

est l'ensemble de tous les n -tuplets (a_1, \dots, a_n) où $a_i \in A_i$. On note $\pi_i : \prod_i A_i \rightarrow A_i$ la *projection sur la i -ème composante*, $\pi_i(a_1, a_2, \dots, a_n) = a_i$.

Une *relation binaire* \mathcal{R} sur un ensemble A est une partie du produit cartésien de A par lui-même, soit $\mathcal{R} \subseteq A^2 = A \times A$. Souvent, on note $x\mathcal{R}y$ pour signifier $(x, y) \in \mathcal{R}$ et, presque par abus, pour tout sous-ensemble Y non vide de A , $x\mathcal{R}Y$ (respecti-

vement $Y\mathcal{R}x$) dénote le fait $\forall y \in Y, x\mathcal{R}y$ (respectivement $y\mathcal{R}x$). La relation est dite :

- *réflexive* si $\forall x \in A, x\mathcal{R}x$;
- *irréflexive* si $\forall x \in A, (x, x) \notin \mathcal{R}$;
- *symétrique* si $\forall x, y \in A, x\mathcal{R}y \Rightarrow y\mathcal{R}x$;
- *antisymétrique* si $\forall x, y \in A, (x, y) \in \mathcal{R} \Rightarrow (y, x) \notin \mathcal{R}$;
- *transitive* si $\forall x, y, z \in A, x\mathcal{R}y \wedge y\mathcal{R}z \Rightarrow x\mathcal{R}z$.

Lorsque \mathcal{R} est réflexive, antisymétrique et transitive alors la relation forme un *ordre partiel* sur l'ensemble A . Si, de plus, tous les éléments de A sont comparables, i.e $\forall x, y \in A, x\mathcal{R}y \vee y\mathcal{R}x$, alors on dit de \mathcal{R} , un *ordre total* sur A . D'autre part, une *relation d'équivalence* est une relation qui est réflexive, symétrique et transitive. Une relation d'équivalence partitionne l'ensemble. En effet, si nous posons $[x] = \{y \in A : x\mathcal{R}y\}$ alors $x \in [x]$. Donc $A = \bigcup_{x \in A} [x]$. De plus, si $(x, y) \notin \mathcal{R}$ alors $[x] \cap [y] = \emptyset$ par transitivité. On appelle $[x]$ la *classe d'équivalence* qui contient x . On définit l'*index* de \mathcal{R} dans A par $|\{[x] : x \in A\}|$, c'est-à-dire le nombre de classes d'équivalence de R dans A .

Une *fonction* f de A vers B et que l'on note $f: A \rightarrow B$ est un sous-ensemble de $A \times B$ tel que $\forall a \in A$ s'il existe $b_1, b_2 \in B$ tels que $(a, b_1), (a, b_2) \in f$ alors $b_1 = b_2$. On écrit souvent $f(a) = b$ lorsque $(a, b) \in f$. Le *domaine* de la fonction f , notée $\text{dom}(f)$, est $\bigcup_{x \in f} \pi_1(x)$ et son *image*, notée $\text{im}(f)$, $\bigcup_{x \in f} \pi_2(x)$. Pour un élément $b \in \text{im}(f)$, l'*image réciproque* $f^{-1}(b) = \{a \in A : f(a) = b\}$ et plus généralement pour un $X \subseteq \text{im}(f)$, $f^{-1}(X) = \bigcup_{x \in X} f^{-1}(x)$. Une *injection* entre A et B est une fonction qui satisfait $\forall a_1, a_2 \in A, f(a_1) = f(a_2) \Rightarrow a_1 = a_2$. Et une *surjection* satisfait $\forall b \in B \exists a \in A, f(a) = b$. Une fonction qui est à la fois une injection et une surjection est une *bijection*.

L'ensemble des parties à k éléments d'un ensemble A est noté

$$\mathcal{P}_k(A) = \{B \subseteq A : |B| = k\}.$$

Il y a $\binom{|A|}{k}$ tels ensembles pour tout A fini. L'ensemble de toutes les parties de A est notée $\mathcal{P}(A)$ et est donc la réunion des $\mathcal{P}_k(A)$ $k = 0, 1, 2, \dots$

1.1.2 Graphes, sous-graphes

Un *graphe orienté* $G = (V, A)$ est un ensemble V muni d'une relation binaire A . Les éléments de V sont appelés *sommets* et ceux de A , *arcs*. Pour un arc $(x, y) \in A$, les sommets x, y constituent ses *extrémités*, et on dit que x *domine* y . Il possède un sens d'orientation, soit de x vers y . Un arc reliant un sommet à lui-même est appelé *boucle*. Si tous les sommets possèdent une boucle, c'est-à-dire la relation est réflexive, alors le graphe est dit *réflexif*.

Un *graphe non-orienté* $G = (V, E)$ est un couple où V est l'ensemble des sommets et $E \subseteq \mathcal{P}_1(V) \cup \mathcal{P}_2(V)$, l'ensemble des *arêtes*, que l'on peut voir comme des arcs sans orientation. Les boucles d'un graphe non-orienté sont les éléments de $E \cap \mathcal{P}_1$ et si ce dernier est égal à V , respectivement \emptyset , alors le graphe non-orienté est réflexif, respectivement *simple*. D'un graphe orienté, on peut alors parler d'un graphe non-orienté sous-jacent, en passant de (x, y) à $\{x, y\}$ pour tous les arcs. Pour un graphe non-orienté, le graphe orienté symétrique correspondant est naturellement celui qui pour chaque arête $\{x, y\}$ admet les deux arcs (x, y) et (y, x) . Le terme graphe sera employé pour désigner l'une ou l'autre des notions, et le contexte devra indiquer laquelle. De même nous nous servons du couple ordonné (x, y) autant pour référer aux arcs qu'aux arêtes. Nous utiliserons les notations $V(G)$, respectivement $E(G)$, pour identifier l'ensemble des sommets du graphe G , respectivement l'ensemble des arcs ou arêtes. On définit l'*ordre* d'un graphe G , que l'on dénote par $|G|$, comme le cardinal de $V(G)$, soit $|G| = |V(G)|$. Ainsi, le graphe est dit *fini* lorsqu'il est d'ordre fini, sinon *infini*. Le *complément* d'un graphe G , noté \overline{G} est le graphe dont l'ensemble des arcs est le complément de celui de G , i.e $\overline{G} = (V(G), V(G)^2 \setminus E(G))$ et dépendamment du contexte, on peut exclure les boucles. Deux graphes simples sur un ensemble de sommets X , d'ordre n , qui sont l'un complément de l'autre sont le *graphe vide* $E_n = (X, \emptyset)$ et le *graphe complet* $K_n = (X, \mathcal{P}_2(X))$. Un graphe orienté est complet si le graphe non-orienté sous-jacent est complet.

Un graphe H est un *sous-graphe* de G , et on écrira $H \leq G$, si $V(H) \subseteq V(G) \cap V(H)^2$ et $E(H) \subseteq E(G)$. Si de plus $E(H) = E(G) \cap V(H)^2$, on dit que H est un sous-graphe *induit* de G sur $V(H)$ et on écrit $H \leq G$. Autrement dit, le sous-graphe induit de G sur un ensemble de sommets X , noté $G \upharpoonright X$, contient tous les autres sous-graphes de G sur X , i.e $\forall H \leq G, V(H) = X \Rightarrow H \leq G \upharpoonright X$. Si H est un sous-graphe non induit de G , alors il existe un arc $(x, y) \in (E(G) \cap V(H)^2) \setminus E(H)$. Un tel arc s'appelle une *corde* de H dans G . Si $V(H) = V(G)$ alors H est un *facteur*, ou *sous-graphe couvrant* de G .

1.1.3 Adjacence, domination

Deux sommets $x, y \in V(G)$ tels que $(x, y) \in E(G)$ sont dits *adjacents* et on écrira $x \sim y$. Les sommets adjacents à $x \in V(G)$ forment son *voisinage*, que l'on note $N_G(x)$. Et plus généralement, on définit $N_G(X) = \bigcup_{x \in X} N_G(x)$ pour un sous-ensemble $X \subseteq V(G)$. Dans le cas où le graphe est orienté, il convient de distinguer entre le voisinage *entrant* $N_G^-(x) = \{y : (y, x) \in E(G)\}$ et le voisinage *sortant* $N_G^+(x) = \{y : (x, y) \in E(G)\}$, avec $N_G(x) = N_G^+(x) \cup N_G^-(x)$. Le voisinage *fermé* de x , $N_G[x]$, consiste en son voisinage et lui-même, soit $N_G[x] = N_G(x) \cup \{x\}$. Une mesure de l'adjacence d'un sommet est justement son *degré*. Le *degré entrant* $d^-(x)$, respectivement *degré sortant* $d^+(x)$, de $x \in V(G)$ est le nombre de ses voisins entrants, respectivement le nombre de ses voisins sortants. Lorsque G n'est pas orienté, le degré d'un sommet est le nombre d'arêtes dont il est l'une des extrémités et en particulier, une boucle compte pour deux arêtes. Un sommet de degré 0 est dit *isolé*, celui de degré 1, *pendant* et celui qui est adjacent à tous les autres sauf lui-même, *universel*. Un sommet de degré entrant, respectivement degré sortant, égal à 0 est une *source*, respectivement *puits*.

Lemme 1 (des poignées de mains). *Pour tout graphe non-orienté G ,*

$$\sum_{x \in G} d(x) = 2|E(G)|.$$

Et pour tout graphe orienté H ,

$$\sum_{x \in H} d^+(x) = \sum_{x \in H} d^-(x) = |E(H)|.$$

Démonstration. Chaque arête (x, y) est comptabilisée deux fois, une pour $d(x)$ et une autre pour $d(y)$. De même, pour chaque arc (x, y) , $x \in N_G^-(y)$ et $y \in N_G^+(x)$. \square

Le *degré minimum*, respectivement *degré maximum*, d'un graphe G sera noté $\delta(G) = \min\{d(x) : x \in V(G)\}$, respectivement $\Delta(G) = \max\{d(x) : x \in V(G)\}$. Si tous les sommets de G ont le même degré k , on dit que G est *régulier*, plus précisément k -régulier. Le graphe complet K_n est par exemple $n - 1$ -régulier.

Un sous-ensemble de sommets $X \subseteq V(G)$ tel que son voisinage fermé (sortant si orienté) revient à tous les sommets de G est appelé un *ensemble dominant* de G . Dit autrement, un sommet de G est soit dans X , soit dominé par un sommet de X . La *nombre de domination* d'un graphe, notée $\gamma(G)$, est alors le minimum des cardinalités des ensembles dominants de G .

Un sous-ensemble $X \subseteq V(G)$ qui induit un sous-graphe vide de G forme un ensemble *indépendant*. En d'autres mots, aucune arête ne relie deux sommets de X dans G . Le *nombre d'indépendance* d'un graphe, noté $\alpha(G)$, est le maximum des cardinalités de ses ensembles indépendants. Par la suite, on peut partitionner les sommets de G en k ensembles indépendants. Si $k \leq |G|$ est le plus petit d'un tel nombre, on dit que le graphe est k -*parti*. Pour $k = 2$, on utilise le terme *biparti*, et $k = 3$, *triparti*.

1.1.4 Connexité

On peut étendre la notion d'adjacence pour former une sorte de concaténation d'adjacences : la *connexité*. Une suite finie de sommets $x_1 x_2 \dots x_k$ où $(x_i, x_{i+1}) \in E(G)$ pour tout $i = 1, 2, \dots, k-1$ forme un *chemin* de longueur $|x_1 x_2 \dots x_k| = k-1$ (le nombre d'arcs) entre les sommets x_1 et x_k . Si $x_1 = x_k$, on parle plutôt de *circuit*. Bien sûr, les termes varient pour les graphes non-orientés pour lesquels on emploie

plutôt *chaîne* et *cycle*. Une chaîne *eulérienne*, respectivement *hamiltonienne* est une chaîne qui passe par tous les arêtes, respectivement tous les sommets du graphe, une et une seule fois. Un graphe possédant un cycle eulérien, respectivement hamiltonien est dit eulérien, respectivement hamiltonien. Maintenant, étant donnés deux sommets, il n'existe pas toujours une chaîne les reliant. Le graphe nul n'en possède aucun pour n'importe quel paire de sommets par exemple. Si pour tous $x, y \in V(G)$ il existe toujours une chaîne de x vers y , le graphe est dit *connexe*. De même s'il existe toujours un chemin entre tout paire de sommets d'un graphe orienté, il est appelé *fortement connexe*, et simplement *connexe* si le graphe non-orienté sous-jacent est connexe. La relation $x \rightsquigarrow y$ s'il existe une chaîne reliant x à y forme une relation d'équivalence. La classe d'équivalence $[x]$ est appelée la *composante connexe* contenant x . Le graphe est donc connexe si l'index de la relation \rightsquigarrow dans $V(G)$ est 1.

Un graphe G est dit k -*connexe* si après avoir retiré n'importe quels $k - 1$ sommets distincts du graphe, il reste encore connexe. Un graphe connexe est donc un graphe 0-connexe. Si $k = 2$, on parle plutôt de graphe *biconnexe*.

La *distance* dans H entre deux sommets x, y de G , notée $d_H(x, y)$ pour $H \leq G$ est la longueur minimale d'une chaîne dont tous le sommets sont dans H reliant x à y . Autrement dit,

$$d_H(x, y) = \min \{|P| : P \text{ est une chaîne dans } H \text{ de } x \text{ vers } y\}.$$

Si aucune chaîne ne relie x et y dans H , on convient de poser $d_H(x, y) = \infty$. Lorsque $H = G$, on peut omettre de spécifier H . La plus grande distance entre deux sommets du graphe représente son *diamètre*, noté $\text{diam}(G)$. Ainsi $\text{diam}(G) = \max \{d(x, y) : x, y \in V(G)\}$.

1.1.5 Homomorphisme, contraction et mineur

Soient G, H deux graphes, une fonction $f: V(G) \rightarrow V(H)$ forme un homomorphisme entre G et H , noté $f: G \rightarrow H$ si

$$\forall x, y \in V(G) (x, y) \in E(G) \Rightarrow (f(x), f(y)) \in E(H).$$

Si, de plus f est une bijection et si f et f^{-1} sont des homomorphismes, alors on dit que f est un *isomorphisme*. On dit alors que G et H sont *isomorphes* et l'on note $G \cong H$. Un isomorphisme d'un graphe sur lui-même est un *automorphisme*. Une *propriété de graphe* P est un ensemble dont les éléments sont des graphes tel que si $G \in P$ et $G \cong H$ alors $H \in P$ aussi.

Soit $f: G \rightarrow H$ un homomorphisme tel que $f: G \rightarrow H \upharpoonright f(V(G))$ est un isomorphisme, alors on dit que f est un *plongement* de G dans H ou encore H *admet* G .

Soit $e = \{x, y\} \in E(G)$ une arête. *Contracter* G le long de e est une opération qui consiste à identifier les sommets x et y en prenant soin d'enlever les arêtes dupliquées et la boucle ainsi créée. On note le graphe résultant G/e . Plus formellement,

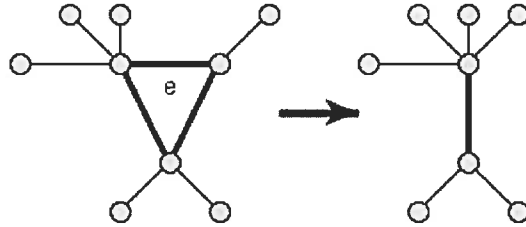
$$\begin{aligned} & - V(G/e) = V(G) \setminus \{x, y\} \cup \{a_e\} \text{ avec } a_e \notin V(G); \\ & - E(G/e) = \left(\bigcup_{f \in E(G) \setminus \{e\}} f/e \right) \setminus \{e\} \text{ où } f/e = \{a_e\} \cup f \setminus e \text{ si } f \cap e \neq \emptyset \text{ et } f/e = f \\ & \text{ sinon.} \end{aligned}$$

La figure montre une telle opération. Un graphe obtenu par contractions successives à partir de G s'appelle une *contraction* de G .

Un graphe H est dit *mineur* de G s'il est obtenu à partir de G par application successive d'opérations parmi :

- suppression d'un sommet et de ses arêtes incidentes ;
- suppression d'une arête ;
- contraction le long d'une arête.

Autrement dit si H est un sous-graphe d'une contraction de G .

FIG. 1.1 – Contraction le long de e

1.1.6 Quelques graphes

La plupart des graphes que l'on considère dans ce mémoire sont des graphes simples non-orientés. Dans cette section nous donnons ceux dont on fera souvent usage. On a déjà rencontré le graphe vide E_n d'ordre n et son complément K_n . Le graphe k -parti complet K_{m_1, m_2, \dots, m_k} est un graphe k -parti tel que n'importe quel pair de sommets n'appartenant pas à un même ensemble indépendant sont adjacents. Les m_i représentent les tailles des k ensembles indépendants. Si tous les m_i sont égaux à m , on écrit plutôt $K_{k; m}$. Le cycle d'ordre n noté C_n est un graphe qui consiste en un unique cycle de longueur n . Un graphe simple connexe qui ne possède pas de cycle est un *arbre*. Un arbre sur n sommets dont exactement deux sont pendants, c'est-à-dire de degré 1, et les $n - 2$ restants de degré 2 est une chaîne de longueur $n - 1$ et que l'on note P_n .

On définit l'*union* de deux graphes G et H , notée $G \cup H$, par $V(G \cup H) = V(G) \cup V(H)$ et $E(G \cup H) = E(G) \cup E(H)$. Si $V(G) \cap V(H) = \emptyset$ alors l'union est *disjointe* et nous la notons $G \sqcup H$. Le *line graphe* d'un graphe G est le graphe $L(G)$ où $V(L(G)) = E(G)$ et $\{x, y\} \in E(L(G))$ si et seulement si x et y possèdent un sommet commun.

1.2 Notions d'analyse

Un espace topologique $\mathcal{T} = (X, T)$ est un ensemble X muni d'une collection $T \subseteq \mathcal{P}(X)$ satisfaisant :

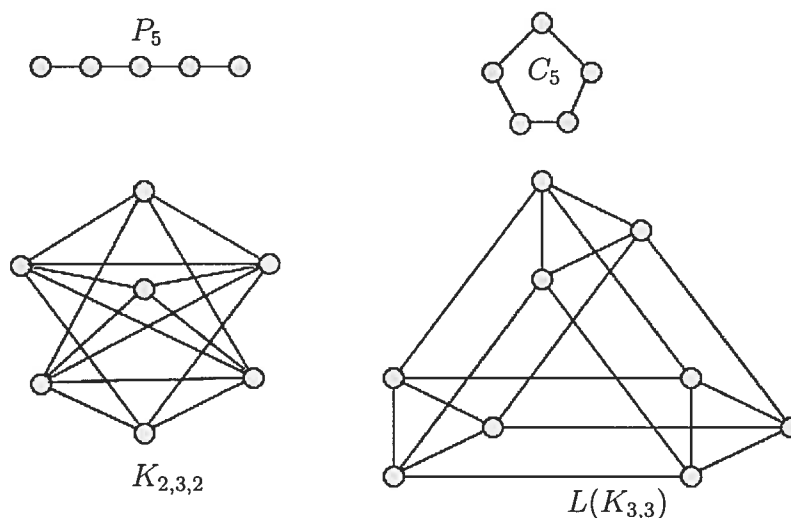


FIG. 1.2 – Quelques graphes

- $\emptyset, X \in T$;
- l'intersection d'un nombre fini d'éléments de T reste dans T ;
- l'union d'un nombre quelconque d'éléments de T appartient aussi à T .

Un élément $A \in T$ est appelé *ouvert* et son complément $X \setminus A$ *fermé*. Pour tout ensemble X , $(X, \{\emptyset, X\})$ et $(X, \mathcal{P}(X))$ forment trivialement des espaces topologiques. Une *base* d'un espace topologique $T = (X, T)$ est un sous-ensemble $B \subseteq T$ tel que tout $A \in T$ est une réunion d'éléments dans B . On dit alors que B *génère* la topologie T . L'ensemble des nombres réels \mathbb{R} muni de la topologie générée par les intervalles ouverts forme un espace topologique. Pour tout $A \subseteq X$, la collection des $A \cap O$ où $O \in T$ forme une topologie sur A , qu'on appelle *la topologie induite*. Soit $C \subseteq X$ non vide et $x \in X$. On dit que x est un *point d'accumulation* de C si pour tout ouvert O contenant x , $O \cap C \setminus \{x\} \neq \emptyset$. Un ensemble fermé F de X doit donc contenir tous ses points d'accumulation car sinon, son complément \overline{F} , qui est un ouvert, doit intersecter F . Ce qui n'est pas possible. On dit que x est un *point intérieur* de C s'il existe un ouvert O qui contient x et contenu dans C .

Soient $\mathcal{T}_1 = (X_1, T_1), \mathcal{T}_2 = (X_2, T_2)$ deux espaces topologiques et $f: X_1 \rightarrow X_2$ une fonction. Si pour tout ouvert $A \in T_2$ l'image inverse $f^{-1}(A) \in T_1$ alors on dit que f est continue. Si de plus f est une bijection telle que f et f^{-1} sont

continues, alors on dit que f est un *homéomorphisme* et on dit de \mathcal{T}_1 et \mathcal{T}_2 qu'ils sont *homéomorphes*.

Considérons le produit cartésien $X_1 \times X_2$. Le produit $\mathcal{T}_1 \times \mathcal{T}_2$ forme une topologie sur $X_1 \times X_2$ et c'est la *topologie produit* standard. De plus, les projections canoniques $p_i: X_1 \times X_2 \rightarrow X_i$ sont continues sur cette topologie.

Un espace topologique $\mathcal{T} = (X, T)$ est dit *compact* si pour tout recouvrement de X par des ouverts, on peut en tirer un qui soit fini. Plus formellement, pour tout $\{A_i\}_{i \in I}$ pour un certain ensemble d'indices I tel que $X = \bigcup_{i \in I} A_i$, il existe $i_1, \dots, i_n \in I, X = \bigcup_{i=1}^n A_i$ avec $n < \aleph_0$. L'image d'un espace compact par une fonction f continue est donc compact puisque les pré-images d'un recouvrement de $f(X)$ induisent un recouvrement de X , duquel on tire un qui est en nombre fini. Un théorème bien connu en analyse [72] caractérise les sous-espaces compacts de \mathbb{R}^n .

Théorème 1 (Heini-Borel). *Un sous-espace de \mathbb{R}^n est compact si et seulement s'il est fermé et borné.*

Soient G un graphe simple non-orienté fini et $\mathcal{T} = (X, T)$ un espace topologique. Supposons que soient donnés une fonction $s: V(G) \rightarrow X$ et pour chaque arête $\{u, v\} \in E(G)$ un homéomorphisme $\alpha_{uv}: [0, 1] \rightarrow Y_{uv}$ avec $Y_{uv} \subseteq X$ et $\alpha_{uv}(0) = s(u), \alpha_{uv}(1) = s(v)$. Si de plus les α_{uv} sont tels que $\alpha_{uv}(t_1) = \alpha_{wz}(t_2)$ seulement si $t_1, t_2 \in \{0, 1\}$ alors la *réalisation* de G dans X (plus précisément dans \mathcal{T}) est l'espace topologique induit par

$$R(G) = \bigcup_{\{u,v\} \in E(G)} \alpha_{uv}([0, 1]) \bigcup s(V(G)).$$

Une réalisation triviale dans $\mathbb{R}^{|G|}$ est celle donnée par

$$\begin{aligned} s(u_1) &= (1, 0, 0, \dots, 0), \\ s(u_2) &= (0, 1, 0, \dots, 0), \\ &\vdots \\ s(u_n) &= (0, 0, \dots, 0, 1) \end{aligned}$$

et

$$\alpha_{uv}(t) = tv + (1 - t)u;$$

si $V(G) = \{u_1, u_2, \dots, u_n\}$. On peut consulter [18] pour une réalisation d'un graphe dans \mathbb{R}^3 . Un graphe qui possède une réalisation dans \mathbb{R}^2 est dit *planaire*.

Nous adoptons les notations asymptotiques suivantes : pour $f: \mathbb{N} \rightarrow \mathbb{R}_+$ et $g: \mathbb{N} \rightarrow \mathbb{R}_+$

- $f \ll g$ ou $f = o(g)$ pour signifier que $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$;
- $f = O(g)$ pour désigner le fait que $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = C$ pour une constante réelle positive C ;
- $f = \Theta(g)$ si $f = O(g)$ et $g = O(f)$.

1.3 Théorie des jeux

Ce qui semblait être au départ une nouvelle approche aux problèmes en économie a très vite suscité assez d'intérêt pour former une discipline à elle seule : la théorie des jeux. John von Neumann et Oskar Morgenstern ont senti que le comportement typique d'un problème en économie s'apparente aux notions mathématiques des jeux de stratégies. Une stratégie n'est autre chose qu'une description complète de comment on doit réagir devant toute circonstance. Depuis la publication de leurs travaux, les applications de cette nouvelle théorie ont trouvé leur place dans de nombreux domaines : science politique, finance, psychologie, sociologie et bien évidemment en mathématiques.

Un théorème général de la théorie des jeux et qui nous importe le plus dans ce mémoire concerne les jeux finis à information complète à somme nulle joués à deux. Sans entrer dans les détails de formalisation, un jeu où joueur A et joueur B alternent leurs coups, est à somme nulle lorsque ce que l'un perd, l'autre le gagne, c'est-à-dire ils ont des intérêts diamétralement opposés. Il est à information complète lorsqu'il n'y a pas de surprise : à tout moment, les deux joueurs sont informés de l'état du jeu.

Théorème 2. *Un jeu à deux personnes fini à information complète et à somme nulle est complètement déterminé.*

Par « complètement déterminé », il faut comprendre qu'un des joueurs possède une stratégie gagnante, c'est-à-dire qu'il existe une façon pour un des deux de se comporter de sorte que l'autre perdra et cela quelque soit ses actions. Ce théorème est une version plus forte du fameux théorème minimax de John von Neumann, qui dit essentiellement que si notre adversaire apprend notre stratégie alors le maximum que l'on peut espérer gagner est le minimum que notre adversaire peut perdre. On peut consulter [80] pour plus de détails. Pour notre part, nous présenterons les grandes lignes de la preuve donnée par Zermelo.

Démonstration. Supposons au contraire qu'aucun des deux n'a de stratégie gagnante. Cela veut dire que pour n'importe quelle stratégie employée par le joueur A, il existe toujours une stratégie du joueur B qui va le mener à la victoire et inversement, quelque soit la stratégie choisie par B, le joueur A en a une qui lui assurera la défaite de B.

Le premier coup de A ne lui permet pas de gagner car B a une stratégie pour éviter cela. De même, le premier coup de B ne lui permet pas de gagner car sinon il a une stratégie gagnante, à savoir attendre que A joue et choisir ce coup victorieux. Maintenant, supposons qu'à leur i -ème coup ni A ni B n'est parvenu à triompher. N'importe quel coup que A va jouer par la suite ne peut être gagnant sinon il a une stratégie gagnante depuis le début. Toutefois, il existe au moins un coup que A peut jouer pour éviter que B gagne sinon c'est B qui a une stratégie gagnante

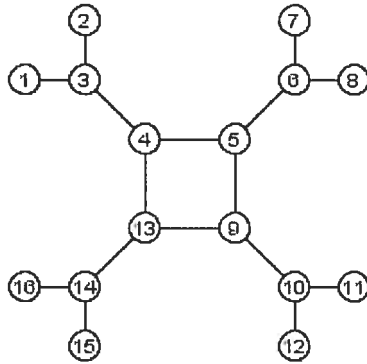
depuis le début. Par induction, le jeu ne peut pas finir, contredisant le fait qu'il est fini. \square

1.4 Jeux de poursuite

1.4.1 Fouille

On considère une réalisation $R(G)$ dans l'espace euclidien \mathbb{R}^3 du graphe non-orienté connexe fini G . Pour garder la notation légère, on s'est permis d'utiliser G à la place de $R(G)$ tout en gardant en tête que lorsqu'on parle d'un sommet u , respectivement d'une arête $\{u, v\}$, il s'agit en réalité de $s(u)$, respectivement $\alpha_{uv}([0, 1])$, où $s: V(G) \mapsto \mathbb{R}^3$ et $\alpha_{uv}: [0, 1] \mapsto \mathbb{R}^3$ avec $\alpha_{uv}(0) = s(u)$ et $\alpha_{uv}(1) = s(v)$. Le déplacement d'un individu dans cette réalisation est représentée par une fonction continue $f: [0, \infty) \mapsto G$. Intuitivement, l'individu se trouve à l'instant t sur un point de l'espace qui est soit un sommet, soit contenu dans une arête du graphe. Pour un entier positif k , on dit qu'il existe une *stratégie de fouille* pour k chercheurs pour G , si ces derniers peuvent se déplacer de sorte à garantir, dans un temps fini, la capture de tout intrus peu importe sa vitesse et son parcours dans le graphe. Plus formellement, une stratégie pour k fouilleurs est une collection f_1, f_2, \dots, f_k de fonctions continues $f_i: [0, \infty) \rightarrow G$, telle que pour tout déplacement de l'intrus $r: [0, \infty) \rightarrow G$, il existe un entier $1 \leq i \leq k$ et un réel $0 \leq t_r$ tels que $f_i(t_r) = r(t_r)$. Nous assignerons le terme *temps de capture* de l'intrus r à t_r . Considérons le graphe Y^2 de la figure 1.3. Une stratégie pour trois chercheurs $\gamma_1, \gamma_2, \gamma_3$ sera la suivante :

- (i) $\gamma_1, \gamma_2, \gamma_3$ se trouvent respectivement sur les sommets 1, 2 et 7 ;
- (ii) γ_1 part du sommet 1 vers 3 et attend γ_2 qui est parti du sommet 2 vers 3 au même moment ;
- (iii) γ_1, γ_2 vont de 3 vers 4 ensemble ;
- (iv) γ_2 va de 4 vers 5 tandis que γ_1 se repose sur 4 ;
- (v) γ_2 se déplace de 5 vers 6 et s'y arrête un instant ; pendant ce temps, γ_3 va de 7 vers 6 et ensuite de 6 vers 8 pour revenir sur 6 ;

FIG. 1.3 – Graphe Y^2

- (vi) γ_2, γ_3 , tous deux, reprennent le chemin vers 5; où ils se séparent, γ_2 va rejoindre γ_1 sur 4 et γ_3 se rend vers 9;
- (vii) γ_1, γ_2 poursuivent leur route de 4 vers 13, où γ_1 se stationne pendant que γ_2 fait l'aller-retour de l'arête (13, 9);
- (viii) γ_1, γ_2 vont de 13 à 14, où ils se délaissent le temps d'un va et vient vers 15 pour l'un et vers 16 pour l'autre;
- (ix) de retour au 14, γ_1, γ_2 vont retrouver γ_3 sur 9, mais continuent leur course vers 10, où γ_1 se dirige vers 11 et γ_2 vers 12.

Un détail important à mentionner est la *récontamination*. Une partie du graphe est réputée contaminée si elle constitue un endroit où l'intrus peut potentiellement se trouver. Une arête qui n'a jamais reçu la visite d'un chercheur, par exemple, est contaminée. Lorsqu'un fouilleur passe par un point, le statut de ce dernier mue à l'état non contaminé. Le problème initial revient alors à trouver une stratégie pour que tous les points du graphe deviennent et restent non contaminés en un temps fini. Une arête devient récontaminée si l'une de ses extrémités est contaminée. Un sommet repasse à l'état souillé si aucun chercheur ne s'y trouve et qu'une de ses arêtes est contaminée. Une façon de voir la récontamination est de considérer le cas

d'un gaz toxique qui se répand dans un tunnel. Une équipe de pompiers sait arrêter la propagation du gaz et le repousser d'un bout à l'autre du corridor. Le gaz peut se disperser vers des endroits où un pompier ne peut l'empêcher. A l'étape (v) de la stratégie précédente, les arêtes (4, 5) et (5, 6) deviennent recontaminées lorsque γ_2 s'est déplacé du sommet 5 vers 6.

