

Titre: Nouveaux résultats sur les arbres, forêts et forêts linéaires
Title: maximum ainsi que sur la distance moyenne dans un graphe.

Auteur: Rim Kilani
Author:

Date: 2010

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Kilani, R. (2010). Nouveaux résultats sur les arbres, forêts et forêts linéaires maximum ainsi que sur la distance moyenne dans un graphe. [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie. <https://publications.polymtl.ca/461/>

Document en libre accès dans PolyPublie

Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/461/>
PolyPublie URL:

Directeurs de recherche: Alain Hertz, Pierre Hansen, & Odile Marcotte
Advisors:

Programme: Mathématiques de l'Ingénieur
Program:

UNIVERSITÉ DE MONTRÉAL

NOUVEAUX RÉSULTATS SUR LES ARBRES, FORÊTS ET FORÊTS LINÉAIRES
MAXIMUM AINSI QUE SUR LA DISTANCE MOYENNE DANS UN GRAPHE

RIM KILANI

DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(MATHÉMATIQUES DE L'INGÉNIEUR)
OCTOBRE 2010

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

NOUVEAUX RÉSULTATS SUR LES ARBRES, FORÊTS ET FORÊTS LINÉAIRES
MAXIMUM AINSI QUE SUR LA DISTANCE MOYENNE DANS UN GRAPHE

présentée par: KILANI Rim

en vue de l'obtention du diplôme de: Philosophiæ Doctor
a été dûment acceptée par le jury d'examen constitué de:

M. ROUSSEAU Louis-Martin, Ph.D., président

M. HERTZ Alain, Doct. ès Sc., membre et directeur de recherche

M. HANSEN Pierre, Doct. Agr., membre et codirecteur de recherche

Mme. MARCOTTE Odile, Ph.D, membre et codirecteur de recherche

M. MÉLOT Hadrien, Ph.D, membre

M. WALSH Timothy Robert, Ph.D, membre

REMERCIEMENTS

Je tiens à remercier sincèrement toutes les personnes qui m'ont aidée à accomplir ce travail.

Mon directeur de recherche, monsieur Alain Hertz pour son aide précieuse, son intérêt, ses corrections qui ont permis l'amélioration de ce travail.

Monsieur Pierre Hansen qui a eu l'idée du sujet de ma thèse, pour son intérêt et ses conseils pertinents.

Madame Odile Marcotte pour ses conseils, ses multiples corrections. De plus, grâce à elle j'ai eu l'opportunité de faire la connaissance de monsieur Pierre et monsieur Alain.

Mes directeurs m'ont guidé, appuyé, conseillé et soutenu financièrement tout au long de ces années. J'admire leur passion pour la recherche et j'apprécie énormément leurs qualités humaines. Merci infiniment.

Je remercie aussi,

David Schindl, son aide était indispensable pour la réalisation de ce travail.

Mustapha Aouchiche, pour sa disponibilité chaque fois que j'avais un problème avec AutoGraphiX.

Serge Bisaillon, pour son aide précieuse quand j'avais des “segmentation fault” dans mon programme.

Tout le personnel du GERAD (Groupe d'Études et de Recherche en Analyse des Décisions), en particulier Francine Benoit, Pierre Girard et tous mes amis au GERAD.

Le président, monsieur Louis-Martin Rousseau et les membres, monsieur Hadrien Mélot et monsieur Timothy Robert Walsh d'avoir accepté de faire partie du jury de ma thèse.

Mes parents, Youssef et Mounira, pour leurs encouragements tout au long de mes études,

mon mari, Faouzi, pour ses encouragements et sa patience, mon frère, toute ma famille en Tunisie et tous mes amis.

Finalement, ma fille Majda, ta joie de vivre illumine chaque instant de ma vie.

RÉSUMÉ

Cette thèse s'inscrit dans le domaine de la recherche scientifique assistée par l'ordinateur. Ce domaine est en pleine expansion depuis quelques années grâce aux récents développements des systèmes informatiques. En théorie des graphes plusieurs systèmes ont été développés afin de générer de nouvelles conjectures entre les différents invariants d'un graphe. Nous citons par exemple les systèmes AutoGraphiX, Graffiti et GraPHedron.

Après un premier chapitre consacré à l'introduction, nous proposons dans le deuxième chapitre une revue de littérature sur les théorèmes connus ainsi que les conjectures qui bornent inférieurement ou supérieurement des invariants de graphe.

Notre intérêt porte essentiellement sur la cardinalité des sous-graphes induits par un bistable, une forêt, une forêt linéaire et un arbre. Rappelons que ces problèmes sont NP-complets.

Fajtlowicz a posé, respectivement pendant les années 1986 et 1992, les conjectures suivantes (générées par Graffiti et listées, parmi plusieurs, dans le document *Written on the Wall*): $\mu(G) \leq \alpha(G)$ (WoW 2) et $\mu(G) \leq \alpha_2(G)/2$ (WoW 747). Où $\mu(G)$ dénote la distance moyenne d'un graphe connexe, $\alpha(G)$ le nombre de stabilité et $\alpha_2(G)$ la cardinalité du plus grand sous-graphe biparti induit.

Il est clair que la deuxième conjecture est plus forte que la première, qui a été prouvée par Chung. Une autre preuve de la première conjecture a été proposée par Dankelmann, qui a aussi donné une description du graphe G qui maximise la distance moyenne pour un ordre donné et une valeur de $\alpha(G)$ donnée.

Dans le troisième chapitre de cette thèse nous prouvons un résultat plus fort que la deuxième conjecture, c'est à dire que $\mu(G)$ est au plus $F(G)/2$, où $F(G)$ dénote *l'ordre maximum d'une forêt induite* de G (rappelons qu'une forêt est un graphe sans cycle).

C'est l'apport important de cette thèse.

Le quatrième chapitre est consacré à faire une étude d'invariants en utilisant le système AutoGraphiX. Nous allons déterminer de nouvelles bornes supérieures sur des différences d'invariants. Les relations que nous allons présenter ont la forme suivante:

$$INV_1(G) - INV_2(G) \leq f(n),$$

on note par $INV_i(G)$ un invariant du graphe G et $f(n)$ une fonction de l'ordre n .

Nous proposons dans le dernier chapitre trois algorithmes constructifs qui tentent de déterminer la plus grande forêt, la plus grande forêt linéaire et le plus grand arbre d'un graphe.

ABSTRACT

There are many discovery systems in graph theory that have been designed to generate new conjectures and explore graphs. For instance the systems AutoGraphiX, Graffiti and GraPHedron.

After the introduction, we present in the second chapter a survey on theorems and conjectures that give bounds on some graph invariants. We are concerned by the order of a largest induced subgraph (bipartite, forest, linear forest, tree). We recall that these problems are NP-hard.

With the help of the Graffiti system, Fajtlowicz conjectured around 1992 [36] that the average distance between two vertices of a connected graph G is at most half the maximum order of an induced bipartite subgraph of G , denoted $\alpha_2(G)$. We prove in the third chapter a strengthening of this conjecture by showing that the average distance between two vertices of a connected graph G is at most half the maximum order of an induced forest, denoted $F(G)$. It is the main result of this thesis. Finally, we conjecture that the average distance between two vertices of a connected graph is at most half the maximum order of an induced linear forest (where a linear forest is a union of paths). Moreover, we characterize the graphs maximizing the average distance among all graphs G having a fixed number of vertices and a fixed value of $F(G)$ or $\alpha_2(G)$.

We present in the next chapter new upper bounds on difference between some graph invariants. The form of these relations is

$$INV_1(G) - INV_2(G) \leq f(n),$$

where $INV_i(G)$ is an invariant of a graph G and $f(n)$ a function of the order n . We used the AutoGraphiX system to generate extremal graphs and find new conjectures.

The last chapter is reserved to an algorithm that tries to find the maximum order of an induced forest, an induced linear forest and an induced tree.

TABLE DES MATIÈRES

REMERCIEMENTS	iii
RÉSUMÉ	v
ABSTRACT	vii
TABLE DES MATIÈRES	ix
LISTE DES TABLEAUX	xii
LISTE DES FIGURES	xiv
CHAPITRE 1 INTRODUCTION	1
CHAPITRE 2 REVUE DE LA LITTÉRATURE	8
2.1 La distance moyenne	9
2.2 Le nombre de stabilité	12
2.3 La cardinalité de la plus grande forêt induite	15
2.4 La cardinalité du plus grand sous-graphe biparti induit	16
2.5 La cardinalité du plus grand arbre induit	17
2.6 Algorithmes de résolution du problème du tranversal de cycles	19
CHAPITRE 3 DISTANCE MOYENNE ET FORêt INDUIITE MAXIMUM .	28
3.1 Double cerf-volant et graphe forêt-critique	29

3.2	Transformations	38
3.3	Le théorème	52
3.4	Résultats connexes	60
 CHAPITRE 4 BORNES SUR DES DIFFÉRENCES ENTRE DES INVARIANTS		63
4.1	Borne pour $\text{Max}(\alpha_2(G) - F(G))$	66
4.1.1	Éditer un problème	67
4.2	Graphes vérifiant $\alpha_2(G) - F(G) = \lfloor \frac{n}{2} \rfloor - 1$	70
4.3	Borne pour $\text{Max}(\alpha_2(G) - LF(G))$	74
4.4	Borne pour $\text{Max}(F(G) - LF(G))$	74
4.5	Borne pour $\text{Max}(F(G) - T(G))$	76
4.6	Graphes vérifiant $\alpha_2(G) = F(G) = LF(G)$	77
4.7	Borne pour $\text{Max}\left(\frac{\alpha_2(G)}{2} - \mu(G)\right)$	79
4.8	Borne pour $\text{Max}\left(\frac{F(G)}{2} - \mu(G)\right)$	84
4.9	Borne pour $\text{Max}\left(\frac{LF(G)}{2} - \mu(G)\right)$	88
 CHAPITRE 5 HEURISTIQUES CONSTRUCTIVES POUR DÉTERMINER LA PLUS GRANDE FORêt, LA PLUS GRANDE FORêt LINéaire ET LE PLUS GRAND ARBRE		93
5.1	Méthodes constructives pour la détermination de la plus grande forêt . .	93
5.2	Résultats expérimentaux	95

5.3 Algorithme constructif pour la détermination de la plus grande forêt linéaire	110
5.4 Algorithme constructif pour la détermination du plus grand arbre	113
CONCLUSION	116
REFERENCES	119

LISTE DES TABLEAUX

Tableau 2.1	. Bornes et graphes extrêmes	19
Tableau 4.1	. Récapitulation des résultats	92
Tableau 5.1	. Résultats obtenus avec le critère 1	96
Tableau 5.2	. Résultats obtenus avec le critère 2	97
Tableau 5.3	. Résultats obtenus avec le critère 3	98
Tableau 5.4	. Résultats obtenus avec le critère 4	99
Tableau 5.5	. Résultats obtenus avec le critère 5	100
Tableau 5.6	. Résultats obtenus avec le critère 6	101
Tableau 5.7	. Résumé des résultats pour les 25 instances	101
Tableau 5.8	. Résultats obtenus par les variantes 1, 2 et 3	105
Tableau 5.9	. Résumé des résultats	106
Tableau 5.10	. Résultats obtenus par les variantes 1 et 2	107
Tableau 5.11	. Résultats obtenus pour les hypercubes	109
Tableau 5.12	. Résultats obtenus par l'algorithme de recherche de la plus grande forêt linéaire induite	111
Tableau 5.13	. Résultats obtenus par l'algorithme de recherche de la plus grande forêt linéaire induite	112
Tableau 5.14	. Résultats obtenus par l'algorithme de recherche du plus grand arbre induit	113

Tableau 5.15	. Résultats obtenus par l'algorithme de recherche du plus grand arbre induit	114
Tableau 5.16	. Récapitulation des résultats	117

LISTE DES FIGURES

Figure 2.1	. Illustration des graphes Δ , Σ , H_k et A_k	20
Figure 3.1	. P_{k-2} avec $k = 6$	29
Figure 3.2	. Illustration d'un double cerf-volant non équilibré noté $G_{4,6,6}$. .	30
Figure 3.3	. Illustration d'un double cerf-volant non équilibré noté $G_{3,7,6}$. .	30
Figure 3.4	. Illustration d'un double cerf-volant équilibré noté $G_{5,6}$	30
Figure 3.5	. Illustration de la transformation no 2	53
Figure 3.6	. Illustration de la transformation no 3	53
Figure 3.7	. Illustration de la transformation no 4	54
Figure 3.8	. Un exemple illustrant les cinq transformations	59
Figure 3.9	. Les trois graphes (6,3)-forêt-extrêmes et leurs complémentaires	61
Figure 3.10	. Un graphe tel que $T(G) = 4$ et $\mu(G) > 2$	61
Figure 3.11	. Illustration des différences entre $\alpha(G)$, $\frac{\alpha_2(G)}{2}$, $\frac{F(G)}{2}$ et $\frac{LF(G)}{2}$. .	62
Figure 4.1	. Énoncé du problème	66
Figure 4.2	. Liste des graphes associés au problème $\text{Max}(\alpha_2(G) - F(G))$	68
Figure 4.3	. Affichage de la relation trouvée	69
Figure 4.4	. Graphes obtenus par AGX en maximisant $\alpha_2(G) - F(G)$ pour $6 \leq n \leq 9$	69
Figure 4.5	. $\alpha_2(G) = 10, F(G) = 8$	72
Figure 4.6	. $\alpha_2(G) = 11, F(G) = 9$	73

Figure 4.7	. $\alpha_2(G) = 10, F(G) = 8$	73
Figure 4.8	. Graphe d'ordre 17 obtenu par AGX en maximisant $F(G) - LF(G)$. On a $F(G) = 17, LF(G) = 13$ et $F(G) - LF(G) = \left\lfloor \frac{17}{4} \right\rfloor = 4$	76
Figure 4.9	. Graphes obtenus par AGX en maximisant $F(G) - T(G)$. . .	77
Figure 4.10	. Graphes obtenus par AGX en maximisant $\alpha_2(G) - \mu(G)$. . .	79
Figure 4.11	. Graphes obtenus par AGX en maximisant $LF(G) - \mu(G)$. .	88
Figure 5.1	. Courbes représentatives des résultats obtenus par les six critères	104
Figure 5.2	. Courbes représentatives des trois variantes	106
Figure 5.3	. Illustration des hypercubes: Q_1, Q_2 et Q_3	108

CHAPITRE 1

INTRODUCTION

«Devant un grand nombre de situations, le mathématicien, comme d'ailleurs le sociologue, l'économiste, le planificateur ou le chimiste a été amené à tracer sur le papier des points (représentant des individus, des nombres, des localités, des opérations, des molécules chimiques, etc.) et des lignes continues reliant certains d'entre eux. Pour raisonner sur de tels schémas indépendamment de leur signification, le mathématicien a convenu d'appeler ces points des sommets et ces lignes des arêtes .» (Claude Berge [64]).

La théorie des graphes est peut-être apparue au dix-huitième siècle, grâce au travail d'Euler, avec le célèbre problème des sept ponts de Königsberg. En effet les habitants de cette ville se demandaient s'ils pouvaient traverser tous les ponts sans passer deux fois par le même pont et revenir par la suite à leur point de départ.

Un des problèmes les plus connus en théorie des graphes est celui du commis-voyageur. Ce problème consiste à trouver la plus courte tournée pour parcourir un ensemble de villes, en ne passant qu'une seule fois par chaque ville. Les applications récentes en théorie des graphes touchent plusieurs domaines, par exemple les transports, les réseaux de communication, la gestion de la production, la chimie, l'économie, etc.

Les graphes constituent donc un moyen simple et intuitif pour modéliser une grande variété de problèmes. Ce domaine s'est développé avec les travaux de plusieurs chercheurs et mathématiciens, par exemple König [13], Menger [1], Cayley [13], Berge[12], Lovász [57] et Erdős [34].

Aujourd’hui la théorie des graphes trouve plusieurs applications dans le domaine de l’informatique. Rappelons qu’en 1976, la preuve du théorème des quatre couleurs a demandé pour la première fois dans l’histoire l’aide d’un ordinateur pour étudier les 1936 cas critiques. En effet grâce à l’émergence de puissants systèmes informatiques, la découverte scientifique assistée par l’ordinateur a connu de nombreux succès.

D’autre part, la génération des conjectures en théorie des graphes peut être assistée par l’ordinateur. Les systèmes Graffiti [30], AutoGraphiX [6][18][19], INGRID [16], GRAPH [23] et GraPHedron [60], entre autres, ont été mis au point afin de générer des conjectures de manière automatisée.

D’après P. Hansen *et al.* [48], les principes sur lesquels se basent ces systèmes informatiques sont les suivants: l’énumération d’un ensemble de graphes ayant certaines propriétés et la manipulation des invariants; en effet chaque système contient une liste d’invariants, ce qui permet un calcul rapide lors de l’exploration interactive des caractéristiques des graphes.

Ces systèmes sont aussi susceptibles de générer des inégalités simples entre les invariants. Par exemple, soit les invariants i_l, i_k, i_m, i_n ; les relations générées peuvent avoir les formes suivantes:

$$i_l \leq i_k, i_l \leq i_k + i_m \text{ ou } i_l + i_k \leq i_m + i_n.$$

Les systèmes peuvent aussi ramener un problème de théorie de graphes à un problème d’optimisation combinatoire.

Le système Graffiti a été développé en 1986 par Siemion Fajtlowicz de l’Université de Houston. Il utilise une base de données de graphes connexes et génère des inégalités entre les différents invariants d’un graphe tels que l’ordre, la distance moyenne, le rayon, le diamètre, le nombre de stabilité, etc.

Au total Graffiti a engendré plus de 900 conjectures. Des contre-exemples ont été trouvés pour certaines de ces conjectures et d'autres ont été prouvées. Les conjectures de Graffiti sont à l'origine de plus de quatre-vingts articles ou thèses de doctorat. L'ensemble de ces conjectures a été répertorié par S. Fajtlowicz dans le recueil «Written on the Wall». Ce recueil est régulièrement mis à jour et il est disponible sur Internet [30].

Le système AutoGraphiX a été développé en 1997 par Gilles Caporossi et Pierre Hansen (GERAD et École des Hautes Études Commerciales de Montréal). Ce système utilise des outils d'optimisation combinatoire, particulièrement la méta-heuristique de recherche à voisinage variable.

Le système AutoGraphiX est capable de trouver des graphes extrêmes, des graphes satisfaisant un certain nombre de contraintes et des contre-exemples à des conjectures. Il peut aussi suggérer de nouvelles conjectures. D'ailleurs AutoGraphiX a trouvé des contre-exemples à certaines des conjectures de Graffiti et a renforcé d'autres conjectures [4][6].

À l'aide du système Graffiti, Fajtlowicz a posé la conjecture suivante, qui porte le numéro 2 dans le document *Written on the Wall*: «Dans tout graphe connexe, la distance moyenne est inférieure ou égale au nombre de stabilité.»

Une conjecture plus faible a d'abord été prouvée par S. Fajtlowicz et W. Waller [35]: «Dans tout graphe connexe, la distance moyenne est inférieure ou égale au nombre de stabilité plus un.»

Par la suite Fan Chung [22] a réussi à prouver la conjecture numéro 2 et plus tard Peter Dankelmann [24] a aussi prouvé la conjecture 2 avec une approche différente de celle de Fan Chung. En effet Dankelmann donne la description du graphe G qui maximise la distance moyenne pour un ordre donné et une valeur du nombre de stabilité donnée.

En 1992, une autre conjecture plus forte que le théorème de Fan Chung a été générée par Graffiti. Cette conjecture porte le numéro 747 dans le recueil de S. Fajtlowicz [30] «Written on the Wall»: «Soit b l'ordre du plus grand sous-graphe induit biparti d'un graphe connexe G . Alors la distance moyenne entre les sommets de G est inférieure ou égale à $b/2$.»

La première motivation du présent travail a été d'étudier cette conjecture qui n'avait jamais été prouvée. De plus comme nous savons que le plus grand sous-graphe biparti induit a une cardinalité supérieure ou égale à la plus grande forêt induite (rappelons qu'une forêt est un graphe sans cycle), nous avons essayé d'améliorer la borne sur la distance moyenne et nous avons posé la conjecture suivante: »: «Soit F l'ordre de la plus grande forêt induite d'un graphe connexe G . Alors la distance moyenne entre les sommets de G est inférieure ou égale à $F/2$.»

Ce travail est organisé de la manière suivante. Le deuxième chapitre est consacré à la revue de littérature, le troisième chapitre est réservé à la preuve d'un résultat plus fort que la conjecture 747. En effet nous allons prouver que pour tout graphe connexe G la distance moyenne entre les sommets de G est inférieure ou égale à $F/2$. Dans le quatrième chapitre nous présentons des bornes supérieures sur des différences d'invariants ainsi que des graphes extrêmes générés par le système AutoGraphiX. Nous présentons dans le cinquième chapitre des algorithmes constructifs qui tentent de déterminer la plus grande forêt induite d'un graphe donné ainsi que la plus grande forêt linéaire induite et le plus grand arbre induit.

L'intérêt que suscite l'étude des conjectures obtenues avec l'assistance de l'ordinateur se résume bien par l'observation suivante: «Les théorèmes célèbres ont d'abord été des conjectures, ne fût-ce que dans l'esprit de ceux qui les ont obtenus» [7].

Nous rappelons dans ce qui suit quelques notions et définitions.

Définitions et notations

Soit $G = (V, E)$ un graphe simple non orienté où V est l'ensemble des sommets et E l'ensemble des arêtes. L'ordre de G est la cardinalité de V . Deux sommets u et v sont *adjacents* si $\{u, v\}$ (ou uv) appartient à E .

Un chemin de longueur ℓ entre u et v est une suite $(u_0 = u, u_1, \dots, u_\ell = v)$ de sommets distincts tels que $u_i u_{i+1} \in E$ pour tout $i \in \{0, 1, \dots, \ell - 1\}$.

Un cycle C de longueur ℓ est une suite $(u_0, u_1, \dots, u_{\ell-1})$ de sommets distincts telle que $u_i u_{i+1} \in E$ pour tout $i \in \{0, 1, \dots, \ell - 1\}$ (où l'addition est modulo ℓ). C est un cycle pair (resp. cycle impair) si ℓ est pair (resp. impair).

La distance entre u et v dans G , qu'on dénote $d_G(u, v)$, est la longueur d'un des plus courts chemins entre u et v . Un graphe G est connexe si pour toute paire de sommets u et v de G , il existe un chemin entre u et v .

Définition 1.0.1. *Un sous-graphe induit de G est un graphe $G' = (V', E')$ tel que V' est un sous-ensemble de sommets de V et E' est l'ensemble de toutes les arêtes qui relient les sommets de V' dans G .*

Définition 1.0.2. *On note $\alpha(G)$ la taille maximale d'un stable de G , c'est-à-dire le nombre maximum de sommets de G qui sont deux à deux non adjacents.*

Définition 1.0.3. *Un graphe biparti G est un graphe dans lequel les sommets peuvent être répartis en deux groupes X et Y et toute arête de G relie un sommet de X à un sommet de Y . On note $\alpha_2(G)$ l'ordre du plus grand sous-graphe biparti induit, appelé aussi *bistable*.*

Propriété 1.0.4. *Pour tout graphe G , on a*

$$\frac{\alpha_2(G)}{2} \leq \alpha(G).$$

Définition 1.0.5. *Un arbre induit est un sous-graphe induit connexe ne contenant aucun cycle. On note par $T(G)$ l'ordre du plus grand arbre induit de G .*

Définition 1.0.6. *Une forêt induite est un sous-graphe induit ne contenant aucun cycle. On note $F(G)$ l'ordre de la plus grande forêt induite.*

Définition 1.0.7. *Une forêt linéaire est une forêt dont tous les sommets ont un degré inférieur ou égal à 2. On note $LF(G)$ l'ordre de la plus grande forêt linéaire.*

Définition 1.0.8. *Soit $d_G(u, v)$ la distance entre les sommets u et v dans G ; alors le diamètre du graphe est défini comme $D(G) = \max_{u, v \in V} d_G(u, v)$.*

Définition 1.0.9. *L'excentricité d'un sommet v de G est définie par:*

$$ecc(v) = \max \{d_G(u, v) / u \in V\}.$$

Définition 1.0.10. *Le rayon de G est le minimum des excentricités, c'est-à-dire*

$$r(G) = \min \{ecc(v) / v \in V\}.$$

Définition 1.0.11. *On note $\mu(G)$ la distance moyenne du graphe G , c'est-à-dire que*

$$\mu(G) = \frac{2}{n(n-1)} \left(\sum_{u, v \in V(G), u \neq v} d_G(u, v) \right).$$

Définition 1.0.12. *On note $\sigma_G(u)$ la transmission d'un sommet u de G , c'est-à-dire que*

$$\sigma(u) = \left(\sum_{v \in V(G), u \neq v} d_G(u, v) \right).$$

Propriété 1.0.13. *Pour tout graphe connexe G , on a*

$$\alpha(G) \geq \frac{\alpha_2(G)}{2} \geq \frac{F(G)}{2}.$$

Preuve. Soit B un bistable maximum de G de taille $\alpha_2(G)$, tel que $B = A_1 \cup A_2$. Sans perte de généralités on peut supposer que $|A_1| \geq |A_2|$, alors le stable maximum de G est au moins égal à $|A_1|$, d'où $\alpha(G) \geq \frac{\alpha_2(G)}{2}$. Comme nous l'avons défini précédemment une forêt induite est un sous-graphe induit ne contenant aucun cycle, donc elle est colorable en deux couleurs, d'où $\alpha_2(G) \geq F(G)$.