Théorème 3. *Pour tout graphe connexe fini G , il existe toujours une stratégie de fouille.*

Démonstration. En effet, une stratégie triviale avec $|G| + 1$ chercheurs est de poster un fouilleur immobile sur chaque sommet et laisser le dernier nettoyer toutes les arêtes du graphe. On peut faire un peu mieux avec $|G|$ chasseurs. On en place un sur chaque sommet. Soit $x \in V(G)$ un sommet de degré maximum. Pour commencer, les $\Delta(G)$ chercheurs qui se trouvent sur $N[x] \setminus \{x\}$ viennent sur x . Si x possède une boucle, on la nettoie en premier. Par la suite, les $\Delta(G)$ chercheurs sur x repartent vers leurs sommets d'origine. Le sommet x et ses arêtes incidentes sont maintenant décontaminés. Celui qui est resté peut maintenant quitter ce sommet pour aller parcourir toutes les arêtes restantes, car il n'y a plus de risque de récontamination.

□

L'écart entre le nombre de chercheurs utilisés dans la preuve et celui de la stratégie trouvée pour le graphe Y^2 motive la quête du nombre minimal de chercheurs pour fouiller un graphe. Pour un graphe G , ce nombre sera noté $sw(G)$. Ainsi, $sw(Y^2) \leq 3$. On verra plus loin que $sw(Y^2) \neq 1, 2$. Le problème de déterminer si $sw(G) \leq k$ pour un k donné est NP-complet pour un graphe en général [76] mais linéaire pour les arbres. De plus, étant donné un entier $k > 0$, on sait construire un arbre T tel que $sw(T) = k$. On remarque que si H est un sous-graphe connexe de G , alors $sw(H) \leq sw(G)$. En effet, la même stratégie, restreinte à H , pour nettoyer G décontaminera H si on prend soin de remplacer le déplacement d'un chercheur de u à v hors H , pour $u, v \in H$, par un déplacement u vers v dans H , étant donné que H est connexe. Donc, si un sous-graphe connexe de G nécessite k fouilleurs au minimum, alors $sw(G) \geq k$. De même, la même stratégie de fouille sur G suffit à

décontaminer une contraction de G le long d'une arête. Donc, si H est un mineur de G alors $\text{sw}(H) \leq \text{sw}(G)$.

En gardant l'analogie du gaz toxique, au moment C où toutes les zones contaminées sont réduites à néant, tout intrus s'y trouvant, quelque soit son déplacement, est garanti capturé. En effet, Parsons a montré dans [81] le théorème suivant :

Théorème 4. *Pour toute stratégie de fouille $F = \{f_i : 1 \leq i \leq k\}$, il existe une constante $C \in \mathbb{R}$ telle que pour toute fonction continue $r : [0, \infty) \rightarrow G$, son temps de capture $t_r \leq C$.*

La démonstration nécessite une définition et un lemme. Soit $X = [0, \infty) \times G$ muni de la topologie produit standard. Pour une fonction continue $r : S \rightarrow G$, où $S \subseteq [0, \infty)$, posons $\text{GR}(r) = \{(t, r(t)) : t \in S\}$ et plus généralement pour une stratégie de fouille $F = \{f_1, f_2, \dots, f_k\}$ sur G , $\text{GR}(F) = \bigcup_{i=1}^k \text{GR}(f_i)$. Dans l'espace topologique X , $\text{GR}(F)$ est un ensemble fermé étant donné que son complément est vide. Maintenant, soient $(t, x), (t', x') \in X$. Définissons la relation binaire \ll_F , par $(t, x) \ll_F (t', x')$ si et seulement si $t < t'$ et il existe une fonction continue $r : [t, t'] \rightarrow G$ telle que $r(t) = x, r(t') = x'$ et $\text{GR}(r) \cap \text{GR}(F) = \emptyset$.

Lemme 2. *Soient $(t, x), (t', x') \in X$ tels que $(t, x) \ll_F (t', x')$. Dénoteons le sous-graphe $G \setminus \{f_1(t), f_2(t), \dots, f_k(t)\}$ par $G(t, F)$ et la composante connexe de $G(t, F)$ contenant le point x par $G(t, x, F)$. Si y est un point de $G(t, x, F)$ et y' un point de $G(t', x', F)$ alors $(t, y) \ll_F (t', y')$.*

Démonstration. L'idée, c'est que la fonction r de la relation $(t, x) \ll_F (t', x')$ fait presque l'affaire sauf autour de y et y' , où il faut s'assurer de trouver un déplacement qui ne rencontre pas les f_i . Soit $r : [t, t'] \rightarrow G$ continue telle que $r(t) = x, r(t') = x'$ et $\text{GR}(r) \cap \text{GR}(F) = \emptyset$. Comme $f_i(t) \neq x$ pour $1 \leq i \leq k$, alors x est un point intérieur de $G(t, x, F)$, et donc il existe un ouvert qui contient x et contenu dans $G(t, x, F)$. Comme r est continue, il existe $\epsilon > 0$ tel que $r([t, t + \epsilon]) \subseteq G(t, x, F)$. Soit P une chaîne de longueur minimale joignant x et y dans $G(t, x, F)$, et posons $H = P \cup r([t, t + \epsilon])$. Clairement, H est connexe et fermé et donc les $f_i(t)$ ne

peuvent pas être des points d'accumulation de H , c'est-à-dire pour tout $1 \leq i \leq k$, il existe $\epsilon_i > 0$ tel que $f_i([t, t + \epsilon_i]) \cap H = \emptyset$. En prenant le minimum des epsilons, $\epsilon_0 = \min \{\epsilon, \epsilon_1, \epsilon_2, \dots, \epsilon_k\}$, les $f_i([t, t + \epsilon_0])$ n'intersectent pas H et $r(t + \epsilon_0) \in H$. On procède de même avec $G(t', x', F)$ et y' pour obtenir δ_0 . La fonction voulue est maintenant, $h: [t, t'] \rightarrow G$ telle que

$$h(s) = \begin{cases} y \hookrightarrow r(t + \epsilon_0) & \text{si } s \in [t, t + \epsilon_0], \\ r(s) & \text{si } s \in [t + \epsilon_0, t' - \delta_0] \\ r(t' - \delta_0) \hookrightarrow y' & \text{si } s \in [t' - \delta_0, t'] \end{cases}$$

où $x \hookrightarrow y$ dénote le déplacement uniforme le long d'une chaîne de x vers y . \square

Preuve du théorème 4. Soit $r: [0, \infty) \rightarrow G$ une fonction continue. On sait que $\text{GR}(r) \cap \text{GR}(F)$ est non-vide et fermé dans X . Ainsi, $\{t : (t, r(t)) \in \text{GR}(F)\} \subset \mathbb{R}$ est non vide et possède une plus grande borne inférieure \tilde{t} . De plus $(\tilde{t}, r(\tilde{t})) \in \text{GR}(F)$ puisque $\text{GR}(r) \cap \text{GR}(F)$ est fermé. Par conséquent, $t_r = \min \{t : (t, r(t)) \in \text{GR}(F)\}$ est bien défini.

Si nous posons $A = \{t_r : r : [0, \infty) \rightarrow G \text{ continue}\}$, il suffit de vérifier que A est borné pour démontrer que $C = \sup A$ est la constante voulue. Supposons le contraire ; pour tout entier positif m , il existe alors une fonction continue $r_m: [0, \infty) \rightarrow G$ pour laquelle $t_{r_m} > m$. En posant $y_m = r_m(0)$ et $z_m = r_m(m)$, on a par définition que $(0, y_m) \ll_F (m, z_m)$. Le sous-graphe $G \setminus \{f_1(0), f_2(0), \dots, f_k(0)\}$ n'a qu'un nombre fini de composantes connexes. Dès lors, une de ces composantes, que l'on notera $C(0, F)$, est telle que $S' = \{m : y_m \in C(0, F)\}$ est infini. Par le lemme 2, pour tout $x_0 \in C(0, F)$, $(0, x_0) \ll_F (m, z_m)$ et cela pour tout $m \in S'$.

Maintenant posons $S = \{(t, x) \in X : \forall t' > t, \exists x' \in G, (t, x) \ll_F (t', x')\}$. L'ensemble S n'est pas vide puisque $(0, x_0) \in S$. Soit alors $(t, x) \in S$, on veut montrer que pour tout $t' > t$, il existe $x_{t'}$ tel que $(t', x_{t'}) \in S$. Pour tout entier $m > t'$, posons $x_m = h_m(t')$. Ainsi, $(t', x_m) \ll_F (m, z_m)$. Par le même fait que $G \setminus \{f_1(t'), f_2(t'), \dots, f_k(t')\}$ ne possède qu'un nombre fini de composantes connexes, une des composantes, $C(t', F)$ est telle que l'ensemble $\{m : x_m \in C(t', F)\}$ est infini. En prenant $x_{t'} \in$

$C(t', F)$, par le lemme 2, $(t', x_{t'}) \ll_F (m, z_m)$ pour une infinité d'entiers m . Et donc, $(t', x_{t'}) \in S$.

A présent, en commençant avec $(0, x_0) \in S$, on sait qu'il existe $(1, x_1) \in S$ tel que $(0, x_0) \ll_F (1, x_1)$. Par induction, si $(m, x_m) \in S$, il existe $(m+1, x_{m+1}) \in S$ tel que $(m, x_m) \ll_F (m+1, x_{m+1})$ avec $m \in \mathbb{N}$. Et voilà, chaque relation $(m, x_m) \ll_F (m+1, x_{m+1})$ définit une fonction continue $h_m: [m, m+1] \rightarrow G$ qui n'intersecte pas $\text{GR}(F)$. En collant bout à bout les h_m , on obtient une fonction continue $h: [0, \infty) \rightarrow G$ telle que $\text{GR}(h) \cap \text{GR}(F) = \emptyset$. Ce qui constitue une contradiction car F est une stratégie de fouille. \square

Si nous revenons un instant sur le graphe Y^2 , on sait maintenant que $\text{sw}(Y^2) \leq 3$. Prouver que $\text{sw}(Y^2) \neq 2$ sera difficile puisqu'il y a tellement de possibilité de déplacements pour deux chercheurs. Une formulation plus discrète de la fouille est la suivante : à chaque tic d'horloge, tous les chercheurs sont localisés sur des sommets, et entre les deux tics, exactement un seul chercheur se dirige vers un sommet voisin. Plus formellement, une *stratégie de fouille combinatoire* pour k chercheurs est une collection f_1, f_2, \dots, f_k de fonctions continues $f_i: [0, \infty) \rightarrow G$ telle que :

- (i) $\forall 1 \leq i \leq k, f_i(t) \in V(G)$ pour tout $t \in \mathbb{N}$;
- (ii) $\forall n \in \mathbb{N}$, il existe un unique indice $1 \leq i_n \leq k$ tel que f_{i_n} demeure constante sur l'intervalle $[n, n+1]$ pour $i \neq i_n$ tandis que f_{i_n} meut de façon uniforme de $f_{i_n}(n)$ vers $f_{i_n}(n+1)$,
- (iii) pour toute fonction continue $r: [0, \infty) \rightarrow G$, il existe un entier $1 \leq i \leq k$ et un réel $0 \leq t_r$ tels que $f_i(t_r) = r(t_r)$.

Le nombre minimum de chercheurs pour lesquels existe une stratégie de fouille combinatoire sera noté $\text{csw}(G)$. Comme une fouille combinatoire n'est qu'une instance particulière d'une fouille, trivialement, $\text{sw}(G) \leq \text{csw}(G)$. Le théorème suivant révèle que la restriction de la fouille combinatoire n'est pas contraignante.

Théorème 5. *Pour tout graphe fini connexe G , $\text{sw}(G) = \text{csw}(G)$.*

Preuve (esquisse). On peut se référer à [10] pour une preuve complète du théorème. Ce qui reste à prouver c'est $\text{csw}(G) \leq \text{sw}(G)$, c'est à dire que pour toute stratégie de fouille, on peut en tirer une qui soit combinatoire. L'idée est la suivante : soit $F = \{f_1, f_2, \dots, f_k\}$ une stratégie de fouille. On sait qu'il existe une constante C à partir de laquelle tout intrus dans le graphe est capturé. Pour chaque f_i , l'ensemble $T_i = \{t : f_i(t) \in V(G) \text{ et } t \in [0, C]\}$ est fermé. Si T_i est non vide, il possède une plus grande borne inférieure t_{i_1} . De même, $T_{i,i_1} = \{t \in T_i : t > t_{i_1} \text{ et } f_i(t) \neq f_i(t_{i_1})\}$ possède une plus grande borne t_{i_2} si il n'est pas vide. A cet instant, les intervalles $[0, t_{i_1}]$, $[t_{i_1}, t_{i_2}]$, $[t_{i_2}, C]$ forment une couverture de $[0, C]$. On continue le processus et ainsi construire une couverture de l'intervalle $[0, C]$ par des fermés ; mais $[0, C]$ étant compact, le processus ne peut se poursuivre qu'un nombre fini de fois. Chaque déplacement f_i donne lieu à des sous-intervalles de $[0, C]$. En les ordonnant linéairement, on élabore une stratégie combinatoire en parcourant par ordre croissant ces sous-intervalles, où chaque i -ème sous-intervalle correspond à un déplacement uniforme d'un chercheur d'un sommet à un autre dans le temps $[i, i + 1]$. L'ordre linéaire à ce moment là est :

- $[a, b] < [c, d]$ si $a < c$,
- $[a, b] < [c, d]$ si $a = c$ et $b < d$,
- $[a, b] < [c, d]$ si $a = c, b = d$ et $i < j$, où i, j sont les indices des chercheurs donnant lieu à $[a, b], [c, d]$ respectivement.

Deux ajustements sont toutefois nécessaires pour que cela fonctionne correctement. Premièrement, $f_i(0)$ n'est pas forcément situé sur un sommet mais sur un point intérieur de l'arête (u, v) . Si $f_i(t_{i_1}) = v$, alors dans la stratégie combinatoire le chercheur i se déplacera uniformément de u à v dans l'intervalle $[j, j + 1]$, où $[0, t_{i_1}]$ est le j -ème intervalle dans l'ordre linéaire.

Ensuite, deux fouilleurs peuvent décontaminer une arête en partant chacun des extrémités et se rencontrer dans un point intérieur de l'arête pour revenir chacun aux sommets par lesquels il sont venus ; et donc aucun d'eux n'a vraiment traversé l'arête en entier. Une pareille situation se traduit par une arête non parcourue dans la stratégie combinatoire obtenue. On se débarrasse de ces situations alors

avant de commencer en commutant les déplacements des deux chercheurs après leur rencontre. Il s'avère que seulement un nombre fini de ces situations nécessite d'être pris en charge. \square

L'introduction de la fouille combinatoire facilite l'argumentation lorsqu'il s'agit de montrer que $csw(G) > k$. En effet, montrer que k chercheurs adoptant une stratégie combinatoire n'y arrivent pas est suffisant. Dans la preuve du théorème 3, la borne $|G|$ est atteinte. En effet, Parsons a noté que $csw(K_n) = n$ pour $n \geq 4$.

Pour de petits entiers, on peut produire des caractérisations de graphe ayant comme nombre minimal de fouilleurs ces entiers. Ici, nous donnerons ces graphes pour $k = 1, 2$. La *réduction homéomorphique* du graphe G est le graphe H obtenu par application de l'opération suivante : pour chaque sommet $u \in V(G)$ tel que $d(u) = 2$ et $N(u) \setminus \{u\}$ non vide, remplacer les arêtes (u, v) et (u, w) par l'arête (v, w) et enlever le sommet u ; et cela jusqu'à ce qu'aucun sommet ne puisse plus être retiré. Si une arête existe déjà entre v et w , la nouvelle ne viendra pas la substituer de sorte qu'on a une arête multiple sur v et w . On appelle alors H *multigraphe*. C'est essentiellement se débarrasser des sommets de degré 2 qui se confondent topologiquement aux points intérieurs d'une arête par contraction. Ainsi, une chaîne suspendue dans G , c'est-à-dire une chaîne dont les sommets sont tous de degré inférieur à 2 dans G , se réduit à une arête ; et un cycle suspendu dans G se réduit à une boucle.

Lemme 3. *Si H est la réduction homéomorphique de G alors $sw(G) = sw(H)$.*

Démonstration. Lorsqu'un fouilleur traverse une arête réduite (u, v) dans H , nous laissons le même fouilleur traverser la chaîne suspendue reliant u à v dans G . Donc $sw(G) \leq sw(H)$.

Lorsqu'un chercheur s'engage dans une chaîne suspendue dans G , cela correspond au même chercheur traversant l'arête réduite correspondante dans H . Donc $sw(H) \leq sw(G)$. \square

Théorème 6. *$sw(G) = 1$ si et seulement si G est une chaîne.*

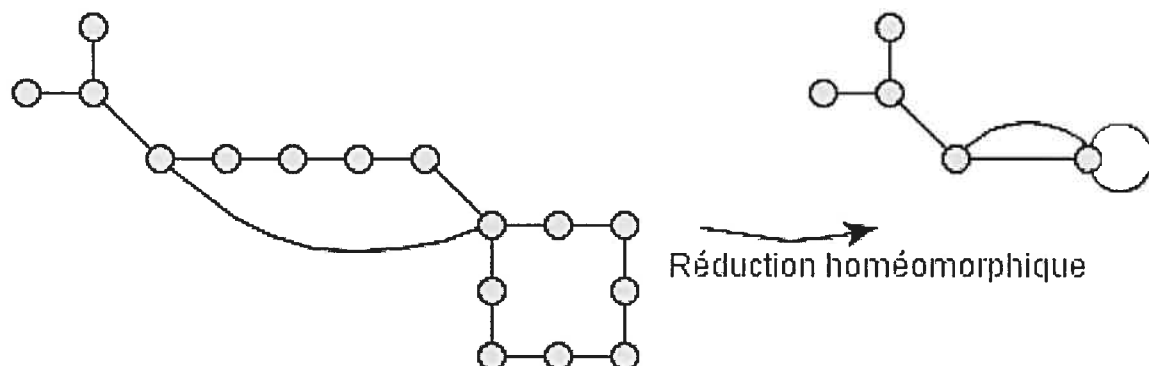
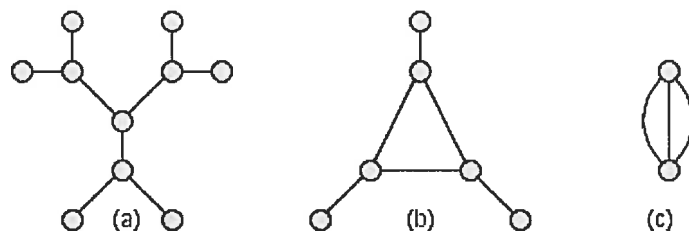


FIG. 1.4 – Réduction homéomorphique

FIG. 1.5 – Mineurs interdits pour $sw(G) \leq 2$

Démonstration. Il est facile à voir qu'un chercheur suffit pour fouiller P_n pour tout $n < \aleph_0$. Pour la partie « seulement si », considérons la dernière arête nettoyée (u, v) d'une fouille combinatoire. Il est clair que $d(v) = 1$, car sinon il y a récontamination et (u, v) ne sera pas la dernière arête nettoyée. Ensuite, pour tout sommet u , on a nécessairement $d(u) \leq 2$; sinon aussitôt qu'il quitte le sommet en empruntant une arête, toutes les autres se récontaminent. Comme, il y a déjà un sommet pendant et $\Delta(G) \leq 2$, il faut au moins un sommet de degré 1 de plus sinon la somme des degrés n'est pas paire; mais il en faut exactement un puisque la réduction homéomorphique ne contient que des sommets de degré 1 et K_2 est le seul graphe connexe de degré maximum 1. \square

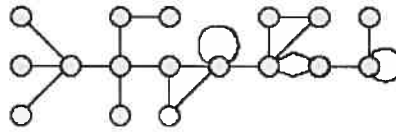
Théorème 7. *Pour tout graphe homéomorphiquement réduit G , les propositions suivantes sont équivalentes :*

- (a) $\text{sw}(G) \leq 2$;
- (b) aucun graphe de la figure 1.5 n'est mineur de G ;
- (c) G est une chaîne $u_1 u_2 \dots u_k$ telle que chaque paire consécutif u_i, u_{i+1} est relié par une ou deux arêtes, et chaque u_i possède un nombre arbitraire de boucles et d'arêtes pendantes.

Démonstration. (a) \rightarrow (b) Par inspection, tous les multigraphes de la figure 1.5 nécessitent plus de 2 fouilleurs. Donc si un de ces graphes est mineur de G , nécessairement $\text{sw}(G) > 2$.

(b) \rightarrow (c) Supposons que (b) tient. Comme le graphe de la figure 1.5c n'est pas un mineur de G , les composantes biconnexes de G sont soit des arêtes, soit des cycles. En effet, la moindre fois que plus de deux chaînes relient deux sommets distincts, on peut, par opérations de contraction d'arêtes, obtenir 1.5c. Et alors, prenons le cas d'un cycle, il ne peut contenir qu'au maximum deux sommets de degré supérieur ou égal à 3 sinon le graphe de la figure 1.5b serait un mineur de G . Par conséquent, les seules composantes biconnexes de G sont soit une arête, soit deux arêtes partageant les mêmes extrémités et soit des boucles. En faisant abstraction des arêtes multiples et des boucles, c'est donc un arbre. Si on prouve qu'en enlevant tous les sommets de degré 1 de G , il ne nous reste qu'une chaîne, en faisant abstraction des boucles et les arêtes multiples encore une fois, alors on aura montré que G est de la forme énoncée en (c). Supposons que cela ne soit pas le cas, il existe alors un sommet dans G qui soit adjacent à trois sommets distincts non-pendants autre que lui-même. Ces sommets sont donc de degré 3 ou plus puisque G ne possède pas de sommets de degré 2; mais alors, le graphe de la figure 1.5a est donc un mineur de G . Contradiction.

(c) \rightarrow (a) Lorsque les deux chercheurs sont sur u_i , le premier y monte la garde (reste immobile) pendant que l'autre nettoie les boucles et les arête pendantes sur

FIG. 1.6 – $sw(G) = 2$

u_i . Par la suite, ils vont sur u_{i+1} en empruntant chacun une arête s'il y en a deux, sinon la même. Ils commencent par u_1 pour terminer sur u_k . Donc $sw(G) = 2$.

□

Cette caractérisation ressemble fortement à celle des graphes pour lesquels $cw(G) \leq 2$ donnée dans [75]. D'ailleurs, il y est montré que $sw(G) \leq cw(G) \leq \lfloor \frac{1}{2}\Delta(G) \rfloor (sw(G) - 1) + 1$. A partir de ce théorème, on sait construire les graphes qui requièrent au plus deux fouilleurs en rajoutant des sommets de degré 2 entre les arêtes. La figure 1.6 donne un exemple de tel graphe. Pour les graphes tels que $sw(G) \leq 3$, on peut consulter [76].

1.4.2 Police-voleur

Dans la fouille combinatoire, un effort de discrétisation a été accompli pour minimiser l'influence des parcours des arêtes dans les stratégies. Ce qui importe, à chaque tic de l'horloge, c'est de savoir quel chercheur emprunte quelle arête. L'intrus, quant à lui, reste totalement libre dans ces mouvements. La question de la visibilité fait toute une différence entre la récontamination ou non d'une arête. En effet, si les chercheurs peuvent apercevoir ce qui se déroule jusqu'aux sommets voisins par exemple, ils peuvent ne pas nettoyer certaines arêtes, notamment celles où les extrémités sont occupées par des fouilleurs. Les stratégies changent d'avantage si la visibilité s'étend à tout le diamètre du graphe. Le jeu du *police-voleur* est un jeu de stratégie à deux tour par tour se jouant sur un graphe G . Le premier joueur, identifié comme la police, commence par placer ses k pions sur les sommets du

graphe. Le second, assimilé au voleur, ensuite installe le sien. Maintenant, à tour de rôle, chacun fait déplacer leurs pions de sommets en sommets voisins. Si une boucle est présente sur un sommet, on peut s'en servir pour rester sur place. Le but, pour les policiers, est de capturer le voleur qui usera de toutes les ruses pour éviter une telle drame. Le jeu est à information parfaite, en ce sens qu'autant le graphe que les positions successives, où qu'elles soient, de chacun sont connus des deux antagonistes. La police gagne si quelque soit la façon dont se comporte le voleur, en un temps fini, possiblement 0, on le capture. Sinon, si le voleur peut adopter des déplacements lui procurant une éternelle évasion, il est déclaré vainqueur. Le jeu est donc à somme nulle puisque de tous les mouvements du voleur, pour que les policiers gagnent, il faut montrer qu'il y aura coup de filet en un temps fini. Pour que le voleur gagne, tous les parcours empruntés par son adversaire ne mènent à aucune capture. Les deux situations sont complémentaires l'une de l'autre. Ainsi, étant donné un graphe et un nombre de policiers k , il existe une stratégie gagnante pour un des joueurs.

Une stratégie gagnante pour k policiers est un couple (I, f) où

- $I \in V(G)^k$ représente les positions initiales des policiers ;
- $f : V(G)^{k+1} \rightarrow V(G)^k$ est telle que pour $(y_1, y_2, \dots, y_k) = f(p_1, p_2, \dots, p_k, v)$, on ait $p_i \sim y_i$ pour $i = 1, 2, \dots, k$ et pour toute suite $(v_i)_{i \in \mathbb{N}}$ de sommets de G telle que $v_i \sim v_{i+1}$ pour tout $i \in \mathbb{N}$, si nous posons $P_0 = I$ et $P_{i+1} = f(P_i \times v_i)$, il existe $j \in \mathbb{N}$ et $1 \leq l \leq k$ tels que $v_j = \pi_l(P_j)$.

Le théorème suivant est une triviale mais nous le mentionnons tout de même étant similaire au théorème 3

Théorème 8. *Pour tout graphe fini G , il existe toujours une stratégie gagnante pour $|G|$ policiers.*

Démonstration. Soient u_1, u_2, \dots, u_n les sommets de G , n étant la cardinalité de G . La stratégie $((u_1, u_2, \dots, u_n), f)$ avec $f : V(G)^{n+1} \rightarrow V(G)^n$ avec $f(x) = (u_1, u_2, \dots, u_n)$ pour $x \in V(G)^{n+1}$ est toujours gagnante. \square

Bien entendu, affecter un policier sur chaque sommet du graphe ne donne aucune chance au voleur même si G est infini. Avec autant d'effectifs, on arrête le voleur avec 0 coup. Encore une fois, l'existence d'une stratégie générique gagnante utilisant un nombre déterminé de policiers suscite l'intérêt pour le nombre minimum de policiers pour gagner sur G . Ce nombre sera noté $sn(G)$, et on dira que G est $sn(G)$ -policier gagnant. Une stratégie évidente qui capturerait le voleur en un tour, tout au plus, consiste à placer un policier sur chaque sommet d'un ensemble dominant ; le graphe au complet en étant un. Ainsi, $sn(G) \leq \gamma(G)$. En effet, un sommet du graphe est soit dans l'ensemble dominant, soit lui est adjacent. Donc, un coup suffit pour la capture si réellement le voleur est sur un sommet adjacent.

Rappelons-nous de la réduction homéomorphique utilisée pour la fouille. Il s'agit d'appliquer successivement des contractions d'arête pour les sommets de degré 2. On peut recourir à une approche similaire pour le jeu de police-voleur. Un sommet u est dit *irréductible* s'il existe un autre sommet $v \neq u$ tel que $N[u] \subseteq N[v]$, autrement s'il existe un sommet qui le couvre autre que lui-même. L'idée est de contracter chaque sommet irréductible le long de l'arête qu'il partage avec son sommet couvrant. Une telle arête sera appelée *arête couvrante*. Bien sûr, pour qu'il existe une arête entre un sommet et celui qui le couvre, il faut qu'il ait une boucle sur les deux. Dorénavant, nous ne considérerons que les graphes réflexifs. Pour le présent jeu, cela signifie qu'un individu peut, quand vient son tour, soit se diriger vers un sommet voisin, soit rester sur place. Un sommet couvert est plutôt un coin à éviter pour les deux partis. En effet, pour la police, il vaut mieux se localiser sur le sommet couvrant étant donnée sa plus grande adjacence ; et pour le voleur, c'est un piège mortel si jamais un policier arrive sur le sommet couvrant. D'où le souhait de contracter le sommet irréductible. Ceci donne une caractérisation des graphes 1-policier gagnant [78].

Théorème 9. *Pour tout graphe réflexif non-orienté fini, $sn(G) = 1$ si et seulement s'il se réduit à un seul sommet par une suite de contractions d'arête couvrante.*

Démonstration. Si G est 1-policier gagnant, si nous prenons une stratégie optimale

du voleur, optimale dans le sens où il n'ira pas se jeter dans la gueule du loup, au dernier tour du voleur avant sa capture, il doit s'être placé sur un sommet u qui est couvert par v ; de sorte qu'en se plaçant sur v , le policier n'a plus qu'à cueillir le voleur à son prochain tour. Le graphe obtenu en contractant G de u vers v , qui est finalement $G - \{u\}$, reste 1-policier gagnant puisque pour toute stratégie gagnante (I, f) de G , si nous posons $X = f^{-1}(u)$ alors (I, g) avec $g : V(G)^2 \rightarrow V(G)$ et

$$g(x) = \begin{cases} f(x) & \text{si } x \notin X \\ v & \text{si } x \in X \end{cases}$$

est une stratégie gagnante pour $G - \{u\}$. On recommence la contraction avec ce nouvel graphe et ainsi de suite. Le graphe étant fini, là où le procédé ne pourra plus s'appliquer c'est lorsqu'il ne reste plus qu'un seul sommet, qui est bien sûr 1-policier gagnant.

Disons $|G| = n$. Numérotons les sommets dans l'ordre de leurs contractions. Ainsi le premier réduit à son sommet couvrant sera marqué u_1 ; le second, u_2 ; le troisième, u_3 ; jusqu'au dernier sommet restant u_n . Le graphe $G \upharpoonright \{u_n\}$ est sûrement 1-policier gagnant. Supposons que $G \upharpoonright \{u_j, u_{j+1}, \dots, u_{n-1}, u_n\}$ est 1-policier gagnant et considérons $G \upharpoonright \{v_{j-1}, v_j, \dots, v_{n-1}, v_n\}$. Il existe $j \leq k \leq n$ tel que v_k couvre v_{j-1} . Soit (I, f) une stratégie gagnante sur $G \upharpoonright \{u_j, u_{j+1}, \dots, u_{n-1}, u_n\}$. Si le voleur restreint son déplacement seulement aux sommets de ce dernier, il est clair qu'appliquer (I, f) suffira pour le capturer. Si, par contre, il emprunte le nouveau sommet v_{j-1} , nous laissons le policier agir selon la stratégie (I, f) , comme si le voleur était sur v_k . Plus précisément, une stratégie gagnante sur $G \upharpoonright \{v_{j-1}, v_j, \dots, v_{n-1}, v_n\}$ est (I, g) avec

$$g(x, y) = \begin{cases} f(x, y) & \text{si } y \neq v_{j-1}; \\ f(x, v_k) & \text{si } y = v_{j-1} \text{ et } f(x, v_k) \neq v_k; \\ v_{j-1} & \text{si } y = v_{j-1} \text{ et } f(x, v_k) = v_k. \end{cases}$$

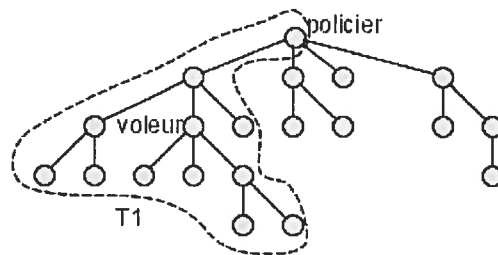


FIG. 1.7 – Police-voleur sur un arbre perché

□

Un sommet pendant dans un graphe réflexif est toujours couvert par son voisin. Ainsi, dans un arbre, on peut contracter chaque feuille vers son parent et ainsi de suite. Il s'en suit qu'un arbre réflexif est 1-policier gagnant. Un arbre fini T_0 est toujours 1-policier gagnant. Le policier peut commencer à un sommet u où il veut, que l'on va considérer par la suite comme la racine de l'arbre. Ceci décompose l'arbre en sous-arbres dont l'intersection est l'unique sommet u . Lorsque le voleur choisira son sommet initial, il appartiendra à un des sous-arbres, disons T_1 , voir figure 1.7. Le policier choisit alors de se déplacer vers son voisin v qui est dans T_1 . Le voleur avance à sa guise. Le sommet v devient la racine de $T_1 \setminus \{u\}$, dont le diamètre est strictement plus petit que celui de T_0 . Le voleur se trouve maintenant dans un des sous-arbres T_2 de T_1 dont l'intersection est v . Le gendarme avance alors vers son voisin dans T_2 . Et ainsi de suite. A chaque étape, on obtient un sous-arbre de diamètre strictement plus petit qui contient à la fois le voleur et le policier. Éventuellement, ce diamètre sera nul, ce qui signifie la capture du voleur. Ceci reste vrai dans le cas d'un arbre infini à condition que son diamètre reste fini.

Dans un graphe complet réflexif, chaque sommet couvre tous les autres. Il se réduit alors à un seul sommet par le précédent processus. D'ailleurs, c'est le seul graphe régulier réflexif qui soit 1-policier gagnant. En effet, soit G un graphe régulier réflexif non complet. Quelque soit le sommet u_0 de départ du policier, il existe un

sommet y_0 qui n'est pas adjacent à u_0 , sinon G est complet. A son premier tour, le policier n'attrapera pas le voleur. Maintenant, supposons que le policier s'est déplacé de x_{i-1} vers x_i et que le voleur soit sur $y_{i-1} \neq x_i$. C'est le tour de ce dernier de se déplacer. Il est clair que $y_{i-1} \notin N_G(x_{i-1})$ sinon le policier l'aurait déjà capturé ; mais alors, x_i ne peut pas couvrir y_{i-1} à cause de la régularité. Il s'en suit qu'il existe un sommet y_i qui ne mettra pas le voleur en danger au prochain coup du policier. Le graphe complet contient beaucoup de triangles, favorisant ainsi les policiers. En effet, Aigner et Fromme dans [6] ont trouvé une borne inférieure pour $\text{sn}(G)$:

Théorème 10. *Si le graphe non réflexif G ne contient pas de triangle ni de carré alors $\text{sn}(G) \geq \delta(G)$.*

Démonstration. Soit G comme dans l'énoncé. Posons $d = \delta(G)$. Le graphe G possède la propriété suivante : $\forall H \subseteq G$ avec $|H| = d - 1$ et $\forall u \in V(G) \setminus V(H)$, il existe $v \in V(G) \setminus V(H)$ tel que $u \sim v$ et $v \not\sim H$. Pour le voir, posons $V(H) = \{v_1, v_2, \dots, v_{d-1}\}$ et $N_G(u) = \{v_{i_1}, \dots, v_{i_k}, u_1, \dots, u_{l-k}\}$, avec $u_i \notin V(H)$. Autrement dit, k voisins de u sont dans H . Le degré minimum étant d , il s'en suit que $d \leq l$ et $l - k \geq 1$. L'inexistence de triangle empêche que tous sommets distincts $x, y \in N(u)$ soient adjacents et l'absence de carré les contraint à n'admettre aucun voisin commun autre que u . Par conséquent, tout $v \in V(H) \setminus \{v_{i_1}, \dots, v_{i_k}\}$ est adjacent à au plus un $u_i, i = 1, 2, \dots, l - k$. Or, $|V(H) \setminus \{v_{i_1}, \dots, v_{i_k}\}| = d - k - 1 < l - k$. Donc au moins un des u_i n'est adjacent à H .

Maintenant, les $d - 1$ policiers commencent sur les sommets $X_1 \subset V(G)$. Il existe un sommet $u_0 \notin X_1$ puisque $\delta(G) = d$. On sait qu'il existe un sommet $u_1 \sim u_0$ et $u_1 \not\sim X_1$. Le voleur adopte u_1 comme sommet initial. Au premier tour des policiers, ils ne captureront pas le fugitif. A présent, supposons que les policiers sont sur les sommets X et le voleur sur u avec $u \notin X$, et c'est au tour du voleur de jouer. La propriété nous dit que u possède un voisin non adjacent à X . Il suffit, pour le voleur, de choisir ce sommet comme prochaine destination. Donc $d - 1$ policiers ne suffisent pas pour capturer le voleur. \square

Des graphes d -réguliers en accord avec les conditions du précédent théorème existent pour tout $d > 0$, voir [10]. Ainsi donc, pour tout entier positif k , il existe un graphe G tel que $\text{sn}(G) \geq k$.

Fouiller un graphe infini avec un nombre fini de chercheurs nécessite un temps infini ; ce qui n'est pas le cas pour le jeu du police-voleur. Pour certains graphes infinis, le graphe complet par exemple, un nombre fini de policiers suffit en optant qu'ils ont une visibilité parfaite. Le théorème 9 ne peut s'appliquer à un graphe infini puisqu'on ne réduira pas le graphe en un nombre fini d'étapes. Une caractérisation qui ne révèle pas la structure du graphe original est donnée dans [78]. Elle a été formulée pour les graphes non-orientés 1-policier gagnants, mais s'étend très bien aux graphes orientés avec un nombre arbitraire de policiers.

Soient G un graphe, possiblement orienté et infini, et k un entier positif non nul, qui sera le nombre de policiers. Construisons un nouveau graphe orienté \tilde{G}_k avec $V(\tilde{G}_k) = V(G)^k$ et pour tous $x = (x_1, x_2, \dots, x_k), y = (y_1, y_2, \dots, y_k) \in V(\tilde{G}_k)$ l'arc $(x, y) \in E(\tilde{G}_k)$ si et seulement si $(x_i, y_i) \in E(G)$ pour $i = 1, 2, \dots, k$. Nous allons définir une série de relations binaires sur $V(\tilde{G}_k)$ de la manière suivante : pour tous $x, y \in V(\tilde{G}_k)$,

- $x \leq_0 y$ si et seulement si $\bigcup_{i=1}^k \pi_i(x) \subseteq \bigcup_{i=1}^k \pi_i(y)$;
- $x \leq_\alpha y$ si et seulement si $\forall u \in \bigcup_{i=1}^k N_G(\pi_i(x)), \exists \beta < \alpha, \exists z \in N_{\tilde{G}_k}(y)$ tels que $(u, u, \dots, u) \leq_\beta z$.

Ici, les indices α, β sont des ordinaux, donc des nombres possiblement infinis. Par définition, si $\alpha < \beta$ alors $\leq_\alpha \subseteq \leq_\beta$. Soit donc α' le plus petit ordinal pour lequel $\leq_{\alpha'} = \leq_{\alpha'+1}$ et posons \leq comme étant $\leq_{\alpha'}$. Si G est fini, $\alpha' \leq n^2 - n$.

Théorème 11. *Un graphe orienté réflexif fortement connexe G est k -policiers gagnant si et seulement $\forall u \in V(G), \forall x \in V(\tilde{G}_k)$ on a $(u, u, \dots, u) \leq x$.*

Démonstration. Supposons que $(u_0, u_0, \dots, u_0) \not\leq x_0$ pour certains $u_0 \in V(G)$ et $x_0 \in V(\tilde{G}_k)$. Assumons pour l'instant que les policiers commencent sur x_0 , c'est à dire le policier Γ_i se trouve sur le sommet $\pi_i(x_0)$ dans G . Dans ce cas, le voleur commence au sommet u_1 de G tel que $x_1 \in N_G(u_0)$ et $\forall x \in N_{\tilde{G}_k}(x_0), (u_1, u_1, \dots, u_1) \not\leq x$

x . Un tel u_1 existe sinon $(u_0, u_0, \dots, u_0) \leq x_0$. En procédant de la sorte, pour tout entier positif j , le voleur est assuré de trouver un sommet $u_{j+1} \in N_G(u_j)$ tel que $\forall x \in N_{\tilde{G}_k}(x_j), (u_{j+1}, u_{j+1}, \dots, u_{j+1}) \leq x$; où u_j et x_j représentent les positions respectives du voleur (G est isomorphe à $\tilde{G}_k \upharpoonright \{(u, u, \dots, u) : u \in V(G)\}$) et des policiers. En particulier, $(u_{j+1}, u_{j+1}, \dots, u_{j+1}) \notin N_{\tilde{G}_k}(x_j)$, et donc le voleur ne sera jamais capturé. Or rien n'oblige les policiers de commencer sur x_0 . Pourtant, s'ils possèdent un sommet stratégique y qui les fera gagner, rien ne les a empêchés de migrer vers ce sommet après avoir commencé sur x_0 et gagner à partir de là puisque le graphe est fortement connexe. Or, notre argument montre que ceci est impossible.

Inversement, soit α' le plus petit ordinal tel que $\leq_{\alpha'} = \leq \alpha' + 1$ et supposons que $\forall u \in V(G), \forall x \in V(\tilde{G}_k)$ on a $(u, u, \dots, u) \leq x$. Les policiers commenceront à un sommet arbitraire $x_0 \in V(\tilde{G}_k)$. Soit u_1 le sommet de départ du voleur. Comme $(u_1, u_1, \dots, u_1) \leq_{\alpha'} x_0$, il existe $x_1 \in N_{\tilde{G}_k}(x_0)$ et $\alpha_1 < \alpha'$ tels que $(u_1, u_1, \dots, u_1) \leq_{\alpha_1} x_1$. De façon générale, après le j -ème tour des policiers, ils sont sur x_j et le voleur sur u_j . Quand ce dernier ira vers un sommet voisin u_{j+1} , il existe $x_{j+1} \in N_{\tilde{G}_k}(x_j)$ tel que $(u_{j+1}, u_{j+1}, \dots, u_{j+1}) \leq_{\alpha_{j+1}} x_{j+1}$ avec $\alpha_{j+1} \leq \alpha_j$. Ainsi, on a obtenu une suite strictement décroissante d'ordinaux. Une telle suite atteindra toujours 0 en un nombre fini de termes. D'où éventuellement, il existe un entier positif l tel que $\alpha_l = 0$, c'est à dire $u_l = \pi_i(x_l)$ pour un certain $1 \leq i \leq k$; autrement dit, le policier Γ_i a capturé le voleur. \square

Le théorème 11 ne jette pas la lumière sur la structure du graphe original G . L'accent est plutôt mis sur le graphe intermédiaire $H = (V(\tilde{G}_k), \leq)$. Néanmoins, il donne un algorithme pour tester si G est k -policiers gagnant. En effet, pour chaque sommet de \tilde{G}_k , on doit déterminer si son voisin dans H contient l'image isomorphe de G . Ce que l'on peut faire, c'est de calculer ces voisinages incrémentalement selon l'ordinal α . On commence par $\alpha = 0, 1, 2, \dots$; mais on sait que pour un graphe fini on doit s'arrêter à $n^2 - n$ si $n = |G|$. Au bout de ce nombre tous les sommets doivent inclure comme voisinage $V(G)$ si G est k -policiers gagnant.

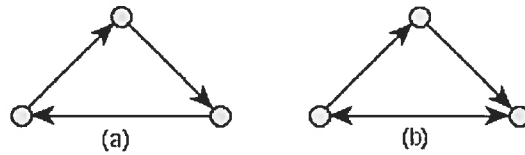


FIG. 1.8 – Triangles avec sens d'orientation

Un algorithme équivalent, a été découvert indépendamment par G. Hahn et G. MacGillivray dans [56]. Il requiert un temps polynomial et donne une stratégie optimale pour les policiers, ainsi que le temps maximal que le voleur peut espérer survivre dans le jeu. On ne connaît pas pour l'instant quel paramètre du graphe influence la durée de la partie. Ce n'est surtout pas le diamètre puisqu'on sait construire, voir [55], étant donné k , un graphe triangulé de diamètre 2 dans lequel le voleur peut jouer au moins k tours.

Aucune caractérisation satisfaisante, dans le sens structural, des graphes non orientés k -policiers gagnants pour $k \geq 2$ n'est connue. On ne peut même pas compter sur la réduction k -policiers à 1-policier dans [56] pour y arriver. L'utilisation indispensable d'arcs au lieu d'arêtes prévient l'application du théorème 9. D'ailleurs, le cas 1-policier gagnant sur un graphe orienté est tout a fait dépourvu de résultat. Considérons le cas du triangle de la figure 1.8. Les deux sont fortement connexes et les graphes non-orientés sous-jacents sont 1-policier gagnants. L'orientation des arcs favorise le voleur dans le premier et le policier dans l'autre. L'argument concernant la couverture de voisinage peut, jusqu'à un certain point, valoir puisqu'au moment précédant la capture, il y a couverture du voisinage sortant ; mais que dire du voisinage entrant ? Et comment réduire le graphe si c'est la bonne façon de procéder ? La réponse reste à découvrir.

CHAPITRE 2

MODÈLES DU WEB

L'apparition d'ordre dans les phénomènes naturels a toujours fasciné l'homme au point de consacrer une part importante des sciences, autant biologiques que physiques, à son sujet. Prenons par exemple la cristallographie, qui a connu un grand essor ces dernières décennies ; l'existence de diverses symétries entre les atomes a facilité l'étude des propriétés des molécules destinées à la pharmacologie. Beaucoup de systèmes nous entourant exhibent des comportements complexes, sinon aléatoires, qui, souvent, se manifestent sous-forme de réseaux. Chaque noeud représente un élément du système et les arêtes, les interactions entre eux. Pensons aux neurones de nos cerveaux, régulations entre protéines et gènes et bien d'autres. Ces organisations en réseaux apparaissent aussi en sciences sociales, où les noeuds correspondent aux individus, organisations ou compagnies ; et les arêtes symbolisent les liens sociaux, organisationnels ou d'affaires. Les citations dans les publications d'articles scientifiques en sont un exemple. Parmi ces phénomènes, un a pris une très grande ampleur, jusqu'à susciter la convoitise du géant Microsoft, et c'est le *Web*. Il s'agit d'une collection ouverte et mondiale de pages d'informations se référant les unes aux autres par des liens hypertextes. Bien qu'il repose sur un support physique (une série de couches de machines et de protocoles) qui est l'Internet, son évolution a défié toute frontière pour atteindre un niveau des plus abstraits. Comprendre la structure du web est devenu un enjeu majeur de la science informatique : moteur de recherche, outils de veille ou encore la sécurité des réseaux. On n'a qu'à compter le nombre de jeunes talents que la compagnie comme Google, celle qui met en ligne un des moteurs de recherche les plus utilisés actuellement, pour s'en convaincre.

Le rôle de toute science est de comprendre pour mieux expliquer et sa récompense réside dans sa capacité de prédire de nouveaux phénomènes. Comme dans toute science expérimentale, une bonne séance d'observations précède toutes expérimentation et validation de prédictions théoriques. Donc, avant même de formuler

quelque modèle que ce soit, une série de mesures s'impose.

2.1 Mesures

Le graphe du web est un énorme graphe orienté \mathcal{W} dont les sommets correspondent aux documents HTML et les arcs, aux liens hypertextes entre eux. Un arc $(u, v) \in E(\mathcal{W})$ veut donc dire qu'il existe une instruction HTML (URL) dans le document u signalant un lien vers v . Les moteurs de recherche tels Google, Lycos, Yahoo, pour ne nommer que ceux-là, construisent incessamment ce graphe. On peut s'y prendre en réalisant un parcours en largeur d'abord à partir de sommets déjà connus. Il faut noter que le caractère ouvert du web fait en sorte que des noeuds et des arcs se font ajouter et enlever constamment de ce graphe, faisant ainsi de \mathcal{W} un graphe dynamique.

Les mesures effectuées concernent la distribution des degrés, les composantes connexes, les noyaux bipartites et le diamètre du web. Les résultats sont ceux de Albert et al. [9], Broder et al. [29] et Kumar et al. [71] que nous reproduisons ici. Les travaux de Albert et al. [9] ont été conduits sur un échantillon de $3,25 \times 10^5$ sommets que possédaient le site de l'Université de Notre Dame, Indiana, en 1999. Ceux de Broder et al. [29] et Kumar et al. [71] ont été plutôt effectués sur le web au complet tel que les moteurs de recherche Altavista et Alexa le représentaient en 1999, et cela avec 2×10^8 noeuds et 1.5 milliards de liens. Actuellement, Google prétend avoir indexé près de 8 milliards de documents web.

2.1.1 Distribution des degrés

Étant donné un graphe G et un entier positif k , on définit la *distribution des degrés* de G par la fraction

$$P_G(k) = \frac{|\{u \in V(G) : \deg(u) = k\}|}{|G|}.$$

Autrement dit, $P_G(k)$ donne la proportion des sommets de degré k de G . On définit de façon similaire la *distribution des degrés entrants* $P_G^-(k)$ et la *distribution des*

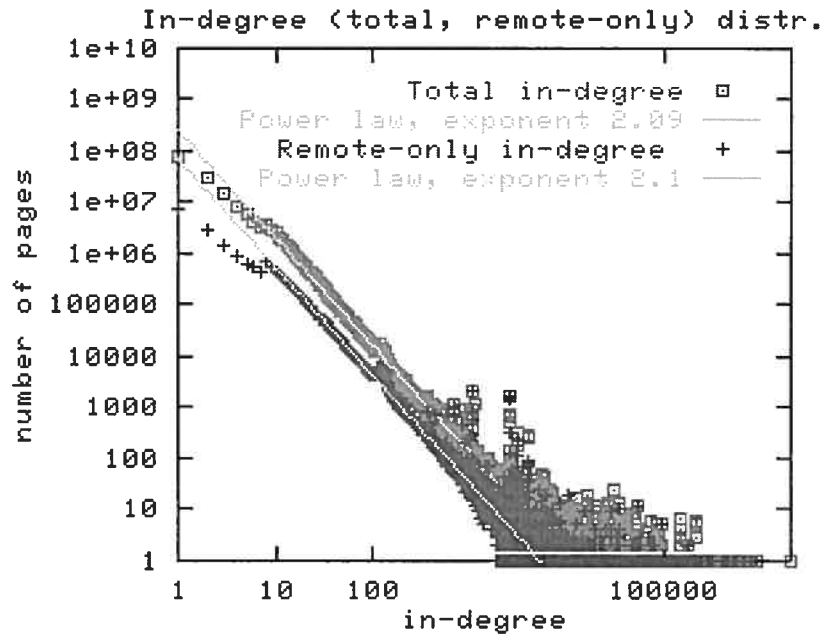


FIG. 2.1 – Distribution des degrés entrants

dégrés sortants $P_G^+(k)$ d'un graphe orienté.

Nous dirons que $P_G(k)$ suit une *loi de puissance* si

$$P_G(k) \sim ck^{-\beta},$$

pour certaines constantes réelles strictement positives β, c .

Les travaux de Albert et al. [9] et ceux de Broder et al. [29] suggèrent tous que les distributions des degrés du web suivent une loi de puissance avec $\beta = 2.1$ pour $P_W^-(k)$ et $\beta = 2.72$ pour $P_W^+(k)$. Les figures 2.1 et 2.2 sont tirées de [29] et montrent en échelle logarithmique les dispersions des pages selon les degrés. Que la distribution des degrés des pages web suive une telle loi n'est pas surprenante. La loi de puissance est présente dans plusieurs graphes incarnant des activités humaines. Elle reflète l'existence de quelques sommets de degré très élevé et plusieurs autres avec un degré moindre. Notons que le ratio $P_G(k)$ est indépendant de la taille du graphe. Cette propriété lui a valu le nom de « scale-free » selon Albert et al. [9]. Le graphe d'appel et le graphe de citations par exemple ont leur coefficient β entre 2.1 et 2.3. Un graphe d'appel est un graphe où les sommets sont les personnes d'une ville et les arêtes, les appels téléphoniques passés entre elles durant une journée.

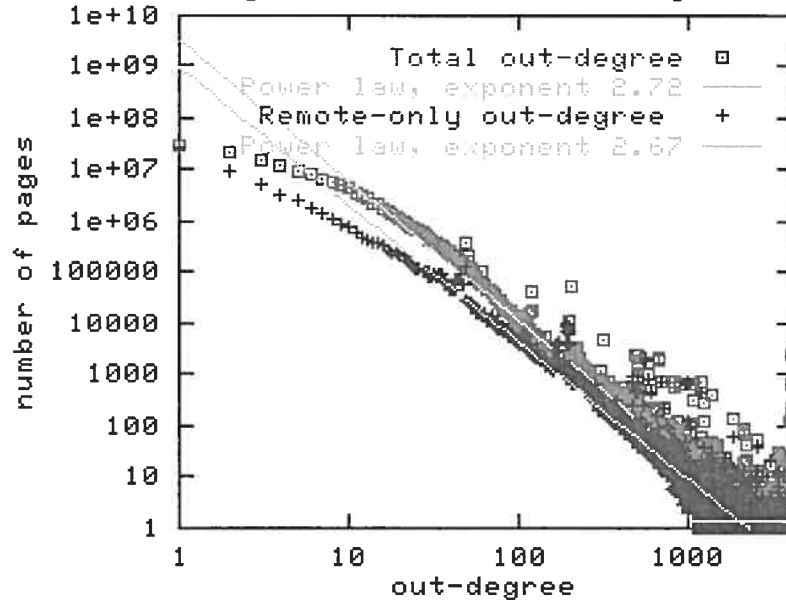


FIG. 2.2 – Distribution des degrés sortants

La loi de puissance exprime la présence d'individus qui reçoivent ou effectuent énormément de coups de fil dans la journée; en quelque sorte, ils font circuler l'information.

2.1.2 Noyaux bipartites

Le web grouille de *communautés*, ou devrait-on dire e-communautés, ces groupes de pages consacrées aux mêmes intérêts. On les retrouve souvent dans les groupes de discussions comme Yahoo, ou dans des répertoires commerciales comme le site d'un concessionnaire automobile par exemple. La nature explicite de leur mise en oeuvre fait en sorte qu'elles sont faciles à trouver. Il suffit d'aller sur les portails qui les hébergent. Pourtant, des communautés moins affichées existent, et sont même plus nombreuses. Les sujets qu'elles traitent sont peut-être trop spécialisés pour attirer les portails, ou faute de moyens ou d'organisation, elles préfèrent rester discrètes.

Kumar et al. [71] ont justement trouvé un moyen pour identifier efficacement ces communautés *implicites*. Les motivations peuvent relever de la sociologie, en étudiant comment évoluent les activités intellectuelles sur le web, de l'économie, en

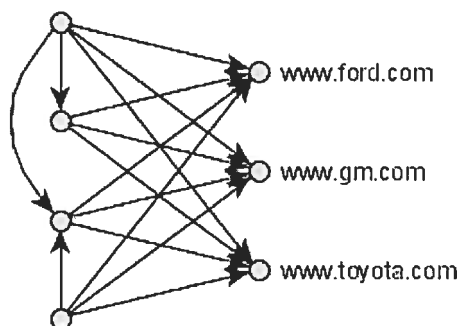


FIG. 2.3 – Noyau biparti

ciblant mieux les publicités ou tout simplement de la convivialité, en aidant les gens à trouver les communautés qui parlent de ceux qui les intéressent. Ils sont partis d'un constat qui veut que les pages d'une même communauté ne se référencent pas entre elles. On peut évoquer pour expliquer cela la compétition, par exemple Chrysler/Ford, la divergence des points de vue ou le simple fait de ne pas connaître l'existence de l'autre. Toutefois, si elles se font référencer, par une page tierce, elles le sont en même temps. Ceci donne lieu aux *noyaux bipartis*, voir figure 2.3. Un noyau biparti est un graphe sur $i + j$ sommets qui contient au moins une clique bipartie $K_{i,j}$. Un noyau biparti est très indicatif d'une communauté émergente. En fait, toute communauté possède son noyau et leur technique consiste à trouver ces noyaux afin d'extraire les communautés. La table 2.1 donne leur résultat sur un échantillon d'une taille de 2.10^8 sommets.

2.1.3 Composantes connexes

Le web n'est pas connexe. Le contraire aurait étonné étant donné la nature ouverte de celui-ci. Si on considère le graphe non-orienté sous-jacent à \mathcal{W} , Broder et al. [29] ont trouvé une composante connexe géante qui représente à elle seule 91% des sommets. Nous y attribuons la notation WCC, pour *weakly connected component*. Un tel constat peut ne pas surprendre puisque la loi de puissance stipule l'existence de sommets de forte adjacence. Ces sommets jouent peut-être le

i	j	Noyaux bipartis
3	3	89565
3	5	70168
3	7	60614
3	9	53567
4	3	29769
4	5	21598
4	7	17754
4	9	15258
5	3	11438
5	5	8062
5	7	6626
5	9	5684
6	3	4854
6	5	3196
6	7	2549
6	9	2141

TAB. 2.1 – Noyaux bipartis

k	1000	100	10	5	4	3
Taille (millions)	177	167	105	59	41	15

TAB. 2.2 – Taille de la plus grosse composante connexe survivant la suppression des sommets de degré entrant au moins k

rôle de pont dans le web. Or ce n'est pas le cas, puisqu'en enlevant des sommets ayant le degré entrant au moins $k = 3, 4, \dots$, le web maintient toujours une bonne connectivité, voir tableau 2.2. Pour $k = 5$ par exemple, la plus grande composante connexe compte pour 30% des sommets restants. Fait étonnant, la distribution des composantes connexes suit une loi de puissance $k^{-2.54}$.

La plus grande composante fortement connexe, quant à elle, occupe 56 millions de sommets, soit 28% de \mathcal{W} et 30% de WCC. On appellera cette composante SCC, pour *strongly connected component*. On peut se demander où sont passés les 70% autres sommets de WCC. Ce dernier étant non orienté, on peut distinguer trois régions de $WCC \setminus SCC$:

- IN : ce sont les sommets desquels existent des chemins menant vers SCC ;

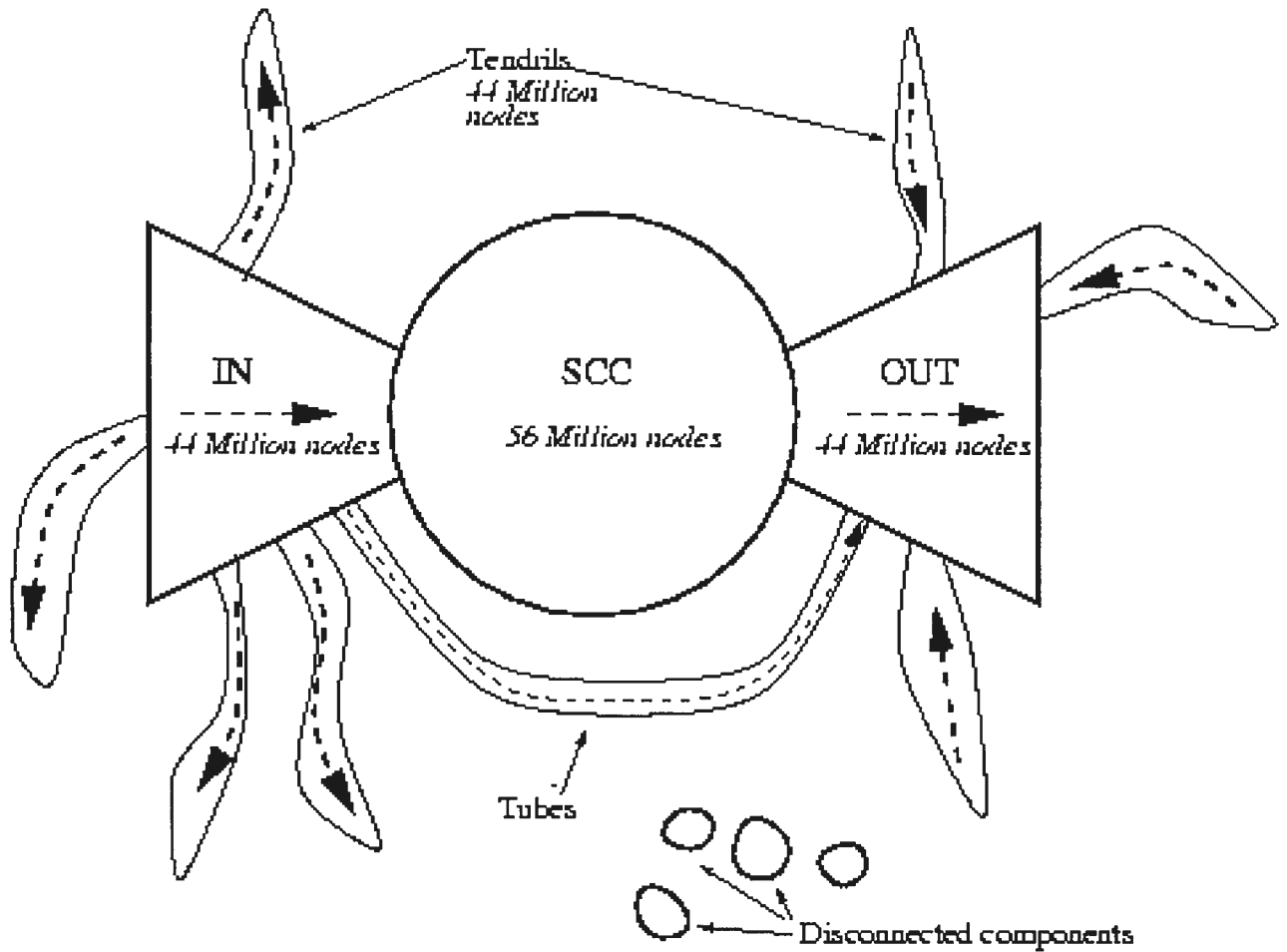


FIG. 2.4 – Connexité du web

- OUT : ce sont les sommets vers lesquels existent des chemins venant de SCC ;
- TENDRIL : ce sont les sommets vers lesquels existent des chemins venant de IN et ceux desquels existent des chemins menant vers OUT.

Le reste des sommets de \mathcal{W} , 10%, étant éparpillé dans d'autres composantes connexes, DISC. La figure 2.4 schématise la connexité du web tel qu'issu de l'étude de Broder et al. [29], tandis que le tableau 2.3 donne la dispersion des sommets dans ces régions.

Région	SCC	IN	OUT	TENDRIL	DISC	Total
Taille (millions)	56	43	43	43	17	202

TAB. 2.3 – Taille des régions du web

2.1.4 Diamètre

Le diamètre du web est infini puisqu'il n'est pas connexe. Albert et al. [9] ont prédit que pour la plupart des paires de sommets $u, v \in V(\mathcal{W})$, la distance orientée de u vers v est en moyenne de 19. Ils ont utilisé, pour cela, un modèle stochastique suivant une loi de puissance. Or, l'image du web donné par Broder et al. [29] nous indique que la probabilité qu'un chemin existe de u à v est non-négligeable seulement lorsque $u \in \text{SCC} \cup \text{IN}$ et $v \in \text{SCC} \cup \text{OUT}$. La probabilité d'un tel évènement est à peu près $24\% = 49\% \times 49\%$. Donc, la plupart du temps, 76%, deux sommets pris au hasard ne sont reliés par aucun chemin. Si un chemin existe, la distance entre deux sommets est en moyenne 16. La distance moyenne entre deux sommets étant

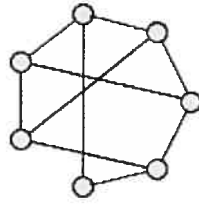
$$L(G) = \sum_{(u,v) \in S} \frac{d(u,v)}{|S|},$$

où S est tel que $(u,v) \in S$ implique $d(u,v) < \infty$. Ceci dit, le web a un petit diamètre relativement à son ordre. De plus, le graphe non-orienté sous-jacent a un diamètre moyen de 7.

2.2 Modèles

La quête d'un bon modèle pour le web a capté l'attention de la communauté scientifique [70], qui va des informaticiens, sociologues, physiciens et bien sûr mathématiciens. Les raisons en sont diverses, on peut citer :

- le besoin d'expliquer les phénomènes observés ;
- l'opportunité de prédire des propriétés jusque là inconnues ;
- la possibilité de formuler et de valider des algorithmes ;
- le forage de données ;

FIG. 2.5 – $G(7, 1/2)$

– la sociologie de la création de contenu.

Nous présentons ici, trois modèles qui ont été proposés dans la littérature. On s'est dispensé, du moins à ce stade ci, d'émettre des critères sur ce qu'est un bon modèle du web pour ne pas nous priver du plaisir que leurs inspections nous procurent. Les modèles sont tous non orientés pour simplifier leur analyse et chronologiquement les modèles prenant en compte l'orientation sont apparus dans la littérature plus tard.