Propriété 1.0.14. *Pour tout graphe connexe G , on a*

$$F(G) \geq T(G) \text{ et } F(G) \geq LF(G).$$

CHAPITRE 2

REVUE DE LA LITTÉRATURE

Depuis une vingtaine d'années, plusieurs logiciels ont été mis au point afin de générer des conjectures sur les graphes. La construction des conjectures se fait automatiquement ou interactivement. Un des programmes importants est le système Graffiti; ce dernier a donné lieu à plus de 900 conjectures.

Les résultats de Graffiti ont suscité l'intérêt de grands chercheurs en théorie des graphes, parmi lesquels nous citons B. Bollobás [14], F. Chung [22], L. Lovász [55], P. Hansen [6][46] et O. Favaron [37][38].

Un grand nombre de conjectures peuvent donc être proposées à la communauté des chercheurs. Dès lors la question qui se pose est la suivante: comment peut-on définir une bonne conjecture?

P. Hansen *et al.* [7] se sont déjà posés cette question. Ils ont proposé quelques critères qui peuvent définir une conjecture intéressante, par exemple la simplicité; en effet les théorèmes célèbres en algèbre et en géométrie sont les plus simples. De plus un résultat attrayant doit forcément contenir des concepts importants de la théorie des graphes tels que la connexité, la coloration, la stabilité, etc. Selon Conway [30] une bonne conjecture doit être surprenante, par exemple une relation entre des invariants qui intuitivement ne sont pas reliés. Un autre critère d'une bonne conjecture est que cette dernière soit la réponse à un problème difficile.

Dans ce qui suit, nous allons passer en revue des théorèmes connus ainsi que des conjectures qui bornent inférieurement ou supérieurement l'un des cinq invariants,

à savoir la distance moyenne, le stable maximum, le bistable maximum, la forêt induite maximale et l'arbre induit maximum.

Dans un deuxième temps, nous allons présenter une synthèse sur les algorithmes qui ont été proposés pour résoudre le problème du complémentaire de la forêt induite maximum. Ce problème revient à déterminer le nombre minimum de sommets à enlever pour supprimer tous les cycles d'un graphe.

Rappelons que dans la plupart des cas, les problèmes de recherche des sous-graphes induits appartiennent à la classe des problèmes NP-difficiles; en effet il semble qu'il n'existe pas d'algorithmes efficaces pour les résoudre.

2.1 La distance moyenne

La distance moyenne a été tout d'abord introduite par Harold Wiener en 1947 [69]. Cet invariant a été alors utilisé dans des applications en chimie. En effet Harold Wiener a considéré l'invariant $W(G) = \frac{1}{2} \sum d_G(u, v)$ représentant la somme de toutes les distances entre deux atomes de carbone dans une molécule (d'où son appellation d'indice de Wiener). La distance moyenne d'un graphe a été présentée pour la première fois par Hosoya en 1971 [51].

En étudiant la conjecture 747 générée par Graffiti, nous essayons de trouver une nouvelle borne sur la distance moyenne entre les sommets d'un graphe non orienté. Beaucoup d'auteurs ont essayé de borner la distance moyenne, citons par exemple J. Plesnik [65], qui a trouvé une borne supérieure pour la distance moyenne d'un graphe G en fonction de l'ordre du graphe:

$$\mu(G) \leq \frac{n+1}{3}.$$

On a l'égalité si et seulement si le graphe G est un chemin d'ordre n .

P. Dankelmann [25] a trouvé une borne plus serrée que celle trouvée par J. Plesnik, en fonction de l'ordre du graphe G et du nombre de domination $\gamma(G)$. Ce nombre est la cardinalité minimale d'un sous-ensemble D de $V(G)$ tel que chaque sommet de G appartient à D ou est adjacent à au moins un sommet de D .

L'inégalité de Dankelmann est

$$\mu(G) \leq \begin{cases} \frac{n+1}{3} - \frac{(n-3\gamma(G))(n-3\gamma(G)+2)(2n+3\gamma(G)-7)}{6n(n-1)} & \text{si } n - \gamma(G) \text{ est pair} \\ \frac{n+1}{3} - \frac{(n-3\gamma(G))(n-3\gamma(G)+2)(2n+3\gamma(G)-7)-9(\gamma(G)-1)}{6n(n-1)} & \text{si } n - \gamma(G) \text{ est impair.} \end{cases}$$

DeLaViña *et al.* [27] ont aussi prouvé une relation entre la distance moyenne et le nombre de domination d'un graphe. En effet pour tout graphe G , on a

$$\gamma(G) \geq \frac{2}{3}\mu(G) - \frac{1}{3}.$$

Récemment P. Dankelmann [26] a trouvé une nouvelle borne pour la distance moyenne en fonction de l'ordre du graphe n , le nombre d'arêtes m et le degré minimum $\delta(G)$:

$$\mu(G) \leq \frac{\left(n - \sqrt{(2m - \delta(G)n)}\right)^2 \left(n + 2\sqrt{(2m - \delta(G)n)}\right)}{(\delta(G) + 1)n^2} + O(1).$$

P. Hansen et M. Aouchiche [8] ont effectué une étude systématique à l'aide du logiciel AutoGraphiX pour générer des conjectures qui relient la distance moyenne à plusieurs invariants d'un graphe connexe.

Ces relations ont la forme suivante:

$$\underline{b}_n \leq \mu \oplus i \leq \bar{b}_n,$$

qui est la Forme 1 d'AutoGraphiX, où l'invariant i est le degré minimum, le degré moyen, le degré maximum, le diamètre, le rayon, la distance moyenne entre sommets, la moyenne des excentricités, le nombre chromatique et le nombre de couplage. Le symbole \oplus représente l'une des opérations $-$, $+$, $/$, $*$.

Pour chaque relation, les auteurs donnent au moins une famille de graphes extrêmes pour lesquels la borne supérieure (resp inférieure) est atteinte. Les auteurs ont prouvé beaucoup de nouvelles relations, par exemple les propositions ci-dessous.

Proposition 2.1.1. *Soit G un graphe connexe d'ordre supérieur ou égal à 3. Alors nous avons*

$$D(G) - \mu(G) \leq \frac{2n - 4}{3}.$$

La borne est atteinte si et seulement si le graphe est un chemin.

Proposition 2.1.2. *Soit G un graphe connexe d'ordre supérieur ou égal à 3. Si r est le rayon du graphe, alors nous avons*

$$\frac{\mu(G)}{r(G)} \leq 2 - \frac{2}{n}.$$

La borne est atteinte si et seulement si le graphe est une étoile.

Proposition 2.1.3. *Soit G un graphe connexe d'ordre supérieur ou égal à 3. Si Δ est le degré maximum du graphe, alors nous avons*

$$\Delta(G) + \mu(G) \leq n + 1 - \frac{2}{n}.$$

La borne est atteinte si et seulement si le graphe est une étoile.

2.2 Le nombre de stabilité

L'étude du nombre de stabilité d'un graphe remonte à plusieurs dizaines d'années. D'ailleurs Claude Berge lui consacre un chapitre dans son livre «Graphes et hypergraphes» [12]. Dans ce chapitre Berge présente quelques bornes pour le nombre de stabilité en fonction de l'ordre du graphe et du nombre d'arêtes. Il prouve, par exemple, que si G est un graphe simple avec n sommets et m arêtes, alors nous avons

$$\alpha(G) \geq \frac{n^2}{2m+n}.$$

On a l'égalité si et seulement si toutes les composantes connexes de G sont des cliques de même cardinalité.

Berge introduit aussi la notion de graphes α -critiques.

Définition 2.2.1. *Un graphe $G(V, E)$ est α -critique si quelle que soit l'arête $e \in E$,*

$$\alpha(G - e) = \alpha(G) + 1.$$

L'exemple le plus simple d'un graphe α -critique est la clique puisque son nombre de stabilité est égal à un, et quelle que soit l'arête qu'on enlève, ce nombre devient deux. D'après Berge cette notion fournit des résultats sur le nombre de stabilité d'un graphe quelconque. Plusieurs résultats et propriétés sont présentés pour cette famille de graphes. Nous citons par exemple les théorèmes suivants:

1. «Dans un graphe α -critique, par deux arêtes adjacentes il passe toujours un cycle élémentaire impair sans cordes»

et

2. «Un graphe α -critique connexe n'a pas de point d'articulation.»

Dans le chapitre 3 nous allons généraliser cette notion de graphes α -critiques à la forêt induite maximum. En effet, nous allons introduire la famille des graphes forêt-critiques et cela va nous aider énormément dans la preuve de la conjecture suivante:

$$\mu(G) \leq \frac{F(G)}{2}.$$

P. Dankelmann [24] a aussi étudié le nombre de stabilité. Il a prouvé une borne inférieure pour la distance moyenne, en fonction de l'ordre du graphe G et du nombre de stabilité $\alpha(G)$:

$$\mu(G) \geq 1 + \frac{\alpha(G)(\alpha(G) - 1)}{n(n - 1)}.$$

Le système Graffiti a généré une borne inférieure pour le nombre de stabilité en fonction du rayon d'un graphe r . En effet pour tout graphe connexe G , on a

$$\alpha(G) \geq r(G).$$

Ce théorème a été prouvé par Fajtlowicz *et al.* [35].

Une autre borne inférieure pour le nombre de stabilité a été aussi générée par Graffiti et prouvée par Fan Chung [22] et plus tard par Dankelmann [24],

$$\mu(G) \leq \alpha(G).$$

Ce théorème célèbre est aussi connu sous le nom de théorème de Fan Chung.

Dans l'article de Aouchiche, Brinkmann et Hansen [5], une série de bornes inférieures et supérieures pour le nombre de stabilité α et un autre invariant usuel de G sont obtenues par AGX2 et prouvées soit automatiquement soit à la main.

Ces relations ont la Forme 1 d'AutoGraphiX que nous avons décrite précédemment. Pour chaque relation, les auteurs donnent au moins une famille de graphes extrêmes pour laquelle la borne supérieure (resp inférieure) est atteinte.

De nombreux résultats obtenus grâce à AutoGraphiX sont mis à jour sur le site <http://www.gerad.ca/~agx>. Citons par exemple le théorème et les propositions suivants.

Théorème 2.2.2. *Pour tout graphe connexe G d'ordre $n \geq 3$, on a*

$$\alpha(G) + \mu(G) \leq n + 1 - \frac{2}{n}.$$

La borne est atteinte si le graphe est une étoile.

Proposition 2.2.3. *Pour tout graphe connexe G d'ordre $n \geq 3$ et de degré moyen $\bar{d}(G)$, on a*

$$\alpha(G) + \bar{d}(G) \leq \begin{cases} \frac{5n-2}{4} & \text{si } n \text{ est pair} \\ \frac{5n-2}{4} + \frac{1}{4n} & \text{si } n \text{ est impair.} \end{cases}$$

Proposition 2.2.4. *Pour tout graphe connexe G d'ordre $n \geq 3$ et de diamètre D , on a*

$$\alpha(G) * D(G) \leq \begin{cases} \frac{n^2-4}{2} & \text{si } n \text{ est pair} \\ \frac{n^2-1}{2} & \text{si } n \text{ est impair.} \end{cases}$$

Proposition 2.2.5. *Pour tout graphe connexe G d'ordre $n \geq 3$ et de diamètre $D(G)$, on a*

$$\alpha(G) - D(G) \geq 1 - \left\lceil \frac{n}{2} \right\rceil.$$

2.3 La cardinalité de la plus grande forêt induite

L'étude de la forêt induite a suscité l'intérêt de plusieurs chercheurs. De ce fait, plusieurs articles ont été publiés sur ce sujet. Des résultats pour la borne inférieure de la forêt induite ont été prouvés, notamment pour des classes particulières de graphes tels que les graphes cubiques et les graphes qui ne contiennent pas de triangles.

Pour plusieurs chercheurs, l'intérêt d'étudier la plus grande forêt induite est relié à l'intérêt d'étudier le complémentaire de la forêt, c'est-à dire un ensemble de sommets qu'on enlève pour obtenir un graphe sans cycle (aussi appelé transversal de cycles).

Le théorème suivant donne une borne inférieure pour la forêt induite en fonction de l'ordre n , du degré maximum $\Delta(G)$ et du nombre de stabilité $\alpha(G)$. Il a été prouvé par Alon, Mubayi et Thomas [3]:

Théorème 2.3.1. *Soit G un graphe connexe. Alors on a*

$$\frac{n - \alpha(G)}{(\Delta(G) - 1)^2} + \alpha(G) \leq F(G).$$

Les auteurs [3] prouvent aussi le résultat suivant:

Corollaire 2.3.2. *Si G est un graphe ne contenant pas de triangle et tel que $\Delta(G) = 3$, alors*

$$F(G) \geq \frac{5n}{8}.$$

Pour les graphes cubiques, une borne inférieure pour la forêt induite a été prouvée par Zheng et Lu [72].

Théorème 2.3.3. *Soit G un graphe connexe cubique qui ne contient pas de cycle de*

longueur trois. Si n est au moins égal à 8, alors

$$F(G) \geq n - \left\lceil \frac{n}{3} \right\rceil.$$

Les conjectures suivantes ont été générées par Graffiti.PC [29] qui est une autre version de Graffiti et elles ont été prouvées par DeLaViña et Waller [28].

Proposition 2.3.4. (Graffiti.PC 47) Soit G un graphe connexe. Alors on a

$$F(G) \geq D(G) + f_1(G) - 1,$$

où $f_1(G)$ est le nombre de sommets de degré 1 dans G .

Proposition 2.3.5. (Graffiti.PC 48) Soit G un graphe connexe. Alors on a

$$F(G) \geq g(G) + f_1(G) - 1,$$

où $g(G)$ est la longueur du plus petit cycle dans G .

Proposition 2.3.6. Soit G un graphe connexe. Alors

$$F(G) \geq c(G) + f_1(G) - 1,$$

où $c(G)$ est la longueur du plus grand cycle dans G .

2.4 La cardinalité du plus grand sous-graphe biparti induit

Plusieurs articles étudient le sous-graphe biparti induit en terme d'arêtes; citons par exemple N. Alon [2]. Néanmoins très peu d'articles étudient l'ordre du plus grand sous-graphe biparti induit. L'un des rares résultats dans ce domaine a été prouvé

par S. Fajtlowicz [36]:

Théorème 2.4.1. *Pour tout graphe connexe G , on a:*

$$r(G) \leq \frac{\alpha_2(G)}{2}$$

où $r(G)$ dénote le rayon de G .

La conjecture suivante, qui a aussi été générée par Graffiti, est toujours ouverte,

Conjecture 2.4.2. *Pour tout graphe G*

$$\alpha_2(G) \geq r(G) + \alpha(G) - 1.$$

Nous constatons que cet invariant, c'est-à-dire $\alpha_2(G)$, n'a pas été beaucoup étudié par les chercheurs en théorie des graphes. C'est peut-être pour cette raison que la conjecture 747, la plus célèbre de Graffiti d'après DeLaViña, était ouverte depuis plus de 15 ans.

2.5 La cardinalité du plus grand arbre induit

Beaucoup de résultats classiques sur les arbres induits ont été prouvés en 1986 par P. Erdős *et al.* [33], dont les résultats suivants.

Théorème 2.5.1. *Pour tout graphe G , on a*

$$T(G) \geq \frac{2n}{m - n + 3},$$

où m est le nombre d'arêtes.

Proposition 2.5.2. *Soit G un graphe connexe. On a $T(G) \leq 3$ si et seulement si*

le complémentaire de G est une union disjointe de graphes bipartis complets et de sommets isolés.

Dans le théorème suivant, les auteurs donnent une borne pour le plus grand arbre induit en fonction du nombre de stabilité.

Théorème 2.5.3. *Pour tout graphe connexe G d'ordre n et pour tout entier m tel que $1 \leq m \leq \frac{n-1}{2}$, on a:*

1. si $\alpha(G) > \frac{(m-1)n}{m} + 1$, alors $T(G) \geq 2m + 1$, et

2. si $\alpha(G) > \frac{(m-1)n}{n} + 1$, alors $T(G) \geq 2m + 2$.

DeLaViña et Waller [28] ont prouvé le résultat suivant.

Proposition 2.5.4. *Pour tout graphe connexe, on a*

$$T(G) \geq \frac{\alpha(G) + 1}{\gamma(G)},$$

où $\gamma(G)$ est le nombre de domination du graphe, c'est-à-dire la cardinalité du plus petit sous-ensemble de sommets A tel que chaque sommet de G est soit inclus dans A , soit adjacent à un sommet de A .

Le système Graffiti.PC [29] a généré de nouvelles conjectures sur l'arbre induit, par exemple la conjecture suivante.

Conjecture 2.5.5. *Pour tout graphe connexe G , on a*

$$T(G) \geq \frac{2r(G)}{\delta(G)},$$

où $\delta(G)$ est le degré minimum.

Récemment, Fox *et al.* [41] ont prouvé les théorèmes suivants:

Théorème 2.5.6. *Soit G un graphe connexe d'ordre n ne contenant pas de triangles.*

Alors on a

$$T(G) \geq \sqrt{n}.$$

Théorème 2.5.7. *Soit G un graphe connexe d'ordre n ne contenant pas de cliques d'ordre k pour $k \geq 4$. Alors on a*

$$T(G) \geq \frac{\log n}{4 \log k}.$$

Nous présentons ici un tableau récapitulatif qui donne pour chaque invariant la borne supérieure \bar{b}_n , la borne inférieure \underline{b}_n et la famille des graphes extrêmes pour chacune des bornes. On note K_n le graphe complet d'ordre n .

G pour \underline{b}_n	\underline{b}_n	Inv.	\bar{b}_n	G pour \bar{b}_n
K_n	1	α	$n - 1$	une étoile ...
K_n	2	α_2	n	une étoile, un chemin, un cycle pair
K_n	2	F	n	une étoile, un chemin
K_n	2	T	n	une étoile, un chemin
K_n	2	LF	n	un chemin

Tableau 2.1 . Bornes et graphes extrêmes

2.6 Algorithmes de résolution du problème du tranversal de cycles

Les sous-graphes induits jouent un rôle important en optimisation combinatoire à cause des résultats théoriques intéressants et de plusieurs applications qui leur sont reliées. Il existe un grand nombre d'articles dans la littérature qui proposent des algorithmes pour résoudre ces problèmes.

Étant donné que la recherche des plus grands sous-graphes induits appartient en général à la classe des problèmes NP-complets, plusieurs méthodes de résolution ont été développées pour des classes particulières de graphes.

Le problème du stable maximum, par exemple, est NP-dur même pour les graphes ne contenant pas de triangles [66]. Cependant, il existe des algorithmes polynomiaux pour certaines classes de graphes, par exemple, les graphes triangulés [42], les graphes ne contenant pas de A_k , H_k ($k \geq 0$) ou de Δ , Σ et H_0 [31][50] (Voir figure 2.1) ainsi que les graphes ne contenant pas de $K_{1,3}$ [62]. Le problème du stable maximum peut alors être considéré comme une généralisation du problème du couplage.

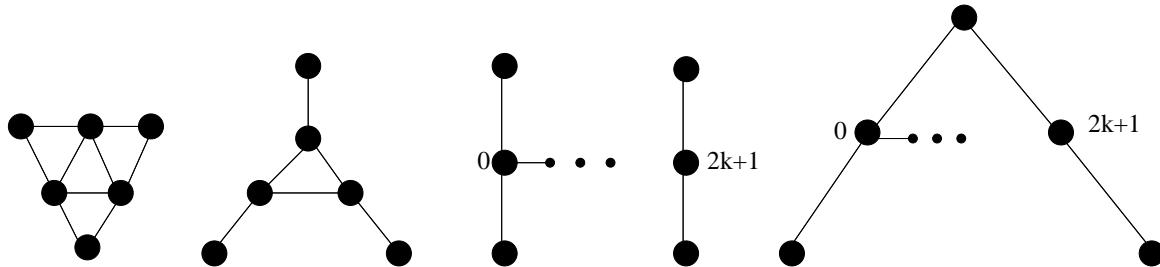


Figure 2.1 . Illustration des graphes Δ , Σ , H_k et A_k

Nous présentons ici une synthèse des algorithmes qui ont été proposés pour résoudre le problème complémentaire de la plus grande forêt, soit le transversal de cycles.

Le problème du transversal de cycles

Résoudre le problème du transversal de cycles revient à déterminer le plus petit ensemble S de sommets de V tel que $(V - S)$ soit un sous-graphe sans cycle. Il y a plusieurs versions de ce problème, selon que le graphe est pondéré, orienté ou non.

Plusieurs méthodes de résolution ont été proposées, par exemple, des métahéuristiques qui sont des méthodes approchées applicables à un grand nombre

de problèmes. Ces méthodes incluent les algorithmes évolutifs, les algorithmes hybrides et les méthodes de voisinage. Des algorithmes d'approximation ont aussi été proposés pour déterminer un transversal de cycles ainsi que des méthodes exactes pour des classes particulières de graphes.

L'étude du problème du transversal de cycles a peut-être commencé dans les années soixante-dix avec les travaux de Yannakakis [70], qui a donné une preuve de la NP-complétude de ce problème. Karp [53] a aussi prouvé la NP-complétude du problème pour les graphes orientés non pondérés. Il y a beaucoup d'applications de cet invariant dans les domaines de la satisfaction des contraintes, de l'intelligence artificielle [40][52] ainsi que la conception des circuits combinatoires. En effet, les cycles dans un circuit électrique peuvent causer un problème car les noeuds peuvent recevoir du courant avant la stabilisation du circuit. D'où la nécessité de poser un sommet enregistreur sur chaque cycle. Toutefois, la vitesse du circuit est proportionnelle aux nombres de ces sommets enregistreurs, l'objectif est alors de les minimiser [40].

Formulation mathématique

Soit G un graphe non pondéré. Le problème du transversal de cycles peut être ramené à un problème de couverture. Soit C_G l'ensemble de tous les cycles de G . Alors déterminer un transversal de cycles revient à résoudre le problème en nombres entiers suivant:

$$(1.1) \text{ Min } \sum_{i=1}^n x_i \text{ s.c.}$$

$$(1.2) \sum_{i \in C_j} x_i \geq 1, C_j \in C_G$$

$$(1.3) x_i \in \{0, 1\}, i \in \{1, \dots, n\}$$

où

$$x_i = \begin{cases} 1 & \text{si le sommet } i \in \text{au transversal de cycles} \\ 0 & \text{sinon.} \end{cases}$$

Si nous considérons un graphe G pondéré, alors l'objectif est de couvrir tous les cycles avec un coût minimum, et le problème peut être formulé de la manière suivante:

$$(1.1) \text{ Min } \sum_{v \in V(G)} w(v)x_v \text{ sous contraintes,}$$

$$(1.2) \sum_{v \in V(C_j)} x_v \geq 1 \text{ pour tout } C_j \in C_G$$

$$(1.3) x_v \in \{0, 1\} \text{ pour tout } v \in V(G).$$

Dans cette formulation $w(v)$ est le poids du sommet v et C_G est l'ensemble de tous les cycles du graphe.

Algorithmes polynomiaux

Pour des classes particulières de graphes, on peut résoudre le problème du transversal de cycles en un temps polynomial.

Les premiers travaux dans ce domaine sont dûs à Shamir [68] qui a donné un algorithme polynomial pour les graphes de flot réductibles, c'est-à-dire les graphes dont il est possible de partitionner les arcs en deux ensembles disjoints, l'un contenant des arcs en avant, l'autre contenant des arcs en arrière. Les arcs en avant forment un graphe acyclique dans lequel chaque noeud peut être atteint à partir du noeud initial du graphe.

Brandstadt et Kratsch [17] étaient les premiers à proposer un algorithme polynomial d'ordre $O(|V|^6)$ pour déterminer le transversal de cycles de coût minimum d'un graphe de permutation; plus tard ce résultat a été amélioré par Liang [56] qui a décrit un algorithme d'ordre $O(|V||E|)$.

Dans ce même cadre, Lung *et al.* [58] ont présenté un algorithme polynomial pour trouver un transversal de cycles de poids minimum dans les graphes d'intervalles. Les sommets i de ces graphes sont représentés par des intervalles I_i et il existe une arête entre i et j si et seulement si $I_i \cap I_j \neq \emptyset$. Les auteurs prouvent que l'ordre de leur algorithme est $O(|V| + |E|)$.