2.2.1 Erdős-Rényi (ER)

On se donne n sommets que l'on numérote $1, 2, \dots, n$. Pour chaque paire non-ordonnée de sommets u, v , on lance une pièce de monnaie et on met $u \sim v$ si on obtient la face. La pièce est biaisée et donne la face avec une probabilité p . Le graphe ainsi obtenu est intuitivement le graphe aléatoire $G(n, p)$, voir figure 2.5. A première vue, il n'y a rien de naïf à considérer le web comme un graphe aléatoire. La nature ouverte de celui-ci, c'est-à-dire le créateur d'un document web peut référencer tout autre document web à sa guise, peut rendre la connectivité dépourvue de structure apparente.

Plus formellement, $G(n, p)$ est un espace de probabilité dont les éléments sont tous les graphes G sur $V(G) = \{1, 2, \dots, n\}$, il y en a $2^{\binom{n}{2}}$, dont les probabilités sont déterminées par $\Pr[u \sim v] = p$ et les $u \sim v$ sont indépendants. Plus précisément, $\Pr\{G\} = p^{|E(G)|} (1-p)^{\binom{n}{2}-|E(G)|}$. On abuse du langage en se référant à cet espace comme étant le *graphe aléatoire* $G(n, p)$.

Maintenant, nous nous intéressons à la distribution des degrés de $G(n, p)$. Qu'un sommet du graphe soit de degré k revient à dire qu'on ait obtenu k faces sur $n - 1$ lancers de la pièce. La probabilité d'un tel événement est ce que la loi *binomiale* donne exactement. Ainsi, elle souligne que la probabilité d'obtenir k succès sur $n - 1$ essais est

$$b(k; n - 1, p) = \binom{n - 1}{k} p^k (1 - p)^{n - 1 - k}. \quad (2.1)$$

Le nombre espéré de succès dans une binomiale est $\lambda = np$, d'où l'on tire $p = \lambda/n$. En ramenant cette quantité dans (2.1), on a

$$\begin{aligned} \lim_{n \rightarrow \infty} b(k; n - 1, p) &= \lim_{n \rightarrow \infty} \binom{n - 1}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n - 1 - k} \\ &= \lim_{n \rightarrow \infty} \underbrace{\frac{n(n - 1) \cdots (n - k + 1)}{n^k}}_{\rightarrow 1} \frac{\lambda^k}{k!} \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\rightarrow \exp(-\lambda)} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\rightarrow 1} \\ &= \frac{\lambda^k}{k!} e^{-\lambda}. \end{aligned}$$

Comme $G(n, p)$ possède n sommets, le nombre espéré de sommets de degré k est $\lambda = n b(k; n - 1, p)$. Par conséquent,

$$P_{ER}(k) \sim \frac{\lambda^k}{k!} e^{-\lambda},$$

qui est une loi de Poisson avec une espérance λ . Ce comportement se généralise même en considérant p comme fonction de n , voir théorème 3.1 de [17]. Cette distribution ne coïncide pas avec celle observée du web. Le degré d'indépendance dans le choix des arêtes dans ce modèle est trop élevé si bien que la distribution des degrés décroît exponentiellement en s'écartant de la moyenne λ . Autrement dit, il n'y a pas de corrélation entre les différentes arêtes. Ce qui prévient l'existence de sommets de degré élevé.

L'analyse du diamètre de $G(n, p)$ révèle qu'à mesure que le graphe grossit, c'est à dire n grandit, le diamètre tend vers 2 comme le mentionne le théorème suivant :

Théorème 12. *Considérons $p = p(n)$ telle que*

$$\lim_{n \rightarrow \infty} p^2 n - 2 \log n = \infty$$

et

$$\lim_{n \rightarrow \infty} n^2(1 - p) = \infty$$

alors

$$\lim_{n \rightarrow \infty} \Pr [\text{diam } G(n, p) = 2] = 1.$$

La preuve du théorème se trouve dans [17], mais on verra dans le prochain chapitre pourquoi c'est vrai lorsque p est constante.

La distribution des degrés n'est pas la seule discordance de ce modèle avec le web. En effet, le web évolue soit parce qu'un sommet s'ajoute, s'enlève, soit un lien change. A chaque fois qu'un sommet s'ajoute, on obtient un graphe qui contient l'ancien. Or dans ce modèle-ci, n est fixe et pour prendre en compte que le nombre sommets est passé à $n + 1$, on recommence le processus depuis le début mais avec $n + 1$ sommets. On utilise souvent le terme *hors ligne* pour de tel modèle.

2.2.2 Attachement préférentiel (BA)

Hormis le non-dynamisme du modèle précédent, un des plus importants aspects qui caractérisent ces réseaux sociaux que ce modèle ignore est le *choix* ou la *préférence*. En effet, que ce soit sur le web ou sur le graphe de citations, un nouveau document a tendance à référencer d'autres documents bien connus ou populaires. Ceci mène à une situation où les sommets de degré élevé, que l'on qualifie de populaires, tendent à attirer de nouveaux voisins plus vite que les autres. Le modèle proposé par Barabási et al. [14] et que nous allons noter (BA), voir figure 2.6, prend en compte l'évolution et l'attachement préférentiel du web :

Évolution : on commence par se donner un graphe H_0 qui n'est que m_0 sommets déconnectés. A l'étape t , on construit H_t à partir de H_{t-1} en ajoutant un nouveau sommet que l'on relie à $m \leq m_0$ sommets existants ;

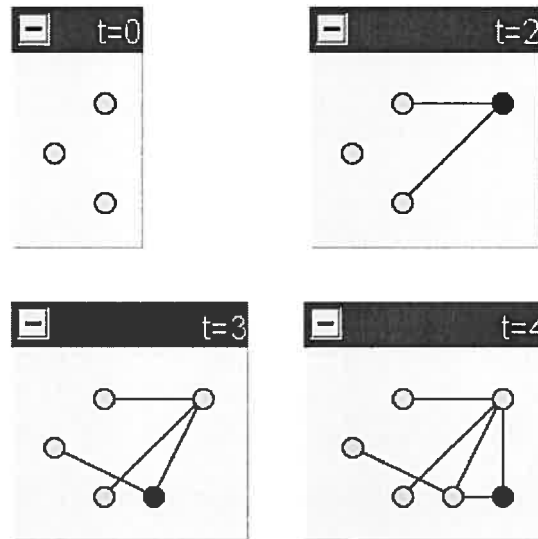


FIG. 2.6 – Attachement préférentiel

Attachement préférentiel : le nouveau sommet u_t ajouté choisit de se relier au sommet existant v selon la probabilité

$$\Pr[u_t \sim v] = \frac{d(v)}{\sum_{x \in H_{t-1}} d(x)}.$$

Ce n'est donc pas un choix uniformément aléatoire mais pondéré par le degré du sommet considéré.

Énoncé de cette façon, le modèle est imprécis. En effet, étant donné qu'au départ, tous les degrés sont nuls, il n'est pas du tout clair comment on commence. De plus, pour $t > 1$, le nombre moyen d'arêtes qui relie un nouveau sommet aux anciens est

$$\sum_{v \in H_{t-1}} \Pr[u_t \sim v] = \sum_{v \in H_{t-1}} \frac{d(v)}{\sum_{x \in H_{t-1}} d(x)} = 1$$

au lieu de m . Barabási et al. [14] ont utilisé un argument heuristique, que nous reprendrons ici d'ailleurs, pour montrer que la distribution des degrés suit une loi

de puissance avec $\beta = 3$. Ce paramètre est en accord avec des simulations qui trouvaient $\beta = 2.9 \pm 0.1$. Une version plus formelle de l'attachement préférentiel est donnée par Bollobás et al. dans [20] et l'analyse de la distribution des degrés concorde bien avec $\beta = 3$. Dorogovtsev et al. [39] donne une généralisation du modèle BA et trouve la forme exacte de la distribution des degrés entrants.

A l'étape t , le modèle contient $t + m_0$ sommets et mt arêtes. Posons $k_i = d(u_i)$ et supposons que k_i est une fonction continue. On peut alors interpréter la probabilité $\Pr[u_t \sim u_i] = k_i / \sum_j k_j$ comme indicatrice de la variation de k_i par rapport à t . Ainsi, on peut supposer

$$\frac{\delta k_i}{\delta t} = C \frac{k_i}{\sum_{j=1}^{m_0+t-1} k_j},$$

pour une certaine constante C que nous allons déterminer. On a que

$$\sum_{j=1}^{m_0+t-1} \frac{\delta k_j}{\delta t} = C \frac{\sum_{j=1}^{m_0+t-1} k_j}{\sum_{i=1}^{m_0+t-1} k_i} = C.$$

Or, $\sum_{j=1}^{m_0+t-1} \delta k_j = m$ et considérant que $\delta t = 1$, on obtient $C = m$. De plus, $\sum_{j=1}^{m_0+t} k_j = 2mt$ et donc,

$$\Pr[u_{t+1} \sim u_i] = \frac{mk_i}{2mt}.$$

En se rappelant toujours que $\Pr[u_{t+1} \sim u_i]$ donne la variation des k_i ,

$$\frac{\delta k_i}{\delta t} = m \frac{k_i}{2mt} = \frac{k_i}{2t}.$$

C'est une équation différentielle. En posant les conditions initiales $k_i(t_i) = m$, où t_i représente l'instant à partir duquel le sommet u_i a été introduit, on obtient la solution

$$k_i(t) = m \sqrt{\frac{t}{t_i}}. \quad (2.2)$$

Là, on aperçoit sans conteste l'intuition derrière ce modèle : les sommets qui acquièrent un nombre relativement élevé de voisins tôt dans le processus continuent

d'accroître leurs degrés plus vite, dû à la sélection des sommets de haut degré de préférence.

Maintenant, on peut récrire la probabilité

$$\Pr[k_i(t) < k] = \Pr[t_i > \frac{m^2 t}{k^2}].$$

En supposant qu'on ajoute les sommets à intervalle de temps régulier, la densité

$$\Pr[t_i = x] = \frac{1}{m_0 + t},$$

qui est la probabilité d'être un des sommets présents. Ainsi,

$$\Pr[t_i \leq t] = \int_0^t \frac{1}{m_0 + t} dy = \frac{t}{m_0 + t}.$$

Par conséquent,

$$\Pr[t_i > \frac{m^2 t}{k^2}] = 1 - \Pr[t_i \leq \frac{m^2 t}{k^2}] = 1 - \frac{m^2 t}{k^2(t + m_0)}.$$

On déduit la distribution

$$P_B(k) = \Pr[k_i = k] = \frac{\delta \Pr[k_i < k]}{\delta k} = \left(\frac{2m^2 t}{m_0 + t} \right) \frac{1}{k^3},$$

qui, avec suffisamment de temps, suit une loi de puissance $\beta = 3$. Intuitivement, c'est le phénomène « des riches s'enrichissant » qui mène aux sommets de haut degré prédit par la loi de puissance.

Suivant le même argument heuristique, Barabási et al. [14] montrent qu'autant le caractère évolutif du modèle que l'attachement préférentiel sont nécessaires pour atteindre cette loi de puissance. Ils ont, pour cela, considéré deux modèles où dans le premier on laisse le graphe évoluer par ajout de sommets successifs, mais dont les sommets d'attaches sont choisis de façon équiprobable. Dans le second modèle, on fixe le nombre de sommets et à chaque étape, on choisit un sommet au hasard que

l'on connecte aux autres avec un mode préférentiel. Dans les deux cas, l'analyse asymptotique ne mène pas à une loi de puissance.

Une analyse du diamètre de ce modèle est donnée par Bollobás et Riordan dans [19], ou Cohen et Havlin [33], et confirme essentiellement l'insignifiance du diamètre du modèle par rapport à son ordre. Ainsi, pour $m \geq 2$, $\text{diam}(BA) = o(\log n / \log \log n)$. Sinon pour $m = 1$, $\text{diam}(BA) = o(\log n)$.

2.2.3 Copie évolutive (BJ)

Le modèle (BA) a été proposé surtout pour modéliser les réseaux biologiques qui exhibent la loi de puissance ; mais ceci dit, le principe de l'attachement préférentiel est tout à fait plausible à l'égard du web. Kumar et al. [69] ont proposé un modèle basé sur le constat suivant : lorsqu'un auteur d'une page apporte son eau dans le moulin du web, chance est qu'il ait identifié un document « prototype » existant duquel il copiera quelques liens en plus des liens qu'il va lui-même introduire. Nous présenterons ici une version plus générale (BJ) formulée par Bonato et Janssen dans [25]. Ainsi le modèle $G(p, \phi, H)$ possède trois paramètres : $p \in (0, 1)$ qui est la probabilité de copier une arête du prototype, $\phi : \mathbb{N} \rightarrow \mathbb{N}$ est une fonction croissante de copie aléatoire et H le graphe initial de départ qui est fini. Le modèle est créé dynamiquement comme suit, voir figure 2.7 :

1. au temps $t = 0$, $G_0 = H$;
2. au temps $t > 0$, on forme G_t à partir de G_{t-1} en lui joignant un nouveau sommet u_t de la manière suivante :
 - (a) on choisit un sommet prototype v de façon uniformément aléatoire dans G_{t-1} ;
 - (b) pour tout $w \in N_{G_t}(v)$ avec probabilité p , et de manière indépendante, on ajoute l'arête (u_t, w) dans G_t ;
 - (c) on choisit aléatoirement et uniformément des sommets $w_1, w_2, \dots, w_{\phi(t)}$ dans G_{t-1} et on ajoute l'arête $(u_t, w_i), i = 1, 2, \dots, \phi(t)$ dans G_t . Ces arêtes sont appelées *liens aléatoires* ;

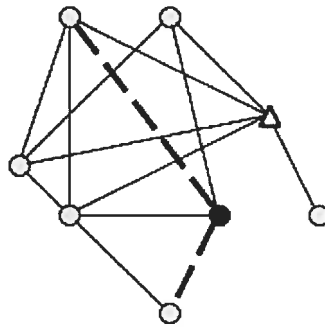


FIG. 2.7 – Copie évolutive : le noeud prototype est en triangle et les liens aléatoires sont en pointillés

(d) enlever les arêtes multiples s'il y en existe.

Lorsque $\phi(t)$ est constante, BJ est la version non-orientée du modèle de Kumar et al. [69], et pour lequel un comportement de loi de puissance avec $\beta = (2-p)/(1-p)$ a été prouvé. Or, c'est la seule forme que $\phi(t)$ peut prendre pour garantir la loi de puissance dans la distribution des degrés, voir [25]. Ce modèle implémente le phénomène « des riches devenant encore plus riches » par le biais de la copie, sans requérir à une connaissance globale des mesures de l'adjacence.

2.3 Discussions

Un bon modèle du web peut s'avérer un précieux outil analytique. Des algorithmes peuvent être vérifiés pour leur convergence ; par le biais de simulation, on peut mesurer leurs performances. Parmi les critères que l'on peut exiger d'un bon modèle du web, on peut sûrement retenir :

- (i) la loi de puissance pour la distribution des degrés, avec $\beta > 2$;
- (ii) le phénomène « petit monde » de Strogatz et Watts : le diamètre est petit par rapport à l'ordre ;
- (iii) localement dense et globalement dispersé : pour tenir compte des communautés ;

(iv) dynamique : le nombre de sommets et d'arêtes évoluent avec le temps.

Ainsi, le modèle ER semble être très désavantagé surtout en ce qui concerne la distribution des degrés. En effet, même en ajoutant un à un les sommets, et ainsi tenir compte du dynamisme, le degré d'indépendance entre les arêtes fait qu'aucun sommet n'accumule assez de degré pour suivre une loi de puissance. Une corrélation entre les arêtes doit exister pour qu'une telle loi puisse apparaître.

Le modèle BA, pour sa part, satisfait (i,ii,iv). L'ensemble de ses arêtes peut être partitionné en m arborescences disjointes. On voit cela par induction. Le nouveau sommet rajoute avec lui m nouvelles arêtes, qui vont se greffer sur les m arborescences précédentes. Une telle structure est trop simple pour contenir les noyaux bipartites. Le modèle BA a le mérite d'être simple et apporte une explication élégante à l'émergence de la loi de puissance : la préférence. Cette notion de préférence quant aux sommets d'attache semble être inhérente à la loi même de la nature ; nature qui a bercé l'homme depuis toujours.

Aucune analyse, à notre connaissance, n'est disponible concernant le diamètre du modèle BJ, autant dans sa forme générale que dans l'expression donnée par Kumar et al. [69]. Il est tout de même étonnant comment la loi de puissance émerge du processus de copie. Notons toutefois que ce processus est à la base des phénomènes des communautés. En effet, deux pages qui traitent d'un même sujet vont souvent pointer vers les mêmes pages pertinentes. Le nombre espéré de noyaux bipartis contenant $K_{i,j}$ au temps t est $t \exp(-i)$, voir [69].

Bien évidemment, d'autres modèles du web ont été proposés dans la littérature, [3, 5, 34]. Certains sont des adaptations de modèles existants, tandis que d'autres adoptent de nouvelles approches. On peut consulter [30] pour plus de détails sur les différents modèles.

CHAPITRE 3

LIMITES DES MODÈLES DU WEB

Le web acquiert de nouveaux noeuds à un rythme effréné. Son caractère évolutif défie toute borne qui essaie de contenir sa croissance. Une structure qui grossit sans limite percevable n'est-elle pas l'idée que l'on se fait de l'infini ? Considérer le graphe du web, ou tout le moins ses modèles, comme étant infini ne constitue peut-être pas seulement un exercice instructif mais un bon moyen pour révéler la structure du web à grande échelle. Étant donnés les processus stochastiques qui définissent nos modèles, la question revient à se demander : qu'arrive-t-il lorsqu'on poursuit ces étapes indéfiniment ? Les premiers à adopter une telle approche pour l'étude du web sont Bonato et Janssen [24] pour le modèle de la copie évolutive. Kleinberg et Kleinberg [67], pour leur part, ont entrepris la même investigation pour le modèle de l'attachement préférentiel. Dans les deux cas, ils ont été guidés par les travaux d'Erdős et Rényi [42] sur le modèle $G(n, p)$ et aboutissent à identifier des propriétés déterministes d'adjacences qui caractérisent, plus ou moins uniquement, ces graphes à l'infini.

Soit une suite imbriquée, ou *chaîne* de graphes $\{G_i\}_{i \in \mathbb{N}}$, c'est à dire $G_i \leq G_{i+1}$. On définit la *limite* de $\{G_i\}_{i \in \mathbb{N}}$ par

$$\lim_{n \rightarrow \infty} G_n = \bigcup_{n \in \mathbb{N}} G_n.$$

Les modèles BA et BJ donnent lieu clairement à des chaînes de graphes. La limite des $G(n, p)$ est construite avec une chaîne de graphes aléatoire de la façon suivante : on commence avec $G_0 = \emptyset$. Supposons que G_n ait déjà été défini, on construit G_{n+1} tel que $V(G_{n+1}) = V(G) \cup \{n+1\}$ et les arêtes de G_{n+1} sont celles de G_n , auxquelles on rajoute, pour chaque $v \in V(G_n)$ et avec probabilité p , l'arête $(n+1, v)$. La

probabilité d'obtenir G_{n+1} de cette façon est

$$p^{e_n}(1-p)^{\binom{n}{2}-e_n} p^{e_{n+1}-e_n}(1-p)^{n-e_{n+1}+e_n} = p^{e_{n+1}}(1-p)^{\binom{n+1}{2}-e_{n+1}},$$

où $e_i = |E(G_i)|$, $i = 1, 2, \dots$. Ce qui est la même que si on l'a pigé directement dans $G(n+1, p)$. La théorie des graphes aléatoires a été fondée par Erdős et Rényi en 1959. Erdős a été le premier à introduire des méthodes probabilistes dans l'étude des propriétés des graphes et cela de manière formelle. Leur approche consiste à approximer les valeurs exactes par des distributions de probabilité appropriées. Lorsque ce qui nous intéresse est le comportement asymptotique d'une propriété, cette méthode peut s'avérer plus efficace que la méthode déterministe, comme l'utilisation du théorème d'énumération de Pólya.

On peut considérer un graphe aléatoire $G(n, p)$ comme un organisme vivant qui grandit en se procurant des arêtes au hasard. Il *évolue* en même temps que $p = p(n)$. Ainsi, à $p = 0$, le graphe n'a aucune arête; à $p = 1/n^2$, les arêtes commencent à apparaître, à $p = 1/n^{1+1/k}$ pour k fixé, les arbres sur $k+1$ arêtes se manifestent et à $p = 1/n$, les cycles surgissent. Prenons juste un petit moment pour nous imprégner du genre d'argument de la théorie des graphes aléatoires. Considérons la propriété $P = \exists u, v : u \sim v$ qui dit simplement qu'il existe une arête. La probabilité $\Pr[G(n, p) \models P] = \binom{n}{2}p$. Et l'on voit bien que si $p = 1/n^2$ alors la probabilité qu'une arête existe dans le graphe tend vers 1. On en saura d'avantage au chapitre 5.

3.1 Limites

Le recours à la limite est un argument bien connu. Le comportement asymptotique est par exemple utilisé en algorithmique comme indicateur de la performance. Nous allons faire appel au même argument pour les modèles du web.

3.1.1 $G(n, p)$

Une des grandes surprises concernant les graphes aléatoires est que $G(n, p)$ est complètement déterministe à grande échelle. Par grande échelle, nous entendons l'infini dénombrable, qui est la limite des constructions de nos modèles. Imaginez que l'on propose à Lotto Québec le jeu $G(n, 1/2)$: chaque client construit son graphe aléatoire et celui qui trouve un graphe isomorphe à celui tiré par Lotto Québec gagne. La probabilité de gagner à ce jeu diminue quand n augmente. Et pourquoi pas, si on est avide de profit, poser $n = \aleph_0$, où \aleph_0 est la cardinalité de \mathbb{N} . Or quelle surprise, tout le monde gagne. Pensons-y un moment, lorsque $G(n, p)$ aura autant de sommets que d'entiers naturels, on aura tellement lancé la pièce de monnaie que toutes les possibilités de connectivité finiront par se trouver dans $G(\aleph_0, p)$. Ceci devrait caractériser la limite. Maintenant, nous considérons la limite de la chaîne de graphes aléatoires décrite plus tôt. Pour éclaircir notre intuition, si on se donne deux sous-ensembles disjoints X, Y d'entiers naturels, avec une certaine probabilité, à une certaine étape i du processus, on a $X, Y \in V(G_i)$ et le nouveau sommet $i + 1$ sera ajouté de sorte que les tirages donnent face sur chaque sommet de X et pile sur chaque sommet de Y . Aussi minuscule, mais non nulle, que cette possibilité soit, le fait de poursuivre une infinité d'étapes nous donne une quasi-certitude qu'elle se réalisera. Plus formellement, considérons la propriété de graphe suivante :

$$\forall X, Y \subseteq V(G) \text{ finis et disjoints, } \exists u \in V(G) \setminus (X \cup Y) \text{ tel que } u \sim X \text{ et } u \not\sim Y.$$

Un sommet $u \in V(G) \setminus (X \cup Y)$ tel que $u \sim X$ et $u \not\sim Y$ est dit *correctement relié aux sous-ensembles* X et Y . Plusieurs noms ont été attribués à cette propriété : f -saturation, existentiellement clos [24] ou restaurant d'Alice [85]. Ce dernier fait référence à une chanson de Arlo Guthrie dont le refrain fredonne : "you can get everything you want at Alice's restaurant" et qui évoque bien l'esprit de la propriété. Pour faire court, nous nous y référons par (AR).

A première vue, il n'est pas du tout clair qu'un graphe possède une telle propriété. Ce qui est sûr, c'est qu'aucun graphe fini ne peut avoir (AR) sinon en posant $X = V(G), Y = \emptyset$, on ne pourra pas trouver un sommet correctement relié à X et à Y . Jusqu'à un certain point, notre intuition a déjà fait la moitié du chemin. Tout repose sur le nombre de fois où on lance la pièce et cette probabilité devient de plus en plus certaine.

Théorème 13. Si $G = \lim_{n \rightarrow \infty} G(n, p)$ avec $p \in (0, 1)$ fixé, alors $\Pr[G \models (\text{AR})] = 1$.

Démonstration. Considérons $G(n, p)$ et $X, Y \subseteq V(G(n, p))$. Disons $|X| = r$ et $|Y| = s$. La probabilité qu'un sommet u soit correctement relié aux sous-ensembles X et Y est $p^r(1-p)^s$, ainsi, la probabilité qu'il ne le soit pas est $1 - p^r(1-p)^s$. Comme les arêtes ont été choisies de façon indépendante, la probabilité qu'aucun des sommets restants ne soit correctement relié à X et Y est $[1 - p^r(1-p)^s]^{n-r-s}$. Cette dernière tend vers 0 à mesure que n tend vers l'infini.

Jusqu'ici, on a juste montré que pour un choix particulier de X et Y , la probabilité de ne pas trouver un sommet correctement relié à X et Y devient arbitrairement petit. Or il y a $\binom{n}{r} \binom{n-r}{s}$ choix pour X et Y . Pour s, r fixés, la probabilité de ne pas trouver un sommet correctement relié à n'importe quels X et Y tels que $|X| = r$ et $|Y| = s$ est inférieure ou égale à $\binom{n}{r} \binom{n-r}{s} [1 - p^r(1-p)^s]^{n-r-s}$. Cette dernière tend aussi vers 0 lorsque n grandit indéfiniment. Il n'y a qu'un nombre dénombrable de valeurs pour r et s . On utilise un lemme élémentaire de la théorie des mesures, qui dit que l'union dénombrable d'ensembles de mesure nulle est de mesure nulle, pour conclure que $\Pr[G \models \neg(\text{AR})] = 0$, c'est à dire $\Pr[G \models (\text{AR})] = 1$. \square

Puisqu'on se permet l'infini dénombrable, on peut toujours rendre un graphe fini existentiellement clos : pour chaque sous-ensemble fini de sommets, on ajoute un nouveau sommet qui lui soit exclusivement adjacent et on recommence. En particulier, on construit une chaîne de graphe $\{G_n\}_{n \in \mathbb{N}}$ avec $G_0 = \emptyset$ et supposons G_n est défini, G_{n+1} est donné par : $V(G_{n+1}) = V(G_n) \cup \{u_X : X \in \mathcal{P}(V(G_n))\}$, $N_{G_{n+1}}(u) = N_{G_n}(u)$ pour $u \in V(G_n)$ et $N_{G_{n+1}}(u_X) = X$. Clairement, la limite de cette chaîne satisfait (AR). Nous voilà donc avec deux graphes, dont un

déterministe et l'autre stochastique, qui satisfont la propriété (AR). Qu'il y en ait deux n'est peut-être pas surprenant. Après tout, on est dans l'infini dénombrable. Ce qui est stupéfiant, c'est que cette propriété est tellement forte qu'elle détermine uniquement le graphe qui la satisfait. Une propriété telle qu'il n'existe qu'un seul graphe dénombrable, à isomorphisme près, qui la satisfait est dite \aleph_0 -catégorique. La propriété (AR) est \aleph_0 -catégorique comme le témoigne le théorème suivant :

Théorème 14. *Deux graphes dénombrables existentiellement clos sont isomorphes.*

Démonstration. Intuitivement, on construit l'isomorphisme par extensions d'isomorphismes de sous-graphe induit. La propriété (AR) de l'image permet d'égaliser l'adjacence d'un nouveau sommet dans le domaine ; et inversement la propriété (AR) du domaine permet de trouver une pré-image pour un sommet laissé pour compte dans l'image.

Soient donc G, H deux graphes dénombrables satisfaisant (AR) et soit $f : U \rightarrow H$ avec $U \subseteq G$ fini, un isomorphisme de sous-graphes induits. Pour un sommet $u \in V(G) \setminus U$, on peut étendre f à u de la façon suivante : on pose $X = N(u) \cap U$ et $Y = U \setminus X$; un bon candidat pour $f(u)$ sera un sommet de H correctement relié aux sous-ensembles $f(X)$ et $f(Y)$. Inversement, pour un sommet $v \in V(H) \setminus f(U)$, si $X = N(v) \cap f(U)$ et $Y = f(U) \setminus X$ alors un bon candidat pour le pré-image de v est un sommet de G correctement relié aux sous-ensembles $f^{-1}(X)$ et $f^{-1}(Y)$.

Maintenant, nous allons recourir à un raisonnement bien connu de la théorie des modèles : l'*aller-retour* (on l'attribue souvent à Cantor, mais c'est Plotkin qui l'a inventé et ensuite Hausdorff l'a répandu). Disons $V(G) = \{u_1, u_2, \dots\}$ et $V(H) = \{v_1, v_2, \dots\}$. Construisons une suite d'isomorphismes finis $\{f_n\}_{n \in \mathbb{N}}$ comme suit : $f_0 = \emptyset$; supposons f_n définie, si n est paire, posons $m = \min \{i : u_i \notin \text{dom}(f_n)\}$ et f_{n+1} l'extension de f_n à u_m . Si n est impaire, nous prenons $m = \min \{i : v_i \notin \text{im}(f_n)\}$ et f_{n+1} l'extension de f_n à v_m . Remarquons que $\{f_n\}$ est une chaîne de graphes et sa limite $f = \bigcup_{n \in \mathbb{N}} f_n$ est l'isomorphisme voulu. En effet, les étapes "aller" nous assurent que tous les u_i sont dans le domaine de f , comme les étapes "retour" nous garantissent que tous les v_i sont dans l'image de f . Le respect de l'adjacence pour

chaque sommet étant au coeur même de chaque extension. \square

On appelle la limite de $G(n, p)$ le *graphe aléatoire* et on le note R , pour “Rado’s graph“, en mémoire à Rado qui en a donné la première construction : pour $u, v \in \mathbb{N}$, $u \sim v$ si le u -ème chiffre du développement en base 2 de v est 1 ou si v -ème chiffre du développement en base 2 de u est 1. Pour deux ensembles disjoints d’entiers naturels X, Y , l’entier $\sum_{x \in X} 2^x + \sum_{y \in Y} 2^{y+z+1}$, avec $z = \max X$, est correctement relié à X et à Y .

Corollaire 1. $R \cong \bar{R}$.

Démonstration. Lorsqu’on remplace les arêtes par des “non-arêtes“ et vice versa dans la propriété (AR), on obtient exactement la même propriété. Donc \bar{R} satisfait (AR).

Une autre façon de le voir est de considérer $G(\mathbb{N}_0, 1/2)$. Il n’y a aucune différence entre une personne qui le construit en mettant une arête lorsqu’il reçoit face, et celle qui, au contraire, en inclut une lorsqu’il récolte pile. \square

Pour deux sommets distincts $u, v \in V(R)$, la propriété (AR) nous garantit un voisin commun w . La distance $d_R(u, v)$ est, donc, toujours inférieur à ou égale à 2. Bien sûr, R n’est pas complet puisque la même propriété (AR) atteste l’existence, pour tout sommet, d’un sommet qui ne lui est pas adjacent. Donc $\text{diam}(R) = 2$. En particulier, R est hautement connexe.

L’infini a été, pendant une certaine période de l’histoire des mathématiques, une source de paradoxes. D’autant plus que l’infini n’est pas unique ; mais l’infini dénombrable nous suffit. Ainsi, n’est-il pas, à première vue, bizarre qu’une partie soit « égale » au tout ? On sait depuis longtemps qu’il y a autant d’entiers naturels que d’entiers naturels paires, bien que ces derniers ne constituent qu’une partie des premiers. Si on partitionne \mathbb{N} en deux sous-ensembles, alors une des parties, au moins, est dénombrable. Toujours selon notre intuition de départ, si on considère R , puisque la partition de ses sommets en deux, disons A et B , laisse un des deux infini, alors le résultat des tirages sur les sommets de la partie infinie devrait déboucher

sur un graphe existentiellement clos. On dit qu'un graphe G vérifie le *principe du pigeonnier* si pour toute partition de ses sommets en deux sous-ensembles A et B , on a $G \upharpoonright A \cong G$ ou $G \upharpoonright B \cong G$. Une propriété moins forte est l'*indivisibilité* [40,49] où l'une des partitions admet R au lieu de lui être isomorphe.