Yannakakis *et al.* [71] a proposé un algorithme polynomial pour les graphes triangulés. Marathe *et al.* [59] ont proposé un algorithme polynomial pour les graphes d'intervalles. En effet, résoudre un problème de transversal de cycles dans ces classes de graphes revient à résoudre un problème de couverture de cliques.

Chang et Liang [20] proposent un algorithme pour calculer le transversal de cycles de coût minimum dans les graphes de cocomparabilité. Rappelons que $G = (V, E)$ est un graphe de comparabilité s'il existe une relation transitive et antisymétrique entre ses sommets. Un graphe de cocomparabilité est le complémentaire d'un graphe de comparabilité.

Les auteurs utilisent la programmation dynamique et se basent sur les deux propriétés suivantes. Soit G un graphe de cocomparabilité.

1. Si G contient un cycle, alors G contient un cycle de longueur 3 ou 4.
2. Il existe une numérotation des sommets ayant la propriété suivante: si $i < k < j$ et $(i, j) \in E$, alors $(i, k) \in E$ ou $(k, j) \in E$.

Ils prouvent que l'ordre de leur algorithme est $O(|V|^2|E|)$.

Chang *et al.* [20] ont aussi donné un algorithme polynomial pour déterminer un transversal de cycles dans des graphes convexes bipartis. Soit $G = (A, B, E)$, G est convexe par rapport à l'ensemble A si les sommets de A reliés à un même sommet $b \in B$ forme un intervalle de A . L'ordre de l'algorithme est $O(|A|^3 + |A|^2|E|)$.

Pour résoudre le problème du transversal de cycles, Levy et Lowe [54] ont proposé des opérations de contraction applicables à certains graphes. Ces opérations réduisent la taille du graphe tout en gardant la valeur du transversal de cycles. Plus tard Jou *et al.* [52] ont aussi proposé des opérations de contraction de graphes pour résoudre ce problème.

L'inconvénient de cet algorithme est qu'il n'est pas optimal pour tous les graphes, mais plusieurs heuristiques et algorithmes d'approximation ont utilisé ce même schéma de réduction.

Pardalos *et al.* [63] proposent un algorithme GRASP pour déterminer le transversal de cycles d'un graphe orienté. Selon Galinier *et al.* [43], GRASP, pour Greedy Randomized Adaptive Search Procedure, est une méthode hybride car elle cherche à combiner les avantages des heuristiques gloutonnes, de la recherche aléatoire et des méthodes de voisinages.

GRASP est un algorithme prometteur qui a été employé avec succès pour de nombreux problèmes d'optimisation tels que des problèmes de télécommunication, d'affectation de tâches, d'ordonnancement et de planification.

Cette méthode est composée de deux phases, une phase constructive et une phase d'amélioration. Dans l'étape de la construction, à chaque itération, un élément est ajouté à la solution courante. Le choix de l'élément ajouté se fait au hasard selon une liste de meilleurs candidats obtenue avec une fonction gloutonne. Cette liste de candidats est notée RCL (Restricted Candidate List) et elle contient les R meilleurs éléments selon un critère fixé.

La deuxième phase de l'algorithme consiste à améliorer les solutions obtenues grâce à une méthode de descente, de recuit simulé ou de recherche tabou.

L'implémentation de GRASP se résume dans les trois étapes suivantes:

1. Générer une solution réalisable avec un algorithme glouton.
2. Appliquer un algorithme de recherche locale.
3. Mettre à jour la meilleure solution.

Dans l'article de Pardalos [63] trois algorithmes gloutons G_A , G_B et G_C sont mis en oeuvre pour générer une solution initiale. Soit $in(i)$ l'ensemble des arcs entrant en i et $out(i)$ l'ensemble des arcs sortant de i ; on choisit le sommet i qui maximise l'un des trois critères suivants.

1. La somme des arcs entrants et sortants, $G_A(i) = |in(i)| + |out(i)|$.
2. Le produit des arcs entrants et sortants, $G_B(i) = |in(i)| * |out(i)|$.
3. Le maximum, $G_C(i) = \max(|in(i)|, |out(i)|)$.

Une recherche locale est appliquée par la suite. L'auteur présente les résultats obtenus sur des graphes aléatoires de 50 à 1000 sommets.

Algorithmes d'approximation

Des algorithmes d'approximation ont été développés pour résoudre le problème du transversal de cycles. La plupart des algorithmes traitent le cas des graphes non orientés. Il semble qu'il soit beaucoup plus difficile de déterminer un algorithme d'approximation pour les graphes orientés.

Graphes non pondérés

Considérons un problème de minimisation P , et I un exemplaire du problème. Soit $A(I)$ la valeur obtenue par l'algorithme A et $OPT(I)$ la valeur optimale. A est une

α -approximation de la valeur optimale si pour tout exemplaire I on a

$$\frac{A(I)}{OPT(I)} \leq \alpha,$$

où $\alpha \geq 1$.

Erdős et Pósa [32] ont proposé un algorithme de ratio $2 \log n$ pour les graphes non orientés. Ce résultat a été amélioré plus tard par Monien *et al.* [61] qui ont présenté un algorithme de ratio $\sqrt{\log n}$.

Mieux encore, Bar-Yehuda *et al.* [10] ont présenté un algorithme d'approximation de ratio $4 - \frac{2}{n}$. Ils utilisent le schéma de contraction proposé par Levy et Lowe [54].

Graphes pondérés

Pour un graphe non orienté pondéré, (G, w) , soit A un algorithme qui retourne le transversal de cycle F_A ; le poids des sommets de F_A est noté $w(F_A)$. Le ratio de performance de A est défini comme $R_A(G, w) = \frac{w(F_A)}{OPT(G, w)}$ où $OPT(G, w)$ est le poids du transversal de cycles le moins lourd.

Bar-Yehuda *et al.* [10] ont développé un algorithme de ratio $\min\{2\Delta^2, 4 \log_2 n\}$ pour les graphes pondérés, où Δ est le degré maximum. Bafna [9] propose un algorithme de ratio 2 pour les graphes non orientés.

Selon Pardalos [63], bien que les algorithmes d'approximation garantissent une solution de bonne qualité, en pratique les méthodes heuristiques donnent de meilleurs résultats.

Algorithmes exacts

Très peu d'articles proposent des algorithmes exacts pour le transversal de cycles.

Parmi ces articles nous citons un algorithme proposé par Cheng *et al.* [21]; c'est un algorithme d'énumération qui utilise le schéma de réduction de graphe.

Nous concluons que, malgré un nombre considérable de travaux effectués sur le transversal de cycles, il y a peu de progrès notés. D'après Festa *et al* [40], ceci est dû au fait que déceler les cycles est une opération coûteuse, ainsi que la recherche locale. Alors pour remédier à ce problème, il faut appliquer des métaheuristiques efficaces.

CHAPITRE 3

DISTANCE MOYENNE ET FORÊT INDUIITE MAXIMUM

À l'aide du système Graffiti , Fajtlowicz a posé, respectivement pendant les années 1986 et 1992, les conjectures suivantes (listées, parmi plusieurs, dans le document *Written on the Wall* [30]): $\mu(G) \leq \alpha(G)$ (WoW 2) et $\mu(G) \leq \alpha_2(G)/2$ (WoW 747). Il est clair que la deuxième conjecture est plus forte que la première, qui a été prouvée par Chung [22]. Une autre preuve de la première conjecture a été proposée par Dankelmann [24], qui a aussi donné une description du graphe G qui maximise la distance moyenne pour un ordre donné et une valeur de $\alpha(G)$ donnée.

Dans ce chapitre nous prouvons un résultat plus fort que la deuxième conjecture, c'est à dire, que $\mu(G)$ est au plus $F(G)/2$, où $F(G)$ dénote *l'ordre maximum d'une forêt induite* de G . L'idée générale de la preuve est la suivante. Soit G un graphe arbitraire. Ce graphe peut être transformé de différentes façons de telle manière que le paramètre $F(G)$ reste fixé mais que la distance moyenne ne diminue pas. Après avoir testé toutes les transformations possibles, on peut montrer que le graphe résultant est le *double cerf-volant*, au moins dans le cas où $F(G)$ est supérieur à 3 (le cas où $F(G)$ est au plus 3 est facile à traiter). Un double cerf-volant, tel qu'il va être défini au début de la section suivante, est un type de graphe extrême qui figure dans la preuve de Dankelmann. Nous utilisons la même notation que Dankelmann [24].

Ce chapitre est organisé de la manière suivante. Dans la Section 3.1 nous commençons par donner des définitions et nous prouvons les lemmes utilisés dans le reste du chapitre. Dans la Section 3.2 nous énonçons et prouvons les lemmes concernant les transformations utilisées dans le théorème. Nous prouvons par la suite ce théorème

dans la Section 3.3 et finalement nous présentons des résultats connexes et d'autres conjectures dans la Section 3.4.

Dans les diverses étapes de la preuve, le système AutoGraphiX (Caporossi *et al.* [18]) a été utilisé afin de trouver des contre-exemples ou poser des conjectures.

3.1 Double cerf-volant et graphe forêt-critique

Dans cette section, nous allons définir le double cerf-volant ainsi que les graphes forêt-critiques et quasi-forêt-critiques. Nous allons donner par la suite plusieurs lemmes et propriétés de ces graphes et de la distance moyenne dans un double cerf-volant.

Définition 3.1.1. Pour $k \geq 4$ et $n_1, n_2 \geq 2$, soit $G_{n_1, n_2, k}$ le graphe obtenu à partir d'un chemin P_{k-2} d'ordre $k-2$ et de deux cliques disjointes G_1 et G_2 d'ordres n_1 et n_2 , respectivement, en identifiant une extrémité de P_{k-2} avec un sommet v_1 de G_1 (noté la base de G_1) et l'autre extrémité de P_{k-2} avec un sommet v_2 de G_2 (noté la base de G_2). $V(G_{n_1, n_2, k})$ est donc une union disjointe de trois ensembles $V(G_1)$, $V(G_2)$ et $V(P_{k-2}) \setminus \{v_1, v_2\}$, et $E(G_{n_1, n_2, k})$ est l'union disjointe des ensembles $E(G_1)$, $E(G_2)$ et $E(P_{k-2})$. Dans le cas particulier où $n_1 = \lfloor \frac{n-k+4}{2} \rfloor$ et $n_2 = \lceil \frac{n-k+4}{2} \rceil$ pour $n \geq k$, $G_{n_1, n_2, k}$ est noté $G_{n, k}$. Pour $n_1, n_2 \geq 2$, $G_{n_1, n_2, 3}$ dénote le graphe obtenu à partir de deux cliques G_1 et G_2 d'ordres n_1 et n_2 , respectivement, en identifiant un sommet de G_1 avec un sommet de G_2 .

Soit G un graphe d'ordre $n = 12$. $G_{n_1, n_2, 6}$ est le graphe obtenu à partir d'un chemin P_{k-2} d'ordre 4 et de deux cliques disjointes



Figure 3.1 . P_{k-2} avec $k = 6$

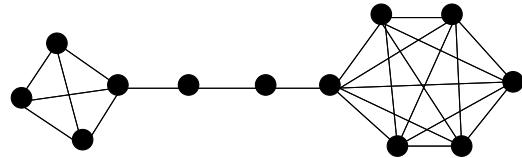


Figure 3.2 . Illustration d'un double cerf-volant non équilibré noté $G_{4,6,6}$.

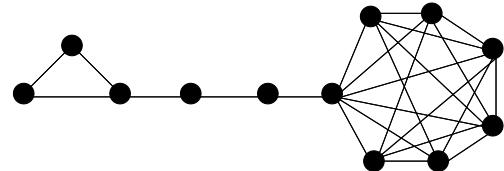


Figure 3.3 . Illustration d'un double cerf-volant non équilibré noté $G_{3,7,6}$.

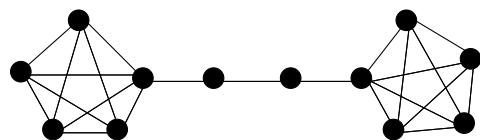


Figure 3.4 . Illustration d'un double cerf-volant équilibré noté $G_{5,6}$.

Remarquons que notre définition est différente de celle de Dankelmann [3]; en effet si k est pair, notre graphe $G_{n,k}$ est le graphe $G_{n,k/2}$ d'après la définition de Dankelmann. Notons aussi que le graphe $G_{2,n_2,k}$ est formé d'une clique d'ordre n_2 et d'un chemin de $k - 1$ sommets et qui ont exactement un sommet en commun. Ce graphe est appelé un *cerf-volant*, et son *extrémité* est le sommet le plus éloigné de la clique; l'extrémité est définie uniquement si n_2 est supérieur à 2. En général, pour $k \geq 4$, un *double cerf-volant* est un graphe de la forme $G_{n_1,n_2,k}$ et un *double cerf-volant équilibré* est un graphe de la forme $G_{n,k}$; ainsi pour $k \geq 4$ un cerf-volant $G_{2,n_2,k}$ est un cas particulier du double cerf-volant. Nous introduisons maintenant une notion qui sera très utile dans la preuve.

Définition 3.1.2. Soit G un graphe connexe.

- . Une arête xy de G est *forêt-critique* si $F(G - xy) = F(G) + 1$.
- . G est *forêt-critique* si chaque arête de G est *forêt-critique*.
- . G est *quasi-forêt-critique* (*qfc* pour abréger) si chaque arête de G est soit un *pont* soit *forêt-critique*.

Lemme 3.1.3. Soit G un graphe connexe.

- . Si G est *forêt-critique* et x est un sommet de G , alors $F(G - x)$ est égal à $F(G)$ et il existe un sous-ensemble U de sommets de $G - x$ tel que $|U| = F(G) - 1$ et $U \cup \{x\}$ induit une forêt dans G .
- . Si G est *quasi-forêt-critique* et x est un sommet de G qui n'est pas l'extrémité d'un *pont*, alors $F(G - x)$ est égal à $F(G)$ et il existe un sous ensemble de sommets U de $G - x$ tel que $|U| = F(G) - 1$ et $U \cup \{x\}$ induit une forêt de G .

Preuve. La preuve de la deuxième proposition est identique à celle de la première; nous allons seulement prouver la première. Soit G un graphe *forêt-critique*, xy une arête de

G et S un sous-ensemble maximum de sommets induisant une forêt de $G - xy$. Comme $F(G - xy)$ est égal à $F(G) + 1$, S contient x et y . Alors $S \setminus x$ est un sous-ensemble de sommets de $G - x$ qui a la cardinalité de $F(G)$ et induit une forêt de $G - x$. Ceci prouve que $F(G - x)$ est égal à $F(G)$. D'où le sous-ensemble $U = S \setminus \{x, y\}$ vérifie la conclusion du théorème. \square

Lemme 3.1.4. *Soit G un graphe quasi-forêt-critique et x un sommet de G qui n'est pas l'extrémité d'un pont. Alors $G - x$ est connexe.*

Preuve. Supposons que $G - x$ n'est pas connexe. Alors $V \setminus \{x\}$ peut être partitionné en deux sous-ensembles non vides V_1 et V_2 tel qu'il n'existe pas d'arête reliant un sommet de V_1 à un sommet de V_2 . Soit H_1 (resp. H_2) le sous-graphe induit par V_1 (resp. V_2) et G_1 (resp. G_2) le sous-graphe induit par $V_1 \cup \{x\}$ (resp. $V_2 \cup \{x\}$). Par le lemme 3.1.3, parmi les sous-ensembles induisant des forêts d'ordre maximal, il existe un sous-ensemble contenant x et un autre ne contenant pas x . Alors nous avons

$$F(G) = F(G_1) + F(G_2) - 1, \quad F(G) = F(H_1) + F(H_2),$$

ce qui implique que soit $F(G_1) = F(H_1) + 1$ et $F(G_2) = F(H_2)$, soit $F(G_1) = F(H_1) + 1$ et $F(G_2) = F(H_2) + 1$. Sans perte de généralité nous supposons que $F(G_1) = F(H_1) + 1$ et $F(G_2) = F(H_2)$ et soit xy une arête de $E(G_1)$. Il est clair que $F(G - xy)$ est au plus $F(H_1) + F(G_2) = F(H_1) + F(H_2) = F(G)$, ce qui contredit l'hypothèse que G est quasi-forêt-critique comme xy est forêt-critique. \square

Il s'en suit par le lemme 3.1.4 que tout graphe forêt-critique est 2-connexe. Le résultat suivant est bien connu (voir Plesnik [6]).

Lemme 3.1.5. *$\sigma(P_n)$ est donné par la formule*

$$\frac{(n-1)n(n+1)}{3}.$$

Lemme 3.1.6. $\sigma(G_{n_1, n_2, k})$ est donné par la formule

$$(n - 3k + 6)(n - k + 3) + \frac{(k - 3)(k - 2)(k - 1)}{3} + 2n_1 n_2 (k - 2) + (n - k + 2)k(k - 3).$$

Preuve. Soit G_1 et G_2 définis selon la définition 3.1.1. D'après le lemme 3.1.5 et l'égalité $n_1 + n_2 = n - k + 4$, il est facile de vérifier les égalités suivantes.

$$\begin{aligned} \sigma(G_{n_1, n_2, k}) &= \sigma(G_1) + \sigma(P_{k-2}) + \sigma(G_2) \\ &\quad + 2 \sum \{d(u, v) \mid u \in V(G_1 - v_1), v \in V(G_2 - v_2)\} \\ &\quad + 2 \sum \{d(u, v) \mid u \in V(G_1 - v_1), v \in V(P_{k-2} - v_1)\} \\ &\quad + 2 \sum \{d(u, v) \mid u \in V(G_2 - v_2), v \in V(P_{k-2} - v_2)\} \\ \\ &= n_1(n_1 - 1) + \sigma(P_{k-2}) + n_2(n_2 - 1) + 2(n_1 - 1)(n_2 - 1)(k - 1) \\ &\quad + 2(n_1 - 1) \sum_{i=2}^{k-2} i + 2(n_2 - 1) \sum_{i=2}^{k-2} i \\ &= n_1(n_1 + n_2 - 1) + \sigma(P_{k-2}) + n_2(n_2 + n_1 - 1) \\ &\quad + 2(n_1 - 1)(n_2 - 1)(k - 1) - 2n_1 n_2 + (n_1 - 1)(k^2 - 3k) \\ &\quad + (n_2 - 1)(k^2 - 3k) \\ &= (n - k + 4)(n - k + 3) + \frac{(k - 3)(k - 2)(k - 1)}{3} + 2n_1 n_2 (k - 2) \\ &\quad - 2(k - 1)(n - k + 3) + (n - k + 2)(k^2 - 3k) \\ &= (n - 3k + 6)(n - k + 3) + \frac{(k - 3)(k - 2)(k - 1)}{3} + 2n_1 n_2 (k - 2) \\ &\quad + (n - k + 2)k(k - 3) \end{aligned}$$

□

Définition 3.1.7. Pour tout graphe G , $\tau(G)$ dénote la transmission maximum d'un sommet de G .

Lemme 3.1.8. Soit G un graphe connexe d'ordre n tel que $F(G) = k$. Alors

$$\tau(G) \leq \tau(G_{2,n-k+2,k}) = (k-1) \left(n - \frac{k}{2} \right).$$

Pour tous n , n_1 et n_2 tels que $2 \leq n_1 \leq n_2$ et $n_1 + n_2 = n - k + 4$, nous avons

$$\tau(G_{n_1,n_2,k}) = (n-1)\frac{k}{2} + (n_2 - n_1) \left(\frac{k}{2} - 1 \right) \geq (n-1)\frac{k}{2}.$$

Preuve. Nous commençons par calculer la transmission maximum dans le graphe $G_{n_1,n_2,k}$, où $2 \leq n_1 \leq n_2$ et $n_1 + n_2 = n - k + 4$. La transmission maximum est la transmission d'un sommet de $G_1 - v_1$. Alors nous avons

$$\begin{aligned} \tau(G_{n_1,n_2,k}) &= (n_1 - 1) + \sum_{i=2}^{k-2} i + (n_2 - 1)(k - 1) \\ &= (n_1 - 1) + \frac{(k-2)(k-1)}{2} - 1 + (n_2 - 1)(k - 1) \\ &= \frac{(k-2)(k-1)}{2} - 1 + (n_1 + n_2 - 2)\frac{k}{2} + (n_2 - n_1) \left(\frac{k}{2} - 1 \right) \\ &= \frac{(k-3)k}{2} + (n - k + 2)\frac{k}{2} + (n_2 - n_1) \left(\frac{k}{2} - 1 \right) \\ &= (n-1)\frac{k}{2} + (n_2 - n_1) \left(\frac{k}{2} - 1 \right). \end{aligned}$$

La dernière inégalité de l'énoncé est vraie car $n_2 \geq n_1$.

Soit G un graphe connexe d'ordre n tel que $F(G) = k$. Soit x un sommet de G de transmission maximum et soit $(x = x_0, x_1, x_2, \dots, x_{p-1})$ le plus court chemin de

longueur maximum. Il est clair que le sous-graphe induit par les sommets du chemin est une forêt, et donc p est au plus k . Comme la distance entre x et un sommet y n'appartenant pas au chemin est au plus $p - 1$, $\sigma_G(x)$ est au plus $\sum_{i=1}^{p-1} i + (n - p)(p - 1)$ qui est inférieur ou égal à $\sum_{i=1}^{k-1} i + (n - k)(k - 1)$. Ainsi nous obtenons

$$\tau(G) \leq \sum_{i=1}^{k-1} i + (n - k)(k - 1) = \sum_{i=1}^{k-2} i + (n - k + 1)(k - 1) = \tau(G_{2,n-k+2,k}).$$

Remplaçons $\sum_{i=1}^{k-1} i$ par $k(k - 1)/2$, finalement nous obtenons

$$\tau(G) \leq \tau(G_{2,n-k+2,k}) = \frac{k(k - 1)}{2} + (n - k)(k - 1) = (k - 1) \left(n - \frac{k}{2} \right).$$

□

Lemme 3.1.9. *La relation $\mu(G_{n,k}) < \mu(G_{n+1,k})$ est vraie pour tout $k \geq 4$.*

Preuve. D'après le lemme 3.1.6 nous savons que

$$\begin{aligned} \sigma(G_{n,k}) &= \sigma(P_{k-2}) + (n - 3k + 6)(n - k + 3) + (n - k + 2)k(k - 3) \\ &\quad + 2 \left\lfloor \frac{n - k + 4}{2} \right\rfloor \left\lceil \frac{n - k + 4}{2} \right\rceil (k - 2) \\ &\leq \sigma(P_{k-2}) + (n - 3k + 6)(n - k + 3) + (n - k + 2)k(k - 3) \\ &\quad + \frac{(n - k + 4)^2}{2}(k - 2). \end{aligned}$$

De plus, $G_{n+1,k}$ est le graphe obtenu en ajoutant le sommet z à la plus petite clique dans $G_{n,k}$. Comme $\sigma(G_{n+1,k})$ est égal à $\sigma(G_{n,k}) + 2\sigma_{G_{n+1,k}}(z)$, nous obtenons

$$\begin{aligned} \sigma(G_{n+1,k}) &= \sigma(G_{n,k}) \\ &\quad + 2 \left(\left\lfloor \frac{n - k + 4}{2} - 1 \right\rfloor + \frac{(k - 2)(k - 1)}{2} + \left\lceil \frac{n - k + 4}{2} - 1 \right\rceil (k - 1) \right) \end{aligned}$$

$$\begin{aligned}
&\geq \sigma(G_{n,k}) + 2 \left(\left(\frac{n-k+4}{2} - 1 \right) k + \frac{(k-2)(k-1)}{2} \right) \\
&= \sigma(G_{n,k}) + (n-k+4)k - 2k + (k-2)(k-1) \\
&= \sigma(G_{n,k}) + nk - k + 2.
\end{aligned}$$

Il s'ensuit que

$$\begin{aligned}
\mu(G_{n+1,k}) - \mu(G_{n,k}) &= \frac{(n-1)\sigma(G_{n+1,k}) - (n+1)\sigma(G_{n,k})}{(n+1)n(n-1)} \\
&\geq \frac{-2\sigma(G_{n,k}) + (n-1)(nk - k + 2)}{(n+1)n(n-1)}.
\end{aligned}$$

Il reste à prouver que le numérateur est strictement positif.

$$\begin{aligned}
&- 2\sigma(G_{n,k}) + (n-1)(nk - k + 2) \\
&\geq -2\sigma(P_{k-2}) - 2(n-3k+6)(n-k+3) - 2(n-k+2)k(k-3) \\
&\quad - (n-k+4)^2(k-2) + (n-1)(nk - k + 2) \\
&= -2\sigma(P_{k-2}) + k^3 - 6k^2 + 11k - 6 = -2\sigma(P_{k-2}) + (k-3)(k-2)(k-1) \\
&= \sigma(P_{k-2}) > 0.
\end{aligned}$$

.

□

Lemme 3.1.10. *Pour tous k , n , n_1 et n_2 tels que $k \geq 4$, $2 \leq n_1 \leq n_2$ et $n_1 + n_2 = n - k + 4$, on a*

$$\mu(G_{n_1,n_2,k}) \leq \frac{k}{2}.$$

Preuve. D'après le lemme 3.1.6 il est clair que $\sigma(G_{n_1,n_2,k})$ est maximum quand $|n_2 - n_1|$ est au plus 1, c'est-à-dire quand G est isomorphe à $G_{n,k}$. Si $n_1 = n_2$, alors

$$\mu(G_{n_1,n_2,k}) \leq \frac{\tau(G_{n_1,n_2,k})}{n-1} = \frac{k}{2},$$

où l'égalité est une conséquence du lemme 3.1.8. Si $n_2 = n_1 + 1$, alors d'après le lemme 3.1.9 on obtient

$$\mu(G_{n_1, n_2, k}) < \mu(G_{n_2, n_2, k}) \leq \frac{k}{2},$$

ce qui complète la preuve. \square

Lemme 3.1.11. *Soit G un graphe connexe et xy un pont de G . Soit G_1 (resp. G_2) la composante connexe de $G - xy$ contenant x (resp. y), et soit n_i l'ordre de G_i pour $i = 1, 2$. Alors on a*

$$\sigma(G) = \sigma(G_1) + \sigma(G_2) + 2n_1\sigma_{G_2}(y) + 2n_2\sigma_{G_1}(x) + 2n_1n_2.$$

Preuve.

$$\begin{aligned} \sigma(G) &= \sum_{u \in G_1} \sum_{v \in G_1} d(u, v) + \sum_{u \in G_2} \sum_{v \in G_2} d(u, v) + 2 \sum_{u \in G_1} \sum_{v \in G_2} d(u, v) \\ &= \sigma(G_1) + \sigma(G_2) + 2 \sum_{u \in G_1} \sum_{v \in G_2} (d(u, x) + d(x, y) + d(y, v)) \\ &= \sigma(G_1) + \sigma(G_2) + 2 \left(\sum_{u \in G_1} d(u, x) \right) \left(\sum_{v \in G_2} 1 \right) + 2d(x, y) \left(\sum_{u \in G_1} 1 \right) \left(\sum_{v \in G_2} 1 \right) \\ &\quad + 2 \left(\sum_{v \in G_2} d(y, v) \right) \left(\sum_{u \in G_1} 1 \right) \\ &= \sigma(G_1) + \sigma(G_2) + 2\sigma_{G_1}(x)n_2 + 2n_1n_2 + 2\sigma_{G_2}(y)n_1. \end{aligned}$$

\square

Corollaire 3.1.12. *Soit G , xy , G_i et n_i définis de la même manière que dans le lemme 3.1.11. Nous supposons qu'il existe un sommet x' de G_1 tel que $\sigma_{G_1}(x') > \sigma_{G_1}(x)$. Soit*

G' le graphe obtenu à partir de G en remplaçant xy par $x'y$. Alors nous avons

$$F(G') = F(G) \text{ et } \sigma(G') > \sigma(G).$$

De plus, si G est qfc, alors G' est aussi qfc.

Preuve. La définition d'un graphe quasi-forêt-critique implique que G' est qfc si G est qfc. De plus, nous avons

$$F(G') = F(G_1) + F(G_2) = F(G)$$

et

$$\begin{aligned} \sigma(G) &= \sigma(G_1) + \sigma(G_2) + 2n_1\sigma_{G_2}(y) + 2n_2\sigma_{G_1}(x) + 2n_1n_2 \\ &< \sigma(G_1) + \sigma(G_2) + 2n_1\sigma_{G_2}(y) + 2n_2\sigma_{G_1}(x') + 2n_1n_2 = \sigma(G'). \end{aligned}$$

□

3.2 Transformations

Les définitions et les résultats de la section précédente suggèrent que dans certains cas, un graphe G peut être transformé de manière à ce que G reste connexe, $F(G)$ ne change pas et $\mu(G)$ ne diminue pas. En l'occurrence, si on enlève de G une arête qui n'est ni forêt-critique ni un pont, alors G reste connexe, la valeur de $F(G)$ ne change pas et $\mu(G)$ ne diminue pas. Cette transformation est la transformation no. 1 du Théorème 3.3.2. De même le corollaire 3.1.12 propose une autre transformation qui a ces propriétés (appelée transformation no 3 dans le théorème 3.3.2). Dans cette section nous allons présenter trois autres transformations.

Le lemme 3.2.2 montre que sous certaines conditions, on peut remplacer un sous-graphe induit par un cerf-volant. Ce changement est appelé la transformation no 2 dans le théorème 3.3.2. Le lemme 3.2.3 montre que sous certaines conditions, on peut déplacer un sommet d'un sous-graphe induit à un autre. Cette transformation est appelée la transformation no. 4 dans le théorème 3.3.2. Les lemmes 3.2.4 et 3.2.5 montrent que sous certaines conditions, on peut enlever un sommet d'un graphe G sans déconnecter G et sans modifier $F(G)$ ou faire diminuer $\mu(G)$. Cette opération est appelée la transformation no. 5 dans le théorème 3.3.2. Le lemme suivant est utilisé dans la preuve du lemme 3.2.2.

Lemme 3.2.1. *Soit x, a, b trois sommets d'un graphe connexe G . Alors*

$$(a) \quad d(a, b) \leq 2F(G) - d(x, a) - d(x, b) - 1, \text{ et}$$

$$(b) \quad \text{si le sommet } a \text{ est adjacent à un sommet } a' \text{ avec } d(x, a') = d(x, a) + 1, \text{ alors}$$

$$d(a, b) \leq 2F(G) - d(x, a) - d(x, b) - 2.$$

Preuve. Soit T l'arbre construit par une fouille en largeur à partir du sommet x . Le chemin partant de x à un autre sommet de T est un plus court chemin. P_0 dénote le chemin entre a et b dans T . Dans ce qui suit, un chemin et l'ensemble de ses sommets vont être notés par le même symbole. Si a est un ancêtre de b , P_0 est le plus court chemin entre a et b et $d(x, b) = d(x, a) + d(a, b)$, alors $F(G)$ est au moins égal à $d(x, b) + 1$. Nous obtenons

$$\begin{aligned} 2F(G) - d(x, a) - d(x, b) - 2 &\geq 2(d(x, b) + 1) - d(x, a) - d(x, b) - 2 \\ &= d(x, b) - d(x, a) \\ &= d(a, b) \end{aligned}$$

si a est un ancêtre de b , ou par symétrie, si b est un ancêtre de a . La relation $2F(G) - d(x, a) - d(x, b) - 2 \geq d(a, b)$ est correcte si a est égal à b . À partir de

maintenant, nous supposons que a n'est pas un ancêtre de b , b n'est pas un ancêtre de a et les sommets x , a et b sont distincts.

Le chemin P_0 passe par l'ancêtre commun le plus proche de a et b , qu'on dénote z . Dans le sous-graphe de G induit par P_0 , nous considérons un plus court chemin de a à b , dénoté P . Notons que si P est différent de P_0 , le chemin P contient une seule arête n'appartenant pas à T , soit l'arête cc' ; dans ce cas, on peut supposer que c appartient au chemin de z à a dans T , et c' au chemin entre z et b dans T . Sans perte de généralité, nous supposons que $d(x, c)$ n'est pas supérieur à $d(x, c')$. Par la construction de T et la définition de c et c' , $d(x, c')$ est égale à $d(x, c)$ ou $d(x, c) + 1$, et si $d(x, c') = d(x, c) + 1$ est vraie, c appartient au plus court chemin qui relie x à b . Si P est identique à P_0 (et donc z appartient à P), c dénote le même sommet que z . Dans tous les cas, soit P' le chemin de x à c dans T .

D'abord nous supposons que z appartient à P , ou que z n'appartient pas à P et $d(x, c')$ est égale à $d(x, c) + 1$. Alors nous avons

$$\begin{aligned}
 d(a, b) &\leq d(a, c) + d(c, b) \\
 &= d(x, a) - d(x, c) + d(x, b) - d(x, c) \\
 &= d(x, a) + d(x, b) - 2d(x, c).
 \end{aligned} \tag{3.1}$$

D'autre part, $P \cup P'$ est une forêt ayant $d(c, a) + d(c, b) + d(x, c) + 1$ sommets et nous avons

$$\begin{aligned}
 F(G) &\geq d(c, a) + d(c, b) + d(x, c) + 1 \\
 &= (d(x, a) - d(x, c)) + (d(x, b) - d(x, c)) + d(x, c) + 1 \\
 &= d(x, a) - d(x, c) + d(x, b) + 1.
 \end{aligned} \tag{3.2}$$

Donc, par (3.1) et (3.2), nous obtenons

$$\begin{aligned} d(a, b) &\leq d(x, a) + d(x, b) + 2(F(G) - d(x, a) - d(x, b) - 1) \\ &= 2F(G) - d(x, a) - d(x, b) - 2. \end{aligned}$$

Les énoncés (a) et (b) sont donc prouvés si $z \in P$ ou $d(x, c') = d(x, c) + 1$. Nous supposons maintenant que $z \notin P$ et $d(x, c') = d(x, c)$. Alors nous avons

$$d(a, b) \leq d(a, c) + d(c, b) \quad (3.3)$$

$$\begin{aligned} &= (d(x, a) - d(x, c)) + (d(x, b) - d(x, c') + 1) \\ &= d(x, a) + d(x, b) - 2d(x, c) + 1. \end{aligned} \quad (3.4)$$

Comme $P \cup P' - \{c'\}$ est une forêt ayant $d(x, a) + d(c', b) + 1$ sommets, nous avons

$$F(G) \geq d(x, a) + d(c', b) + 1 \quad (3.5)$$

$$\begin{aligned} &= d(x, a) + (d(x, b) - d(x, c')) + 1 \\ &= d(x, a) + d(x, b) - d(x, c) + 1. \end{aligned} \quad (3.6)$$

De (3.4) et (3.6), nous obtenons

$$\begin{aligned} d(a, b) &\leq d(x, a) + d(x, b) + 2(F(G) - d(x, a) - d(x, b) - 1) + 1 \\ &= 2F(G) - d(x, a) - d(x, b) - 1. \end{aligned}$$

Pour prouver (b), il suffit de prouver que (3.3) ou (3.5) est une inégalité stricte s'il existe un sommet a' adjacent à a tel que $d(x, a') = d(x, a) + 1$. Alors nous supposons qu'un tel sommet existe et que (3.3) est satisfaite à égalité. Ceci implique que $d(a, b) = |P| - 1$. Soit B la forêt $P \cup P' - \{c'\}$, composée de deux chemins disjoints. D'abord observons que si a' appartenait à B , alors il ferait partie du chemin de c' à b dans T , et alors P ne

serait pas un plus court chemin de a à b dans G . Donc $|B \cup \{a'\}|$ est plus grand que $|B|$.

Nous affirmons que $B \cup \{a'\}$ induit une forêt. Si le sous-graphe induit par $B \cup \{a'\}$ contient un cycle, ce cycle est formé de a' et de tous les sommets du chemin entre c'' et c''' (incluant c'' et c'''), où c'' et c''' sont deux sommets différents de c' sur le chemin de c' à b . Sans perte de généralité, nous supposons que c''' est plus proche de b que c'' . Alors le chemin (a, a', c''') est plus petit que la portion de chemin de P entre a et c''' , ce qui contredit le fait que $d(a, b)$ est égale à $|P| - 1$. Alors $B \cup \{a'\}$ induit une forêt et (3.5) est stricte. Ceci complète la preuve de (b) dans le cas où $z \notin P$ et $d(x, c') = d(x, c)$ est vraie. \square

Dans le reste de cette section nous considérons un graphe G contenant un pont xy . Dans chacun des lemmes suivants, G_i et n_i sont définis de la même manière que dans le lemme 3.1.11.

Lemme 3.2.2. *Soit G le graphe contenant un pont xy . Nous supposons que $n_1 \leq n_2$ et G_1 n'est ni une clique ni un cerf-volant d'extrémité x . Soit G' le graphe obtenu en remplaçant G_1 par le cerf-volant $G'_1 = G_{2,n_1-F(G_1)+2,F(G_1)}$, le sommet y étant relié à l'extrémité x' de G'_1 dans G' . Alors*

$$F(G') = F(G) \text{ et } \sigma(G) < \sigma(G').$$

De plus, si G est qfc, alors G' est aussi qfc.

Preuve. Notons que $F(G_1)$ (et donc n_1) sont au moins égaux à 3 car G_1 n'est pas une clique. Comme un cerf-volant est qfc, alors G' est qfc si G est qfc. Nous avons aussi

$$F(G) = F(G_1) + F(G_2) = F(G'_1) + F(G_2) = F(G').$$

Nous dénotons par $d(u, v)$ la distance entre u et v s'ils appartiennent au graphe G , et

$d'(u, v)$ s'ils appartiennent au graphe G' . Nous remarquons que $d(u, v) = d'(u, v)$ pour chaque paire $\{u, v\}$ de sommets dans G_2 . Nous devons montrer que $\sigma(G') - \sigma(G) > 0$. Considérons un sous-ensemble $V'_2 = \{y_1, y_2, \dots, y_{n_1}\}$ de n_1 sommets dans G_2 . Par le lemme 3.1.8 nous avons $\sigma_{G_1}(x) \leq \sigma_{G'_1}(x')$. D'où, pour chaque sommet v de G_2 , nous avons

$$\begin{aligned}\sigma_G(v) &= \sigma_{G_2}(v) + n_1 d(v, x) + \sigma_{G_1}(x) \\ &\leq \sigma_{G_2}(v) + n_1 d'(v, x) + \sigma_{G'_1}(x') = \sigma_{G'}(v).\end{aligned}$$

Alors il nous reste à prouver que

$$\sum_{v \in V'_2} (\sigma_{G'}(v) - \sigma_G(v)) + \sum_{v \in G'_1} \sigma_{G'}(v) - \sum_{v \in G_1} \sigma_G(v) > 0.$$

Les sommets de G'_1 sont dénotés $x'_1, x'_2, \dots, x'_{F(G_1)-1}$ de telle manière que $d(x', x'_i) = i-1$, et les autres sommets de G'_1 , ceux qui forment une clique avec $x'_{F(G_1)-1}$, sont notés $x'_{F(G_1)}, \dots, x'_{n_1}$. Soit le sous-ensemble de sommets A de G_1 vérifiant les propriétés suivantes

- . si $d(a, b) \leq 2F(G_1) - d(x, a) - d(x, b) - 2$ pour toute paire de sommets a et b dans G_1 , alors soit A un sous ensemble de $F(G_1) - 1$ sommets dans G_1 tel que $d(x, a) \leq d(x, v)$ pour tout $a \in A$ et $v \notin A$ (notez que $x \in A$).
- . s'il existe deux sommets a et b dans G_1 tels que $d(a, b) = 2F(G_1) - d(x, a) - d(x, b) - 1$, nous définissons P et P' comme dans la preuve du lemme 3.2.1 (b), et A comme $(P \cup P') - \{a, b\}$. Nous avons montré que $|A| = F(G_1) - 1$ et $d(u, v) \leq 2F(G_1) - d(x, u) - d(x, v) - 2$ pour tout $u \in A$ et tout $v \in G_1$ (car chaque sommet u dans A a un voisin u' tel que $d(x, u') = d(x, u) + 1$).

Dans les deux cas, notons les sommets de A $x_1, x_2, \dots, x_{F(G_1)-1}$ de manière à ce que $i < j$ lorsque $d(x, x_i) < d(x, x_j)$. Nous remarquons que $G_1[A]$ est connexe (mais pas nécessairement un arbre) et x_1 est le sommet x . Les sommets de $V(G_1) \setminus A$ sont $x_{F(G_1)}, \dots, x_{n_1}$. L'inégalité que nous voulons prouver est stricte, mais dans le reste de la preuve nous considérons l'inégalité relaxée.

$$\begin{aligned}
0 &\leq \sum_{i=1}^{n_1} (\sigma_{G'}(y_i) - \sigma_G(y_i) + \sigma_{G'}(x'_i) - \sigma_G(x_i)) \\
&= \sum_{i=1}^{n_1} \left(\sum_{v \in G'_1} d'(y_i, v) - \sum_{v \in G_1} d(y_i, v) \right) + \sum_{i=1}^{n_1} \left(\sum_{v \in G'_1} d'(x'_i, v) - \sum_{v \in G_1} d(x_i, v) \right) \\
&\quad + \sum_{i=1}^{n_1} \left(\sum_{v \in G_2} d'(x'_i, v) - \sum_{v \in G_2} d(x_i, v) \right).
\end{aligned}$$

Considérons les trois termes tour à tour. Le premier terme peut être écrit de la manière suivante.

$$\sum_{i=1}^{n_1} \left(\sum_{v \in G'_1} d'(y_i, v) - \sum_{v \in G_1} d(y_i, v) \right) = \sum_{i=1}^{n_1} \left(\sum_{v \in G'_1} d'(x, v) - \sum_{v \in G_1} d(x, v) \right) = n_1(\sigma_{G'_1}(x) - \sigma_{G_1}(x)).$$

Pour le deuxième terme, nous utilisons le fait que le chemin P_n de n sommets a la distance moyenne maximale parmi tous les graphes connexes d'ordre n . Nous savons par le Lemme 3.2.1 et le choix de A que $d(x_i, x_j) \leq 2F(G_1) - d(x, x_i) - d(x, x_j) - 2$ pour tout i tel que $1 \leq i < F(G_1)$ et tout j tel que $F(G_1) \leq j \leq n_1$ et que $d(x_i, x_j) \leq 2F(G_1) - d(x, x_i) - d(x, x_j) - 1$ pour tous i et j tels que $F(G_1) \geq i$ et $j \leq n_1$. Ainsi nous avons les relations suivantes.

$$\begin{aligned}
& \sum_{i=1}^{n_1} \left(\sum_{v \in G'_1} d'(x'_i, v) - \sum_{v \in G_1} d(x_i, v) \right) \\
&= \sum_{i=1}^{n_1} \left(\sum_{j=1}^{F(G_1)-1} (d'(x'_i, x'_j) - d(x_i, x_j)) + \sum_{j=F(G_1)}^{n_1} (d'(x'_i, x'_j) - d(x_i, x_j)) \right) \\
&= \sum_{i=1}^{F(G_1)-1} \sum_{j=1}^{F(G_1)-1} (d'(x'_i, x'_j) - d(x_i, x_j)) + \sum_{i=1}^{F(G_1)-1} \sum_{j=F(G_1)}^{n_1} (d'(x'_i, x'_j) - d(x_i, x_j)) \\
&\quad + \sum_{i=F(G_1)}^{n_1} \sum_{j=1}^{F(G_1)-1} (d'(x'_i, x'_j) - d(x_i, x_j)) + \sum_{i=F(G_1)}^{n_1} \sum_{j=F(G_1)}^{n_1} (d'(x'_i, x'_j) - d(x_i, x_j)) \\
(*) &\geq \sigma(P_{F(G_1)-1}) - \sigma(G_1[A]) \\
&\quad + \sum_{i=1}^{F(G_1)-1} \sum_{j=F(G_1)}^{n_1} ((F(G_1) - i) - (2F(G_1) - d(x, x_i) - d(x, x_j) - 2)) \\
&\quad + \sum_{i=F(G_1)}^{n_1} \sum_{j=1}^{F(G_1)-1} ((F(G_1) - j) - (2F(G_1) - d(x, x_i) - d(x, x_j) - 2)) \\
&\quad + \sum_{i=F(G_1)}^{n_1} \sum_{j=F(G_1)}^{n_1} (1 - (2F(G_1) - d(x, x_i) - d(x, x_j) - 1)) \\
&\geq \sum_{i=1}^{F(G_1)-1} (n_1 - F(G_1) + 1)(2 - F(G_1) - i) + \sum_{i=1}^{F(G_1)-1} \sum_{j=F(G_1)}^{n_1} (d(x, x_i) + d(x, x_j)) \\
&\quad + \sum_{j=1}^{F(G_1)-1} (n_1 - F(G_1) + 1)(2 - F(G_1) - j) + \sum_{i=F(G_1)}^{n_1} \sum_{j=1}^{F(G_1)-1} (d(x, x_i) + d(x, x_j)) \\
&\quad + (n_1 - F(G_1) + 1)^2(2 - 2F(G_1)) + \sum_{i=F(G_1)}^{n_1} \sum_{j=F(G_1)}^{n_1} (d(x, x_i) + d(x, x_j)) \\
&= 2(n_1 - F(G_1) + 1) \left((F(G_1) - 1)(2 - F(G_1)) - \frac{F(G_1)(F(G_1) - 1)}{2} \right) \\
&\quad + \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} (d(x, x_i) + d(x, x_j)) - \sum_{i=1}^{F(G_1)-1} \sum_{j=1}^{F(G_1)-1} (d(x, x_i) + d(x, x_j))
\end{aligned}$$

$$\begin{aligned}
& + (n_1 - F(G_1) + 1)^2 (2 - 2F(G_1)) \\
= & (n_1 - F(G_1) + 1)(F(G_1) - 1)(4 - 2F(G_1) - F(G_1)) \\
& - 2(n_1 - F(G_1) + 1)^2 (F(G_1) - 1) + 2n_1 \sigma_{G_1}(x) - 2(F(G_1) - 1) \sigma_{G_1[A]}(x) \\
\geq & (n_1 - F(G_1) + 1)(F(G_1) - 1)(4 - 3F(G_1) - 2n_1 + 2F(G_1) - 2) \\
& + 2n_1 \sigma_{G_1}(x) - 2(F(G_1) - 1) \sum_{i=1}^{F(G_1)-2} i \\
= & (n_1 - F(G_1) + 1)(F(G_1) - 1)(2 - F(G_1) - 2n_1) \\
& + 2n_1 \sigma_{G_1}(x) - 2(F(G_1) - 1) \left(\frac{(F(G_1) - 1)(F(G_1) - 2)}{2} \right) \\
= & 2n_1 \sigma_{G_1}(x) + (F(G_1) - 1)(-2n_1^2 + n_1 F(G_1) + F(G_1)^2 - 3F(G_1) + 2) \\
& - (F(G_1) - 1)(F(G_1) - 1)(F(G_1) - 2) \\
= & 2n_1 \sigma_{G_1}(x) + (F(G_1) - 1)(-2n_1^2 + n_1 F(G_1)) \\
= & 2n_1 \sigma_{G_1}(x) - 2n_1^2 F(G_1) + n_1 F(G_1)^2 + 2n_1^2 - n_1 F(G_1) \\
= & 2n_1 \sigma_{G_1}(x) - 2n_1 \sigma_{G'_1}(x)
\end{aligned}$$

Le troisième terme peut être écrit de la manière suivante.

$$\begin{aligned}
& \sum_{i=1}^{n_1} \left(\sum_{v \in G_2} d'(x'_i, v) - \sum_{v \in G_2} d(x_i, v) \right) = \sum_{i=1}^{n_1} n_2 (d'(x'_i, x'_1) - d(x_i, x)) \\
= & n_2 (F(G_1) - 1) \left(n_1 - \frac{F(G_1)}{2} \right) - n_2 \sigma_{G_1}(x) \\
\geq & n_1 (F(G_1) - 1) \left(n_1 - \frac{F(G_1)}{2} \right) - n_1 \sigma_{G_1}(x) \\
= & n_1 \sigma_{G'_1}(x) - n_1 \sigma_{G_1}(x).
\end{aligned}$$

Nous déduisons des valeurs des trois termes que $\sigma(G') - \sigma(G)$ est positive. Si $G_1[A]$ n'est pas un chemin de $F(G_1) - 1$ sommets, alors $\sigma(P_{F(G_1)-1}) > \sigma(G_1[A])$ est vraie et

l'inégalité suivant (*) est stricte. Si $G_1[A]$ est un tel chemin, alors il existe un sommet $v \in A$ tel que $d(x, v) < F(G_1) - 2$ et v a au moins trois sommets voisins dans G_1 (sinon G_1 est un cerf-volant). Soit $v' \notin A$ un voisin de v . Alors $d(v, v') = 1$ et $d(x, v') \leq F(G_1) - 2$, et nous avons

$$\begin{aligned} & 2F(G_1) - d(x, v) - d(x, v') - 2 \\ & \geq 2F(G_1) - (F(G_1) - 3) - (F(G_1) - 2) - 2 = 3 > d(v, v') \end{aligned}$$

L'inégalité (*) est alors stricte, ce qui prouve que l'énoncé $\sigma(G) < \sigma(G')$ est vrai. \square

Lemme 3.2.3. *Soit G un graphe qfc contenant un pont xy tel que y a au moins trois voisins. Nous supposons que $\sigma_{G_2}(y)$ est égal à $\max_{v \in G_2} \sigma_{G_2}(v)$. Alors*

(a) *y a un voisin z qui n'est pas l'extrémité d'un pont,*

(b) *Si G_1 est une clique d'ordre au moins 2 ou le graphe $G_{2,n_1-k+2,k}$ pour un entier $k \geq 3$, et si G' est défini comme le graphe obtenu à partir de G en enlevant z de G_2 et en ajoutant un nouveau sommet noté z' à la clique de G_1 , nous avons alors*

$$F(G') = F(G) \text{ and } \sigma(G') - \sigma(G) \geq 2k(n_2 - n_1 + k - 3),$$

où k dénote $F(G_1)$.