Proposition 1. *R satisfait le principe du pigeonnier.*

Démonstration. Soient A, B une partition de $V(R)$. Supposons que A ne soit pas isomorphe à R , c'est-à-dire qu'il existe $S, T \subseteq A$ finis disjoints tels qu'aucun sommet de ne soit correctement relié à S et T . Soient $U, V \subseteq B$ finis disjoints et considérons $X = S \cup U, Y = T \cup V$. Comme R est existentiellement clos, il existe $u \in V(R)$ correctement relié aux sous-ensembles X et Y . Comme A et B forment une partition de $V(R)$, ou bien $u \in A$, ou bien $u \in B$. Or, si $u \in A$ alors u est correctement relié à S et à T , ce qui est une contradiction. Donc $u \in B$ et par conséquent, B est existentiellement clos. \square

Vu selon notre intuition probabiliste, le principe du pigeonnier semble être naturel pour R , voir lui est inhérent. Il se trouve, [31], que R soit le seul graphe dénombrable, à isomorphisme près, non trivial qui le respecte :

Proposition 2. *Les graphes dénombrables satisfaisant le principe du pigeonnier sont le graphe nul, le graphe complet et R .*

Démonstration. Pour les graphes complet et nul, c'est trivial. Supposons que G respecte le principe du pigeonnier et n'est ni complet, ni nul et encore moins R . Considérons la partition de $V(G)$ suivante : $A = \{u : N_G[u] = V(G)\}, B = V(G) \setminus A$. Donc $G \cong B$ puisque $G \upharpoonright A$ est complet. Aucun sommet de G n'est donc universel. De même, la partition $A = \{u : N_G[u] = \emptyset\}, B = V(G) \setminus A$ implique que G ne possède aucun sommet isolé.

Comme G ne satisfait pas (AR), il existe $X, Y \subseteq V(G)$ finis disjoints auxquels aucun sommet n'est correctement relié. De plus, parmi de tels sous-ensembles, prenons les de telle sorte que $m = |X| + |Y|$ soit minimal. Remarquons que $m > 1$

puisque $m = 1$ annonce l'existence d'un sommet universel. On peut alors partitionner $X \cup Y$ en deux sous-ensembles non vides A et B . Pour chaque sommet $u \in V(G) \setminus (A \cup B)$, ou bien u n'est pas correctement relié aux sommets de A , ou (non exclusif) bien u n'est pas correctement relié aux sommets de B . Si nous mettons les éléments de A et les sommets qui ne leur sont pas correctement reliés dans un même ensemble U et le reste dans V , alors U, V forment une partition de $V(G)$. De plus $G \upharpoonright U, G \upharpoonright V$ ne sont pas existentiellement clos et pour chacun d'eux la taille des sous-ensembles qui violent (AR) est plus petite que m . Donc, par minimalité de m , ni $G \upharpoonright U$, ni $G \upharpoonright V$ n'est isomorphe à G . \square

Il découle du principe du pigeonnier que retirer un nombre fini de sommets de R ne l'altère en rien. Un graphe G tel que $G \cong G - u$, pour tout $u \in V(G)$ est dit *inépuisable*; $G - u$ étant le graphe $G \upharpoonright (V(G) \setminus \{u\})$. Le graphe R est donc inépuisable. Plus encore, si nous fixons un sommet particulier $u \in V(R)$, on peut toujours partitionner les sommets de R en $N(u)$ et $V(R) \setminus N(u)$. Un de $R \upharpoonright N(u)$ et de $R \upharpoonright (V(R) \setminus N(u))$, au moins, est donc isomorphe à R . Il s'avère qu'ils le sont tous les deux, et par conséquent, enlever une copie de R dans R laisse un graphe isomorphe à R :

Proposition 3. *Soient $X, Y \subseteq R$ finis et disjoints. Le sous-graphe induit par*

$$Z = \{u : u \sim X, u \not\sim Y\}$$

est isomorphe à R .

Démonstration. Soient $U, V \subseteq Z$ finis disjoints. Le sommet $u \in V(R)$, qui est correctement relié aux sous-ensembles $X \cup U$ et $Y \cup V$, appartient à Z puisque $u \sim X$ et $u \not\sim Y$. En même temps, $u \sim U$ et $u \not\sim V$; donc, Z satisfait (AR). \square

Notre intuition, qui veut que lancer la pièce de monnaie une infinité de fois aboutit à la propriété (AR), suggère que le graphe induit par un nombre infini de sommets de R est existentiellement clos. On se pose la question : existe-t-il un sous-graphe induit infini de R qui ne lui est pas isomorphe ? Au premier abord, la réponse

est surprenante : non seulement il en existe, mais ils y sont tous. Imaginez, tirer à pile ou face une infinité de fois et n'obtenir que pile à tous les coups. Effectivement, posons $E_0 = \emptyset$ et supposons E_n est un graphe vide alors $E_{n+1} = R \upharpoonright V(E_n) \cup \{u\}$ avec u le sommet de R correctement relié à \emptyset et à $V(E_n)$ est un graphe vide. Donc $\lim_{n \rightarrow \infty} E_n$ est un sous-graphe induit infini nul de R . Le graphe R étant son propre complément, le graphe complet infini est aussi induit dans R . Au fond, la propriété (AR) permet d'induire n'importe quel graphe dénombrable dans R de la même façon. En particulier, R contient des noyaux bipartis de toute taille. Un graphe qui admet tous les graphes dénombrables est dit \aleph_0 -universel.

Proposition 4. *R admet tout graphe dénombrable.*

Démonstration. C'est l'étape "aller" seulement de la preuve du théorème 14. Soit G un graphe dénombrable et disons $V(G) = \{u_1, u_2, u_3, \dots\}$. Supposons que la fonction $f_n : \{u_1, u_2, \dots, u_n\} \rightarrow R$ est un isomorphisme de sous-graphes induits. Si nous posons $X = \{u_1, \dots, u_n\} \cap N_G(u_{n+1})$ et $Y = N_G(u_{n+1}) \setminus X$ alors on peut étendre f_n à u_{n+1} pour donner f_{n+1} avec $f(u_{n+1}) = v$, où v est le sommet de R correctement relié à $f_n(X)$ et à $f_n(Y)$. En commençant avec $f_0 = \emptyset$, la limite $f = \cup f_n$ est l'isomorphisme de sous-graphes induits voulu. \square

Même si R est \aleph_0 -universel, ce n'est pas tous les graphes dénombrables qui peuvent être un sous-graphe couvrant de R . Le graphe K_{\aleph_0} ne l'est pas par exemple. Un sous-graphe couvrant ne doit pas contenir d'arête là où le supergraphe n'en possède pas. Si nous examinons justement la propriété (AR), la partie " $u \sim X$ " spécifie où les arêtes se trouvent et la partie " $u \not\sim Y$ " stipule où elles ne doivent pas se situer. Si un graphe dénombrable respecte cette dernière partie, il est fort probable qu'il soit un graphe couvrant de R . C'est effectivement le cas :

Proposition 5. *Un graphe G sur \mathbb{N} est, à isomorphisme près, un sous-graphe couvrant de R si et seulement si pour tout sous-ensemble fini $Y \subseteq V(G)$, il existe $u \in V(G) \setminus Y$ tel que $v \not\sim Y$.*

Démonstration. On sait déjà que pour tout $Y \subseteq V(R)$, il existe $u \in V(R)$ qui n'est pas adjacent à Y . Comme G est un sous-graphe couvrant de R , $N_G(u) \cap Y \subseteq N_R(u) \cap Y = \emptyset$.

Considérons un homomorphisme $f : H \rightarrow R$ avec $H \subseteq G$ fini. Pour un sommet $u \in V(G) \setminus V(H)$, on peut étendre f à u en utilisant la propriété (AR) de R , avec $f(u) = v$ où v est le sommet de R correctement relié à $f(N_G(u) \cap V(H))$ et à $f(V(H) \setminus N_G(u))$. De même pour un sommet $v \in V(R) \setminus f(H)$, il existe un sommet $z \in V(G) \setminus H$ non adjacent à H . On peut alors étendre l'homomorphisme f à z en posant $f(z) = v$.

Maintenant, posons $f_0 = \emptyset$ et supposons que f_n est déjà définie. Si n est pair, nous prenons $m = \min \{i : i \notin \text{dom}(f_n)\}$ et f_{n+1} sera l'extension de f_n à m . Si n est impair, nous prenons $m = \min \{i : i \notin \text{im}(f_n)\}$ et f_{n+1} sera l'extension de f_n à m . La limite $f = \bigcup_{n \in \mathbb{N}} f_n$ est telle que $\text{dom}(f) = V(G)$ et $\text{im}(f) = V(R)$. De plus $f(E(G)) \subseteq E(R)$. \square

De même, tout graphe dénombrable tel que tout sous-ensemble fini de sommets possède un voisin commun est un supergraphe de R .

3.1.2 BA(m, t)

Rappelons-nous que le modèle de l'attachement préférentiel a un paramètre m qui est le degré du sommet nouvellement ajouté. Il joue certainement un rôle quantitatif puisque le nombre d'arêtes à l'étape t est mt . L'élément de surprise dans ce modèle est que ce paramètre intervient aussi à un niveau plus qualitatif. Pour $m = 1, 2$, l'issue du processus stochastique à l'infini est unique, à isomorphisme près. Lorsque $m \geq 3$, deux processus lancés en parallèle donneront, avec une probabilité non nulle, deux graphes non isomorphes.

Dans le chapitre précédent, on a considéré les graphes du modèle BA comme non-orientés, mais il est clair qu'ils possèdent une orientation naturelle, à savoir, les arêtes associées à un nouveau sommet sont sortantes. Cette distinction s'avère parfois utile pour décrire le type d'isomorphisme de la limite. Lorsque nous construi-

sons une chaîne de graphes $G_0 \subseteq G_1 \subseteq G_2 \subseteq \dots$ par le processus aléatoire BA, sa limite $G = \lim G_n$ n'est qu'une instance parmi les autres graphes limites possiblement obtenus ainsi. On considère alors $\text{BA}(m, \aleph_0)$ comme étant l'espace de probabilité donné par la distribution de ces graphes dénombrables.

Proposition 6. *Pour tous $G, H \in \text{BA}(1, \aleph_0)$, avec probabilité 1, $G \cong H$ comme graphe orienté.*

Démonstration. Disons $V(G) = \{u_0, u_1, \dots\}$, $V(H) = \{v_0, v_1, \dots\}$ où l'indice indique l'étape à laquelle le sommet a été introduit. Par construction, les sommets initiaux u_0, v_0 sont des puits, c'est-à-dire $d^+(u_0) = d^+(v_0) = 0$, et pour tous les autres sommets, le degré sortant est 1. L'équation (2.2) du chapitre précédent indique que pour tout $x \in V(G) \cup V(H)$, $d^-(x) = \infty$. Cette équation a été déduite d'une manière heuristique, mais un résultat similaire [67], soustrait quant à lui de manière plus formelle, lui est en accord.

Maintenant, pour un entier naturel d , considérons $G_d = \{u : d_G(u, u_0) \leq d\}$, où la distance est une distance orientée de u vers u_0 , et de façon similaire H_d . Il est clair que $G \upharpoonright G_0 \cong H \upharpoonright H_0$. Les sous-ensembles de sommets G_1 et H_1 , sont indépendants dans leurs graphes respectifs et sont infinis dénombrables. Par conséquent, on a $G \upharpoonright G_1 \cong H \upharpoonright H_1$ de façon naturelle. Supposons donc que f soit un isomorphisme entre $G \upharpoonright G_d$ et $H \upharpoonright H_d$. On peut étendre f aux sommets de $G_{d+1} \setminus G_d$ comme suit : pour chaque $u \in G_{d+1} \setminus G_d$, il existe $v \in G_d$ tel que $N_G^+(u) = \{v\}$. Le voisinage entrant $N_G^-(v)$ étant dénombrable et indépendant, le sommet u doit apparaître à l'indice i lorsqu'on l'énumère. Le voisinage entrant $N_H^-(f(v))$ étant lui aussi dénombrable et indépendant, il suffit de prendre comme image de u le i -ème élément de ce voisinage lorsqu'on l'énumère. Donc $G \upharpoonright G_{d+1} \cong H \upharpoonright H_{d+1}$. \square

Pour le cas $m = 2$, considérons les propriétés de graphe infini dénombrable suivantes :

1. il existe deux sommets u_0, u_1 tel que $d^+(u_0) = d^+(u_1) = 0$, et pour tout $v \neq u_0, u_1$, $d^+(v) = 2$;

2. pour tout pair de sommets u, v , il existe une infinité de sommets w tel que $N^+(w) = \{u, v\}$;
3. tous les chemins (chaîne orientée) du graphe sont finis, donc le graphe ne possède pas de circuit.

Proposition 7. *Deux graphes orientés infinis dénombrables G, H qui vérifient les propriétés (1) à (3) sont isomorphes.*

Démonstration. Nous allons recourir à la technique “aller-retour“ pour cela. Notons d’abord que l’on peut énumérer les sommets de G de manière à ce que les voisins sortants d’un sommet se trouvent lister avant lui. En effet, récursivement, on pose u_0 le sommet qui ne possède pas de voisin sortant. On choisit un sommet arbitraire $v \in V(G) \setminus \{u_0, u_1, \dots, u_n\}$ et on rajoute tous les sommets que l’on rencontre en suivant les arcs sortants. En un temps fini, on doit rencontrer un puit sinon on vient de rencontrer un chemin infini. Et ainsi de suite. On obtient alors l’énumération voulue de cette façon. On en fait de même pour les sommets de H pour parvenir à v_0, v_1, \dots .

Maintenant, nous commençons avec $f_0(u_0) = v_0, f_0(u_1) = v_1$ et supposons qu’il existe une fonction $f_n : G \upharpoonright U \rightarrow H \upharpoonright V$ qui soit un isomorphisme de sous-graphes induits où $U \subseteq V(G)$ et $V \subseteq V(H)$ tels que pour tous $u \in U$, respectivement $v \in V$, le voisinage $N_G^+(u) \subseteq U$, respectivement $N_H^+(v) \subseteq V$. C’est clairement vrai pour f_0 avec $U = \{u_0, u_1\}$ et $V = \{v_0, v_1\}$.

Si n est pair, soit u_j le sommet d’indice minimal dans $V(G) \setminus U$. De la façon dont on a énuméré les sommets de G , on sait que ses deux voisins sortants u_k, u_l sont tels que $k, l < j$ et donc $u_k, u_l \in U$. Un bon candidat pour l’image par extension de f_n à u_j est un sommet v de $V(H) \setminus V$ dont $N_H^+(v) = \{f_n(u_k), f_n(u_l)\}$. La propriété (2) de H garantit l’existence d’un tel sommet. Il y a deux points à souligner ici : d’abord, $U \cup \{u_j\}$ et $V \cup \{v\}$ respectent le fait que les voisinages sortants de leurs sommets restent à l’intérieur de U et V , on respecte donc notre hypothèse d’induction ; ensuite, par ce même fait, $N_G^-(u_j) \cap U = \emptyset$ et $N_H^-(v) \cap V = \emptyset$ et donc f_{n+1} construite comme étant l’extension de f_n à u_j est bien un isomorphisme de

sous-graphes induits.

Si n est impair, de la même façon, on prend v_j comme étant le sommet d'indice minimal dans $V(H) \setminus V$ et ses voisins sortants $v_k, v_l \in V$. Une pré-image de v_j serait donc un sommet u dans $V(G) \setminus U$ tel que $N_G^+(u) = \{f_n^{-1}(v_k), f_n^{-1}(v_l)\}$.

L'isomorphisme cherché est alors $f = \cup f_n$. \square

Proposition 8. *Soit $G \in \text{BA}(2, \aleph_0)$ alors pour tous $u, v \in V(G)$, avec probabilité 1, il existe une infinité de sommets $w \in V(G)$ tel que $N_G^+(w) = \{u, v\}$.*

Démonstration (esquisse). Intuitivement, par l'équation (2.2), la probabilité que le sommet x_t introduit à l'étape t possède comme voisin sortant u et v est

$$\frac{d(u)}{\sum_{x \in G_t} d(x)} \frac{d(u)}{\sum_{x \in G_t} d(x)} = \frac{4t}{t_u \cdot t_v \cdot 4t \cdot 4t} = \frac{1}{4t_u \cdot t_v \cdot t}.$$

Par conséquent la probabilité qu'aucun sommet antérieur à l'étape t_0 n'ait comme voisin sortant u et v est

$$\prod_{t=t_0}^{\infty} \left(1 - \frac{1}{4t_u t_v t}\right)$$

qui tend vers 0. Bien sûr l'hypothèse de l'indépendance dans le choix des arêtes est fautive et du coup on se trouve à ne plus travailler avec la distribution de l'attachement préférentiel, mais ce n'est que l'intuition derrière le fait. D'ailleurs une version plus complète de [67] devrait établir l'expression formelle de cette probabilité. \square

Ces deux propositions nous mènent directement à :

Théorème 15. *Il existe un graphe orienté infini dénombrable R_2^∞ tel qu'avec probabilité 1, tout graphe $G \in \text{BA}(2, \aleph_0)$ est isomorphe à R_2^∞ .*

Démonstration. Par construction, G possède deux puits et tous ses autres sommets, ceux introduits étape par étape, sont de degré sortant 2. Donc G satisfait (1). La proposition précédente atteste qu'il vérifie (2). Par construction encore, pour un sommet u de G introduit à l'étape t , le chemin le plus long que l'on peut prendre

à partir de u est celui qui mène vers un des puits et sa longueur est bornée par t .
Donc G remplit (3). \square

Comme pour le graphe aléatoire R , on peut procéder à une construction déterministe de R_2^∞ comme une limite : on pose G_0 le graphe orienté avec deux sommets u_0, u_1 de degré sortant 0, avec des sommets infinis dénombrables, chacun avec deux arêtes sortantes pointant vers u_0 et u_1 . Étant donné G_n , nous construisons G_{n+1} en lui adjoignant pour tout pair de sommets $u, v \in V(G_n)$ un nombre infini dénombrable de sommets, chacun avec deux arêtes sortantes pointant vers u et v . On pose alors

$$R_2^\infty = \bigcup_{i=0}^{\infty} G_i.$$

Pour $m \geq 3$, deux tirages indépendants dans $\text{BA}(m, \aleph_0)$ ont une probabilité non nulle de ne pas être isomorphes. L'idée est la suivante : la probabilité que $G \in \text{BA}(m, \aleph_0)$ admette N copies disjointes de K_4 appartient à l'intervalle $(0, 1)$, où $N > N_0$ et N_0 est le nombre espéré de copies distinctes de K_4 dans G . Or la propriété $P = "G$ admet N copies distinctes de $K_4"$ est une propriété de graphe. Par conséquent, certains graphes dans $\text{BA}(m, \aleph_0)$ vérifient P et d'autres ne la satisfont pas. On peut consulter [67] pour plus de détails.

3.1.3 BJ(p, ϕ, H)

Le recours à l'infini dans ces constructions stochastiques a permis, dans la plupart des cas, de dégager des propriétés déterministes respectées par leurs limites ; et cela conformément à notre intuition. Pour la copie évolutive, à mesure que l'on va à l'infini, la probabilité qu'un sommet soit copié tend vers 1. De plus, on peut s'arranger pour trouver une copie qui respecte certaine adjacence que l'on s'est fixée. Plus formellement, considérons la

Propriété (B) : Un graphe G a la propriété (B) si pour tout $u \in V(G)$, pour tout $X \subseteq N_G(u)$ fini et tout $Y \subseteq V(G) \setminus X$, il existe $v \neq u \in V(G)$ correctement relié aux sous-ensembles X et Y .

Posons $\text{BJ}_\infty(p, \phi, H)$ l'espace de probabilité donné par la distribution des graphes limites obtenus par le processus aléatoire $\text{BJ}(p, \phi, H)$.

Théorème 16. *Pour tout $G \in \text{BJ}_\infty(p, \phi, H)$, $\Pr[G \models (B)] = 1$.*

Démonstration. La preuve fonctionne de la même façon que celle du théorème 13. Soient donc $u \in V(G)$, $X \subseteq N_G(u)$ fini et $Y \subseteq V(G) \setminus X$ fini. Posons m comme étant le plus petit entier tel que u, X, Y soient contenus dans $V(G_m)$. Disons $|G_m| = m'$, $|X| = i$, $|N_G(u) \cap Y| = j$ et $k_t = \min \{\phi(t), |Y|\}$. Si $t \geq 1$, posons u_{m+t} le nouveau sommet ajouté pour construire G_{m+t} à partir de G_{m+t-1} . La probabilité de choisir u comme source de la copie est

$$\frac{1}{m' + t};$$

la probabilité de ne pas choisir un des sommets de Y comme lien aléatoire est

$$1 - \frac{k_{m'+t}}{m' + t}.$$

La probabilité que u_{m+t} soit correctement joint à X et à Y est alors

$$\frac{1}{m' + t} \left(1 - \frac{k_{m'+t}}{m' + t}\right) p^i (1 - p)^j.$$

Par conséquent, la probabilité qu'aucun sommet de G_{m+t} soit correctement relié aux sous-ensembles X et Y est au plus

$$\prod_{l=1}^t \left(1 - \frac{1}{m' + l} \left(1 - \frac{k_{m'+l}}{m' + l}\right) p^i (1 - p)^j\right),$$

qui tend vers 0 lorsque $t \rightarrow \infty$.

Puisque l'union dénombrable d'ensembles de mesure nulle est de mesure nulle, on conclut que $\Pr[G \not\models (B)] = 0$. \square

Le processus aléatoire de la copie évolutive génère donc des graphes ayant la propriété (B). Considérons le processus menant à une chaîne de graphes suivant :

- on se donne un graphe fini G_0 ,

– pour $n \geq 1$, on construit G_{n+1} à partir de G_n par une application de (P1) suivie de l'application d'un nombre fini de (P2) où

(P1) pour chaque $u \in V(G_n)$ et chaque $X \subseteq N_{G_n}(u)$ fini, on crée un nouveau sommet v_X avec $N_{G_{n+1}}(v_X) = X$.

(P2) on choisit un sous-ensemble fini $Y \subseteq V(G_n)$, on ajoute un nouveau sommet w_Y de sorte que $N_{G_{n+1}}(w_Y) = Y$.

La chaîne de graphe obtenue ainsi sera appelée une (B) -chaîne. Une (B) -chaîne pour laquelle aucune application n'est intervenue (P2) est dite *pure* sinon *composite*. Une (B) -chaîne satisfait évidemment la propriété (B).

Proposition 9. *Soient un graphe infini dénombrable G et $H \leq G$ fini. Si G possède la propriété (B) alors il existe une (B) -chaîne $\{G_n : n \in \mathbb{N}\}$ avec $G_0 = H$ telle que $G = \lim_{n \rightarrow \infty} G_n$.*

Démonstration. Soit u_0, u_1, u_2, \dots une énumération des sommets de G . Sans perte de généralité, on peut supposer que $u_0 \in V(H)$. On va construire une chaîne de graphes finis telle que $u_n \in V(G_n)$ et $G_n \leq G$ pour tout $n \in \mathbb{N}$.

Posons $G_0 = H$ alors $u_0 \in V(G_0)$ et $G_0 \leq G$. Supposons alors que G_n est fini tel que $u_n \in V(G_n)$ et $G_n \leq G$. Pour chaque $u \in V(G_n)$, et chaque $X \subseteq N_{G_n}(u)$, posons $Y = V(G_n) \setminus X$. Comme G a la propriété (B), il existe $v_X \in V(G) \setminus V(G_n)$ correctement relié aux sous-ensembles X et Y . Soit $Z = \{v_X : X \subseteq N_{G_n}(u)\}$ et définissons $G_{n+1} = G \upharpoonright (V(G_n) \cup Z \cup \{u_{n+1}\})$. On vérifie facilement que les G_n forment une chaîne de graphes. Le fait que $G_n \leq G$ pour tout $n \in \mathbb{N}$ implique $G = \lim_{n \rightarrow \infty} G_n$. Remarquons que rajoutés les sommets de Z dans G_n revient à appliquer le processus (P1). Si $u_{n+1} \notin Z$, son rajout est équivalent à utiliser (P2). \square

On a donc comme corollaire :

Corollaire 2. *Un graphe infini dénombrable G a la propriété (B) si et seulement si G est limite d'une (B) -chaîne.*

Maintenant, pour répondre à une question qui a dû nous titiller depuis longtemps, considérons les deux (B) -chaînes pures suivantes : $\{G_n : n \in \mathbb{N}, G_0 = E_4\}$

et $\{H_n : n \in \mathbb{N}, H_0 = K_4\}$. Les limites de ces deux chaînes ne sont pas isomorphes. Il s'avère, voir [24], qu'il y a exactement \aleph_0 limites de (B)-chaînes pures non isomorphes, et 2^{\aleph_0} (B)-graphes composites non isomorphes. La propriété (B) n'est donc pas \aleph_0 -catégorique.

La propriété (AR) qui définit R est plus forte que (B) puisque (AR) implique (B). La question que l'on se pose est : peut-on générer R par la copie évolutive ? Ou quelle restriction s'impose dans la construction d'une (B)-chaîne pour obtenir (R) comme limite ? Pour pouvoir y répondre, considérons la modification de la propriété (B) suivante : pour $n \in \mathbb{N}$ on définit

Propriété (B, n) : Un graphe G a la propriété (B, n) si pour tout $u \in V(G)$, pour tout $X \subseteq N_G(u)$ fini, pour tout $Y \subseteq V(G) \setminus X$ et tout $Z \subseteq V(G) \setminus (N_G(u) \cup Y)$ avec $|Z| \leq n$, il existe $v \neq u \in V(G)$ correctement relié aux sous-ensembles $X \cup Z$ et Y .

Autrement dit, le nombre de liens aléatoires est borné par n , fixé d'avance. Remarquons que (B)=(B,0).

Théorème 17. *Fixons un entier $n > 0$. Si G satisfait (B, n) alors $G \cong R$.*

Démonstration. Par le théorème 14, il suffit de montrer que G possède la propriété (AR). Soit donc $X, Y \subseteq V(G)$ finis disjoints. Étant donné que X soit fini, soit X_1, X_2, \dots, X_r une partition de X en parts de taille au plus n . Prenons n'importe quel sommet $u \notin X \cup Y$. Par application de (B, n) sur $U = N_G(u) \cap X_1$, $V = Y$ et $W = X_1 \setminus N_G(u)$, il existe un sommet $v_1 \notin U \cup V$ tel que $X_1 \subseteq N(v_1)$. Par la propriété (B, n) avec $U = X_1, V = \emptyset, W = X_2$, il existe $v_2 \notin X_1$ tel que $X_1 \cup X_2 \subseteq N(v_2)$. Par la propriété (B, n), $U = X_1 \cup X_2, V = \emptyset, W = X_3$, il existe $v_3 \notin X_1 \cup X_2$ tel que $X_1 \cup X_2 \cup X_3 \subseteq N(v_3)$. Et ainsi de suite jusqu'au sommet v_r tel que $X \subseteq N_G(v_r)$. Une dernière application de (B, n) avec $U = X, V = Y$ et $W = \emptyset$, on obtient un sommet correctement relié à X et à Y . \square

Si nous nous tournons vers les noyaux bipartites. On a dit que c'est à la base des phénomènes des communautés sur le Web. Le théorème suivant confirme l'existence de ces noyaux dans le modèle de la copie évolutive.

Théorème 18. *Si G non nul satisfait (B) alors G admet K_{\aleph_0, \aleph_0} .*

Démonstration. Soit U_1 , un sous-graphe de G constitué d'une seule arête. Il admet $K_{1,1}$. Maintenant, supposons que U_n est un sous-graphe induit de G isomorphe à $K_{n,n}$. Soient $A = \{x_i : 1 \leq i \leq n\}$ et $B = \{y_i : 1 \leq i \leq n\}$ deux ensembles indépendants tels que $V(U_n) = A \cup B$. Par la propriété (B), il existe un sommet $x_{n+1} \in V(G) \setminus V(U_n)$ correctement relié à B et à A . Puisque $A \cup \{x_{n+1}\} \subseteq N_G(y_1)$, par la propriété (B), il existe un sommet $y_{n+1} \in V(U_n) \cup \{x_{n+1}\}$ correctement relié à $A \cup \{x_{n+1}\}$ et à B . Dès lors, $U_{n+1} = G \upharpoonright (V(U_n) \cup \{x_{n+1}, y_{n+1}\})$ est une clique bipartite contenant U_n . La limite de la chaîne $U_1 \leq U_2 \leq \dots$ est isomorphe à K_{\aleph_0, \aleph_0} . \square

3.2 Jeu de poursuite

Maintenant nous allons nous intéresser au jeu du police-voleur sur les limites des modèles du Web. Poursuivre un intrus sur un graphe infini ne veut pas forcément signifier que l'on perd d'avance. En fait, il y a une infinité de graphes infinis où un nombre fini de policiers suffit pour capturer le voleur. Le graphe complet dénombrable en est un exemple trivial. On peut aussi penser à un bouquet infini de cycles de longueur fini. On a déjà mentionné le cas d'un arbre infini de profondeur finie. Toutefois, notre intuition semble indiquer que la proportion de tels graphes doit être négligeable.

Théorème 19. *Pour tout entier k , R n'est pas k -policiers gagnant.*

Démonstration. Les k policiers commencent par choisir leurs sommets de départ, disons $U_0 = \{u_{01}, u_{02}, \dots, u_{0k}\}$. Par la propriété (AR), il existe un sommet $v_0 \neq u_i$ correctement relié aux sous-ensembles \emptyset et U_0 . Le voleur choisit v_0 alors comme sommet de départ. Ainsi, au premier tour des policiers, le voleur ne sera pas capturé. Supposons maintenant que les policiers sont sur U_n et le voleur sur $v_{n-1} \notin U_n$ et c'est à son tour de jouer. Par la propriété (AR), il existe un sommet $v_n \in V(G) \setminus (U_n \cup \{v_{n-1}\})$ correctement relié à $\{v_{n-1}\}$ et à U_n . En se déplaçant vers

v_n , le voleur est garanti de ne pas être coffré au prochain coup des policiers. Par induction donc, le voleur peut s'enfuir indéfiniment. \square

Le graphe R est de diamètre 2 et malgré cela, un nombre arbitraire fini de policiers ne permettra pas de capturer le voleur. Rappelons-nous que R est la limite de $G(n, p)$. Et ce dernier est un espace de probabilité. On peut interpréter le théorème précédent comme suit : presque tous les graphes sont voleurs gagnants étant donné $k < \aleph_0$ policiers lancés à sa poursuite. L'utilisation de mot "presque" fait référence au fait que la probabilité tend vers 1 lorsque l'ordre du graphe grossit.

Tout nombre fini est insignifiant devant l'infini. Ainsi, un nombre fini de policiers est négligeable devant $|R|$. Une question intéressante qui s'impose est : quelle proportion des sommets du graphe les policiers devraient-ils occuper pour gagner dans R . La propriété (AR) n'indique rien à ce propos ; mais rappelons-nous qu'un graphe de $G(\aleph_0, 1/2)$ est isomorphe à R . Pour un entier non nul d , si nous posons un policier sur les sommets $d \cdot [n] = \{dx : 0 \leq x \leq n\}$, la probabilité que le sommet du voleur soit adjacent à un des policier est $\frac{2^n - 1}{2^n}$, qui tend vers 1 lorsque $n \rightarrow \infty$. Donc avec une densité de présence $1/d$, les policiers sont assurés de gagner en un coup. Or d est arbitraire, c'est-à-dire la densité des policiers nécessaire pour maintenir l'ordre est 0.

Le voleur a beaucoup d'options d'évasion devant un nombre fini de policiers lorsque tous les sommets du graphe lui sont accessibles. Qu'arrive-t-il lorsqu'il est assiégé sur une partie H de R seulement ? Si $|H|$ est fini, la propriété (AR) nous dit qu'il existe un sommet w adjacent à tous les sommets de $N_H(v) \cup \{u, v\}$ où v , respectivement u , est le sommet actuel du voleur, respectivement du policier, et c'est au tour de ce dernier de jouer. En allant sur w , le policier a cerné le voleur. Restreindre le voleur à un sous-graphe fini ne nécessite qu'un seul policier pour gagner. On peut même considérer des sous-graphes induits infinis de R . En effet, on sait que R admet des sous-graphes infinis localement fini, c'est-à-dire, dont le voisinage induit de chaque sommet est fini. Soit donc $H \leq R$ un graphe localement fini. Le voleur est restreint aux sommets de H . Puisque $N_H(v)$ est fini pour tout

$v \in V(H)$, le seul policier nécessaire pour gagner, qui est maintenant sur un sommet u de R , utilise la propriété (AR) pour trouver un sommet w correctement relié à $N_H(v) \cup \{u\}$ et à \emptyset . Comme $w \in N_R(u)$ couvre v , le voleur est acculé.

Considérons maintenant une chaîne de graphes $\{G_n : n \in \mathbb{N}\}$ obtenue aléatoirement par le processus $BJ(p, \vec{0}, K_1)$, où $\vec{0}$ est la fonction $\phi(t) = 0$ pour tout t . Si u_n est le nouveau sommet rajouté dans G_{n-1} pour obtenir G_n , on sait que $N_{G_n}(u_n) \subseteq N_{G_n}(u)$ pour un certain $u \in V(G_{n-1})$. En effet, on ne se permet pas des liens aléatoire pour u_n mais seulement des liens copies. Comme $G_0 = K_1$ est policier gagnant, il s'en suit par induction que G_n est policier gagnant pour tout $n \in \mathbb{N}$. Peut-on alors conclure que sa limite est aussi policier gagnant ?

Théorème 20. *Si G dénombrable possède la propriété (B), alors $\text{sn}(G) > k$ pour tout $k \in \mathbb{N}$.*

Démonstration. Pour toute position initiale des policiers $Y_0 = \{u_{00}, u_{01}, \dots, u_{0k}\}$, il existe un sommet v_0 correctement relié à \emptyset et à Y_0 . Le voleur choisit donc v_0 comme sommet de départ. Au premier tour des policiers, le voleur n'a pas à s'inquiéter. Maintenant supposons que les policiers viennent de se placer aux sommets $Y_n = \{u_{n0}, u_{n1}, \dots, u_{nk}\}$ et le voleur se trouve sur $v_{n-1} \notin Y_n$. Il existe un sommet v_n correctement relié à $\{v_{n-1}\}$ et à Y_n . En effet, $\{v_{n-1}\} \subseteq N_G(x)$ pour $x \in N_G(v_{n-1})$. S'il n'existe pas de tel x , alors le voleur se trouve déjà à un endroit sécuritaire. En allant sur v_n , le voleur court toujours. \square

On sait que $G = \lim_{n \rightarrow \infty} G_n$ possède la propriété (B) et donc G n'est pas policier gagnant. Intuitivement, on aurait répondu de façon affirmative à la précédente question puisque tous les graphes de la chaîne sont policiers gagnants. Preuve que l'infini est à manipuler avec précaution.

CHAPITRE 4

GRAPHES HOMOGÈNES

Une structure \mathcal{S} est dite *homogène* si à chaque fois que deux sous-structures finies $\mathcal{S}_1, \mathcal{S}_2$ de \mathcal{S} sont isomorphes, il existe un automorphisme de \mathcal{S} qui envoie \mathcal{S}_1 à \mathcal{S}_2 . Une notion plus forte est celle d'une structure *ultrahomogène* dans laquelle, l'isomorphisme entre les sous-structures s'étend à un automorphisme de la structure au complet. Un exemple typique de structure ultrahomogène est un espace vectoriel, principalement dû au fait qu'un ensemble de vecteurs indépendants peut être étendu pour former une base. On peut remonter au moins à Cantor pour l'émergence de cette notion ; en 1895 il a fait l'observation que (\mathbb{Q}, \leq) est ultrahomogène. Pour tous rationnels tels que $x_1 < x_2 < \dots < x_n$ et $y_1 < y_2 < \dots < y_n$, il existe une permutation $\pi : \mathbb{Q} \rightarrow \mathbb{Q}$ préservant l'ordre telle que $\pi(x_i) = y_i$ pour $i = 1, 2, \dots, n$. En effet,

$$\pi(q) = \begin{cases} q + y_1 - x_1 & \text{si } q < x_1 \\ y_i + \frac{(q-x_i)(y_{i+1}-y_i)}{x_{i+1}-x_i} & \text{si } y_i \leq q \leq y_{i+1} \\ q + y_n - x_n & \text{si } q > x_n \end{cases}$$

est un tel automorphisme. Notons qu'il est important que les sous-structures soient de cardinalité finie. En effet, si nous prenons $\mathcal{S}_1 = \{-1 + \frac{1}{2^n}, 1 - \frac{1}{2^n} : n \in \mathbb{N}\}$ et $\mathcal{S}_2 = \mathbb{Z}$, il est clair qu'elles sont isomorphes ; mais aucun automorphisme de (\mathbb{Q}, \leq) ne peut envoyer \mathcal{S}_1 vers \mathcal{S}_2 puisque la première est bornée (par -1 et 1) tandis que la seconde non.

Le logicien Fraïssé a ensuite pris (\mathbb{Q}, \leq) comme prototype pour étudier les structures ultrahomogènes ; mais il a plutôt employé le terme "homogène" pour y référer. Pour des raisons historiques peut-être, certains auteurs [31, 40, 49, 58] ont gardé cet usage. Bien que ce chapitre ne traite que des graphes ultrahomogènes, nous avons maintenu le titre ainsi pour la même raison. La caractérisation des

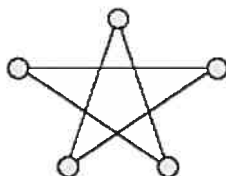


FIG. 4.1 – Un graphe ultrahomogène

graphes non-orientés ultrahomogènes a été donnée par Sheehan [84], Gardiner [52] pour les graphes finis et par Lachlan et Woodrow [73] pour les graphes infinis dénombrables.

Plus formellement, un graphe G est dit *homogène* si pour tous $H, K \subseteq V(G)$ finis tels que $G \upharpoonright H \cong G \upharpoonright K$, il existe un automorphisme $f : G \rightarrow G$ tel que $f(H) = K$. Il est dit *ultrahomogène* si pour tous $H, K \subseteq V(G)$ finis et tout isomorphisme $g : G \upharpoonright H \rightarrow G \upharpoonright K$, il existe un automorphisme $f : G \rightarrow G$ tel que $f|_H = g$. Un graphe ultrahomogène est donc homogène. La propriété d'homogénéité est une propriété de graphe, c'est à dire, tout graphe isomorphe à un graphe homogène est homogène.

Proposition 10. G est ultrahomogène si et seulement si \overline{G} est ultrahomogène.

Démonstration. Soient $H, K \subseteq V(G)$ finis tels que $\overline{g} : \overline{G} \upharpoonright H \rightarrow \overline{G} \upharpoonright K$ est un isomorphisme. Soit la fonction $g : H \rightarrow K$ telle que $g(u) = \overline{g}(u)$. Pour une arête $(u, v) \in E(G) \cap H^2$, par définition du complément d'un graphe, on a que $(u, v) \notin E(\overline{G}) \cap H^2$ et donc $(\overline{g}(u), \overline{g}(v)) = (g(u), g(v)) \notin E(\overline{G}) \cap K^2$ parce que \overline{g} est un isomorphisme; ce qui implique $(g(u), g(v)) \in E(G) \cap K^2$. De la même façon, une non-arête dans $G \upharpoonright H$ est préservée par g dans $G \upharpoonright K$ et par conséquent, $g : G \upharpoonright H \rightarrow G \upharpoonright K$ est un isomorphisme. Il existe alors un automorphisme $f : G \rightarrow G$ tel que $f|_H = g$. Or $g = \overline{g}$ donc $\overline{g} = f|_H$. La fonction $\overline{f} : V(G) \rightarrow V(G)$ avec $\overline{f}(u) = f(u)$ est automorphisme de \overline{G} puisque pour toute arête $(u, v) \in E(\overline{G})$,

par complémentarité $(u, v) \notin E(G)$ et $(\bar{f}(u), \bar{f}(v)) = (f(u), f(v)) \notin E(G)$, c'est-à-dire $(\bar{f}(u), \bar{f}(v)) \in E(\bar{G})$. Le même principe fonctionne pour le respect des non-arêtes de \bar{G} . \square

L'ultrahomogénéité est une propriété autant globale que locale. En effet,

Proposition 11. *Si G est ultrahomogène et $u \in V(G)$ alors $G \upharpoonright N_G(u)$ est ultrahomogène.*

Démonstration. Soient donc $H, K \subseteq N_G(u)$ finis et $h : N_G(u) \upharpoonright H \rightarrow N_G(u) \upharpoonright K$ un isomorphisme de sous-graphes induits. Posons $H' = H \cup \{u\}$ et $K' = K \cup \{u\}$. Nous avons un isomorphisme naturel $h' : N_G(u) \upharpoonright H' \rightarrow N_G(u) \upharpoonright K'$ avec $h'(u) = u$ et $h'|_H = h$. Il existe alors un automorphisme f de G qui étend h' . La restriction f à $N_G(u)$ est alors un automorphisme puisque $f(N_G(u)) = N_G(u)$, dû au fait que $f(u) = u$. \square

Étant donné donc un graphe ultrahomogène G , on peut construire une séquence de graphes ultrahomogènes imbriqués par $G \geq G \upharpoonright N_G(v_0) \geq G \upharpoonright N_G(v_1) \geq \dots$ où $v_i \in N_G(v_{i-1})$. Si G est fini, cette séquence est finie et le plus petit élément de la séquence doit être $t \cdot K_1$.

4.1 Cas fini

Le graphe complet K_n et par conséquent le graphe vide E_n sont trivialement ultrahomogènes. De même une union disjointe de copies de K_n est ultrahomogène et par conséquent, son complément qui est multi-parti complet est ultrahomogène. Un graphe G est dit *distance-transitive* si pour tous $u, v, w, z \in V(G)$ tels que $d_G(u, v) = d_G(w, z)$, il existe un automorphisme g de G tel que $g(u) = w$ et $g(v) = z$. Autrement dit, $AUT(G)$ agit transitivement sur les classes de distance. Si nous fixons un sommet u , alors tous les sommets à distance d de u ont leurs voisinages isomorphes.

4.1.1 Caractérisation

Sheehan [84] a donné la caractérisation, quoiqu'incomplète, des graphes finis ultrahomogènes suivante :

Théorème 21. *Un graphe ultrahomogène fini G satisfait une des assertions suivantes :*

- (i) $G \cong t \cdot K_{1+k}$, $t \geq 1$;
- (ii) $G \cong C_5$;
- (iii) $G \cong K_{k,k}$;
- (iv) G est distance-transitive, connexe, de diamètre 2 et de circonférence 3.

Si ce n'est que la partie (iv), ce théorème caractérise explicitement les graphes ultrahomogènes finis. Ce théorème a quelque peu créé une confusion dans la littérature lorsqu'on a cru qu'il s'appliquait aux graphes homogènes. Toutefois, ce malentendu n'en était pas finalement un puisque Gardiner [52] a montré que c'est bien le cas. La différence entre homogène et ultrahomogène réside dans la partie (iv). Ainsi,

Théorème 22. *Un graphe ultrahomogène fini G est isomorphe à un des graphes suivants :*

- (i) $t \cdot K_{1+k}$, $t \geq 1$;
- (ii) C_5 ;
- (iii) $K_{t,k}$, $t \geq 2$ où $K_{t,k} = \underbrace{K_{k,k,\dots,k}}_t$;
- (iv) $L(K_{3,3})$, voir figure 1.2.

4.1.2 Jeu de poursuite

En inspectant un à un les graphes du théorème 22, on a

Proposition 12. *Si G est un graphe ultrahomogène connexe fini alors $\text{sn}(G) \leq 2$.*

Démonstration. 1. Le seul graphe connexe de (i) est quand on a une seule copie de K_{1+k} et c'est policier gagnant.

2. Le cycle C_5 est 2-policiers gagnant.
3. Un policier ne peut pas gagner sur un graphe multiparti $K_{t,2+r}$, car le voleur n'a qu'à se placer à chaque fois dans le même ensemble indépendant qu'occupe le policier. Deux policiers suffisent autrement. Un policier gagne sur $K_{t,1}$.
4. $L(K_{3,3})$ est 2-policiers gagnant.

□

Pour les graphes ultrahomogènes non connexes, il faut autant de policiers que de composantes connexes.

4.2 Cas infini

Le graphe vide E_{\aleph_0} étant ultrahomogène, son complément K_{\aleph_0} l'est aussi. On a aussi vu que des copies disjointes de K_n sont ultrahomogènes. Il suffit d'en faire une infinité dénombrable de copies pour obtenir un graphe ultrahomogène infini. On a déjà rencontré la situation où on a étendu un isomorphisme à tout le graphe infini dénombrable à partir d'un isomorphisme de sous-graphes finis ; et c'était lorsque nous avons montré que deux graphes dénombrables existentiellement clos sont isomorphes. Ainsi,

Proposition 13. *R est ultrahomogène.*

Démonstration. Soient $H, K \subseteq V(R)$ finis et $h : R \upharpoonright H \rightarrow R \upharpoonright K$ un isomorphisme. Posons $f_0 = h$. Soit $x_0 = \min \{x \in R : x \notin \text{dom}(f_0)\}$, on peut étendre f_0 à x_0 en lui associant comme image le sommet correctement relié à $f_0(\text{dom}(f_0) \cap N_R(x_0))$ et à $f_0(\text{dom}(f_0) \setminus N_R(x_0))$. Notons f_1 ce nouvel isomorphisme. Soit alors $x_1 = \min \{x \in V(R) : x \notin \text{im}(f_1)\}$. On peut étendre l'image de f_1 pour inclure x_1 en lui associant comme pré-image le sommet correctement relié à $f_1^{-1}(\text{im}(f_1) \cap N_R(x_1))$ et à $f_1^{-1}(\text{im}(f_1) \setminus N_R(x_1))$. Et ainsi de suite. L'automorphisme $f = \cup f_n$ de R est tel que $f \upharpoonright H = h$. □

On peut tirer de cette technique « aller-retour », une propriété qui généralise (AR) et en même temps caractérise les graphes ultrahomogènes.

Théorème 23. *G dénombrable est ultrahomogène si et seulement si pour tout $U \subseteq V(G)$ fini, pour tout isomorphisme h de G avec $\text{dom}(h) = U$ et tout $x \notin V(G) \setminus \text{dom}(h)$, il existe un isomorphisme f de G tel que $\text{dom}(f) = U \cup \{x\}$ et $h = f|_U$.*

Démonstration. La direction « \Rightarrow » est évidente.

Remarquons d'abord que la condition permet d'étendre l'isomorphisme h à un sommet qui n'est pas dans son image en considérant h^{-1} . Maintenant supposons que cette condition tienne. Soient $H, K \subseteq V(G)$ finis et $h : G \upharpoonright H \rightarrow G \upharpoonright K$ un isomorphisme. Posons $f_0 = h$ et soit $x_0 = \min \{x \in V(G) : x \notin \text{dom}(f_0)\}$, il existe donc f_1 qui étend f_0 à x_0 . Maintenant, prenons $x_1 = \min \{x \in V(G) : x \notin \text{im}(f_1)\}$. Il existe donc f_2 qui étend f_1 de sorte que $x_1 \in \text{im}(f_2)$. Ainsi de suite. L'automorphisme cherché est $\cup f_n$. \square

4.2.1 Amalgamation

Nous avons défini la limite de graphes comme étant l'union dénombrable d'une chaîne. A ce moment là, chaque terme de la chaîne est un sous-graphe induit de la limite. On peut généraliser la notion de limite pour considérer les graphes qui sont isomorphes à un sous-graphe induit de la limite. Soit G , un graphe. On définit l'âge de G comme étant la classe des graphes finis admis, c'est-à-dire isomorphes à un sous-graphe induit, par G et que l'on note $\Sigma(G)$. Par exemple, $\Sigma((\mathbb{Q}, \leq))$ est la classe de tous les ordres linéaires finis.

Fraïssé a identifié la propriété qui lie l'âge d'une structure et l'ultrahomogénéité. Une classe de graphes finis \mathcal{C} possède la propriété d'*amalgamation* si pour tous $U, V_1, V_2 \in \mathcal{C}$, pour tous $f_1 : U \rightarrow V_1, f_2 : U \rightarrow V_2$ isomorphismes de sous-graphes induits, il existe $W \in \mathcal{C}$ et $g_1 : V_1 \rightarrow W, g_2 : V_2 \rightarrow W$ isomorphismes de sous-graphes induits tels que $g_1 \circ f_1 = g_2 \circ f_2$. Intuitivement, cela veut dire que deux graphes de \mathcal{C} qui contiennent une même copie d'un autre peuvent-être collés ensemble de telle sorte qu'ils coïncident à cette partie commune. Le graphe résultant appartient encore à \mathcal{C} .

Proposition 14. *Si G infini dénombrable est ultrahomogène alors $\Sigma(G)$ possède la propriété d'amalgamation.*

Démonstration. Soit G infini dénombrable et ultrahomogène. Prenons $U, V_1, V_2 \in \Sigma(G)$ et sans perdre de généralité, supposons $U \subseteq V_1$, $f_2 : U \rightarrow V_2$ isomorphisme de sous-graphes induits. Puisque $V_2 \in \Sigma(G)$, il existe un isomorphisme de sous-graphes induits g qui emmène V_2 dans G . Soit alors g' l'automorphisme de G qui étend g . Prenons $g_1 = g'|_{V_1}$ et $g_2 =$ l'identité sur V_2 . Il suffit de poser maintenant $W = V_2 \cup g_1(V_1)$. On a que $g_2 \circ f_2 = g' \circ \subseteq$. \square

Comme R est ultrahomogène infini dénombrable, son âge possède donc la propriété d'amalgamation. On sait que R admet tous les graphes finis. Par conséquent, $\Sigma(R)$ est la classe de tous les graphes finis.

Réciproquement, nous avons le résultat suivant :

Proposition 15. *Soit \mathcal{C} une classe de graphes finis close sous l'isomorphisme et la prise de sous-graphes induits. Supposons que \mathcal{C} possède la propriété d'amalgamation et qu'elle contienne des graphes arbitrairement grands, alors il existe un graphe ultrahomogène infini dénombrable G tel que $\Sigma(G) = \mathcal{C}$.*

Démonstration. On peut consulter [49, 90] pour la preuve. \square

Les graphes ultrahomogènes et les classes des graphes finis sont intimement liés comme le montre le théorème d'unicité suivant :

Théorème 24. *Soient G, H deux graphes ultrahomogènes infinis dénombrables. Alors $G \cong H$ si et seulement si $\Sigma(G) = \Sigma(H)$.*

Démonstration. Soit $f : G \rightarrow H$ un isomorphisme et prenons $K \in \Sigma(G)$. Il existe un isomorphisme de sous-graphes induits $g : K \rightarrow G$. Ainsi, $f \circ g$ plonge K dans H , c'est-à-dire $K \in \Sigma(H)$ et donc $\Sigma(G) \subseteq \Sigma(H)$. De la même façon, $\Sigma(H) \subseteq \Sigma(G)$.

Réciproquement, supposons que $\Sigma(G) = \Sigma(H)$. Soient u_0, u_1, \dots , respectivement v_0, v_1, \dots , l'énumération du sommet de G , respectivement H . Comme G et H ont le même âge, H admet $G \upharpoonright \{u_0\}$. Cela donne lieu à un isomorphisme local

f_0 entre G et H dont le domaine est $\{u_0\}$. Par le théorème 23, on peut étendre f_0 à v_0 pour obtenir g_0 tel que $\text{dom}(g_0) = \{u_0\}$ et $\text{im}(g_0) = \{v_0\}$. Ainsi de suite et l'isomorphisme cherché est l'union des f_i . \square

Lorsqu'une classe \mathcal{C} possède la propriété d'amalgamation, alors l'unique graphe ultrahomogène infini G tel que $\Sigma(G) = \mathcal{C}$ est appelé la *limite de Fraïssé* de \mathcal{C} . La graphe de Rado R est donc la limite de Fraïssé de tous les graphes finis, (\mathbb{Q}, \leq) , la limite de Fraïssé de tous les ordres linéaires finis, et K_{\aleph_0} est la limite de Fraïssé de toutes les cliques finies.

Jusqu'ici, nous avons rencontré $t \cdot K_n$, $K_{t;n}$ avec $n + t = \aleph_0$, et R comme graphes ultrahomogènes dénombrables. Pour $n \geq 2$ fixé, considérons la classe \mathcal{C}_n des graphes finis qui n'admettent pas K_n . Clairement, \mathcal{C}_n est clos sous l'isomorphisme. Un sous-graphe induit d'un graphe de cette classe n'admettra pas évidemment K_n et donc appartient à \mathcal{C}_n . De plus, \mathcal{C}_n contient des graphes arbitrairement grands puisqu'il suffit de rajouter un sommet isolé à un graphe appartenant à \mathcal{C}_n pour obtenir un graphe plus grand et appartenant toujours à \mathcal{C}_n . Soient $U, V_1, V_2 \in \mathcal{C}_n$ et $f_1 : U \rightarrow V_1, f_2 : U \rightarrow V_2$ des isomorphismes de sous-graphes induits. On sait que $f_1(U) \cong f_2(U)$ par $f = f_2 \circ f_1^{-1}$. Construisons le graphe W comme suit :

- $V(W) = V(V_2) \cup [V(V_1) \setminus V(f_1(U))]$;
- pour tous $u, v \in V(W)$:
 1. si $u, v \in V(V_2)$ alors $(u, v) \in E(W)$ si et seulement si $(u, v) \in E(V_2)$;
 2. si $u, v \in V(V_1)$ alors $(u, v) \in E(W)$ si et seulement si $(u, v) \in E(V_1)$;
 3. si $u \in V(V_1), v \in V(V_2)$ alors $(u, v) \in E(W)$ si et seulement si $v \in f_2(U)$ et $(u, f^{-1}(v)) \in E(V_1)$.

Autrement dit, on vient de coller V_1 sur V_2 de sorte à ce que la copie de U dans V_1 coïncide avec celle de V_2 et la question est de savoir si $W \in \mathcal{C}_n$. Si K_n se trouve tout d'un coup admis dans W , il ne peut pas l'être dans $W \upharpoonright V_2$ puisque W admet V_2 . Il ne peut pas non plus être admis dans $W \upharpoonright [V(V_1) \setminus V(f_1(U)) \cup V(V_2) \setminus V(f_2(U))]$ puisque ce graphe est vide. Il doit donc être admis dans $W \upharpoonright V(V_1) \cup V(f_2(U))$. Or ce dernier est isomorphe à V_1 . Contradiction. Par conséquent, $W \in \mathcal{C}_n$. Il

existe donc un graphe ultrahomogène infini dénombrable $\mathcal{K}_n\mathcal{F}$ dont l'âge est C_n . Hormis le complément de ce dernier, y-a-t-il d'autres graphes homogènes infinis dénombrables? Il n'y a que \aleph_0 graphes homogènes infinis dénombrables non-orientés et Lachlan et Woodrow [73] les ont caractérisés :

Théorème 25. *Soit G un graphe homogène infini dénombrable non-orienté. Alors G est isomorphe à un des graphes suivants :*

1. $t \cdot K_n$ avec $t, n \leq \aleph_0$ avec $n + t = \aleph_0$;
2. les compléments des graphes en (1)
3. la limite de Fraïssé des graphes finis qui n'admettent pas K_n pour $n \geq 3$ noté H_n (pour Henson graph) ;
4. les compléments des graphes en (3) ;
5. R .

Quels sont ces graphes en (3) et (4) du théorème 25? Pour un entier $n \geq 3$, considérons la propriété (A_n) de graphe suivante : un graphe G satisfait (A_n) si

- (i) G n'admet pas K_n ;
- (ii) pour tous sous-ensembles finis disjoints $U, V \subseteq V(G)$ tels que $G \upharpoonright U$ n'admet pas K_{n-1} alors il existe $u \in V(G) \setminus (U \cup V)$ correctement relié à U et à V .

Un graphe qui satisfait (A_n) est nécessairement infini puisque sinon en posant $U = \emptyset$ et $V = V(G)$, on a une contradiction.

Proposition 16. *Pour un entier $n \geq 3$, soient G un graphe satisfaisant (A_n) et H un graphe fini qui n'admet pas K_n . Pour un sommet $u \in V(H)$, si $f : H \setminus \{u\} \rightarrow G$ est un isomorphisme de sous-graphes induits alors on peut étendre f à u de telle sorte qu'il reste un isomorphisme de sous-graphes induits.*

Démonstration. On sait que $f(N_H(u))$ n'admet pas K_{n-1} sinon, $H \upharpoonright (N_H(u) \cup \{u\})$ admet K_n . Il existe donc un sommet v dans G correctement relié à $f(N_H(u))$ et $f(H \setminus (N_H(u) \cup \{u\}))$. Il suffit donc de poser $f(u) = v$ pour étendre f à u . \square

La propriété (A_n) est très similaire à la propriété (AR) qui définit R . D'ailleurs, elle aussi est \aleph_0 -catégorique comme le montre le corollaire qui suit :

Corollaire 3. *Deux graphes dénombrables G et H satisfaisant (A_n) sont isomorphes.*

Démonstration. On utilise la même technique d'aller-retour employée dans la preuve du théorème 14. On pose $f_0 = \emptyset$ et supposons que f_{k-1} est définie. Si k est pair, on pose $u_k = \min V(G) \setminus \text{dom}(f_{k-1})$. Comme $G \upharpoonright (\text{dom}(f_{k-1}) \cup \{u_k\})$ n'admet pas K_n , par la proposition 16, on peut étendre f_{k-1} à u_k . Nous prenons alors f_k comme étant cette extension. Si k est impair, on pose $u_k = \min V(H) \setminus \text{im}(f_{k-1})$, et on peut étendre f_{k-1}^{-1} à u_k . On prend alors $f_k = f_{k-1} \cup \{(f_{k-1}^{-1}(u_k), u_k)\}$. L'isomorphisme entre G et H est alors $f = \cup f_k$. \square

Par la proposition 5, un graphe G satisfaisant (A_n) est isomorphe à un sous-graphe couvrant de R et en même temps, R admet G . Construisons à cet effet un graphe U satisfaisant (AR) de la manière suivante : on se donne une énumération $\{P_n : n \in \mathbb{N}\}$ de tous les sous-ensembles finis de \mathbb{N} , où chaque sous-ensemble apparaît une infinité de fois, et une suite infinie d'entiers naturels $v_0 < v_1 < \dots$ telle que $v_n > \max(P_n)$. Ensuite, on pose $V(U) = \mathbb{N}$ et $E(U) = \{\{u, v_n\} : u \in P_n, n \in \mathbb{N}\}$. C'est la même construction que celle qu'on a proposée pour R à ce détail près que le sommet u correctement relié aux sous-ensembles X, Y est tel que $u > \max X$. Ce graphe U satisfait alors la propriété (A) suivante : pour tout $X \subset V(U)$ fini, il existe $u \in V(U)$ arbitrairement grand tel que $X = \{v : v < u \text{ et } v \sim u\}$. Le graphe U est isomorphe à R , et en général un graphe qui satisfait (AR) ne satisfait pas (A). Considérons maintenant la famille $\{G_n\}_{n \geq 3}$ de sous-graphes induits de U où

$$V(G_n) = \{m \in \mathbb{N} : \forall A \subset \mathbb{N} \text{ fini, } m = \max A \Rightarrow U \upharpoonright A \not\cong K_n\}.$$

Les G_n forment une chaîne de graphes et à mesure que n augmente, de plus en plus de sommets de U se retrouvent dans G_n si bien que la limite de cette chaîne est U .

Proposition 17. *Pour tout $n \geq 3$, G_n satisfait (A_n) .*

Démonstration. Bien sûr, G_n n'admet pas K_n par choix de ses sommets. Soient alors X, Y deux sous-ensembles finis disjoints de $V(G_n)$ tels que $G_n \upharpoonright X$ n'admet pas K_{n-1} . Or U satisfait (A), il existe donc $u > \max(X \cup Y)$ tel que l'ensemble $X = \{v : v < u \text{ et } v \sim u\}$. Comme $U \upharpoonright X = G_n \upharpoonright X$ n'admet pas K_{n-1} , tout sous-ensemble $A \subset V(U)$ tel que $u = \max A$ ne peut admettre K_n puisque tous les voisins de u qui lui sont inférieurs sont dans X . Il en découle donc que $u \in V(G_n)$. \square

Corollaire 4. *Pour tout $n \geq 3$, G_n est ultrahomogène.*

Démonstration. Comme G_n satisfait (A_n) , la conclusion découle de l'application de la proposition 16 et du théorème 23. \square

Puisque G_n est ultrahomogène, connaître son âge va nous permettre de le classer parmi la liste du théorème 25. Ainsi,

Proposition 18. *Pour tout $n \geq 3$,*

$$\Sigma(G_n) = \{H : H \text{ graphe fini n'admettant pas } K_n\}.$$

Démonstration. Comme G_n n'admet pas K_n , tout sous-graphe induit fini de G_n n'admet pas non plus K_n . Soit alors H un graphe fini n'admettant pas K_n et soient u_1, u_2, \dots, u_k ses sommets. Comme $f_0 = \emptyset$ est un isomorphisme de sous-graphes induits, par la proposition 16, on peut étendre f_0 à u_1 et obtenir $f_1 : H \upharpoonright \{u_1\} \rightarrow G_n$ un isomorphisme de sous-graphes induits. Des applications successives de la même proposition permet alors d'admettre H dans G_n . \square

Les graphes en (3) et (4) du théorème 25 sont donc G_n et \overline{G}_n pour $n \geq 3$ en vertu du théorème 24. On a déjà vu que le voisinage d'un sommet d'un graphe ultrahomogène est aussi ultrahomogène. La propriété (A_n) avec $U = \{u\}, V = N_{G_n}(u)$ montre que le voisinage de tout sommet de G_n est infini. Il doit donc figurer dans la liste du théorème 25. En fait,

Proposition 19. *Pour $n \geq 3$ et tout $v \in V(G_n)$, $G_n \upharpoonright N_{G_n}(v) \cong G_{n-1}$.*

Démonstration. Soient $X, Y \subset N_{G_n}(u)$ finis disjoints tels que X comme sous-graphe induit de $G_n \upharpoonright N_{G_n}(u)$ n'admet pas K_{n-1} . Il existe alors un sommet v de G_n correctement relié à $X \cup \{u\}$ et à Y ; mais alors, $v \in N_{G_n}(u)$. Par conséquent, $G_n \upharpoonright N_{G_n}(u)$ satisfait (A_{n-1}) . \square

4.2.2 Jeu de poursuite

Parmi les graphes ultrahomogènes infinis dénombrables, notons que $\text{sn}(K_{\aleph_0}) = 1$ et $\text{sn}(R) = \aleph_0$. La propriété d'homogénéité ne détermine donc pas à elle seule le nombre de policiers pour gagner sur le graphe puisque k copies disjointes de K_{\aleph_0} est k -policiers gagnant pour tout $k > 0$. Par ailleurs, $\text{sn}(\overline{t \cdot K_n} = K_{t,n}) = 2$ même si $t + n = \aleph_0$. En nous référant à la liste du théorème 25, il nous reste à considérer le jeu de police-voleur sur H_n et son complément.

Proposition 20. *Pour tout $n \geq 3$, $\text{sn}(H_n) = \aleph_0$.*

Démonstration. Considérons un nombre fini k de policiers. Étant donnée leurs positions initiales $X_0 \subset V(H_n)$, par la propriété (A_n) , il existe $v_0 \in V(H_n)$ non adjacent à X . Au premier tour des policiers, le voleur ne sera donc pas capturé. Maintenant, supposons que les policiers soient situés sur X_i et le voleur sur $v_i \notin X_i$. Comme $H_n \upharpoonright \{v_i\}$ n'admet pas K_{n-1} , il existe $v_{i+1} \notin X_i \cup \{v_i\}$ tel que $v_i \sim v_{i+1}$ et $v_i \not\sim X_i$. En choisissant v_{i+1} comme prochain sommet, le voleur évite une arrestation au prochain coup des policiers. \square

Ainsi donc, un nombre fini de policiers ne suffit pas pour capturer le voleur dans H_n . Pour le graphe aléatoire infini dénombrable R , on a eu recours à un modèle probabiliste pour nous convaincre que la proportion des sommets que les policiers doivent occuper pour gagner peut être rendue aussi petite que l'on veut. On peut aussi constater ce fait avec une construction directe d'un graphe satisfaisant (AR) , en particulier le graphe U de la section précédente. En effet, étant donnée une partie finie P_n de \mathbb{N} , il suffit de placer un policier sur v_n . Or $P_n \cup \{v_n\} \subseteq P_i$ pour un certain i et de ce fait, il suffit finalement de placer un policier sur v_i . Étant

donné donc $X \subseteq \mathbb{N}$, un ensemble dominant de U , on peut en construire un autre qui n'a que la « moitié » d'éléments en partitionnant $X = \cup X_i$ en sous-ensembles de deux éléments. Pour chaque $X_i = \{x_{1i}, x_{2i}\}$, on peut trouver un v_j tel que v_j est adjacent à

$$\bigcup_{k=1,2} \{w : w < x_{ki}, w \sim x_{ki}\} \cup \{x_{ki}\}.$$

On peut donc trouver un ensemble dominant dans R dont la densité est aussi petite que l'on veut ; le nombre de domination $\gamma(G(n, p))$ étant $O(\log n)$ [89]. De même, si on construit un graphe qui satisfait (A_n) directement par la propriété elle-même, on voit que H_n possède aussi un ensemble dominant de densité aussi petite que l'on veut.