Preuve. Pour prouver (a), nous supposons que chaque voisin de y est l'extrémité du pont. Soit y_1, \dots, y_r les voisins de y dans G_2 . Si chaque arête yy_i pour $i \in \{1, 2, \dots, r\}$ est un pont, alors soit n'_i le nombre de sommets de la composante connexe de $G - \{yy_i\}$ contenant y_i . Nous avons

$$\sigma_{G_2}(y) = \sigma_{G_2}(y_i) + (n'_i - 1) - \sum_{j \neq i} n'_j \leq \sigma_{G_2}(y_i) + n'_i - n'_j - 1$$

pour toutes les paires $\{i, j\}$. Comme $\sigma_{G_2}(y) = \max_{v \in G_2} \sigma_{G_2}(v)$, nous avons $n'_i \geq n'_j + 1$ pour toutes les paires $\{i, j\}$, ce qui est impossible. Alors au moins deux voisins de y , par exemple, y_1 et y_2 , sont tels que yy_1 et yy_2 ne sont pas des ponts. Comme y_i (pour $i = 1, 2$) est l'extrémité d'un pont, il a un voisin z_i tel que $G_2 - \{y_iz_i\}$ est non connexe. Dénotons par G'_1 (resp. G'_2) la composante connexe de $G_2 - \{y_1z_1\}$ (resp. $G_2 - \{y_2z_2\}$) contenant z_1 (resp. z_2). Soit n'_1 (resp. n'_2) l'ordre de G'_1 (resp. G'_2). Nous supposons sans perte de généralité que $n'_1 \leq n'_2$. Considérons un sommet w de G'_1 tel que $d(z_1, w)$ est maximum. Nous montrons que $\sigma_{G_2}(w) - \sigma_{G_2}(y) > 0$, ce qui contredit l'égalité $\sigma_{G_2}(y) = \max_{v \in G_2} \sigma_{G_2}(v)$.

Il est clair que $d(w, v) \geq d(y_1, v) + 1 \geq d(y, v)$ pour tout $v \neq y$ dans $G_2 - G'_1 - G'_2$, et que $d(w, y) \geq 2$. Nous avons

$$\begin{aligned} \sigma_{G_2}(w) - \sigma_{G_2}(y) &= \sum_{v \in G'_1} (d(w, v) - d(y, v)) + \sum_{v \in G'_2} (d(w, v) - d(y, v)) \\ &\quad + \sum_{v \in G_2 - G'_1 - G'_2} (d(w, v) - d(y, v)) \\ &\geq \sum_{v \in G'_1} (d(w, v) - d(y, v)) + \sum_{v \in G'_2} (d(w, v) - d(y, v)) + 2. \end{aligned}$$

Comme $d(w, v)$ est au moins 1 pour tout $v \neq w$ dans G'_1 , $d(y, v)$ est au plus $d(w, z_1) + 2$ pour chaque $v \in G'_1$, et $d(y, v)$ est au plus $d(y_1, v) = d(w, v) - d(w, z_1) - 1$ pour chaque v dans G'_2 , nous avons

$$\begin{aligned} \sigma_{G_2}(w) - \sigma_{G_2}(y) &\geq \sum_{v \in G'_1} (1 - (d(w, z_1) + 2)) - 1 \\ &\quad + \sum_{v \in G'_2} (d(y, v) + d(w, z_1) + 1 - d(y, v)) + 2 \\ &= n'_1(-d(w, z_1) - 1) + n'_2(d(w, z_1) + 1) + 1 \\ &= (n'_2 - n'_1)(d(w, z_1) + 1) + 1 \end{aligned}$$

$$> 0.$$

Pour prouver (b), nous supposons que G_1 est une clique d'ordre au moins 2 ou le graphe $G_{2,n_1-k+2,k}$ pour un $k \geq 3$. Comme G est qfc et z n'est pas l'extrémité d'un pont, nous avons $F(G_2 - z) = F(G_2)$ par le lemme 3.1.3. D'où

$$F(G') = F(G_1 + z') + F(G_2 - z) = F(G_1) + F(G_2) = F(G).$$

Pour tous les sommets $v, w \neq z'$, nous avons

$$d_{G'}(v, w) = d_{G' - \{z'\}}(v, w) = d_{G - \{z\}}(v, w) \geq d_G(v, w).$$

Nous obtenons les résultats suivants, qui sont valides que G_1 soit une clique ou un cerf-volant.

$$\begin{aligned} \frac{\sigma(G') - \sigma(G)}{2} &= \frac{1}{2} \left(\sum_{v \neq z', w \neq z'} d_{G'}(v, w) + 2\sigma_{G'}(z') - \sum_{v \neq z, w \neq z} d_G(v, w) - 2\sigma_G(z) \right) \\ &\geq \sigma_{G'}(z') - \sigma_G(z) \\ &= (n_1 - k + 2) + \sum_{i=2}^{k-1} i + (n_2 - 1)k + \sigma_{G_2 - \{z\}}(y) \\ &\quad - \left(\sigma_{G_2}(z) + \sum_{i=2}^k i + (n_1 - k + 1)(k + 1) \right) \\ &= n_1 - k + 2 + \frac{k(k - 1)}{2} - 1 + n_2k - k + \sigma_{G_2 - \{z\}}(y) \\ &\quad - \sigma_{G_2}(z) - \frac{k(k + 1)}{2} + 1 - n_1k - n_1 + k^2 - 1 \\ &= (n_2 - n_1)k + k^2 - 3k + 1 + \sigma_{G_2 - \{z\}}(y) - \sigma_{G_2}(z) \end{aligned}$$

Par le choix de y , nous avons $\sigma_{G_2-\{z\}}(y) \geq \sigma_{G_2}(y) - 1 \geq \sigma_{G_2}(z) - 1$. D'où

$$\begin{aligned} \frac{\sigma(G') - \sigma(G)}{2} &\geq (n_2 - n_1)k + k^2 - 3k + 1 - 1 \\ &= (n_2 - n_1 + k - 3)k. \end{aligned}$$

□

Lemme 3.2.4. *Soit G un graphe qfc de n sommets dans lequel chaque pont xy a une extrémité x de degré 1 et l'autre extrémité y telle que $\sigma_{G-x}(y) = \max_{v \in G-x} \sigma_{G-x}(v)$. Si G n'est pas un chemin, alors*

(a) *G contient un seul pont xy (vérifiant ces conditions), et*

(b) *il existe au moins un sommet z dans G qui n'est pas l'extrémité d'un pont et tel que $\sigma_G(z) \leq \frac{\sigma(G)}{n}$.*

Preuve. Pour prouver (a), nous supposons que G contient deux ponts xy et $x'y'$ avec x et x' de degré 1. Si $y = y'$, alors $n > 3$ (comme G n'est pas un chemin), et nous obtenons

$$\begin{aligned} \sigma_{G-x'}(y) &= \sigma_G(y) - 1 = \sigma_G(x) - (n - 1) = (\sigma_{G-x'}(x) + 2) - (n - 1) \\ &= \sigma_{G-x'}(x) - n + 3 < \sigma_{G-x'}(x), \end{aligned}$$

ce qui contredit l'hypothèse $\sigma_{G-x'}(y) = \max_{v \in G-x'} \sigma_{G-x'}(v)$.

Si $y \neq y'$, alors nous supposons (sans perte de généralité) que $\sigma_G(y) \geq \sigma_G(y')$. D'autre part la relation $d(x, x') \leq n - 2$ est vraie (sinon G serait un chemin), et nous obtenons

$$\begin{aligned} \sigma_{G-x'}(y') &= \sigma_G(y') - 1 < \sigma_G(y) \leq \sigma_G(y) + (n - 2) - d(x, x') \\ &= \sigma_G(x) - d(x, x') = \sigma_{G-x'}(x). \end{aligned}$$

Ceci contredit l'hypothèse $\sigma_{G-x'}(y') = \max_{v \in G-x'} \sigma_{G-x'}(v)$.

Pour prouver (b) nous supposons que G contient un pont; sinon on peut choisir z tel que $\sigma_G(z) = \min_{v \in G} \sigma_G(v)$. Alors soit xy l'unique pont dans G , avec x un sommet de degré 1, et soit S l'ensemble des sommets z de G tels que $\sigma_G(z) \leq \frac{\sigma(G)}{n}$. Il reste à prouver que S contient un sommet $z \neq x, y$. Supposons le contraire. Alors y appartient à S car $\sigma_G(x) = \sigma_G(y) + n - 2 > \sigma_G(y)$ et S n'est pas vide. Il s'ensuit que $\sigma_G(w) \geq \sigma_G(y) + 1$ pour tout $w \neq y$. Donc pour tout voisin $z \neq x$ de y , nous avons $\sigma_G(z) = \sigma_G(y) + 1$ car

$$\sigma_G(z) \geq \sigma_G(y) + 1 = \sigma_{G-x}(y) + 2 \geq \sigma_{G-x}(z) + 2 = \sigma_G(z).$$

Soit z un voisin de y . Alors nous avons

$$\begin{aligned} \sigma(G) &= \sigma_G(x) + \sigma_G(y) + \sigma_G(z) + \sum_{w \neq x, y, z} \sigma_G(w) \\ &\geq (\sigma_G(y) + n - 2) + \sigma_G(y) + (\sigma_G(y) + 1) + (n - 3)(\sigma_G(y) + 1) \\ &= n\sigma_G(y) + 2n - 4. \end{aligned}$$

Comme on a $n \geq 4$ (car sinon G serait un chemin), nous obtenons $\frac{\sigma(G)}{n} \geq \sigma_G(y) + 2 - \frac{4}{n} \geq \sigma_G(y) + 1 = \sigma_G(z)$, ce qui contredit l'hypothèse que z n'appartient pas à S . \square

Lemme 3.2.5. *Soit G un graphe qfc de n sommets. Supposons qu'il existe un sommet z de G qui n'appartient pas à un pont et tel que $\sigma_G(z) \leq \frac{\sigma(G)}{n}$. Alors $F(G - z)$ est égal à $F(G)$ et $\mu(G - z)$ est au moins $\mu(G)$.*

Preuve. Comme G est qfc et z n'appartient pas à un pont, le lemme 3.1.3 entraîne que

$F(G - z)$ est égal à $F(G)$. De plus, les relations suivantes sont faciles à vérifier.

$$\begin{aligned}\mu(G - z) &= \frac{\sigma(G - z)}{(n - 1)(n - 2)} = \frac{\sigma(G) - 2\sigma_G(z)}{(n - 1)(n - 2)} \\ &\geq \frac{\sigma(G) - 2\frac{\sigma(G)}{n}}{(n - 1)(n - 2)} = \frac{\frac{(n-2)\sigma(G)}{n}}{(n - 1)(n - 2)} \\ &= \frac{\sigma(G)}{n(n - 1)} = \mu(G).\end{aligned}$$

□

3.3 Le théorème

Nous allons maintenant prouver que les doubles cerfs-volants sont des graphes extrêmes quand $F(G)$ est au moins égal à 4.

Définition 3.3.1. Soit deux entiers n et k tels que $n \geq k$. Nous disons que le graphe G est (n, k) -forêt-extrême s'il est d'ordre n , $F(G)$ est égal à k , et G a la plus grande distance moyenne parmi tous les graphes H d'ordre n tels que $F(H) = k$.

Théorème 3.3.2. Soit G un graphe connexe de n sommets. Alors

$$\mu(G) \leq \frac{F(G)}{2}.$$

De plus, si $F(G)$ est au moins égal à 4, G est $(n, F(G))$ -forêt-extrême si et seulement si G est un graphe de la forme $G_{n,k}$.

Preuve. Nous appliquons les transformations suivantes au graphe G , en ordre de priorités décroissantes. En d'autres termes, si la transformation no i (pour tout i) peut être utilisée, elle sera appliquée avant la transformation no $(i + 1)$. Le graphe transformé

est noté G' et son ordre n' . Ainsi, au début de l'algorithme, G' est initialisé à G et n' à n , l'ordre de G . L'algorithme s'arrête si aucune transformation ne peut être appliquée. Notons que chaque fois que nous nous référons au pont xy de G' dans cette preuve, G_1 et G_2 dénotent les composantes connexes du graphe $G' - xy$ et l'ordre de G_1 est au plus l'ordre de G_2 (c'est-à-dire que $n_1 \leq n_2$). Voici les transformations.

1. Si G' contient une arête e qui n'est pas forêt-critique et qui n'est pas un pont, alors nous remplaçons G' par $G' - e$.
2. Si G' contient un pont xy tel que $3 \leq n_1 \leq n_2$ et G_1 n'est ni une clique ni le graphe $G_{2,n_1-F(G_1)+2,F(G_1)}$ avec x comme extrémité, nous remplaçons G_1 dans G' par $G_{2,n_1-F(G_1)+2,F(G_1)}$ avec x comme extrémité.

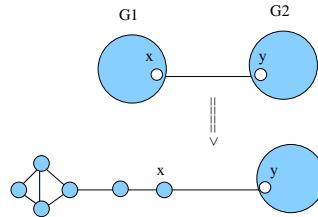


Figure 3.5 . Illustration de la transformation no 2

3. S'il y a un pont xy tel que $\sigma_{G_2}(y) \neq \max_{w \in G_2} \sigma_{G_2}(w) = \sigma_{G_2}(y')$, nous remplaçons l'arête xy dans G' par l'arête xy' .

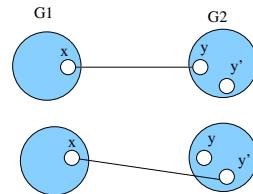


Figure 3.6 . Illustration de la transformation no 3

4. Supposons que G' n'est pas un double cerf-volant équilibré et contient un pont xy tel que $\deg_{G'}(y)$ est au moins 3, $n_2 - n_1 + F(G_1) - 3$ est positif et G_1 est une

clique d'ordre au moins 2 ou le graphe $G_{2,n_1-F(G_1)+2,F(G_1)}$. Alors on supprime un voisin de y dans le graphe G' qui n'est pas l'extrémité d'un pont et on l'ajoute à la clique de G_1 .

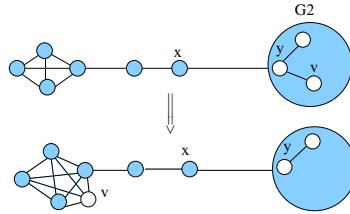


Figure 3.7 . Illustration de la transformation no 4

5. Si G' est d'ordre au moins 3 et contient un sommet z qui n'est pas l'extrémité d'un pont et tel que $\sigma_{G'}(z)$ est au plus $\frac{\sigma(G')}{n'}$, alors on enlève z de G' et l'ordre devient égal à $n' - 1$.

Si $F(G)$ est égal à 1 (c'est à dire, le graphe original G est un sommet isolé), aucune transformation ne sera appliquée. Si $F(G)$ est égal à 2, G est une clique, donc un graphe forêt-critique, et l'algorithme ci-dessus va appliquer la transformation no 5 jusqu'à l'obtention du graphe K_2 , c'est-à-dire un chemin de longueur 1. Si $F(G)$ est au plus 2, la relation $\mu(G) \leq \frac{F(G)}{2}$ est trivialement vérifiée.

Si $F(G)$ est égal à 3, la transformation no 1 est appliquée jusqu'à ce que l'algorithme produise un graphe qfc G' . À ce stade, G' est un graphe forêt-critique ou bien il contient un pont. Si G' est forêt-critique, la transformation no 5 sera appliquée par l'algorithme. Si G' contient un pont xy , alors G_1 est un sommet isolé (x) et G_2 est une clique. Alors si G' n'est pas un chemin de trois sommets, le lemme 3.2.4 implique que l'algorithme va appliquer la transformation no 5 et produire P_3 , le chemin de trois sommets. Remarquons que comme les transformations nos 1 et 5 ne diminuent pas $\mu(G')$ et

gardent $F(G')$ fixé, alors nous avons

$$\mu(G) \leq \mu(G') = \mu(P_3) = \frac{4}{3} < \frac{3}{2} = \frac{F(G)}{2}.$$

Ainsi, dans le reste de la preuve, nous supposons que $F(G)$ est au moins égal à 4. Nous allons prouver plus loin que l'algorithme s'arrête après un nombre fini d'itérations, en considérant le résultat après chaque transformation. Nous commençons par prouver que le graphe final de l'algorithme est un double cerf-volant équilibré. Nous allons montrer qu'à moins que G' (qui est identique à G au début de l'algorithme) soit un double cerf-volant équilibré, au moins une transformation peut être appliquée à G' .

Supposons que G' n'est pas un double cerf-volant équilibré et les transformations nos 1, 2 et 3 ne peuvent pas être appliquées. Nous affirmons que la transformation no 4 ou la transformation no 5 peut être appliquée. Observons que si G' est isomorphe au graphe $G_{m_1, m_2, k}$, où $m_1 \leq m_2 - 2$, la transformation no 4 peut être appliquée en utilisant le pont incident avec la clique d'ordre m_2 . Nous pouvons supposer que G' n'est pas un double cerf-volant. G' est un graphe qfc car la transformation no 1 ne peut être appliquée.

Il y a deux cas à traiter. Si $F(G_1)$ est égal à 1 pour chaque pont xy , la transformation no 4 ne peut être appliquée mais le lemme 3.2.4 implique que la transformation no 5 peut être appliquée. D'autre part, si $F(G_1) \geq 2$ est vrai pour au moins un pont xy , G_1 doit être une clique ou un cerf-volant (sinon la transformation no 2 peut être appliquée). Notons que n_1 est strictement inférieur à n_2 ; en effet, si n_1 et n_2 sont égaux, la transformation no 2 peut être appliquée à G' (qui n'est pas un double cerf-volant) en considérant le pont yx (y prend la place de x). Donc on obtient $n_2 - n_1 + F(G_1) - 3 \geq 0$.

Si $\deg_{G'}(y)$ est au moins égal à 3, la transformation no 4 peut alors être appliquée. Si $\deg_{G'}(y)$ est strictement inférieur à 3, soit y' le sommet dans G_2 le plus proche de y

parmi tous les sommets vérifiant $\deg_{G'}(y') \geq 3$. Un tel sommet doit exister car G_2 n'est pas un chemin (sinon G' est un double cerf-volant). Soit x' le prédecesseur de y' sur le chemin de y à y' . Considérons le pont $x'y'$ au lieu de xy . Alors $n'_1 < n'_2$ (où n'_1 et n'_2 dénotent les ordres des composantes de $G' - x'y'$), sinon la transformation no 2 pourrait être appliquée en considérant le pont $y'x'$ (y' jouant le rôle de x). La transformation no 4 peut donc être appliquée.

Nous avons prouvé que si l'algorithme s'arrête, alors le graphe G' obtenu est un double cerf-volant équilibré. Nous allons maintenant montrer que l'algorithme s'arrête à la suite d'un nombre fini de transformations. En effet, les transformations nos 1, 2 et 3 font augmenter $\mu(G')$ strictement, et la transformation no 5 fait diminuer l'ordre de G' de 1. Considérons maintenant l'effet de la transformation no 4. D'abord observons que lorsque cette transformation est appliquée, n_2 est strictement supérieur à n_1 (sinon la transformation no. 2 peut être appliquée en inversant les rôles de x et y).

La transformation no 4 fait augmenter strictement la valeur de $\mu(G')$ si n_2 est au moins égal à $n_1 + 2$ ou $F(G_1)$ est au moins égal à 3. Supposons que n_2 est égal à $n_1 + 1$ et $F(G_1)$ est égal à 2, et soit z le sommet déplacé de G_2 à G_1 par la transformation no 4. Soit G'' le graphe résultant de cette transformation. Comme G_2 n'est pas une clique et G' est qfc, $F(G_2 - z)$ est au moins 3.

- . Si $G_2 - z$ n'est pas qfc, la transformation no 1 est appliquée à G'' , et $\mu(G')$ augmente strictement.
- . Si $G_2 - z$ est qfc mais n'est pas un cerf-volant avec y comme extrémité, la transformation no 2 est appliquée à G'' et $\mu(G')$ augmente strictement.
- . Si $G_2 - z$ est un cerf-volant et y est l'extrémité, la transformation no 4 est appliquée à G'' en utilisant yx comme pont. La valeur de $\mu(G')$ augmente strictement car $F(G_2 - z)$ est au moins 3 et la composante $G'' - yx$ contenant x a un sommet de

plus que la composante contenant y .

Nous observons que $\mu(G')$ va augmenter strictement dans tous les cas. Ainsi l'algorithme s'arrête, après un nombre fini d'itérations, car à la suite d'au plus deux itérations consécutives, $\mu(G')$ augmente strictement ou l'ordre de G' diminue strictement. De plus, si G' n'est pas identique à G , alors soit $n' < n$ ou $\mu(G') > \mu(G)$.

Si $F(G)$ est au moins 4, le graphe produit par l'algorithme est clairement $G_{n',F(G)}$ et on a $\mu(G) \leq \mu(G_{n',F(G)})$. L'inégalité étant stricte si $n = n'$ et $G \neq G_{n,F(G)}$. De plus, nous savons par le lemme 3.1.9 que nous avons $\mu(G_{n',F(G)}) < \mu(G_{n,F(G)})$ si $n' < n$. Nous concluons que nous avons $\mu(G) \leq \mu(G_{n,F(G)})$, l'inégalité est stricte si $G \neq G_{n,F(G)}$. Finalement, nous savons d'après le Lemme 3.1.10 que $\mu(G_{n,F(G)})$ est au plus $\frac{F(G)}{2}$, ce qui complète la preuve. \square

Les seuls graphes (n, k) -forêt-extrêmes tels que $k < 3$ sont des cliques de n sommets. Nous caractérisons ci-dessous les graphes $(n, 3)$ -forêt-extrêmes.

Théorème 3.3.3. *Un graphe connexe G de n sommets vérifiant $F(G) = 3$ est un graphe $(n, 3)$ -forêt-extrême si et seulement si \overline{G} est l'union de deux étoiles.*

Preuve. Si G est connexe et $F(G) \leq 3$, alors \overline{G} est P_4 -libre (c'est-à-dire qu'il ne contient pas un sous-graphe induit isomorphe à P_4), ne contient pas de cycle et n'est pas connexe. En effet,

- . \overline{G} est P_4 -libre (sinon G contient un P_4 , ce qui implique que $F(G) > 3$);
- . \overline{G} ne contient pas C_3 , sinon l'union de C_3 avec un sommet de G (qui existe car G est connexe) induit une forêt de 4 sommets dans G ;
- . \overline{G} ne contient pas C_4 , sinon C_4 induit une forêt de quatre sommets dans G ;

. \overline{G} n'a pas de cycle supérieur à 4 car il est P_4 -libre.

D'où \overline{G} est une forêt P_4 -libre, c'est-à-dire une forêt dans laquelle chaque composante est une étoile. Alors \overline{G} ne peut être connexe (sinon G ne serait pas connexe).

Réiproquement, si \overline{G} est P_4 -libre, ne contient aucun cycle et n'est pas connexe, alors G est connexe et $F(G) \leq 3$. En effet, G est connexe car le complémentaire d'une forêt non connexe est connexe, et $F(G)$ est au plus 3 (car sinon G contiendrait un $\overline{K_3}$, un P_4 , ou un $2K_2$, ce qui implique que \overline{G} contient un triangle, un P_4 ou un carré).

Nous concluons que G est connexe et $F(G) = 3$ si et seulement si \overline{G} est l'union de $\ell \geq 2$ étoiles, dont au moins une contient deux sommets. Comme $\sigma(G) = n(n-1) + 2(n-\ell)$ dans de tels graphes, la distance moyenne est maximale si et seulement si $\ell = 2$.

□

Exemple: Nous présentons dans la suite un exemple d'un graphe connexe de douze sommets auquel nous appliquons les transformations définies précédemment. Notons qu'après chacune des transformations la distance moyenne augmente et la taille de la forêt maximum demeure égale à 8.

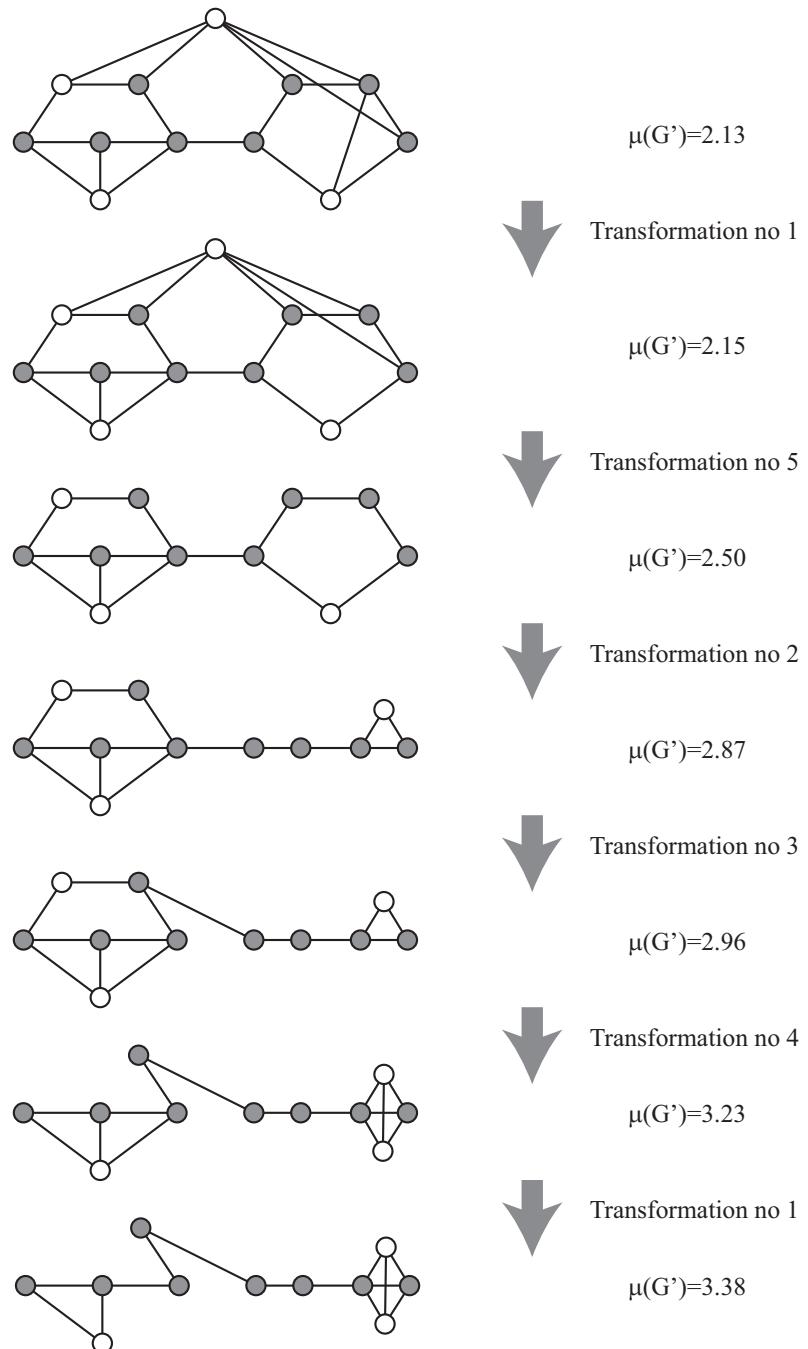


Figure 3.8 . Un exemple illustrant les cinq transformations

3.4 Résultats connexes

Définition 3.4.1. Soit n et k deux entiers tels que $n \geq k$. Le graphe G est (n, k) - α_2 -extrême s'il contient n sommets, $\alpha_2(G)$ est égale à k , et il a une distance moyenne maximale parmi tous les graphes d'ordre n contenant un sous-graphe biparti maximum d'ordre k .

Corollaire 3.4.2. Soit G un graphe connexe (n, k) - α_2 -extrême. Si k est égal à 1 ou 2, alors G est une clique d'ordre n . Si k est égal à 3, alors G est une clique d'ordre $n - 1$ plus une arête pendante. Si k est au moins 4, alors G est le graphe $G_{n, \alpha_2(G)}$.

Preuve. Soit $A_{\alpha_2}(n, k)$ l'ensemble des graphes G de n sommets tels que $\alpha_2(G) = k$, et soit $B_F(n, k)$ l'ensemble des graphes G de n sommets tels que $F(G) \leq k$. Comme $F(G) \leq \alpha_2(G)$ pour tout graphe G , nous avons $A_{\alpha_2}(n, k) \subseteq B_F(n, k)$. Dans la section précédente nous avons prouvé que les graphes appartenant à $B_F(n, k)$ de distance moyenne maximale sont de la forme $G_{n, k}$ si $k > 3$, le complémentaire de l'union de deux étoiles si $k = 3$, et une clique si $k < 3$.

- . Comme $G_{n, k} \in A_{\alpha_2}(n, k)$ nous concluons que $G_{n, k}$ est l'unique graphe (n, k) - α_2 -extrême pour $k > 3$.
- . Pour $k = 3$, soit $G(a, b)$ le complémentaire de deux étoiles, l'une ayant a sommets et l'autre ayant b sommets. Si a et b sont strictement supérieurs à 1, alors G contient un carré, donc $\alpha_2(G(a, b)) \geq 4$. Comme $G(1, b) \in A_{\alpha_2}(n, 3)$ nous concluons que $G(1, n - 1)$ est l'unique graphe $(n, 3)$ - α_2 -extrême .
- . Comme la clique d'ordre n appartient à $A_{\alpha_2}(n, k)$ pour $k < 3$, nous concluons que la clique est l'unique graphe $(n, 3)$ - α_2 -extrême pour $k < 3$.

□

Les trois graphes $(6, 3)$ -forêt-extrêmes sont représentés sur la première ligne de la Figure 3.1, et leurs complémentaires sont sur la deuxième ligne. Le premier graphe est l'unique graphe $(6, 3)$ - α_2 -extrême.

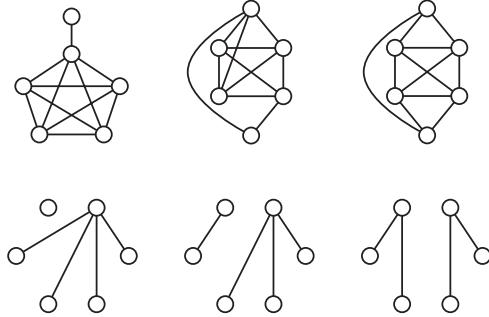


Figure 3.9 . Les trois graphes $(6,3)$ -forêt-extrêmes et leurs complémentaires

Nous pouvons essayer de généraliser ce résultat. Par exemple, soit $T(G)$ la taille du plus grand arbre induit dans G . Comme $T(G) \leq F(G)$ pour tout graphe G , on peut essayer de prouver que $\mu(G) \leq \frac{T(G)}{2}$ pour tout graphe G . Ceci n'est cependant pas vrai. En effet, comme le montre le graphe de la Figure 3.2, nous avons $\mu(G) = \frac{91}{45} > 2 = \frac{T(G)}{2}$.

De la même façon, nous pourrions être amenés à penser que $\mu(G) \leq \frac{\alpha_k(G)}{k}$ pour $k > 2$. Cependant, le plus petit contre-exemple est la clique de deux sommets, où $\mu(G) = 1 > \frac{2}{k} = \frac{\alpha_k(G)}{k}$ pour tout $k > 2$.

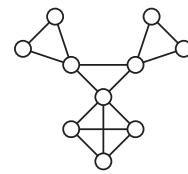


Figure 3.10 . Un graphe tel que $T(G) = 4$ et $\mu(G) > 2$

Soit $LF(G)$ le nombre maximum de sommets dans une forêt linéaire induite, c'est-à-dire, un sous-graphe induit de G qui est une union de chemins. Observons que $LF(G) \leq F(G)$ pour chaque graphe G . Nous proposons une nouvelle conjecture, qui, si elle est vraie, sera plus forte que le Théorème 3.3.2.

Conjecture 3.4.3. Soit G un graphe connexe de n sommets. Alors

$$\mu(G) \leq \frac{LF(G)}{2}.$$

Pour illustrer la différence entre les résultats connus et la nouvelle conjecture, considérons le graphe de la Figure 3. Nous avons

$$\mu(G) = \frac{15}{7} < 3 = \frac{LF(G)}{2} < \frac{7}{2} = \frac{F(G)}{2} < 4 = \frac{\alpha_2(G)}{2} < 6 = \alpha(G)$$

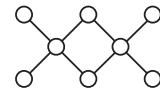


Figure 3.11 . Illustration des différences entre $\alpha(G)$, $\frac{\alpha_2(G)}{2}$, $\frac{F(G)}{2}$ et $\frac{LF(G)}{2}$

CHAPITRE 4

BORNES SUR DES DIFFÉRENCES ENTRE DES INVARIANTS

Dans ce chapitre nous allons effectuer des comparaisons systématiques entre les invariants suivants: la distance moyenne, le stable maximum, le bistable, la forêt induite, l'arbre induit et la forêt linéaire induite maximum.

En premier lieu, on va déterminer des bornes supérieures sur les différences entre ces invariants. Les relations que nous allons présenter ont la forme suivante:

$$INV_1(G) - INV_2(G) \leq f(n)$$

On note par $INV_i(G)$ un invariant du graphe G parmi ceux cités ci-dessus. $f(n)$ est une fonction de l'ordre n du graphe et c'est la meilleure borne supérieure possible, dans le sens où elle est atteinte pour toutes les valeurs de n , sauf pour des petites valeurs. Tous les résultats prouvés dans ce chapitre sont valables pour tout graphe connexe G .

Nous présentons également des exemples divers de graphes extrêmes générés par le système AutoGraphiX et nous allons essayer de trouver des classes de graphes vérifiant des égalités entre les cardinalités des sous-graphes induits maximaux.

Étant donné que nous savons que pour tout graphe connexe G d'ordre n on a les résultats suivants:

$$\frac{\alpha_2(G)}{2} - \mu(G) \geq 0$$

et

$$\frac{F(G)}{2} - \mu(G) \geq 0,$$

nous allons proposer, dans la deuxième partie, des bornes supérieures à ces relations, ainsi qu'une borne supérieure pour la différence $\frac{LF(G)}{2} - \mu(G)$.

Pour réaliser ce travail nous avons utilisé le système AutoGraphiX. L'idée est de considérer le problème de trouver la borne recherchée comme un problème d'optimisation. Dans notre cas c'est un problème de maximisation. Théoriquement, le problème se formule comme suit:

$$\max_{G \in \mathcal{G}_n} (INV_1(G) - INV_2(G)),$$

où \mathcal{G}_n désigne la famille des graphes connexes d'ordre n .

La première étape d'AutoGraphiX consiste à résoudre heuristiquement le problème ci-dessus pour différentes valeurs de n . La méthode d'optimisation (la recherche de graphes extrémaux) d'AutoGraphiX est basée sur la *métaheuristique de recherche à voisnages variables* [46].

Les graphes obtenus lors de la résolution du problème de minimisation ou de maximisation ne sont pas nécessairement optimaux, car le système utilise une heuristique. Cependant, les graphes optimaux sont très souvent obtenus pour $n \leq 12$.

La famille de graphes extrémaux obtenue est exploitée, dans une deuxième étape, de trois manières pour obtenir des conjectures [19]:

- (a) une méthode numérique, basée sur les mathématiques de l'analyse en composantes principales, qui fournit une base de relations affines entre invariants;

- (b) une méthode géométrique qui consiste à déterminer la fermeture convexe de l'ensemble des graphes extrémaux obtenus considérés comme des points dans l'espace des invariants;
- (c) une méthode algébrique, qui consiste à reconnaître la (ou les) famille(s) de graphes extrémaux obtenus et à substituer les expressions des invariants apparaissant dans l'objectif en fonction de n , puis à simplifier.

À l'issue de ces étapes, une conjecture est formulée, puis sa preuve est faite.

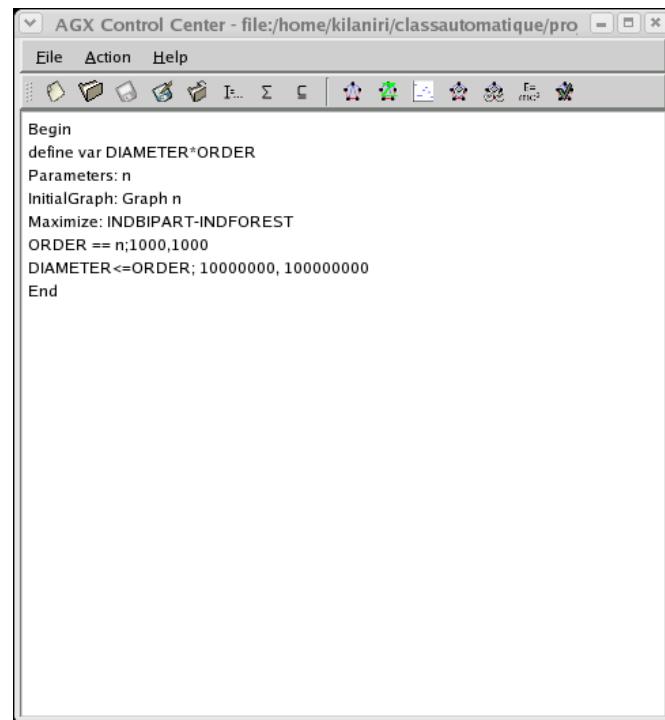
PREMIÈRE PARTIE

4.1 Borne pour $\text{Max}(\alpha_2(G) - F(G))$

Dans cette section nous allons essayer de trouver la borne supérieure sur la différence entre les ordres du bistable maximum et de la forêt induite maximum. Nous utilisons AGX pour réaliser ce travail.

AutoGraphiX: Génération de conjectures

La première étape dans la génération de conjectures est d'éditer notre problème. Voici la syntaxe du problème $\text{Max}(\alpha_2(G) - F(G))$.



The screenshot shows the AGX Control Center window with the following content:

```

AGX Control Center - file:/home/kilaniri/classautomatique/pro
File Action Help
[File icons] [Edit icons] [Graph icons] [Statistics icons] [Graph View icons] [Help icons]
Begin
define var DIAMETER*ORDER
Parameters: n
InitialGraph: Graph n
Maximize: INDBIPART-INDFOREST
ORDER == n;1000,1000
DIAMETER<=ORDER; 10000000, 100000000
End

```

Figure 4.1 . Énoncé du problème

4.1.1 Éditer un problème

Un fichier problème est un fichier ".def", édité une première fois en utilisant un éditeur de texte de base comme *Nedit* et *Emacs*. Les instructions du problème sont comprises entre les deux mots-clés *Begin* et *End* ; chacun d'eux occupe une ligne. Les deux champs, avant *Begin* et après *End* sont réservés aux commentaires de l'usager. Les composantes du champs des instructions sont:

- *Define*: une partie optionnelle qui offre à l'usager la possibilité de définir un nouvel invariant en combinant ceux déjà programmés dans AGX et/ou ceux définis (à l'aide de *Define* même).
- *Parameters*: Les symboles des paramètres du problème, dans l'exemple c'est l'ordre n .
- *InitialGraph*: Tout processus de recherche locale nécessite une solution de départ, pour AutoGraphiX, il s'agit d'un graphe initial. *InitialGraph* est suivi d'un mot-clé correspondant au graphe initial choisi (on peut avoir la liste complète via la documentation). Dans notre problème le graphe initial choisi est un graphe quelconque d'ordre n .
- *Minimize* ou *Maximize*: est suivi de la fonction objectif. Dans l'exemple l'objectif est de maximiser $\alpha_2(G) - F(G)$.
- *Constraint* (optionnelle) suivi d'une liste de contraintes, chacune se terminant avec ":" puis une pénalité constante et une autre linéaire. Dans l'exemple, une première contrainte impose à l'ordre de prendre la valeur du paramètre n et une deuxième contrainte le diamètre à être fini: une façon d'assurer la connectivité du graphe.

L'étape suivante consiste à lancer le processus d'optimisation. Pour l'utilisateur, il s'agit de choisir la cardinalité des graphes à générer ainsi que le nombre de graphes. Comme nous l'avons dit précédemment, AutoGraphiX réussit le plus souvent à générer des graphes optimaux pour $n \leq 12$. Une fois cette étape terminée, le système affiche la liste des graphes associés au problème.

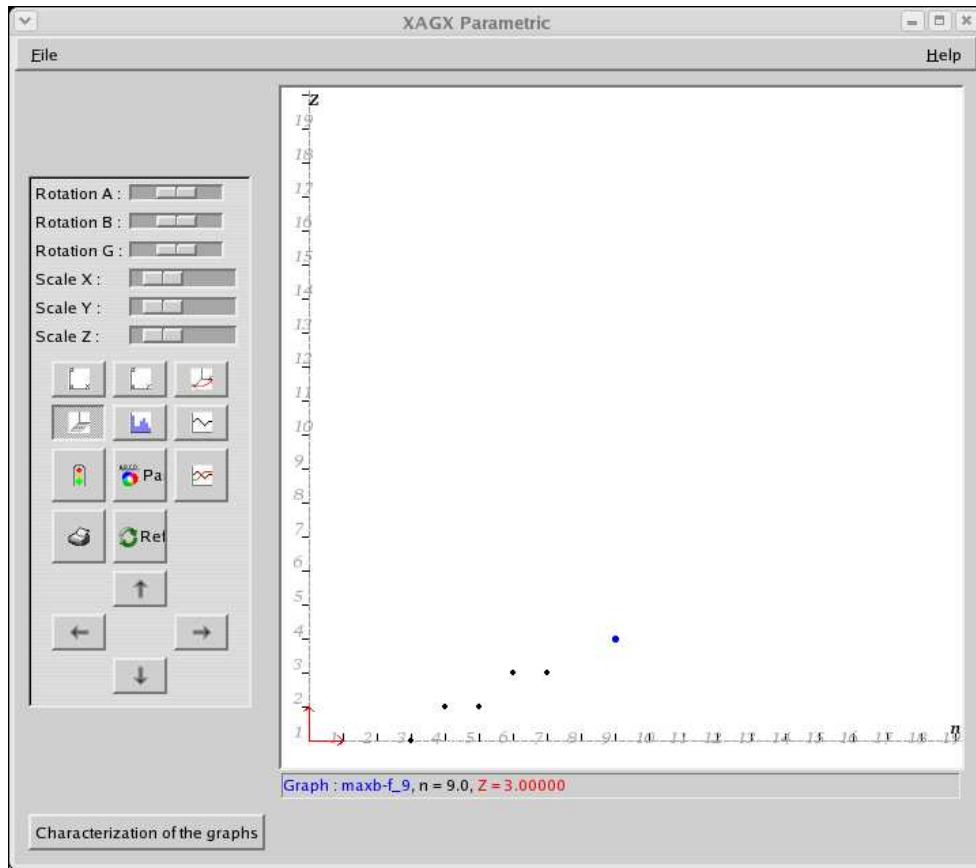


Figure 4.2 . Liste des graphes associés au problème $\text{Max}(\alpha_2(G) - F(G))$

AutoGraphiX affiche les résultats de la recherche des graphes extrêmes sous forme de points (représentants les graphes) dans un espace à deux ou trois dimensions. Dans notre cas une dimension représente la fonction objective Z et une autre représente l'ordre n .

Au cours de la dernière étape, l'usager va choisir des invariants de la liste ou bien de nouveaux invariants définis au début du problème et le système va essayer de trouver une relation qui satisfait tous les graphes extrêmes comme nous l'avons expliqué précédemment. Nous sélectionnons les invariants α_2 , F et $\lfloor \frac{n}{2} \rfloor$. AutoGraphiX réussit à générer la

conjecture suivante pour les graphes extrémaux:

$$F(G) = -\left\lfloor \frac{n}{2} \right\rfloor + \alpha_2(G) + 1$$

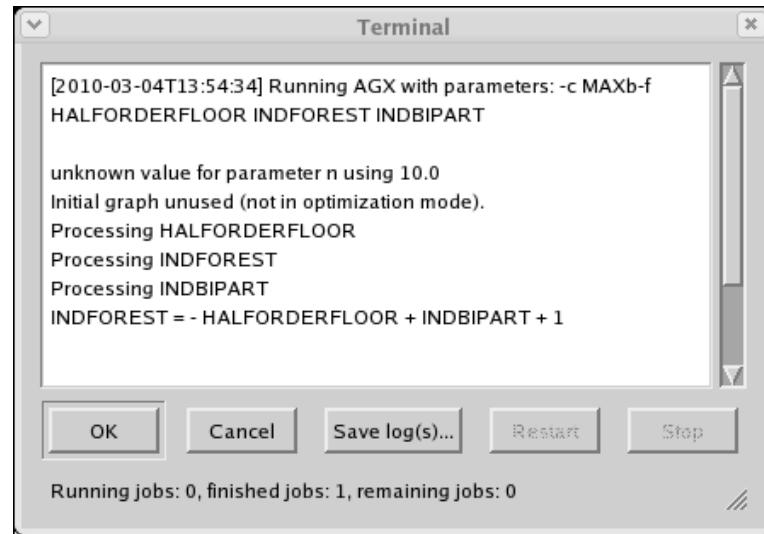


Figure 4.3 . Affichage de la relation trouvée

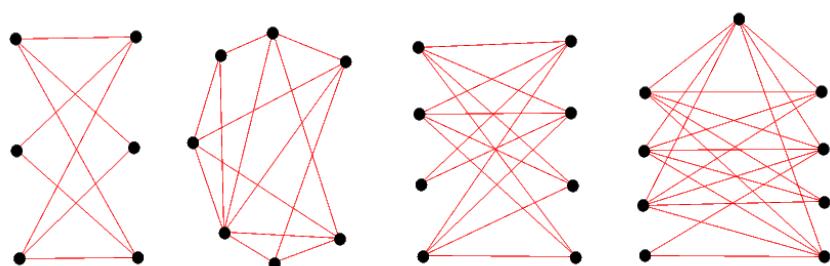


Figure 4.4 . Graphes obtenus par AGX en maximisant $\alpha_2(G) - F(G)$ pour $6 \leq n \leq 9$

Nous allons prouver dans la suite ce résultat.

Proposition 4.1.1. *Soit G un graphe connexe d'ordre n . Alors on a*

$$\alpha_2(G) - F(G) \leq \left\lfloor \frac{n}{2} \right\rfloor - 1.$$

Preuve. G est un graphe connexe d'ordre n . Soit B un bistable maximum de G de taille $\alpha_2(G)$, tel que $B = A_1 \cup A_2$.

Nous supposons que $|A_1| \geq |A_2|$.

Pour tout $v \in A_2$, il est clair que $A_1 \cup \{v\}$ est une forêt. Donc

$$F(G) \geq \left\lceil \frac{\alpha_2(G)}{2} \right\rceil + 1.$$

Il s'ensuit que

$$\alpha_2(G) - F(G) \leq \alpha_2(G) - \left\lceil \frac{\alpha_2(G)}{2} \right\rceil - 1 = \left\lfloor \frac{\alpha_2(G)}{2} \right\rfloor - 1 \leq \left\lfloor \frac{n}{2} \right\rfloor - 1.$$

La dernière inégalité provient du fait que

$$\alpha_2(G) \leq n.$$

□

4.2 Graphes vérifiant $\alpha_2(G) - F(G) = \lfloor \frac{n}{2} \rfloor - 1$

Dans ce paragraphe nous allons essayer de donner une description des graphes extrêmes vérifiant $\alpha_2(G) - F(G) = \lfloor \frac{n}{2} \rfloor - 1$. Malheureusement, nous n'avons pas

réussi à caractériser ces graphes mais nous allons présenter des conditions nécessaires qu'un tel graphe doit vérifier.

Le corollaire suivant est une conséquence de la preuve de la proposition 4.1.1.

Corollaire 4.2.1. *Si $\alpha_2(G) - F(G) = \lfloor \frac{n}{2} \rfloor - 1$, alors $\lfloor \frac{\alpha_2(G)}{2} \rfloor = \lfloor \frac{n}{2} \rfloor$ et $F(G) = \lceil \frac{\alpha_2(G)}{2} \rceil + 1$.*

Pour les Graphes vérifiant $\alpha_2(G) - F(G) = \lfloor \frac{n}{2} \rfloor - 1$, nous allons considérer les cas suivants.

1er cas: Si n est pair, alors $\alpha_2(G) = n$ donc le graphe G est un graphe biparti.

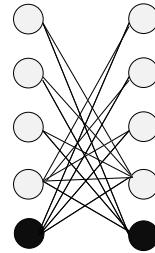
D'après le corollaire, on a $F(G) = \lceil \frac{\alpha_2(G)}{2} \rceil + 1$ alors G est un graphe biparti équilibré vérifiant les conditions suivantes:

- $G = (V_1, V_2, E)$ tel que $|V_1| = |V_2| = \frac{n}{2}$.
- Pour tous $x, y \in V_1$, on a $N_{V_2}(x) \cap N_{V_2}(y) \geq 2$.

où $N_{V_2}(x)$ est l'ensemble des sommets de V_2 qui sont adjacents à x . En effet s'il existe $x, y \in V_1$, tels que $N_{V_2}(x) \cap N_{V_2}(y) < 2$, alors $F(G) \geq \frac{n}{2} + 2$, contradiction.