Maintenant, considérons \overline{H}_n . Le complément de (ii) de la propriété (A_n) dit alors que pour tous $X, Y \subset V(\overline{H}_n)$ finis disjoints tels que $H_n \upharpoonright \overline{X}$ n'admet pas K_{n-1} , il existe un sommet u dans \overline{H}_n tel que $u \sim Y$ et $u \not\sim X$. On peut l'utiliser pour analyser le jeu de police-voleur sur \overline{H}_n . Ainsi,

Proposition 21. *Pour tout $n \geq 3$, $\text{sn}(\overline{H}_n) = n - 1$.*

Démonstration. Supposons qu'on ait seulement $n - 2$ policiers pour patrouiller \overline{H}_n . Disons qu'ils commencent en X_0 . Comme $H_n \upharpoonright X_0$ ne peut jamais admettre K_{n-1} , il existe donc un sommet u_0 non-adjacent à X_0 . Le voleur choisit ce sommet comme sommet de départ. Au premier tour des policiers, le voleur est à l'abri. Ainsi si les policiers sont sur X_i et le voleur sur $u_i \notin X_i$ et c'est au tour du voleur de jouer, il existe un sommet u_{i+1} correctement relié à $\{u_i\}$ et à X_i . Le voleur peut alors se déplacer sans souci puisqu'au tour des policiers il sera encore libre.

Maintenant si $n - 1$ policiers feront la ronde sur \overline{H}_n , la question qui se pose en premier concerne les sommets de départ. Notons d'abord que H_n n'admet pas K_n mais admet bel et bien K_{n-1} . Plaçons alors les $n - 1$ policiers sur X tel que $H_n \upharpoonright X \cong K_{n-1}$. Le voleur devra alors trouver un sommet u tel que $N_{\overline{H}_n}(u) \cap X = \emptyset$ pour éviter la capture au premier coup ; mais cela ne veut-il dire que u est adjacent à tous les sommets de X dans H_n ? Or ceci est impossible car H_n n'admet pas K_n . En d'autres termes, le nombre de domination $\gamma(\overline{H}_n) = n - 1$. \square

Le graphe $G = \overline{H}_{k+1}$ est donc un exemple de graphe infini connexe non trivial tel que $\text{sn}(G) = k$. Remarquons que dans ces graphes ultrahomogènes infinis dénombrables, soit le voleur peut jouer un coup et, dans ce cas, il est gagnant, soit il est capturé dès le départ.

CHAPITRE 5

GRAPHES ALÉATOIRES

Considérons la propriété de graphe $S_k = "G \text{ est } k\text{-policiers gagnant}"$. On a vu dans le chapitre 3 que $\lim_{n \rightarrow \infty} \Pr [G(n, p) \models S_k] = 0$ pour tout $k > 0$ et tout $p \in (0, 1)$ fixé. Dans le jargon de la théorie des graphes aléatoires, on dira que presque tous les graphes sont voleurs gagnants. Une des occupations majeures de la théorie des graphes aléatoires est l'étude du comportement d'une propriété dans un modèle de graphe aléatoire. Un *modèle de graphe aléatoire* est un espace de probabilité dont les éléments sont des graphes. On peut citer comme exemple de modèle $\mathcal{G}_n(m)$, l'espace de tous les graphes sur $[n]$ ayant m arêtes et où chaque graphe apparaît avec la même probabilité $1/\binom{N}{m}$ avec $N = \binom{n}{2}$. Il y a aussi bien sûr le modèle $G(n, p)$ et d'ailleurs, Bollobàs [17] montre que lorsque m est proche de pn , les deux modèles sont pratiquement les mêmes. Si tous les graphes d'un modèle sont tous d'ordre n , on dit que le modèle est d'ordre n . Ce chapitre sera consacré à l'étude de la propriété S_k dans le modèle $G(n, p)$.

Soit \mathcal{G}_n un modèle de graphe aléatoire d'ordre n . Une propriété de graphe P dans \mathcal{G}_n est donc un ensemble de graphes de \mathcal{G}_n tel que si $G, H \in \mathcal{G}_n$ sont isomorphes et $G \in P$ alors $H \in P$ aussi. Ainsi $\Pr [P] = \sum_{G \in P} \Pr [G \models P]$ a la signification naturelle : la probabilité qu'un graphe de \mathcal{G}_n satisfait la propriété P . Ainsi, si $\lim_{n \rightarrow \infty} \Pr [P] = 1$, on dit P tient *presque sûrement* dans \mathcal{G}_n ou encore *presque tous les graphes de \mathcal{G}_n ont P* . De même, si $\lim_{n \rightarrow \infty} \Pr [P] = 0$, on dit que P ne tient *presque jamais* dans \mathcal{G}_n ou encore *presque aucun graphe de \mathcal{G}_n n'a P* . Une propriété P de graphe est dite *monotone* si pour tout $H \in P$ et tout $G \in \mathcal{G}_n$, $G < H$ implique que $G \in P$.

5.1 Évolution Erdős-Rényi

Dans cette section, nous allons nous familiariser avec le point de vue donné par Erdős et Rényi dans leur papier original [42]. Le graphe aléatoire $G(n, p)$ évolue « du vide au plein » à mesure que p croît. Pour assister à cette évolution, la probabilité $p = p(n)$ est maintenant prise comme une fonction de n et non plus comme constante fixée. La question principale qui se pose est alors : comment est-ce qu'un graphe typique de $G(n, p)$ change relativement à la variation de la distribution de probabilité ? Pour plusieurs propriétés, on remarque cette espèce de « phase de transition » : jusqu'à ce que p n'atteigne un certain seuil, presque aucun graphe de $G(n, p)$ ne possède la propriété en question ; et au-delà du seuil, soudainement presque tous les graphes la satisfont. Nous allons résumer ici les étapes marquantes de cette évolution.

5.1.1 Le vide

Supposons que $p \ll n^{-2}$, on a déjà vu que $\Pr[\exists u, v \ u \sim v] = \frac{n^2 - n}{2}p$ et par conséquent, presque aucun graphe de $G(n, p)$ n'a d'arête. La limite de $G(n, p)$ est donc le graphe vide infini dénombrable, qui est unique à isomorphisme près. La propriété " $\forall u, v \ u \not\sim v$ " est \aleph_0 -catégorique.

5.1.2 Les arbres

Par le temps où $p \in \Theta(n^{-2})$, les arêtes apparaissent, mais restent dispersées. Considérons la propriété $P = \exists u, v, w \ u \sim v \wedge u \sim w$. Autrement dit, le graphe contient des arêtes possédant un sommet commun. La probabilité

$$\Pr[P] = 3 \binom{n}{3} p^2.$$

Par conséquent, les arêtes à sommet commun, qui ne sont que des arbres sur 3 sommets, émergent à $p \in \Theta(n^{-3/2})$. Lorsque $p \in \Theta(n^{-1-1/k})$, les arbres sur $k + 1$ se

manifestent. En effet, la probabilité d'avoir un arbre sur $k + 1$ noeuds est

$$(k + 1)! \binom{n}{k + 1} p^k.$$

Qu'est-ce qui arrive pendant la période où $n^{-1-1/k} \ll p(n) \ll n^{-1-\frac{1}{k+1}}$? Avec probabilité élevée,

- les composantes connexes de $G(n, p)$ seront de taille strictement plus petite que $k + 2$, pour la simple raison que l'arbre étant le plus petit graphe connexe sur les $k + 2$ sommets ;
- il n'y a aucun cycle. En effet, pour $m < k + 2$, la probabilité d'avoir un m -cycle est $\binom{n}{m} p^m$ qui tend vers 0 ;
- pour tout r , tout arbre sur au plus $k + 1$ sommets apparaît au moins r fois comme composante connexe.

Maintenant, quand $n^{-1-\epsilon} \ll p(n) \ll n^{-1}$ pour tout $\epsilon > 0$, les cycles n'y figurent toujours pas et tous les arbres finis apparaissent comme composante connexe au moins r fois pour tout r . La probabilité que $G(n, p)$ contienne une clique de taille $d > 1$ est

$$\binom{n}{d} p^{\frac{d(d+1)}{2}} = O(n^d p^{\frac{d(d+1)}{2}}).$$

Par conséquent, pour $n^{-2/(d-1)} \ll p(n)$, avec probabilité élevée $G(n, p)$ admet K_d .

5.1.3 Composante géante

Lorsque $p \in \Theta(n^{-1})$, les cycles apparaissent, le graphe devient non planaire. Si $p \sim c/n$ avec $0 < c < 1$, les composantes connexes sont soit des arbres, soit des arbres avec une arête de plus. La plus grande composante connexe est un arbre d'ordre $\Theta(\log n)$. Si $p = n^{-1}$, la plus grande composante connexe est d'ordre $n^{2/3}$. Si $p \sim c/n$ avec $c > 1$, on assiste aux fusions des composantes connexes dont la plupart sont des arbres. La plus grande atteint maintenant les $\Theta(n)$ sommets.

A $\frac{1}{n} \ll p(n) \ll \frac{\log n}{n}$, il y a encore des sommets de petit degré, ce qui empêche l'apparition de graphe plus complexe localement. Pour tout k , le nombre d'arêtes

sur k sommets est au plus k , soit un cycle. En effet, la probabilité d'avoir $k + 1$ arêtes sur k sommets est $\binom{n}{k} \binom{m}{k+1} p^{k+1} = o(1)$ avec $m = \binom{k}{2}$.

5.1.4 Connexité

Considérons la propriété $P = “\forall u \exists v u \sim v”$ qui dit tout simplement qu'aucun sommet n'est isolé, c'est-à-dire $\neg P = “\exists u \forall v u \not\sim v”$. La probabilité $\Pr[\neg P] = n(1-p)^{n-1}$. D'où si $p = \frac{\log n}{n}$,

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr[\neg P] &= \exp\left(\lim_{n \rightarrow \infty} (n-1) \log(n - \log n)\right) \\ &= \exp\left(\lim_{n \rightarrow \infty} \frac{\log(n - \log n)}{(n-1)^{-1}}\right) \\ &= \exp\left(\lim_{n \rightarrow \infty} \frac{(1 - 1/n)/(n - \log n)}{-(n-1)^{-2}}\right) \\ &= \exp\left(-\lim_{n \rightarrow \infty} \frac{(n-1)^2(1-1/n)}{n(1 - \log n/n)}\right) \\ &= 0. \end{aligned}$$

Par conséquent, avec probabilité élevée, $G(n, p)$ perd les arbres sur un sommet lorsque $p \in \Omega(\log n/n)$. De plus, [17], si $p = (\log n + c + o(1))/n$ alors la probabilité

$$\Pr[\text{“est connexe”}] = e^{-e^{-c}}$$

pour tout $c \in \mathbb{R}$. A $p = \sqrt{\log n/n}$, chaque paire de sommets possède un voisin commun.

L'étude de ces phases de transition constitue encore un élément central de la théorie des graphes aléatoires.

5.2 Logique du premier ordre

Soit P la propriété qui veut que le graphe possède au moins $\binom{n}{2} n^{-\pi}$ arêtes. Sa probabilité est $p^{\binom{n}{2} n^{-\pi}}$ et on en déduit qu'à $p = n^{-\pi}$, avec probabilité élevée $G(n, p)$ possède P . Bien qu'on a pu déterminer la phase de transition pour cette propriété,

il y a en elle un aspect qui déplaît, à savoir qu'elle n'est pas du tout naturelle. Plutôt que de débattre sur ce que peut constituer une propriété naturelle, concentrons-nous, du moins dans cette section, sur des propriétés que l'on peut exprimer dans un langage particulier : la logique du premier ordre.

5.2.1 Langage

La logique du premier ordre est une notion de base pour les logiciens. Ici, nous allons considérer le langage du premier ordre pour les graphes. L'alphabet du langage consiste en

1. une source inépuisable de symboles que l'on nomme *variables* $:x, y, z, \dots$
Dans notre langage, une variable correspond à un sommet ;
2. deux relations binaires $=$ (égalité) et \sim (adjacence) ;
3. les quantificateurs *universel* \forall et *existantiel* \exists qui s'appliquent seulement sur les variables ;
4. les connecteurs booléens usuels $:\vee, \wedge, \neg, \Rightarrow$ (même si \wedge, \neg suffisent).

La forme générale d'une expression dans ce langage est le *prédicat* que l'on peut définir récursivement. Pour tout pair de variables x, y , $x = y$ et $x \sim y$ sont des prédicats. Pour tous prédicats P, Q , on a que $\neg P, P \vee Q, P \wedge Q, P \Rightarrow Q$ sont des prédicats et pour toute variable x , les expressions $\forall x P$ et $\exists x P$ sont aussi des prédicats. Une variable x dans un prédicat P est dite *liée* si une des quantifications $\forall x$ ou $\exists x$ apparaît dans P , sinon elle est dite *libre*. Les prédicats qui contiennent une ou plusieurs variables libres sont appelé *formules*. Les prédicats ne contenant que des variables liées sont les propriétés. Le *rang de quantification* d'un prédicat P , notée $\text{qd}(P)$, est le nombre maximal de quantificateurs imbriqués dans P et se définit inductivement par :

- $\text{qd}(x = y) = 0$ et $\text{qd}(x \sim y) = 0$ pour toutes variables x, y ;
- $\text{qd}(\neg P) = \text{qd}(P)$;
- $\text{qd}(P \oplus Q) = \max\{\text{qd}(P), \text{qd}(Q)\}$ où $\oplus = \vee, \wedge$;
- $\text{qd}(\forall x P) = \text{qd}(\exists x P) = 1 + \text{qd}(P)$.

Par exemple $\text{qd}(\exists v v = v \wedge \forall x \exists y x \sim y) = 2$.

La sémantique du langage est celle de la théorie des graphes habituelle. Un ensemble \mathcal{T} de propriétés est appelé *théorie* et un graphe G pour lequel chaque propriété de \mathcal{T} est vraie lorsque les quantifications prennent leurs valeurs dans $V(G)$ constitue un *modèle* pour \mathcal{T} .

Voyons quelques exemples (les parenthèses et les crochets sont là pour des raisons de lisibilité) :

- il existe un triangle : $\exists x \exists y \exists z [\neg x = y \wedge \neg y = z \wedge \neg z = x \wedge x \sim y \wedge y \sim z \wedge z \sim x]$;
- est complet : $\forall x \forall y \neg x = y \Rightarrow x \sim y$;
- diamètre 2 : $\forall x \forall y \exists z [\neg x = y \wedge \neg y = z \wedge \neg z = x \wedge \neg x \sim y \Rightarrow (x \sim z \wedge z \sim y)]$;
- aucun sommet isolé : $\forall x \exists y [\neg x = y \wedge x \sim y]$.

Remarquons d'abord l'usage d'un nombre dans ces exemples. On n'a aucune difficulté à exprimer le fait de contenir un k -cycle, K_k ou être de diamètre k si k est donné avant même la transcription. Par contre, si k n'est pas connu ce n'est pas évident. Par exemple, peut-on exprimer le fait de contenir un cycle? Si on commence par construire une propriété qui est la disjonction de toutes les propriétés $P_k = \text{"contient un } k\text{-cycle"}$ pour $k = 3, 4, \dots$, on obtient un prédicat de longueur infinie. Ce qui n'est pas permis par notre langage.

Ensuite, le domaine de quantification est l'ensemble des sommets du graphe et on ne peut pas quantifier un sous-ensemble de sommets. L'ordre du graphe n'est même pas accessible au langage. Ainsi on ne peut pas exprimer par exemple le fait de contenir un ensemble indépendant de taille la moitié de l'ordre du graphe ou tout simplement le fait que le graphe est d'ordre pair.

Se questionner sur la puissance d'expression du langage s'impose alors. Est-ce que les propriétés intéressantes des graphes sont exprimables en logique du premier ordre, en particulier les propriétés S_k ? Pour répondre à cette question, nous avons besoin d'un outil bien connu des logiciens : le jeu d'Ehrenfeucht-Fraïssé. Ce qu'il y a de remarquable dans ce jeu c'est le lien qu'il établit entre la logique et le combinatoire. Ici, on l'énoncera dans le contexte particulier de la théorie des graphes.

5.2.2 Jeu de Ehrenfeucht-Fraïssé

C'est un jeu à somme nulle et à information complète qui se pratique à deux. Le premier joueur sera appelé le *saboteur* et le second l'*imitateur*. On dispose de deux graphes G_1, G_2 et de $2r$ jetons que l'on va numéroter $u_1, u_2, \dots, u_r, v_1, v_2, \dots, v_r$. Le jeu commence en distribuant les jetons u_i , respectivement v_i , au saboteur, respectivement à l'imitateur. Au coup i , le saboteur pose le jeton u_i sur un des sommets de G_j , où $j = 1$ ou 2 , et l'imitateur place v_i sur un sommet de $G_{i+1 \pmod{2}}$. Un sommet peut recevoir plus d'un jeton. Ils jouent ainsi tant et aussi longtemps qu'ils ont de jetons. A la fin, qui gagne? L'imitateur gagne si pour tout $1 \leq i \leq j \leq r$, $v_i \sim v_j$ si et seulement si $u_i \sim u_j$ et $u_i = u_j$ si et seulement si $v_i = v_j$. Par $u_i \sim u_j$, nous voulons bien sûr entendre le sommet occupé par u_i est adjacent à celui occupé par u_j . Autrement dit, l'imitateur gagne si $G_1 \upharpoonright \{x_1, \dots, x_r\} \cong G_2 \upharpoonright \{y_1, \dots, y_r\}$ où les x_i , respectivement y_i , sont les sommets de G_1 , respectivement G_2 , qui ont reçu un jeton. On dira que le jeu de Ehrenfeucht-Fraïssé sur G_1, G_2 avec r coups, noté $\text{EHR}(G_1, G_2, r)$, est *imitateur gagnant* si l'imitateur a une stratégie qui lui permettra de gagner quelque soit les choix du saboteur. Sinon, on dira que $\text{EHR}(G_1, G_2, r)$ est *saboteur gagnant*.

Prenons un exemple en considérant ce jeu sur P_9 et P_8 où $P_i = ([i], \leq)$, la chaîne sur i éléments. Comme le nombre d'indépendance $\alpha(P_9) = 5$, le saboteur peut choisir tous les 5 sommets indépendants de P_9 et l'imitateur ne pourra pas égaliser, donc $\text{EHR}(P_9, P_8, 5)$ est *saboteur gagnant*. Pour $r < 5$, $\text{EHR}(P_9, P_8, r)$ est *imitateur gagnant*.

5.2.3 Expressivité

L'intuition derrière le jeu de Ehrenfeucht-Fraïssé est la suivante : l'imitateur gagne lorsque le saboteur n'arrive pas à montrer par choix de r sommets que G_1, G_2 ne sont pas isomorphes. Pensons-y, étant donné un jeu imitateur gagnant, pour tout sous-graphe induit de G_1 ayant r sommets, on peut en trouver un dans G_2 , qui lui est isomorphe et vice versa. Or, puisque le domaine de quantification de

notre langage du premier ordre est seulement l'ensemble des sommets, l'ensemble $\{G : G \models P\}$ où P est une proposition du premier ordre forme une propriété de graphes, c'est-à-dire, isomorphisme invariant. Une propriété de rang de quantification r conserve donc la même valeur de vérité lorsqu'on passe du modèle G_1 vers G_2 . C'est le résultat fondamental concernant le jeu de Ehrenfeucht-Fraïssé. Pour l'énoncer clairement, nous avons besoin de relever quelques faits.

Pour $r \geq 1$, soit la relation binaire \equiv_r sur les graphes définie par $G_1 \equiv_k G_2$ si et seulement si $\text{EHR}(G_1, G_2, k)$ est imitateur gagnant. Clairement, \equiv_k est réflexive et symétrique. Soient alors G_1, G_2, G_3 des graphes tels que $G_1 \equiv_k G_2$ et $G_2 \equiv_k G_3$ et considérons le jeu $J_3 = \text{EHR}(G_1, G_3, r)$. L'imitateur va jouer $J_1 = \text{EHR}(G_1, G_2, k)$ et $J_2 = \text{EHR}(G_2, G_3, k)$ en parallèle pour obtenir sa stratégie. Dans le jeu J_3 , si le saboteur place son i -ème jeton u_i dans G_1 , on peut considérer le même coup dans le jeu J_1 et l'imitateur sait où placer son jeton v_i dans G_2 pour ne pas perdre. En jouant dans J_2 avec un coup du saboteur sur le sommet occupé par v_i dans G_2 , l'imitateur réplique avec w_i dans J_2 et J_3 . De même si le saboteur pose son jeton dans G_3 pour le jeu J_3 , l'imitateur peut riposter en regardant dans J_2 puis dans J_1 . A la fin du jeu, la question est de savoir si $G_1 \upharpoonright \{x_1, \dots, x_r\} \cong G_3 \upharpoonright \{z_1, \dots, z_r\}$, où les x_i , respectivement z_i , sont les sommets occupés par un jeton de G_1 , respectivement G_3 . Or, $G_1 \upharpoonright \{x_1, \dots, x_r\} \cong G_2 \upharpoonright \{y_1, \dots, y_r\}$ et $G_2 \upharpoonright \{y_1, \dots, y_r\} \cong G_3 \upharpoonright \{z_1, \dots, z_r\}$, où les y_i étant les sommets occupés par un jeton de G_2 , puisqu'à chaque fois, l'imitateur a utilisé sa stratégie gagnante. Par transitivité de \cong , on a donc $G_1 \equiv_r G_3$.

Pour un $r \geq 1$ fixé donc tout graphe appartient à une classe d'équivalence. Prenons $r = 1$. Les classes d'équivalence de \equiv_1 sont les graphes réflexifs, les graphes irreflexifs et les graphes non réflexifs qui contiennent des boucles sur certains sommets. Il est clair qu'une propriété dont les relations entre variables sont toutes de la forme $x = x$ ou $x \sim x$ garde la même valeur de vérité dans chacune de ces classes. C'est un cas particulier d'un résultat plus général.

Théorème 26. *Pour chaque classe d'équivalence $[G]$ de \equiv_r , il existe une propriété P telle $\text{qd}(P) = r$ et $H \in [G]$ si et seulement si $H \models P$. Autrement dit, on peut*

exprimer une classe d'équivalence \equiv_r en logique de premier ordre.

Démonstration. Pour $r = 1$, la classe des graphes réflexifs est définie par " $\forall x x \sim x$ "; celles des graphes non-réflexifs par " $\forall x \neg x \sim x$ " et la classe restante par " $\exists x x \sim x \wedge \exists x \neg x \sim x$ ".

Supposons l'assertion vraie pour r et considérons le jeu $\text{EHR}(G_1, G_2, r + 1)$. Si $G_1 \equiv_{r+1} G_2$ alors $G_1 \equiv_r G_2$. Il existe donc une propriété P de rang de quantification r qui définit la classe \equiv_r . Soient x_1, \dots, x_{r+1} les sommets occupés par des jetons dans G_1 durant une partie où l'imitateur a gagné. La propriété $Q = P \wedge \exists a_1 \exists a_2 \dots \exists a_{r+1} \phi$, où ϕ est une formule telle que $\phi(x_1, x_2, \dots, x_r)$ décrit l'adjacence des x_i dans G_1 , est vraie autant dans G_1 que dans G_2 puisque $G_1 \equiv_{r+1} G_2$. Si $G_1 \not\equiv_{r+1} G_2$ alors soit $G_1 \not\equiv_r G_2$ et dans ce cas $G_1 \models \neg P$ et $G_2 \models \neg P$, soit $G_1 \equiv_r G_2$ et dans ce cas l'imitateur n'a pas pu égaliser le dernier coup du saboteur. Or cela veut dire que G_1 et G_2 sont en désaccord sur ϕ et donc $G_2 \not\models Q$. \square

Notons que l'index de \equiv_r , c'est-à-dire le nombre de classe d'équivalence, est fini, voir [85]. L'observation clé est qu'il n'existe qu'un nombre m fini de propriétés, disons ϕ_1, \dots, ϕ_m de rang de quantification r et donc tout graphe doit satisfaire une des

$$\begin{aligned} \psi_1 &= \phi_1 \wedge \phi_2 \wedge \dots \wedge \phi_m, \\ \psi_2 &= \neg\phi_1 \wedge \phi_2 \wedge \dots \wedge \phi_m, \\ &\vdots \\ \psi_{2^m} &= \neg\phi_1 \wedge \neg\phi_2 \wedge \dots \wedge \neg\phi_m. \end{aligned}$$

Maintenant, on peut énoncer le théorème principal du jeu de Ehrenfeucht.

Théorème 27. *Pour tout $r \geq 1$, $G_1 \equiv_r G_2$ si et seulement pour toute propriété P telle que $\text{qd}(P) = r$ on a $G_1 \models P$ si et seulement $G_2 \models P$.*

Démonstration. " \Leftarrow " C'est le corollaire du théorème précédent puisqu'une des propriétés qui décrivent les classes d'équivalence de \equiv_r sera vraie en même temps sur G_1 et G_2 .

“ \Rightarrow ” Nous allons plutôt montrer que si $G_1 \models P$ et $G_2 \models \neg P$ pour une certaine propriété de rang de quantification r alors $G_1 \not\equiv_r G_2$. Illustrons avec le cas $r = 3$, mais la technique marche pour tout r .

Posons $P = \forall x_1 \exists x_2 \forall x_3 \phi(x_1, x_2, x_3)$ telle que $G_1 \models P$ et $G_2 \models \neg P$, où ϕ est une formule qui ne contient pas de quantificateur. On va exhiber une stratégie gagnante du saboteur. Comme G_2 satisfait $\neg P = \exists x_1 \forall x_2 \exists x_3 \neg \phi(x_1, x_2, x_3)$, le saboteur commence en choisissant un sommet v_1 de G_2 tel que $G_2 \models \forall x_2 \exists x_3 \neg \phi(v_1, x_2, x_3)$. L'imitateur répond en choisissant un sommet v_1 dans G_1 . Dû au quantificateur universel $\forall x_1$ dans P , $G_1 \models \exists x_2, \forall x_3 \phi(u_1, x_2, x_3)$. Ensuite, le saboteur opte pour un sommet u_2 de G_1 pour lequel $G_1 \models \forall x_3 \phi(u_1, u_2, x_3)$. L'imitateur riposte avec un sommet v_2 dans G_2 et encore un fois à cause de $\forall x_2$ dans $\neg P$, $G_2 \models \exists x_3 \neg \phi(v_1, v_2, x_3)$. Le dernier coup du saboteur sera un sommet v_3 de G_2 tel que $G_2 \models \neg \phi(v_1, v_2, v_3)$. Qu'importe le choix u_3 de l'imitateur dans G_1 , on sait que $G_1 \models \phi(u_1, u_2, u_3)$. Par conséquent, $G_1 \upharpoonright \{u_1, u_2, u_3\} \not\equiv G_2 \upharpoonright \{v_1, v_2, v_3\}$ puisque ϕ est une propriété de graphe, étant exprimée en logique du premier ordre.

Si une telle P n'existe pas, on peut prendre $\phi(a_1, a_2, a_3) = \neg a_1 \sim a_3 \wedge \neg a_1 \sim a_1$ et P caractérise donc le graphe nul. Dans ce cas, si $G_1 \models P$ et $G_2 \models P$ alors $\text{EHR}(G_1, G_2, r)$ est imitateur gagnant.

□

Le corollaire suivant de ce théorème est ce qui nous importe le plus car il nous donne un outil pour prouver qu'une propriété de graphe n'est pas exprimable en logique du premier ordre.

Corollaire 5. *Soit P une propriété de graphe et supposons que pour tout $r \geq 1$ il existe deux graphes G_1, G_2 tels que $G_1 \equiv_r G_2$ et $G_1 \models P$ mais $G_2 \models \neg P$. Alors P n'est pas exprimable en logique du premier ordre.*

Pour illustrer ce corollaire, prenons la propriété P qui veut que l'ordre du graphe soit pair. Pour tout $r \geq 1$, prenons $G_1 = \overline{K_{2r}}$ et $G_2 = \overline{K_{2r+1}}$. Il est clair que

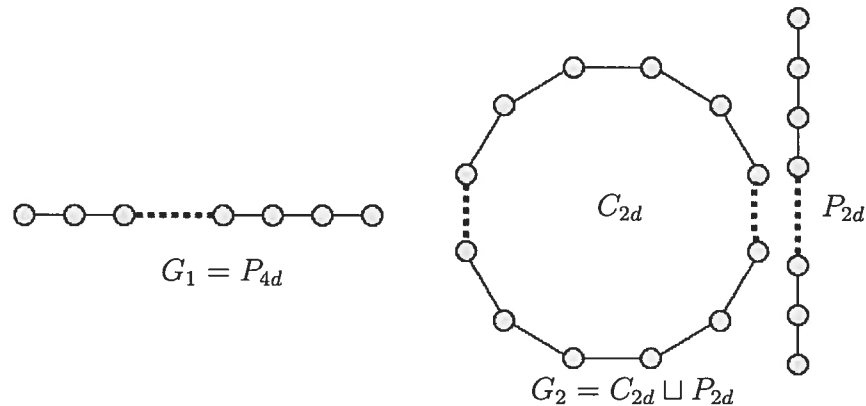


FIG. 5.1 – Jeu Ehrenfeucht-Fraïssé imitateur gagnant

l'imitateur gagne $\text{EHR}(G_1, G_2, r)$ et $G_1 \models P$ tandis que $G_2 \models \neg P$. La propriété P n'est donc pas exprimable en logique du premier ordre.

Prenons maintenant la propriété $S_k = \text{“}G \text{ est } k\text{-policiers gagnant”}$. Les caractérisations que l'on connaît du jeu requièrent un peu plus que la logique du premier ordre. Par exemple, l'appartenance à un voisinage. Néanmoins, cela ne prouve pas qu'il n'existe pas une façon d'exprimer S_k en logique du premier ordre. Pour $k \geq 2$, on peut montrer qu'il n'y a aucune façon.

Proposition 22. *Pour tout $k \geq 1$, S_k n'est pas exprimable en logique du premier ordre.*

Démonstration. Pour $k = 1$, soient $r \geq 1$ et $d \geq 3^{r-1}$ et posons $G_1 = P_{4d}$ et $G_2 = P_{2d} \sqcup C_{2d}$, voir figure 5.1. L'idée est que :

- avec r coups, le saboteur ne peut former une chaîne reliant les deux sommets de degré 1 de G_2 . L'imitateur peut donc « couper » la chaîne P_{2d} de G_2 en deux les faire correspondre aux deux extrémités de G_1 .
- avec r coups, le saboteur ne peut montrer que C_{2d} de G_2 est un cycle. On peut donc le « couper » et le faire correspondre à une partie de G_1 .

Au i -ème coup, l'imitateur regarde parmi les sommets précédemment marqués s'il y en a un qui soit à une distance $r - i$ du nouveau sommet choisi par le saboteur. Si tel est le cas, l'imitateur marque alors dans l'autre graphe le sommet qui soit à

la même distance et dans le même sens du sommet correspondant, c'est-à-dire de même indice. Sinon, deux cas sont à considérer :

- le saboteur marque un sommet dans G_1 . S'il est à distance inférieure à $r - i$ des extrémités de G_1 , on marque le sommet de même distance dans P_{2d} . Sinon, tout au plus, $i - 1$ sommets ont reçu des jetons dans C_{2d} . On peut donc y trouver un sommet qui est assez éloigné ($> r - i$) des sommets déjà marqués. L'imitateur choisit un tel sommet au hasard.
- le saboteur marque un sommet dans G_2 . Si ce sommet est dans P_{2d} et est à distance inférieure à $r - i$ de ses extrémités, l'imitateur posera un jeton à distance égale de l'extrémité correspondante de G_1 . Sinon, on peut toujours trouver un sommet de G_1 qui soit suffisamment éloigné ($> r - i$) des sommets déjà marqués. L'imitateur choisit un tel sommet au hasard.