• Par symétrie, on a aussi $N_{V_1}(x) \cap N_{V_1}(y) \geq 2$, pour tous $x, y \in V_2$, où $N_{V_1}(x)$ est l'ensemble des sommets de V_1 qui sont adjacents à x . Les conditions décrites ci-dessus sont des conditions nécessaires pour que le graphe G vérifie $\alpha_2(G) - F(G) = \lfloor \frac{n}{2} \rfloor - 1$; en revanche elles ne sont pas des conditions suffisantes. La figure 4.5 est un exemple dans lequel l'ordre de la forêt induite maximum (les sommets blancs) est égal à 8 et l'ordre du bistable maximum est égal à 10, donc $\alpha_2(G) - F(G) \neq \lfloor \frac{n}{2} \rfloor - 1$.

2 ème cas: Si n est impair, on distingue les deux sous-cas suivants:

Figure 4.5 . $\alpha_2(G) = 10, F(G) = 8$

a/ $\alpha_2(G) = n$:

Dans ce cas, G est un graphe biparti. D'après le corollaire, nous avons $F(G) = \frac{n+1}{2} + 1$. Donc le graphe G est équilibré et il vérifie les conditions suivantes:

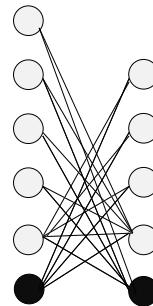
- $G = (V_1, V_2, E)$ tel que $|V_1| = |V_2| + 1$.
- Pour tous $x, y, z \in V_1$, on a $V_2 \cup \{x, y, z\}$ contient un cycle.
- Pour tous $x, y \in V_2$, on a $N_{V_1}(x) \cap N_{V_1}(y) \geq 2$.

Rappelons que $N_{V_1}(x)$ dénote l'ensemble des sommets de V_1 qui sont adjacents à x .

Tel qu'indiqué précédemment, ces conditions sont nécessaires pour que le graphe G vérifie $\alpha_2(G) - F(G) = \lfloor \frac{n}{2} \rfloor - 1$; cependant elles ne sont pas suffisantes. La figure 4.6 est un exemple où l'ordre de la forêt induite maximum (représentée par les sommets blancs) est égal à 9 et l'ordre du bistable maximum est égal à 11; donc $\alpha_2(G) - F(G) \neq \lfloor \frac{n}{2} \rfloor - 1$.

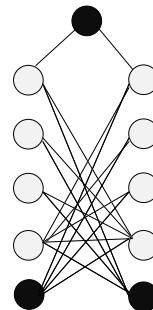
b/ $\alpha_2(G) = n - 1$:

Si $\alpha_2(G) = n - 1$, il existe un sommet $v \in V$ tel que $G - v$ soit un graphe biparti. D'autre part, d'après le corollaire, nous avons $F(G) = \frac{n-1}{2} + 1$, et donc le sous-graphe induit par $G - v$ est équilibré. Posons $H = G - v$. Alors nous obtenons $H = (V_1, V_2, E)$

Figure 4.6 . $\alpha_2(G) = 11, F(G) = 9$

tel que $|V_1| = |V_2| = \frac{n-1}{2}$.

Comme $n-1$ est pair, il s'ensuit que des conditions similaires à celles décrites au 1er cas s'appliquent à ce graphe. Voici un exemple où ces trois conditions sont vérifiées mais $\alpha_2(G) - F(G) = 2 \neq \left\lfloor \frac{n}{2} \right\rfloor - 1 = 4$.

Figure 4.7 . $\alpha_2(G) = 10, F(G) = 8$

Par la suite, nous allons donner les bornes supérieures sur les différences $\alpha_2(G) - LF(G)$ et $F(G) - LF(G)$.

4.3 Borne pour **Max** ($\alpha_2(G) - LF(G)$)

Proposition 4.3.1. *Soit G un graphe connexe d'ordre n . Alors on a*

$$\alpha_2(G) - LF(G) \leq \left\lfloor \frac{n}{2} \right\rfloor$$

Preuve. G est un graphe connexe d'ordre n . Soit B un bistable maximum de G de taille $\alpha_2(G)$, tel que $B = A_1 \cup A_2$. On suppose que $|A_1| \geq |A_2|$. Il est clair que A_1 est une forêt linéaire induite du graphe G . On a donc

$$LF(G) \geq \left\lceil \frac{\alpha_2(G)}{2} \right\rceil.$$

Alors

$$\alpha_2(G) - LF(G) \leq \alpha_2(G) - \left\lceil \frac{\alpha_2(G)}{2} \right\rceil.$$

Comme

$$\alpha_2(G) \leq n,$$

on obtient

$$\alpha_2(G) - LF(G) \leq \left\lfloor \frac{n}{2} \right\rfloor.$$

4.4 Borne pour **Max** ($F(G) - LF(G)$)

Proposition 4.4.1. *Soit G un graphe connexe d'ordre n . Alors on a*

$$\max F(G) - LF(G) \leq \left\lfloor \frac{n}{4} \right\rfloor.$$

Preuve. Soit H une forêt induite maximum de G . On peut supposer que H contient

au moins 4 sommets car sinon H est également une forêt linéaire et on a donc $F(G) = LF(G)$. Nous allons déterminer un sous-graphe induit H' de H qui est une forêt linéaire. Nous procérons comme suit. Soit f une feuille de H et soit f' une feuille de H à distance maximale de f . Notons p le père de f' . Observons que p est de degré au moins 2 car H contient au moins 4 sommets et tous les voisins de p sont des feuilles sauf éventuellement le sommet le plus proche de p sur la chaîne reliant p à f .

- Si p est de degré 2, nous insérons f' dans H' et ôtons f' de H .
- Si p est de degré 3, nous insérons p et les deux feuilles adjacentes à p (différentes de f) dans H' et ôtons p et ses 3 voisins de H .
- Si p est de degré au moins 4, nous insérons toutes les feuilles adjacentes à p (différentes de f) dans H' et ôtons ces feuilles ainsi que p de H .

Ce processus est répété jusqu'à ce que H soit vide. La forêt linéaire H' ainsi construite vérifie $|H'| \geq \frac{3}{4}|H|$. On a donc, $\frac{3}{4}F(G) = \frac{3}{4}|H| \leq |H'| \leq LF(G)$, ce qui signifie que

$$F(G) - LF(G) \leq F(G) - \left\lceil \frac{3}{4}F(G) \right\rceil \leq \left\lfloor \frac{n}{4} \right\rfloor.$$

□

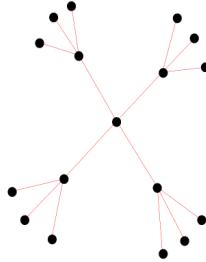


Figure 4.8 . Graphe d'ordre 17 obtenu par AGX en maximisant $F(G) - LF(G)$. On a $F(G) = 17$, $LF(G) = 13$ et $F(G) - LF(G) = \left\lfloor \frac{17}{4} \right\rfloor = 4$

4.5 Borne pour $\text{Max } (F(G) - T(G))$

Nous avons essayé de trouver la borne supérieure sur la différence $F(G) - T(G)$. Le système AutoGraphiX a généré les graphes extrêmes pour $5 \leq n \leq 8$. (Voir figure 4.9)

Nous posons la conjecture suivante.

Conjecture 4.5.1. *Soit G un graphe connexe d'ordre n . Alors on a*

$$F(G) - T(G) \leq n - 2\lfloor \sqrt{n} \rfloor.$$

Malheureusement nous n'avons pas réussi à prouver ce résultat.

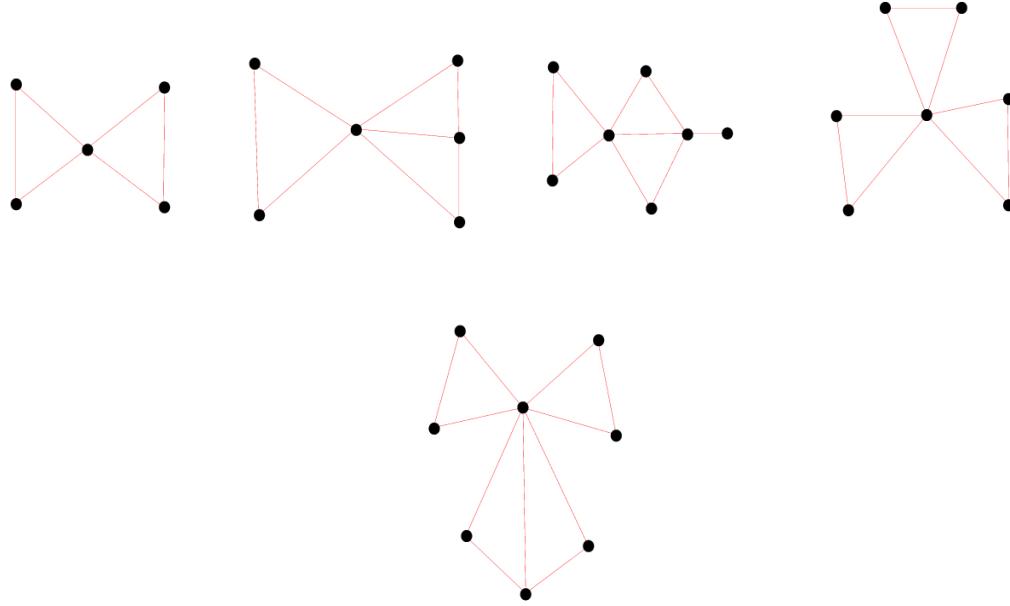


Figure 4.9 . Graphes obtenus par AGX en maximisant $F(G) - T(G)$

4.6 Graphes vérifiant $\alpha_2(G) = F(G) = LF(G)$

Nous avons essayé de trouver des classes de graphes vérifiant l'égalité $\alpha_2(G) = F(G) = LF(G)$. En utilisant le système AGX nous avons obtenus les familles de graphes suivantes: les graphes complets, les graphes formés par un chemin et une clique tel que chaque sommet de la clique est relié à tous les sommets du chemin, les cycles impairs, et les graphes formés par un stable et un chemin tel que chaque sommet du stable est relié à tous les sommets du chemin.

Proposition 4.6.1. *Soit G un graphe connexe. Si $\alpha_2(G) = 3$, alors*

$$\alpha_2(G) = F(G) = LF(G).$$

Preuve. Si $\alpha_2(G) = 3$, alors le sous-graphe induit par les sommets du bistable maximum de G n'est pas un cycle et nous concluons que

$$\alpha_2(G) = F(G) = 3.$$

Comme $F(G) = 3$, alors le sous-graphe induit par la forêt maximum ne contient pas un sommet de degré supérieur à 2; donc cette forêt est une forêt linéaire.

D'où le résultat $\alpha_2(G) = F(G) = LF(G)$.

□

DEUXIÈME PARTIE

Dans cette partie nous allons borner supérieurement les différences $\left(\frac{\alpha_2(G)}{2} - \mu(G)\right)$, $\left(\frac{F(G)}{2} - \mu(G)\right)$ et $\left(\frac{LF(G)}{2} - \mu(G)\right)$ et nous allons présenter des familles de graphes extrêmes pour lesquelles la borne est atteinte.

4.7 Borne pour $\text{Max } \left(\frac{\alpha_2(G)}{2} - \mu(G)\right)$

D'après le système AutoGraphiX les graphes qui maximisent la différence $\left(\frac{\alpha_2(G)}{2} - \mu(G)\right)$ sont les graphes bipartis complets équilibrés. Dans la suite nous allons prouver ce résultat.

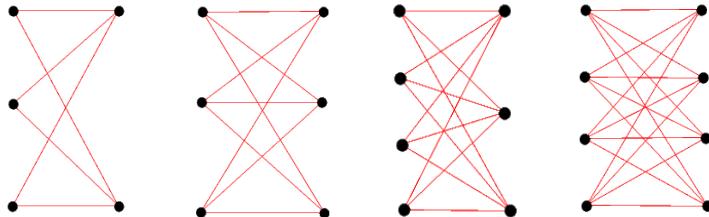


Figure 4.10 . Graphes obtenus par AGX en maximisant $\alpha_2(G) - \mu(G)$

Lemme 4.7.1. *Soit G un graphe connexe d'ordre n , tel que $\alpha_2(G) = k$. Alors on a*

$$\mu(G) \geq \mu(H)$$

où H est le graphe d'ordre n formé par un biparti complet équilibré d'ordre k et une clique d'ordre $n - k$, tel que chaque sommet de la clique est relié à tous les sommets du biparti.

Preuve. Soit A_1 et A_2 les deux stables d'un biparti maximum de G tel que $\alpha_2(G) = k = |A_1 \cup A_2|$. Nous supposons que $|A_1| \geq |A_2|$.

Soit G' le graphe obtenu à partir de G en ajoutant les arêtes entre toutes les paires $\{x, y\}$ telles que $\{x, y\} \not\subseteq A_i$, ($i = 1, 2$). Il est clair que $\mu(G) \geq \mu(G')$.

Si $|A_2| + 1 \geq |A_1| \geq |A_2|$, alors G' est isomorphe à H et on a bien

$$\mu(G) \geq \mu(H).$$

Nous supposons maintenant que $|A_1| > |A_2| + 1$. Soit s un sommet de A_1 . On a

$$\sum_{u,v \in G'} d(u, v) = \sum_{u,v \in G'; u,v \neq s} d(u, v) + \sigma_{G'}(s)$$

On note que $\sigma_{G'}(s)$ est égale à $|A_2| + 2(|A_1| - 1) + (n - k)$. Soit G'' le graphe obtenu en déplaçant le sommet s de A_1 à A_2 . La transmission du sommet s devient :

$$\sigma_{G''}(s) = (|A_1| - 1) + 2|A_2| + (n - k).$$

Nous remarquons que la différence $\sigma_{G'}(s) - \sigma_{G''}(s)$ est plus grande que 0 car $|A_1| - |A_2| - 1 > 0$. Nous répétons successivement cette opération jusqu'à l'obtention d'un biparti complet équilibré (c'est-à-dire tel que $|A_2| + 1 \geq |A_1|$). Étant donné que le graphe obtenu est isomorphe à H et que la transmission du sommet déplacé va toujours diminuer, nous avons

$$\mu(G) \geq \mu(G') \geq \mu(H)$$

□

Théorème 4.7.2. Soit G un graphe connexe d'ordre n . Alors on a

$$\frac{\alpha_2(G)}{2} - \mu(G) \leq \begin{cases} \frac{n}{2} - \frac{3n-4}{2(n-1)} & \text{si } n \text{ est pair} \\ \frac{n}{2} - \frac{3n-4}{2(n-1)} - \frac{1}{2n(n-1)} & \text{si } n \text{ est impair.} \end{cases}$$

La borne est atteinte par un biparti complet équilibré.

Preuve. Soit B le graphe biparti complet équilibré d'ordre n . Il est facile de vérifier les relations suivantes:

$$\mu(B) = \begin{cases} \frac{3n-4}{2(n-1)} & \text{si } n \text{ est pair} \\ \frac{3n-4}{2(n-1)} + \frac{1}{2n(n-1)} & \text{si } n \text{ est impair.} \end{cases}$$

Soit G un graphe connexe d'ordre n . Si $\alpha_2(G) = n$, on sait par le lemme 4.7.1 que

$$\mu(G) \geq \mu(B),$$

et on a donc

$$\frac{\alpha_2(G)}{2} - \mu(G) \leq \frac{n}{2} - \mu(B),$$

ce qui prouve le théorème dans ce cas.

Si $\alpha_2(G) = k < n$, considérons le graphe H qui est un biparti complet équilibré d'ordre k relié à une clique d'ordre $n - k$ tel que chaque sommet de la clique est relié à tous les sommets du biparti complet. Nous avons

$$\frac{\alpha_2(G)}{2} - \mu(G) \leq \frac{\alpha_2(H)}{2} - \mu(H).$$

Il nous suffit donc de montrer que

$$\frac{\alpha_2(H)}{2} - \mu(H) \leq \frac{n}{2} - \mu(B).$$

Comme par hypothèse nous avons

$$\frac{\alpha_2(H)}{2} \leq \frac{n-1}{2},$$

il suffit de prouver que

$$\mu(H) + \frac{1}{2} \geq \mu(B).$$

Si k est impair on a

$$\begin{aligned} \mu(H) &= \frac{\frac{3k^2-4k+1}{2} + (n-k)(n-k-1) + 2k(n-k)}{n(n-1)} \\ &= \frac{n^2 + \frac{k^2}{2} - k + \frac{1}{2} - n}{n(n-1)}. \end{aligned}$$

Si k est pair on a

$$\begin{aligned} \mu(H) &= \frac{\frac{3k^2-4k}{2} + (n-k)(n-k-1) + 2k(n-k)}{n(n-1)} \\ &= \frac{n^2 + \frac{k^2}{2} - n - k}{n(n-1)}. \end{aligned}$$

On a donc

$$\mu(B) \leq \frac{3n-4}{2(n-1)} + \frac{1}{2n(n-1)} \text{ pour tout } n$$

et

$$\mu(H) = \frac{n^2 + \frac{k^2}{2} - n - k}{n(n-1)} \text{ pour tout } n.$$

On déduit que

$$\begin{aligned}
\frac{1}{2} - (\mu(B) - \mu(H)) &\geq \frac{1}{2} - \left(\frac{3n^2 - 4n + 1}{2n(n-1)} - \frac{n^2 + \frac{k^2}{2} - n - k}{n(n-1)} \right) \\
&= \frac{\frac{n(n-1)}{2}}{n(n-1)} - \left(\frac{\frac{3n^2}{2} - 2n - n^2 - \frac{k^2}{2} + n + k + \frac{1}{2}}{n(n-1)} \right) \\
&= \frac{\frac{n^2}{2} - \frac{n}{2} - \frac{n^2}{2} + n + \frac{k^2}{2} - k - \frac{1}{2}}{n(n-1)} \\
&= \frac{\frac{n}{2} + \frac{k^2}{2} - k - \frac{1}{2}}{n(n-1)} \geq 0,
\end{aligned}$$

ce qui prouve le résultat. □

4.8 Borne pour $\text{Max} \left(\frac{F(G)}{2} - \mu(G) \right)$

D'après le système AGX, les graphes qui maximisent la différence $\frac{F(G)}{2} - \mu(G)$ sont les étoiles. Nous allons prouver ce résultat.

Définition 4.8.1. *Le graphe $S_{n,k}$ est le graphe composé d'une étoile de k sommets et d'une clique de $(n-k)$ sommets, tel que chaque sommet de la clique est relié à chaque sommet de l'étoile. Remarque: Il est facile de vérifier que $\mu(S_{n,n}) = 2 - \frac{2}{n}$.*

Définition 4.8.2. *Soit $\mathcal{T}_{n,k}$ l'ensemble des graphes composés d'un arbre de k sommets et d'une clique de $(n-k)$ sommets, tels que chaque sommet de la clique est relié à chaque sommet de l'arbre.*

Lemme 4.8.3.

$$\sigma(T) = n(n-1) + (k-1)(k-2), \forall T \in \mathcal{T}_{n,k}, k < n.$$

Preuve. $T \in \mathcal{T}_{n,k}$ est composé d'un arbre A de k sommets et d'une clique Q de $(n-k)$ sommets, tels que chaque sommet de la clique est relié à chaque sommet de l'arbre. Donc deux sommets de A sont à distance une ou deux unités. Nous avons alors

$$\begin{aligned} \sigma(T_{n,k}) &= \sigma(A) + \sigma(Q) + \sum_{u \in A, v \in Q} d(u, v) \\ &= 2(k-1) + 2(k-1)(k-2) + (n-k)(n-k-1) + 2k(n-k) \\ &= (k-1)(k-2) + 2k-2 + (k-1)(k-2) + (n-k)(n+k-1) \\ &= (k-1)(k-2) + (2k-2+k^2-2k-k+2) + (n^2+nk-n-nk-k^2+k) \\ &= (k-1)(k-2) + n^2 - n \\ &= n(n-1) + (k-1)(k-2). \end{aligned}$$

□

Notons $t_{n,k} = n(n-1) + (k-1)(k-2)$.

Lemme 4.8.4. *Soit G un graphe connexe de n sommets et m arêtes, tel que $F(G) = n$.*

Si $G \neq S_{n,n}$, alors on a

$$\sigma(G) > \sigma(S_{n,n}).$$

Preuve. Comme G est un arbre d'ordre n tel que $F(G) = n$, G appartient à $\mathcal{T}_{n,n}$. Soit m' le nombre d'arêtes du graphe $S_{n,n}$; il est clair que $m = m'$.

Soit $S_i(G) = \{(x, y) | x, y \in V, d_G(x, y) \geq i\}$. Nous avons $S_i(S_{n,n}) = \emptyset$ pour $i \geq 3$, donc

$$\frac{\sigma(G)}{2} - \frac{\sigma(S_{n,n})}{2} = \sum_{i=3}^{n-1} |S_i(G)| \geq 0.$$

Si $G \neq S_{n,n}$, alors $|S_3(G)| > 0$ et $\sigma(G) > \sigma(S_{n,n})$.

Lemme 4.8.5. *Soit G un graphe connexe de n sommets et m arêtes, tel que $F(G) = k$ et $k < n$. Alors*

$$\sigma(G) \geq t_{n,k}.$$

Preuve. Soit F une forêt maximum de G , et soit $T = (V', E')$ le graphe obtenu à partir de G en ajoutant des arêtes de manière à ce que F devienne un arbre A , chaque sommet de $T - A$ est relié à chaque sommet de A , et les sommets de $T - A$ forment une clique. Il est clair que $m' \geq m$, où m' est le nombre d'arêtes de T et que T appartient à $\mathcal{T}_{n,k}$.

Soit $S_i(G) = \{(x, y) | x, y \in V, d(x, y) \geq i\}$. La transmission de tout graphe connexe H peut s'obtenir par la formule suivante:

$$\frac{\sigma(H)}{2} = \sum_{i=1}^{n-1} |S_i(H)|$$

$$\begin{aligned}
&= |S_1(H)| + |S_2(H)| + \sum_{i=3}^{n-1} |S_i(H)| \\
&= \frac{n(n-1)}{2} + \left(\frac{n(n-1)}{2} - m \right) + \sum_{i=3}^{n-1} |S_i(H)|.
\end{aligned}$$

Nous avons $\sum_{i=3}^{n-1} |S_i(T)| = 0$. Donc

$$\frac{\sigma(G)}{2} - \frac{\sigma(T)}{2} = (m' - m) + \sum_{i=3}^{n-1} |S_i(G)| \geq 0.$$

Il en résulte que

$$\sigma(G) \geq \sigma(T) = t_{n,k}.$$

□

Théorème 4.8.6. Soit G un graphe connexe d'ordre $n \geq 3$. Alors on a

$$\frac{F(G)}{2} - \mu(G) \leq \frac{n}{2} + \frac{2}{n} - 2.$$

La borne est atteinte par une étoile d'ordre n .

Preuve.

Soit G un graphe connexe d'ordre n . Si $F(G) = n$, alors d'après le lemme 4.8.4, on a

$$\frac{F(G)}{2} - \mu(G) \leq \frac{n}{2} - \mu(S_{n,n}) = \frac{n}{2} + \frac{2}{n} - 2.$$

Sinon, si $F(G) = k$ et $k < n$, alors d'après le lemme 4.8.5, on a

$$\sigma(G) \geq \sigma(T), \forall T \in \mathcal{T}_{n,k}, k < n.$$

Or

$$\sigma(S_{n,k}) = \sigma(T), \forall T \in \mathcal{T}_{n,k}, k < n.$$

Alors

$$\frac{F(G)}{2} - \mu(G) \leq \frac{k}{2} - \mu(S_{n,k}).$$

Soit le graphe $S_{n,k+1}$ obtenu à partir de $S_{n,k}$ en déplaçant un sommet s de la clique à l'étoile, c'est-à-dire qu'on enlève les arêtes de la forme $\{s, x_i\}$, où x_i est un sommet de l'étoile différent du sommet central.

Nous prouvons que

$$\frac{k}{2} - \mu(S_{n,k}) \leq \frac{k+1}{2} - \mu(S_{n,k+1}).$$

Ainsi, en répétant ce processus on aura

$$\frac{k}{2} - \mu(S_{n,k}) \leq \frac{n}{2} - \mu(S_{n,n}).$$

Il faut donc prouver que

$$\frac{1}{2} - \mu(S_{n,k+1}) + \mu(S_{n,k}) \geq 0$$

Nous avons

$$\begin{aligned} \frac{1}{2} - (\mu(S_{n,k}) - \mu(S_{n,k+1})) &= \frac{1}{2} - \left(\left(1 + \frac{k(k-1)}{n(n-1)}\right) - \left(1 + \frac{(k-1)(k-2)}{n(n-1)}\right) \right) \\ &= \frac{1}{2} - \frac{2(k-1)}{n(n-1)} = \frac{n(n-1) - 4(k-1)}{2n(n-1)} = \frac{n^2 - n - 4k + 4}{2n(n-1)}. \end{aligned}$$

Or $k < n$, donc $n^2 - n - 4k + 4 \geq n^2 - n - 4n + 4 = n^2 - 5n + 4 \geq 0$ pour $n > 3$.