On a donc que $G_1 \equiv_r G_2$; [85] donne des conditions suffisantes qui renforcent cette affirmation. Comme G_1 est 1-voleur gagnant et G_2 ne l'est pas, alors la classe des graphes 1-policiers gagnants ne sont pas définissables en logique du premier ordre. Pour $k > 1$, il suffit d'ajouter $k - 1$ sommets isolés à chacun des G_1 et G_2 . \square

Notons que G_1 est connexe et G_2 ne l'est pas. Par conséquent, la connexité n'est pas non plus une propriété exprimable en logique du premier ordre. De plus, G_2 contient un cycle et G_1 n'en possède aucun. Donc, il n'y a pas de prédicat en logique du premier ordre qui peut exprimer le fait de contenir un cycle.

Étant donné un graphe G d'ordre n , on peut toujours trouver une expression en logique du premier ordre qui le caractérise à savoir

$$\exists x_1 \exists x_2 \cdots \exists x_n (x_i \neq x_j \text{ pour } i \neq j) \wedge \forall (x_{n+1} x_{n+1} = x_1 \vee \cdots x_{n+1} = x_n) \wedge \\ (x_i \sim x_j \text{ si et seulement si } (u_i, u_j) \in E(G),$$

avec $V(G) = \{u_1, \dots, u_n\}$. Bien sûr on a employé une sorte de raccourci dans le prédicat précédent, mais on comprend bien que tout s'exprime en logique du premier ordre. Ce n'est pas forcément la meilleure façon de « décrire » G . Parmi

toutes les façons de représenter G en logique du premier ordre, il y en a sûrement une qui est de rang de quantification minimum. On peut alors considérer

$$D(G) = \min \{ \text{qd}(P) : G \models P \wedge H \not\models P \text{ pour tout } H \text{ non isomorphe à } G \}$$

comme mesure de la complexité de G . La formulation donnée plus haut montre que $D(G) \leq n + 1$. En général, cette borne est la meilleure possible. Une approche probabilistique dans [64] donne $D(G(n, p)) = O(\log n)$ pour $0 < p < 1$ fixée.

La question de l'expressivité d'un problème dans un langage donné est un thème bien étudié en logique. D'ailleurs c'est une forme de complexité, appelée *complexité descriptive* et dont Fagin a mis à jour le lien avec la complexité de calcul. En effet, Fagin [43] a donné une description purement logique, sans aucune notion de machine, de la classe NP. Il a prouvé que les propriétés sur des structures finies qui sont décidables dans NP sont précisément celles qui sont définissables en logique existentielle du second ordre, c'est-à-dire, exprimable par une proposition de la forme $\exists R_1 \cdots \exists R_n \phi$ où R_i sont des variables dont le domaine de quantification est l'ensemble des relations et ϕ une expression en logique du premier ordre. Par la suite, il a été prouvé [61, 88] que la classe P coïncide avec celle des structures finies définissables avec la logique du point fixe. Par conséquent, S_k est exprimable en logique du point fixe.

5.3 Loi Zéro-Un

Intuitivement, la loi zéro-un dit que dans un gros espace de probabilité, tout évènement raisonnable est soit très improbable, soit très probable. Par très improbable nous entendons avec probabilité très proche de 0 comme très probable veut dire avec probabilité très proche de 1. Plus formellement, soient $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$ une suite de graphes aléatoires et $\{P_n\}_{n \in \mathbb{N}}$ une suite d'évènements où P_i se produit dans \mathcal{G}_i . On dit que le couple (\mathcal{G}_n, P_n) satisfait la *loi Zéro-Un* si

$$\lim_{n \rightarrow \infty} \Pr [\mathcal{G}_n \models P_n] = 0 \text{ ou } 1. \quad (5.1)$$

Que peut-on dire lorsque ces évènements sont des propriétés de la logique du premier ordre ? Et comme si cela ne risque pas de nous surprendre, le jeu d'Ehrenfeucht-Fraïssé fournit un outil qui permet d'identifier ces gros espaces qui rendent les évènements « du premier ordre » raisonnables.

Théorème 28. *Soient $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$ une suite de graphes aléatoires et P une propriété de graphe en logique du premier ordre. Alors (\mathcal{G}_n, P) satisfait la loi Zéro-Un si et seulement si pour tout $r \geq 1$*

$$\lim_{n, m \rightarrow \infty} \Pr[\mathcal{G}_n \equiv_r \mathcal{G}_m] = 1. \quad (5.2)$$

Démonstration. Supposons que (5.1) ne tient pas pour une certaine propriété P . Il existe alors $\epsilon > 0$ tel que pour tout entier N ,

- $m > N$ implique $\Pr[\mathcal{G}_m \models P] > \epsilon$, sinon la limite sera nulle ;
- $n > N$ implique $\Pr[\mathcal{G}_n \models P] < 1 - \epsilon$, sinon la limite sera 1.

La probabilité que $\mathcal{G}_m \models P$ et que $\mathcal{G}_n \models \neg P$ est donc au moins ϵ^2 puisque les deux évènements sont indépendants. Or si tel est le cas, par le théorème 27, $\text{EHR}(\mathcal{G}_m, \mathcal{G}_n, \text{qd}(P))$ est saboteur gagnant et donc (5.2) ne tient pas.

Réciproquement, supposons (5.1) tient pour toute propriété du premier ordre. Pour $r \geq 1$ arbitraire mais fixé, par le théorème 26 on sait que pour chaque classe d'équivalence $[G]$ de \equiv_r il existe une propriété du premier ordre P_G qui la caractérise. Les P_G partitionnent l'espace de probabilité et

$$\sum_{[G]} \Pr[\mathcal{G}_n \models P_G] = 1. \quad (5.3)$$

Posons $p_G = \lim_{n \rightarrow \infty} \Pr[\mathcal{G}_n \models P_G]$. Les classes d'équivalence n'étant qu'en nombre fini, on peut porter l'équation 5.3 à la limite et obtenir $\sum_{[G]} p_G = 1$. Or par hypothèse, $p_G = 0$ ou 1. Et comme les classes d'équivalence sont deux à deux disjointes, exactement une seule des p_G est 1. Si $[G_0]$ est cette classe alors pour tout $\epsilon > 0$, il existe N tel que pour tout $n > N$, la probabilité $\Pr[\mathcal{G}_n \models P_{G_0}] = 1 - \epsilon$. Et donc pour $n, m > N$, la probabilité que $\mathcal{G}_n \in [G_0]$ et $\mathcal{G}_m \in [G_0]$ est $(1 - \epsilon)^2$. Comme $\epsilon \rightarrow 0$,

alors $\lim_{n,m \rightarrow \infty} \Pr [\mathcal{G}_n \equiv_r \mathcal{G}_m] = 1$. \square

Une application directe de ce théorème est le cas où $\mathcal{G}_n = G(n, 1/2)$. Dans cet espace de probabilité, tous les graphes sont équiprobables, si bien que $\Pr [\mathcal{G}_n \models P]$ donne exactement la proportion de graphes d'ordre n qui satisfont P . On sait déjà que $\lim_{n,m \rightarrow \infty} \Pr [\mathcal{G}_n \cong \mathcal{G}_m] = 1$ alors $\lim_{n,m \rightarrow \infty} \Pr [\mathcal{G}_n \equiv_r \mathcal{G}_m] = 1$ pour tout $r \geq 1$. Donc, pour toute propriété P du premier ordre, soit tous les graphes ont P , soit aucun ne la satisfait lorsque l'ordre du graphe grandit indéfiniment. C'est le résultat bien connu de Fagin [44].

Corollaire 6. Soit $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$ une suite de graphes aléatoires où \mathcal{G}_i est d'ordre i telle que

$$\lim_{n \rightarrow \infty} \Pr [\mathcal{G}_n \text{ est } \aleph_0\text{-catégorique}] = 1$$

alors (\mathcal{G}_n, P) satisfait la loi Zéro-Un pour toute propriété du premier ordre P .

Démonstration. Pour tout $\epsilon > 0$, il existe N tel que si $n > N$ alors

$$\Pr [\mathcal{G}_n \text{ est } \aleph_0\text{-catégorique}] = 1 - \epsilon.$$

Donc pour $n, m > N$, $\Pr [\mathcal{G}_n \cong \mathcal{G}_m] = \Pr [\mathcal{G}_n \equiv_r \mathcal{G}_m] = 1 - \epsilon$ pour tout $r \geq 1$. \square

La loi Zéro-Un n'est pas une loi triviale dans ces espaces de probabilités. En effet, une propriété qui ne satisfait pas la loi Zéro-Un dans la suite d'espaces de probabilité $G(n, 1/2)$ est par exemple $P = “|E(G)| \text{ est pair}”$, puisque la moitié de l'espace a un nombre pair d'arêtes. Cette propriété n'est certainement pas du premier ordre. Si nous prenons $\mathcal{G}_n = G_n(2n)$ alors la loi Zéro-Un est respectée.

Revenons à la propriété S_k . On sait qu'elle vérifie la loi Zéro-Un si on considère les espaces $G(n, p)$ pour $0 < p < 1$ fixé. Bien que S_k soit bien tranchée pour les graphes ultrahomogènes connexes finis, une suite d'espaces de probabilité pour laquelle presque tous les graphes sont ultrahomogènes connexes ne constitue pas obligatoirement un terrain dans lequel S_k satisfait la loi Zéro-Un. En effet, prenons un modèle \mathcal{G}_n de graphe aléatoire d'ordre n pour le graphe ultrahomogène infini

dénombrable G_{k+1} n'admettant pas K_{k+1} du chapitre précédent. Une suite d'espaces de probabilité qui contiennent, et ce de façon équiprobable, les graphes de \mathcal{G}_n et leurs compléments ne satisfait pas la loi Zéro-Un pour S_k puisque la moitié des graphes satisfont S_k et le reste non.

5.4 Seuil

Revenons à l'évolution des graphes aléatoires $G(n, p)$. A quel moment dans leur évolution est-ce que la propriété S_k apparaît? Les propriétés en logique du premier ordre va nous servir de prototype dans cette quête. Par le théorème 28, si $G(n, p(n))$ satisfait (AR) alors on a un phénomène de loi Zéro-Un non seulement pour les propriétés en logique du premier ordre mais aussi pour S_k . Il se trouve alors [17] que

Théorème 29. *Soit $p(n)$ telle que pour tout $\epsilon > 0$, $n^{-\epsilon} \ll p(n)$ et $n^{-\epsilon} \ll 1 - p(n)$, alors presque tous les graphes de $G(n, p(n))$ satisfont (AR).*

En ce qui nous concerne donc, si $p(n)$ ne tend pas vers 0 plus vite que n'importe $n^{-\epsilon}$ et $p(n)$ ne tend pas vers 1 plus vite que $1 - n^{-\epsilon}$, alors $G(n, p)$ est voleur gagnant. D'autre part, on sait aussi que pour $p(n) \ll n^{-2}$, $G(n, p)$ est vide et donc voleur gagnant. Par conséquent, pour $1 - p(n) \ll n^{-2}$, $G(n, p)$ est complet et donc 1-policier gagnant. Si $G(n, p)$ devait satisfaire S_k avec $k \geq 2$, alors cela doit se faire lorsque $p(n)$ tend vers 1 plus vite que $1 - n^{-\epsilon}$ pour $0 < \epsilon < 2$ mais moins vite que $1 - n^{-2}$. Ce qui arrive quand même assez tard dans l'évolution et de surcroît sur un petit laps de « temps ».

La probabilité que $G(n, p)$ possède un ensemble dominant de taille k est

$$\binom{n}{r} (1 - (1 - p)^r)^{n-r} \geq (1 - (1 - p)^r)^{n-r}$$

Cette probabilité tend vers 1 pour $p(n) = 1 - n^{-\epsilon}$ avec $\frac{1}{k} < \epsilon$. Toutefois, cela ne veut pas forcément dire que $k - 1$ policiers ne suffisent pas pour gagner.

CONCLUSION

Dans ce mémoire, nous avons souvent flirté avec l'infini ; ce qui peut constituer une contradiction avec l'aspect pratique que représente le jeu de poursuite. La théorie des graphes aléatoires est étrange dans ce sens puisque l'étude d'une propriété requiert souvent ce passage à la limite. Le choix du graphe du Web pour amorcer ce travail représente notre volonté de renouer avec le côté pratique du jeu. Seulement, l'étude du comportement asymptotique de ses modèles ne nous a pas laissé le choix que d'y rester jusqu'à en prendre goût. D'ailleurs, la majeure partie de notre contribution réside dans la considération des graphes infinis dans le jeu de poursuite et ce sous la loupe de la théorie des graphes aléatoires.

Modéliser le web constitue un domaine où la recherche reste très active. D'ailleurs, $G(n, p)$ le premier modèle proposé à cet effet n'a pas fait le poids devant les mesures prises directement du web, notamment en ce qui concerne la distribution des degrés. Contrairement à notre impression au départ, la loi de puissance que l'on observe dans les données réelles n'est pas celle qui est la plus difficile à prendre en compte. Le défi réside dans la compréhension de la sociologie de la création de contenu. Le modèle de l'attachement préférentiel va d'ailleurs dans cette direction en incorporant le facteur préférence. Ce dernier n'explique toutefois pas le phénomène des communautés perçues sur le web dont une des caractéristiques est la copie de liens existants. L'analyse d'un modèle n'est pas souvent chose facile, surtout que la plupart sont décrites de manière stochastique. Un modèle est d'autant plus agréable à étudier qu'à manipuler lorsqu'il en découle des propriétés déterministes. L'approche de Bonato et Janssen [25] pour identifier une propriété faible d'adjacence sert de précurseur dans ce sens. Elle nous a permis par exemple de trouver une chaîne de graphes policiers gagnants mais dont la limite ne l'est pas. On pense entre autres à la limite du modèle de la copie évolutive.

L'incursion dans l'univers de l'infini est inévitable dans l'étude du web. La croissance à un rythme effréné de celui-ci constitue en dehors de la curiosité intellectuelle la principale raison. Un des aspects étonnants qui ressort de l'étude

asymptotique des modèles du web est que selon certaines valeurs initiales des paramètres, le modèle peut devenir totalement déterministe. L'exemple éloquent en est sans doute $G(n, p)$. On est tous d'accord à présent que ce dernier est loin d'être le modèle idéal du web, mais cette défaite de l'aléatoire face à l'infini fascine. Ne doit-on pas sursauter lorsqu'on entend que presque tous les graphes sont homogènes ? On a vu que cette propriété est très restrictive, la caractérisation de tels graphes le montre, et pourtant à grande échelle la proportion de graphes homogènes prend le dessus.

L'homogénéité est une des propriétés des structures très étudiées en théorie des modèles. En ce qui concerne le jeu de poursuite, ce fût contre toute attente que le complément des graphes de Henson \overline{H}_k soient $(k - 1)$ -policiers gagnants. C'est un bel exemple de graphes infinis connexes sur lesquels un nombre fini de policiers gagnent, tout en maintenant une propriété d'adjacence faible non triviale. Toutefois, c'est la présence d'un ensemble dominant de taille $k - 1$ qui en est la raison, ce qui les rend moins attrayant pour étudier le jeu de police-voleur. D'ailleurs, Dolinka et Mudrinski [60] montrent que la seule rétraction sur ces graphes est l'identité.

Que peut-on dire de la complexité d'un graphe k -policiers gagnant ? On peut donner un premier sens à cette question en considérant la longueur minimale d'une expression en logique du premier ordre équivalente à S_k . Il s'est avéré que celle-ci est forcément infinie et cette étude de la complexité descriptive du jeu de police-voleur constitue une partie essentielle de notre contribution. Une deuxième interprétation serait le moment où des graphes satisfaisant S_k apparaissent dans l'évolution de $G(n, p)$. Une propriété est de plus en plus complexe lorsqu'elle nécessite une probabilité d'une arête plus grande. Toutefois, lorsque p approche trop vite de 1, on perd en complexité puisque le graphe devient complet. On a donc vu que les graphes k -policiers gagnants arrivent relativement tard dans cette évolution et ce pendant un moment relativement court. On peut remarquer dans l'évolution de $G(n, p)$ que sur une très longue période, le graphe est homogène. En effet, il passe du vide à l'existentiellement clos pour aboutir au complet. Bien sûr, il y a un certain intervalle séparateur dans ces transitions, mais on se demande

si les graphes de Henson, et par conséquent son complément, n'apparaissent pas dans cette évolution. Si tel est le cas, $G(n, p)$ est peut-être toujours homogène tout au long de son évolution et on soupçonne que les graphes qui sont k -policiers gagnants contiennent pour la plupart un ensemble dominant de taille k ; comme le fait que $G(n, p)$ est existentiellement clos pour la plupart du temps est révélateur du fait que presque tous les graphes sont voleurs gagnants.

La théorie des graphes aléatoires n'a pas été d'un grand secours dans l'étude des graphes k -policiers gagnants comme ne l'a été la caractérisation des graphes k -policiers gagnants du théorème 11. La réduction k -policiers ver 1 policier dans [56] peut aider si on sait caractériser une version modifiée du jeu dans lequel le voleur est restreint à un sous-graphe induit. La question reste ouverte.

Une variation du jeu de police-voleur qui commence à faire son chemin consiste à limiter la portée des visions des policiers. Au lieu d'une vue parfaite qui s'étend sur tout le diamètre du graphe, chaque policier ne peut apercevoir que les sommets à une distance $0 \leq d \leq \text{diam}(G)$. Si $\text{sn}_d(G)$ est le nombre minimal de policiers pour gagner à ce jeu alors il est clair que $\text{sn}(G) \leq \text{sn}_d(G) \leq \text{sw}(G)$ pour tout $0 \leq d \leq \text{diam}(G)$. La version qui s'apparente le plus à la fouille est donc celle avec $d = 0$. Des résultats concernant $\text{sn}_d(G)$ pour $d = 0, 1$ sont rapportés dans [87] lorsque G est un arbre ou est unicyclique. Une question intéressante serait de savoir si le problème $\text{sn}_0(G) \leq k$ est NP-difficile. En d'autres termes, est-ce que la difficulté du problème de fouille réside dans la cécité des chercheurs ou plutôt dans la possibilité de mouvements de l'intrus? On peut essayer d'exprimer la propriété k -policiers gagnant pour $d = 0$ en logique du point fixe si c'est possible.

BIBLIOGRAPHIE

- [1] L. Adamic and B. Huberman, *Power-law distribution of the world wide web*, Science **287** (2000), no. 2115a.
- [2] M. Adler, H. Racke, and N. Sivadasan, *Randomized pursuit-evasion in graphs*, Proceedings of the International Colloquium on Automata, Languages and Programming (ICALP) (2002).
- [3] W. Aiello, F. Chung, and L. Lu, *A random graph model for massive graphs*, Proceedings of the 32nd annual ACM symposium on Theory of computing, 2000, pp. 171–180.
- [4] ———, *A random graph model for power law graphs*, Proc. 41st Symposium on Foundations of Computer Science (FOCS), 2000, pp. 171–180.
- [5] ———, *Random evolution in massive graphs*, Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science, 2001, pp. 510–519.
- [6] M. Aigner and M. Fromme, *A game of cops and robbers*, Discrete Applied Mathematics **8** (1984), no. 1, 1–12.
- [7] R. Albert, A. Barabási, and H. Jeong, *The diameter of the world wide web*, Nature **401** (1999), 130–131.
- [8] R. Albert, A. L. Barabási, and H. Jeong, *Scale-free characteristics of random networks : The topology of the world wide web*, Physica A **281** (2000), 69–77.
- [9] R. Albert, H. Jeong, and A. Barabasi, *Diameter of the world wide web*, Nature **401** (1999), 130–131.
- [10] B. Alspach, X. Li, and B. Yang, *Searching graphs and directed graphs*, University of Regina (2004).
- [11] R. Anstee and M. Farber, *On bridged graphs and cop-win graphs*, J. Comb. Theory (Series B) **44** (1988), 22–28.
- [12] J. Balogh, B. Bollobás, and D. Weinreich, *Measures on monotone properties of graphs*, Discrete Applied Mathematics **116** (2002), no. 1-2, 17–36.

- [13] J. Bang-Jensen and G. Gutin, *Digraphs : Theory, algorithms and applications*, Springer, 2001.
- [14] A. L. Barabási, R. Albert, and H. Jeong, *Mean-field theory for scale-free random networks*, *Physica A* (1999), no. 272, 173–187.
- [15] L. Barrière, P. Flocchini, P. Fraigniaud, and N. Santoro, *Capture of an intruder by mobile agents*, Proceedings of the fourteenth annual ACM symposium on Parallel algorithms and architectures, ACM Press, 2002, pp. 200–209.
- [16] B. Bollobás, *The evolution of random graphs*, Transactions of the American Mathematical Society **286** (1984), no. 1, 257–274.
- [17] ———, *Random graphs*, Academic Press Inc., London, 1985.
- [18] ———, *Modern graph theory*, Springer-Verlag, 1998.
- [19] B. Bollobás and O. Riordan, *The diameter of a scale-free random graph*, *Combinatorica* **24** (2004), no. 1, 5–34.
- [20] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnady, *The degree sequence of a scale-free random graph process*, *Random Structures and Algorithms* **18** (2001), 279–290.
- [21] A. Bonato and K. Cameron, *On an adjacency property of almost all graphs*, *Discrete Mathematics* **231** (2001), 103–119.
- [22] A. Bonato, P. J. Cameron, D. Deli, and S. Thomassé, *Generalized pigeonhole properties of graphs and oriented graphs*, *European Journal of Combinatorics* **23** (2002), no. 3, 257–274.
- [23] A. Bonato and D. Delić, *A note on orientations of the infinite random graph*, *European Journal of Combinatorics* **25** (2004), 921–926.
- [24] A. Bonato and J. Janssen, *Infinite limits of copying models of the web graph*, *Internet Mathematics* **1** (2003), 193–213.
- [25] ———, *Limit behaviour of models for the web graph and other self-organizing networks*, citeseer.ist.psu.edu/661533.html, 2004.

- [26] R. Breisch, *An intuitive approach to speleotopology*, Southwestern Covers VI (1967), no. 5, 72–78.
- [27] G. Brightwell and P. Winkler, *Gibbs measures and dismantlable graphs*, J. Comb. Theory (Series B) **78** (2000), 141–166.
- [28] A. Broder, R. Krauthgamer, and M. Mitzenmacher, *Improved classification via connectivity information*, Symposium on Discrete Algorithms, 2000, pp. 576–585.
- [29] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, *Graph structure in the web : Experiments and models*, Proceedings of the 9 International World Wide Web Conference, 2000.
- [30] G. Caldarelli, P. De Los Rios, L. Laura, S. Leonardi, and S. Millozzi, *A study of stochastic models for the web graph*, Technical Report 04-03, dipartimento di Informatica e Sistemistica, Universita' di Roma la Sapienza", (2003).
- [31] P. J. Cameron, *The random graph*, The mathematics of Paul Erdős (1996), 331–351.
- [32] N. E. Clarke, *A game of cops and robber played with partial information*, 2004.
- [33] R. Cohen and S. Havlin, *Scale-free networks are ultrasmall*, Physical Review Letters **90** (2003), no. 5.
- [34] C. Cooper and A. M. Frieze, *A general model of the web graph*, Proceedings of ESA (2001), 500–511.
- [35] M. D. Davis, *Game theory*, Basic Books Inc, 1970.
- [36] N. D. Dendris, L. M. Kirousis, and D. M. Thilikos, *Fugitive-search games on graphs and related parameters*, Theoretical Computer Science **172** (1997), no. 1-2, 233–254.
- [37] A. Devillers, *Classification of some homogenous and ultrahomogenous structures*, Ph.D. thesis, Université libre de Bruxelles, 2002.
- [38] S. Dill, S. R. Kumar, K. S. McCurley, S. Rajagopalan, D. Sivakumar, and A. Tomkins, *Self-similarity in the web*, Proceedings of 27th International Conference on Very Large Data Bases, 2001, pp. 69–78.

- [39] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Structure of growing networks with preferential linking*, Physical Review Letters **85** (2000), no. 21, 4633–4636.
- [40] M. El-Zahar and N. Sauer, *The indivisibility of the k_n -free graphs*, Journal of combinatorial theory **B** (1989), no. 42, 161–170.
- [41] J. A. Ellis, I. H. Sudborough, and J. S. Turner, *The vertex separation and search number of a graph*, Information and Computation **113** (1994), no. 1, 50–79.
- [42] P. Erdős and A. Rényi, *On the evolution of random graphs*, Publications of the Math. Inst. of the Hungarian Academy of Science **5** (1960), 17–61.
- [43] R. Fagin, *Generalized first-order spectra and polynomial-time recognizable sets*, Complexity of Computation (R. M. Karp, ed.), SIAM-AMS, vol. 7, 1974, pp. 43–73.
- [44] ———, *Probabilities in finite models*, Journal of Symbolic Logic **41** (1976), no. 1, 50–58.
- [45] C. Faloutsos, M. Faloutsos, and P. Faloutsos, *On power-law relationships of the internet topology*, Proceedings of ACM SIGCOMM, 1999.
- [46] F. V. Fomin and P. A. Golovach, *Graph searching and interval completion*, SIAM Journal on Discrete Mathematics **13** (2000), no. 4, 454–464.
- [47] F. V. Fomin, D. Kratsch, and H. Müller, *On the domination search number*, Discrete Appl. Math. **127** (2003), no. 3, 565–580.
- [48] P. Frankl, *On a pursuit game on cayley graphs*, Combinatorica **7** (1987), no. 1, 67–70.
- [49] R. Fraïssé, *Theory of relations*, Studies in logic and the foundations of mathematics, vol. 118, North-Holland, 1986.
- [50] E. Friedgut and J. Bourgain, *Sharp thresholds of graph properties, and the k -sat problem*, Journal of the American Mathematical Society **12** (1999), no. 4, 1017–1054.

- [51] E. Friedgut and G. Kalai, *Every monotone graph property has a sharp threshold*, Proceedings of the American Mathematical Society, vol. 124, 1996, pp. 2993–3002.
- [52] A. Gardiner, *Homogenous graphs*, Journal of combinatorial theory **B** (1976), no. 20, 94–102.
- [53] A. Goldstein and E. Reingold, *The complexity of pursuit on a graph*, Theoretical Computer Science (Math Games) **143** (1995), 93–112.
- [54] L. Guibas, J. C. Latombe, S. M. LaValle, D. Lin, and R. Motwani, *A visibility-based pursuit-evasion problem*, International Journal of Computational Geometry and Applications **9** (1999), no. 4-5, 471–493.
- [55] G. Hahn, F. Laviolette, N. Sauer, and R. E. Woodrow, *On cop-win graphs*, Discrete Mathematics **258** (2002), 27–41.
- [56] G. Hahn and G. MacGillivray, *A characterisation of k -cop-win graphs and digraphs*, à paraître dans Discrete Mathematics.
- [57] Y. O. Hamidoune, *On a pursuit game on cayley digraphs*, European journal of combinatorics **8** (1967), 289–295.
- [58] C. W. Henson, *A family of countable homogeneous graphs*, Pacific journal of mathematics **38** (1971), no. 1, 69–83.
- [59] J. P. Hespanha, H. J. Kim, and S. Sastry, *Multiple-agent probabilistic pursuit-evasion games*, Proceedings of the 38th IEEE Conference of Decision and Control, vol. 3, 1999, pp. 2432–2437.
- [60] I. Dolinka and N. Mudrinski, *On retracts of henson graphs and their complements*.
- [61] N. Immerman, *Languages that capture complexity classes*, SIAM Journal of Computing **16** (1987), no. 4, 760–778.
- [62] V. Isler, S. Kannan, and S. Khanna, *Randomized pursuit-evasion with limited visibility*, In Proc. of ACM-SIAM Symposium on Discrete Algorithms (SODA) (2004), 1053–1063.

- [63] H. J. Kim, S. Sastry, O. Shakernia, D. Shim, and R. Vidal, *Probabilistic pursuit-evasion games : Theory, implementation and experimental evaluation*, IEEE Transactions on Robotics and Automation (2002).
- [64] J. H. Kim, O. Pikhurko, J. Spencer, and O. Verbitsky, *How complex are random graphs in first order logic ?*, Random Structures and Algorithms **26** (2005), no. 1–2, 119–145.
- [65] N. G. Kinnersley, *The vertex separation number of a graph equals its path-width*, Information Processing Letters **42** (1992), no. 6, 345–350.
- [66] J. M. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, and A. S. Tomkins, *The Web as a graph : Measurements, models and methods*, Lecture Notes in Computer Science **1627** (1999), 1–18.
- [67] R. Kleinberg and J. Kleinberg, *Isomorphism and embedding problems for infinite limits of scale-free graphs*, Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms, January 2005.
- [68] P. G. Kolaitis, H. J. Prömel, and B. L. Rothschild, *k_{l+1} -free graphs : asymptotic structure and 0 – 1 law*, Transactions of the American Mathematical Society **303** (1987), no. 2, 637–671.
- [69] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal, *Stochastic models for the web graph*, Proceedings of the 41st Annual Symposium on Foundations of Computer Science, 2000, pp. 57–65.
- [70] ———, *The Web as a graph*, Proc. 19th ACM SIGACT-SIGMOD-AIGART Symp. Principles of Database Systems, PODS, ACM Press, 15–17 2000, pp. 1–10.
- [71] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins, *Trawling the web for emerging cyber-communities*, In Proceedings of the 8th World Wide Web Conference, May 1999, pp. 403–415.
- [72] J. Labelle and A. Mercier, *Introduction à l'analyse réelle*, 1993.

- [73] A. H. Lachlan and R. E. Woodrow, *Countable ultrahomogenous undirected graphs*, Transactions of the American Mathematical Society **262** (1980), no. 1, 51–94.
- [74] A. S. LaPaugh, *Recontamination does not help to search a graph*, Journal of the ACM **40** (1993), 224–245.
- [75] F. S. Makedon and I. H. Sudborough, *On minimizing width in linear layouts*, Discrete Applied Mathematics **23** (1989), 243–265.
- [76] N. Megiddo, S. L. Hakimi, M. R. Garey, D. S. Johnson, and C. H. Papadimitriou, *The complexity of searching a graph*, Journal of the ACM **35** (1988), no. 1, 18–44.
- [77] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Random graphs with arbitrary degree distribution and their applications*, Physical Review E **64** (2001), no. 026118.
- [78] R. Nowakowski and P. Winkler, *Vertex-to-vertex pursuit in a graph*, Discrete Math **43** (1983), 235–239.
- [79] D. Osthus, H. J. Prömel, and A. Taraz, *For which densities are random triangle-free graphs almost surely bipartite?*, Combinatorica **23** (2003), no. 1, 105–150.
- [80] G. Owen, *Game theory*, W. B. Saunders Company, 1968.
- [81] T. D. Parsons, *Pursuit evasion in a graph*, Theory and Applications of Graphs (D.R. Lick Y. Alavi, ed.), Springer, 1976, pp. 426–441.
- [82] H. J. Prömel, T. Schickinger, and A. Steger, *On the structure of clique-free graphs*, Random Structures and Algorithms **19** (2001), no. 1, 37–53.
- [83] A. Quilliot, *A short note about pursuit games played on a graph with a given genus*, Journal of Combinatorial Theory, Series B **38** (1985), no. 1, 89–92.
- [84] J. Sheehan, *Smoothly embeddable subgraphs*, J. London Math. Soc. (2) **9** (1974), 212–218.
- [85] J. Spencer, *The strange logic of random graphs*, vol. 22, Springer-Verlag, 2001.

- [86] I. Suzuki and M. Yamashita, *Searching for a mobile intruder in a polygonal region*, SIAM Journal of Computing **21** (1992), no. 5, 868–888.
- [87] A. Tang, *Zero-visibility cops and robber*, Master's thesis, Dalhousie University, 2004.
- [88] M. Vardi, *The complexity of relational query languages*, In Proceedings of 14th ACM Symposium on Theory of Computing, ACM Press, 1982, pp. 137–146.
- [89] Ben Wieland and Anant P. Godbole, *On the domination number of a random graph.*, Electronic Journal of Combinatorics **8** (2001), no. 1.
- [90] R. E. Woodrow, *There are four ultrahomogenous graphs without triangles*, Journal of combinatorial theory **B** (1979), no. 27, 168–179.