Si $n = 3$, alors $k = 2$ donc $n^2 - n - 4k + 4 = 9 - 3 - 8 + 4 = 2 > 0$. D'où le

résultat. □

4.9 Borne pour $\text{Max} \left(\frac{LF(G)}{2} - \mu(G) \right)$

Définition 4.9.1. *Le graphe $P_{n,k}$ est le graphe composé d'un chemin de k sommets et d'une clique de $(n - k)$ sommets, tel que chaque sommet de la clique est relié à chaque sommet du chemin.*

Remarque: *Il est facile de vérifier que $\mu(P_{n,n-1}) = \frac{2n^2-6n+6}{n(n-1)}$.*

D'après le système AGX, les graphes qui maximisent la différence $\frac{LF(G)}{2} - \mu(G)$ sont les graphes $P_{n,n-1}$. Nous allons prouver ce résultat.

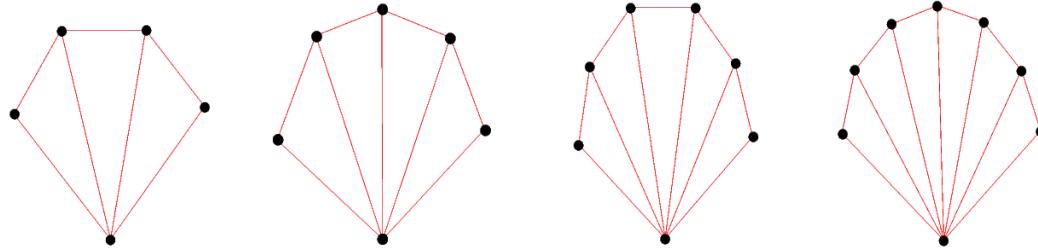


Figure 4.11 . Graphes obtenus par AGX en maximisant $LF(G) - \mu(G)$

Lemme 4.9.2. *Nous avons*

$$\sigma(P_{n,k}) = 2(k-1)^2 + (n-k)(n+k-1).$$

Preuve. Soit G_1 le chemin de k sommets et G_2 la clique de $(n - k)$ sommets. On a

$$\sigma(P_{n,k}) = \sum_{u \in G_1} \sum_{v \in G_1} d(u, v) + \sum_{u \in G_2} \sum_{v \in G_2} d(u, v) + 2 \sum_{u \in G_1} \sum_{v \in G_2} d(u, v)$$

$$= 2(k-1)^2 + (n-k)(n-k-1) + 2k(n-k)$$

$$= 2(k-1)^2 + (n-k)(n+k-1).$$

□

Lemme 4.9.3. Soit G un graphe connexe de n sommets et m arêtes tel que $LF(G) = k$ et $k < n$. Alors

$$\sigma(G) \geq \sigma(P_{n,k}),$$

et on a égalité uniquement si $G = P_{n,k}$.

Preuve. Soit LF une forêt linéaire maximum de G , et soit $P_{n,k} = (V', E')$ le graphe obtenu à partir de G en ajoutant des arêtes de manière à ce que LF devienne un chemin A et chaque sommet de $(P_{n,k} - A)$ soit relié à chaque sommet de A , et les sommets de $(P_{n,k} - A)$ forment une clique. Il est clair que $m' \geq m$, où m' est le nombre d'arêtes de $P_{n,k}$.

Soit $S_i(G) = \{(x, y) | x, y \in V, d_G(x, y) \geq i\}$. On a $S_i(P_{n,k}) = \emptyset$ pour $i \geq 3$. Donc

$$\frac{\sigma(G)}{2} - \frac{\sigma(P_{n,k})}{2} = (m' - m) + \sum_{i=3}^{n-1} |S_i(G)| \geq 0$$

Il en résulte que

$$\sigma(G) \geq \sigma(P_{n,k}).$$

On a l'égalité uniquement si $m = m'$. □

Dans la suite nous allons déterminer la borne sur la différence $\frac{LF(G)}{2} - \mu(G)$. Soit G un graphe connexe d'ordre n . Si $n = 3$, il est facile de vérifier que la différence $\frac{LF(G)}{2} - \mu(G)$ est maximisé par un chemin sur 3 sommets puisqu'on a alors $\frac{LF(G)}{2} - \mu(G) = \frac{3}{2} - \frac{4}{3} = \frac{1}{6}$ alors que cette différence vaut zéro pour la clique sur 3 sommets.

Théorème 4.9.4. Soit G un graphe connexe d'ordre $n > 3$. Alors on a

$$\frac{LF(G)}{2} - \mu(G) \leq \frac{n-1}{2} - \frac{2n^2 - 6n + 6}{n(n-1)}.$$

La borne est atteinte par $P_{n,n-1}$.

Preuve. Soit G un graphe connexe d'ordre n . Si $LF(G) = n$, alors G est un chemin et donc

$$\mu(G) > \mu(P_{n,n-1}).$$

Montrons que

$$\frac{LF(G)}{2} - \mu(G) \leq \frac{LF(P_{n,n-1})}{2} - \mu(P_{n,n-1}).$$

On a

$$\frac{LF(G)}{2} - \mu(G) = \frac{n}{2} - \frac{n+1}{3}$$

et

$$\frac{LF(P_{n,n-1})}{2} - \mu(P_{n,n-1}) = \frac{n-1}{2} - \frac{2n^2 - 6n + 6}{n(n-1)}.$$

Alors on a

$$\begin{aligned} \left(\frac{LF(P_{n,n-1})}{2} - \mu(P_{n,n-1}) \right) - \left(\frac{LF(G)}{2} - \mu(G) \right) &= \left(\frac{n-1}{2} - \frac{2n^2 - 6n + 6}{n(n-1)} \right) - \left(\frac{n}{2} - \frac{n+1}{3} \right) \\ &= \frac{n+1}{3} - \frac{1}{2} - \frac{2n^2 - 6n + 6}{n(n-1)} \\ &= \frac{n(2n^2 - 15n + 37) - 36}{6n(n-1)} > 0 \text{ pour } n > 3. \end{aligned}$$

Sinon, si $LF(G) = k$ et $k < n$, alors d'après le lemme 4.9.3, on a

$$\sigma(G) \geq \sigma(P_{n,k}).$$

Si $k = n - 1$, on a $LF(G) - \mu(G) \leq \frac{n-1}{2} - \mu(P_{n,n-1})$ et l'égalité n'a lieu que si $G = P_{n,n-1}$. Sinon, considérons le graphe $P_{n,k+1}$ obtenu à partir de $P_{n,k}$, avec $k \leq n-2$, en déplaçant un sommet s de la clique au chemin, c'est-à-dire qu'on enlève les arêtes $\{s, x_i\}$ pour un sommet x_i du chemin.

Nous prouvons que

$$\frac{k}{2} - \mu(P_{n,k}) < \frac{k+1}{2} - \mu(P_{n,k+1}),$$

ce qui impliquera que

$$\frac{k}{2} - \mu(P_{n,k}) < \frac{n-1}{2} - \mu(P_{n,n-1}).$$

Il nous faut donc prouver que

$$\frac{1}{2} - \mu(P_{n,k+1}) + \mu(P_{n,k}) \geq 0.$$

Nous avons

$$\begin{aligned} \frac{1}{2} - (\mu(P_{n,k}) - \mu(P_{n,k+1})) &= \frac{1}{2} - \left(\frac{2(k-1)^2 + (n-k)(n+k-1)}{n(n-1)} - \frac{2k^2 + (n-k-1)(n+k)}{n(n-1)} \right) \\ &= \frac{1}{2} - \left(\frac{(2k^2 - 4k + 2) + (n-k)(n+k) - (n-k) - 2k^2 - (n-k)(n+k) + (n+k)}{n(n-1)} \right) \\ &= \frac{1}{2} - \left(\frac{2 - 2k}{n(n-1)} \right) = \frac{n(n-1) - (4 - 4k)}{2n(n-1)} \end{aligned}$$

$$= \frac{n^2 + 4k - n - 4}{2n(n-1)} > 0 \text{ pour } n > 2.$$

Ceci prouve le résultat. □

Conclusion

Nous résumons à la fin de ce chapitre les étapes de notre travail. Tout d'abord, écrire le problème d'optimisation, dans notre cas un problème de maximisation sur les différences des invariants. AutoGraphiX va essayer par la suite de trouver les graphes extrêmes relatifs au problème. La recherche des graphes extrêmes est basée sur la *méta-heuristique de recherche à voisinages variables* [47]. C'est la partie la plus délicate du travail, car le système peut ne pas réussir à trouver les bons graphes. Dans ce cas nous pouvons aider AGX interactivement. L'étape suivante consiste à trouver une conjecture générée par le système ou déduite par l'utilisateur; en dernier lieu on essaie de prouver cette conjecture. Voici un tableau récapitulatif des résultats prouvés dans ce chapitre (toutes les bornes sont atteintes).

Borne inférieure	$Inv_1 - Inv_2$	Borne supérieure
0	$\alpha_2 - F$	$\lfloor \frac{n}{2} \rfloor - 1$
0	$\alpha_2 - LF$	$\lfloor \frac{n}{2} \rfloor$
0	$F - LF$	$\lfloor \frac{n}{4} \rfloor$
0	$\frac{\alpha_2}{2} - \mu$	$\frac{n}{2} - \frac{3n-4}{2(n-1)}$
0	$\frac{F}{2} - \mu$	$\frac{n}{2} + \frac{2}{n} - 2$
?	$\frac{LF}{2} - \mu$	$\frac{n-1}{2} + \frac{2n^2-6n+6}{n(n-1)}$
0	$F - T$?

Tableau 4.1 . Récapitulation des résultats

CHAPITRE 5

HEURISTIQUES CONSTRUCTIVES POUR DÉTERMINER LA PLUS GRANDE FORêt, LA PLUS GRANDE FORêt LINÉAIRE ET LE PLUS GRAND ARBRE

Dans ce chapitre nous allons présenter les heuristiques que nous avons développées afin de déterminer la plus grande forêt induite, la plus grande forêt linéaire induite et le plus grand arbre induit dans un graphe connexe. Ce sont des algorithmes constructifs qui n'entrent bien sûr pas en compétition avec des mét heuristicques (par exemple une recherche tabou ou une méthode de descente) qui sont susceptibles de produire des solutions de bien meilleure qualité. Le but de ce chapitre est de démontrer que de simples algorithmes constructifs permettent d'obtenir des solutions de qualité raisonnable.

Nos algorithmes constructifs peuvent être décrits comme suit. À chaque étape de la construction, un élément est ajouté à la solution partielle courante; le choix de l'élément ajouté se fait selon une fonction qui mesure la qualité de l'insertion d'un sommet additionnel dans la solution partielle.

5.1 Méthodes constructives pour la détermination de la plus grande forêt

Pour déterminer une forêt dans un graphe G , nous partons du sous-graphe induit dans G par tous les sommets de degré 1 ou 0. Ces sommets font en effet partie de toute forêt de grandeur maximale. Nous ajoutons ensuite un sommet à chaque itération, en prenant garde de ne créer aucun cycle. Le processus s'arrête dès qu'aucun sommet supplémentaire ne peut être ajouté. Soit C_1, \dots, C_r les composantes connexes de la forêt courante

: si v est adjacent à au moins deux sommets de C_i , v n'est pas un sommet insérable. Nous sommes donc en mesure de construire une liste contenant tous les sommets qui peuvent être ajoutés à F sans créer de cycle.

Soit V_F l'ensemble des sommets de la forêt F , soit V_I l'ensemble des sommets qui peuvent être ajoutés à F sans créer de cycle, et soit G' le sous-graphe de G induit par V_I . Nous choisissons le sommet à ajouter à F selon l'un des six critères suivants:

Critère 1 Nous choisissons au hasard un sommet $v \in V_I$.

Critère 2 Nous choisissons un sommet $v \in V_I$ de degré minimum dans G .

Critère 3 Nous choisissons un sommet $v \in V_I$ de degré minimum dans G' .

Critère 4 Pour chaque sommet $v \in V_I$, nous calculons le nombre r_v de sommets qui pourront toujours être ajoutés à la forêt une fois que v est ajouté, et nous choisissons un sommet qui maximise r_v .

Critère 5 Nous choisissons un sommet ayant le moins de voisins dans V_F .

Critère 6 Pour chaque sommet $v \in V_I$, nous calculons le nombre r_v de sommets qui pourront toujours être ajoutés à la forêt une fois que v est ajouté. Si plusieurs sommets ont la même valeur maximale r_v , nous en choisissons un de degré minimum dans G' .

Pour chacun des 6 critères mentionnés ci-dessus, il se peut que plusieurs sommets puissent être choisis, auquel cas nous en prenons un au hasard.

5.2 Résultats expérimentaux

Graphes aléatoires

Pour évaluer les performances de nos heuristiques constructives, nous les avons testées sur un ensemble de graphes aléatoires ayant 30, 40, 50, 60 et 70 sommets et une densité variant entre 0.1 et 0.9, par pas de 0.2, ce qui donne un total de 25 graphes.

Nous comparons les résultats de nos algorithmes avec les valeurs optimales obtenues par un algorithme exact développé par David Schindl. Voici un pseudo-code de cet algorithme exact dans lequel $\text{Forêt}(X, Y, Z)$ est une procédure qui détermine la plus grande forêt dans G qui contient tous les sommets de Y et éventuellement quelques sommets de X . Chaque sommet $x \in X$ est tel que $Y \cup \{x\}$ induit une forêt dans G et Z représente la meilleure borne inférieure connue sur la taille maximale $F(G)$ d'une forêt dans G .

Algorithme énumératif pour la plus grande forêt

1. Déterminer une forêt à l'aide d'un algorithme constructif. Poser B égal au nombre de sommets de cette forêt. B est donc une borne inférieure sur la taille $F(G)$ de la plus grande forêt dans G .
2. Poser $F(G) = \text{Forêt}(V(G), \emptyset, B)$.

Procédure $\text{Forêt}(\text{Sous-ensemble de sommets } V', \text{Forêt } F, \text{Borne } B)$:

1. Si $|F| + |V'| \leq B$, **retourner** B .
2. Choisir un sommet $v \in V'$.

Poser $F_{avec} := F \cup \{v\}$. Si $|F| = B$ alors poser $B := B + 1$.

Poser $V_{avec} := V' \cup \{v\}$ et $V_{sans} := V' - \{v\}$.

Ôter de V_{avec} tous les sommets v' tels que $V_{avec} \cup \{v'\}$ n'induit pas une forêt dans G .

3. **retourner** $\max(\text{Forêt}(V_{avec}, F_{avec}, B), \text{Forêt}(V_{sans}, F, B))$

Nous avons effectué cinq essais de chaque algorithme pour chacune des 25 instances. Dans les colonnes des tableaux, nous donnons les valeurs de la pire solution trouvée, de la meilleure ainsi que la valeur moyenne. Nous donnons également le temps moyen (en secondes) nécessaire pour obtenir ces solutions. L'avant-dernière colonne des tableaux correspond à la valeur optimale fournie par l'algorithme exact décrit ci-dessus. Nous présentons dans la dernière colonne les écarts moyens (en %) par rapport à une solution optimale.

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Algorithme exacte	Écart moyen
RAND30.1	22	25	23.4	0.001	26	10
RAND30.3	9	12	10.8	0.003	15	28
RAND30.5	8	8	8	0.004	10	20
RAND30.7	4	6	5.6	0.004	7	20
RAND30.9	3	4	3.8	0.007	5	24
RAND40.1	26	27	26.6	0.003	31	14.1
RAND40.3	12	15	14.4	0.002	17	15.2
RAND40.5	10	10	10	0.004	11	9.09
RAND40.7	5	6	5.8	0.01	8	27.5
RAND40.9	4	5	4.2	0.007	5	16
RAND50.1	29	32	30.8	0.005	34	9.4
RAND50.3	13	14	13.2	0.01	18	26.6
RAND50.5	7	9	8.4	0.01	13	35.3
RAND50.7	5	7	6.2	0.01	8	22.5
RAND50.9	4	4	4	0.02	7	42.8
RAND60.1	33	35	33.8	0.01	40	15.5
RAND60.3	14	15	14.8	0.01	20	26
RAND60.5	9	10	9.8	0.02	13	24.6
RAND60.7	6	7	6.6	0.02	9	26.6
RAND60.9	5	5	5	0.03	6	16.6
RAND70.1	30	38	34.8	0.02	44	20.9
RAND70.3	14	17	16.4	0.03	22	25.45
RAND70.5	10	10	10	0.03	14	28.57
RAND70.7	6	7	6.4	0.04	9	28.88
RAND70.9	5	5	5	0.04	6	16.66

Tableau 5.1 . Résultats obtenus avec le critère 1

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Algorithme exacte	Écart moyen
<i>RAND</i> 30.1	25	25	25	0.004	26	3.84
<i>RAND</i> 30.3	14	14	14	0.004	15	6.66
<i>RAND</i> 30.5	8	10	9.2	0.004	10	8
<i>RAND</i> 30.7	6	6	6	0.006	7	14.28
<i>RAND</i> 30.9	5	5	5	0.09	5	0
<i>RAND</i> 40.1	30	31	30.6	0.004	31	1.29
<i>RAND</i> 40.3	15	17	15.8	0.005	17	7.05
<i>RAND</i> 40.5	11	11	11	0.01	11	0
<i>RAND</i> 40.7	6	6	6	0.01	8	25
<i>RAND</i> 40.9	4	5	4.2	0.006	5	16
<i>RAND</i> 50.1	32	33	32.4	0.009	34	3.52
<i>RAND</i> 50.3	15	16	15.6	0.01	18	13.33
<i>RAND</i> 50.5	10	11	10.2	0.01	13	16.92
<i>RAND</i> 50.7	5	7	6.2	0.02	8	22.5
<i>RAND</i> 50.9	5	7	6.2	0.02	7	11.42
<i>RAND</i> 60.1	39	39	39	0.02	40	2.5
<i>RAND</i> 60.3	19	19	19	0.04	20	5
<i>RAND</i> 60.5	11	12	11.2	0.03	13	13.84
<i>RAND</i> 60.7	6	7	6.8	0.05	9	24.44
<i>RAND</i> 60.9	5	5	5	0.04	6	16.66
<i>RAND</i> 70.1	40	42	40.8	0.02	44	7.27
<i>RAND</i> 70.3	18	18	18	0.05	22	18.18
<i>RAND</i> 70.5	13	13	13	0.05	14	7.14
<i>RAND</i> 70.7	5	5	5	0.06	9	44.44
<i>RAND</i> 70.9	4	5	4.4	0.06	6	26.66

Tableau 5.2 . Résultats obtenus avec le critère 2

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Algorithme exacte	Écart moyen
<i>RAND</i> 30.1	24	26	25.2	0.002	26	3.07
<i>RAND</i> 30.3	14	15	14.4	0.002	15	4
<i>RAND</i> 30.5	8	10	9.4	0.001	10	6
<i>RAND</i> 30.7	6	7	6.4	0.001	7	8.57
<i>RAND</i> 30.9	5	5	5	0.003	5	0
<i>RAND</i> 40.1	30	31	30.4	0.001	31	1.93
<i>RAND</i> 40.3	16	17	16.6	0.002	17	2.35
<i>RAND</i> 40.5	10	11	10.8	0.003	11	1.81
<i>RAND</i> 40.7	6	6	6	0.002	8	25
<i>RAND</i> 40.9	4	5	4.2	0.004	5	16
<i>RAND</i> 50.1	31	33	32.2	0.004	34	5.29
<i>RAND</i> 50.3	15	17	16	0.002	18	11.11
<i>RAND</i> 50.5	10	11	10.6	0.003	13	18.46
<i>RAND</i> 50.7	7	7	7	0.001	8	12.5
<i>RAND</i> 50.9	5	7	6	0.001	7	14.28
<i>RAND</i> 60.1	37	39	38.2	0.003	40	4.5
<i>RAND</i> 60.3	18	20	18.6	0.003	20	7
<i>RAND</i> 60.5	10	12	11.2	0.005	13	13.84
<i>RAND</i> 60.7	8	9	8.6	0.005	9	4.44
<i>RAND</i> 60.9	5	5	5	0.005	6	16.66
<i>RAND</i> 70.1	41	42	41.4	0.006	44	5.9
<i>RAND</i> 70.3	19	21	19.4	0.004	22	11.81
<i>RAND</i> 70.5	12	13	12.6	0.005	14	10
<i>RAND</i> 70.7	7	8	7.8	0.006	9	13.33
<i>RAND</i> 70.9	5	5	5	0.004	6	16.66

Tableau 5.3 . Résultats obtenus avec le critère 3

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Algorithme exacte	Écart moyen
<i>RAND</i> 30.1	25	26	25.6	0.004	26	1.53
<i>RAND</i> 30.3	13	15	14.2	0.003	15	5.33
<i>RAND</i> 30.5	8	10	9.4	0.004	10	6
<i>RAND</i> 30.7	6	7	6.6	0.004	7	5.71
<i>RAND</i> 30.9	5	5	5	0.004	5	0
<i>RAND</i> 40.1	30	31	30.4	0.006	31	1.93
<i>RAND</i> 40.3	15	17	16.4	0.006	17	3.52
<i>RAND</i> 40.5	10	11	10.8	0.005	11	1.81
<i>RAND</i> 40.7	6	6	6	0.004	8	25
<i>RAND</i> 40.9	4	5	4.6	0.003	5	8
<i>RAND</i> 50.1	32	33	32.4	0.003	34	4.7
<i>RAND</i> 50.3	15	17	16.2	0.002	18	10
<i>RAND</i> 50.5	10	11	10.6	0.004	13	18.46
<i>RAND</i> 50.7	7	7	7	0.002	8	12.5
<i>RAND</i> 50.9	5	7	6.4	0.003	7	8.57
<i>RAND</i> 60.1	37	40	38.8	0.006	40	3
<i>RAND</i> 60.3	18	20	19.2	0.004	20	4
<i>RAND</i> 60.5	10	12	11	0.003	13	15.38
<i>RAND</i> 60.7	8	9	8.8	0.002	9	2.22
<i>RAND</i> 60.9	5	5	5	0.003	6	16.66
<i>RAND</i> 70.1	41	43	42	0.003	44	4.54
<i>RAND</i> 70.3	19	20	19.8	0.005	22	10
<i>RAND</i> 70.5	12	14	12.6	0.005	14	10
<i>RAND</i> 70.7	7	8	7.8	0.005	9	13.33
<i>RAND</i> 70.9	5	5	5	0.005	6	16.66

Tableau 5.4 . Résultats obtenus avec le critère 4

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Algorithme exacte	Écart moyen
<i>RAND</i> 30.1	25	26	25.2	0.004	26	3.07
<i>RAND</i> 30.3	13	15	14.2	0.002	15	5.33
<i>RAND</i> 30.5	8	10	9.4	0.004	10	6
<i>RAND</i> 30.7	6	7	6.6	0.002	7	5.71
<i>RAND</i> 30.9	5	5	5	0.002	5	0
<i>RAND</i> 40.1	30	31	30.6	0.003	31	1.29
<i>RAND</i> 40.3	16	17	16.8	0.003	17	1.17
<i>RAND</i> 40.5	11	11	11	0.002	11	0
<i>RAND</i> 40.7	6	6	6	0.005	8	25
<i>RAND</i> 40.9	4	5	4.6	0.002	5	8
<i>RAND</i> 50.1	32	33	32.6	0.002	34	4.11
<i>RAND</i> 50.3	15	17	16.2	0.004	18	10
<i>RAND</i> 50.5	10	11	10.6	0.003	13	18.46
<i>RAND</i> 50.7	7	7	7	0.003	8	12.5
<i>RAND</i> 50.9	5	7	6	0.005	7	14.28
<i>RAND</i> 60.1	38	39	38.8	0.006	40	3
<i>RAND</i> 60.3	18	20	19	0.003	20	5
<i>RAND</i> 60.5	10	12	10.8	0.004	13	16.92
<i>RAND</i> 60.7	7	9	8.4	0.003	9	6.66
<i>RAND</i> 60.9	5	5	5	0.003	6	16.66
<i>RAND</i> 70.1	40	42	41	0.004	44	6.81
<i>RAND</i> 70.3	19	21	19.8	0.003	22	10
<i>RAND</i> 70.5	12	14	13	0.002	14	7.14
<i>RAND</i> 70.7	7	8	7.6	0.002	9	15.55
<i>RAND</i> 70.9	5	5	5	0.005	6	16.66

Tableau 5.5 . Résultats obtenus avec le critère 5

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Algorithme exacte	Écart moyen
RAND30.1	25	26	25.6	0.004	26	1.53
RAND30.3	13	15	14	0.005	15	6.66
RAND30.5	8	10	9.4	0.004	10	6
RAND30.7	6	7	6.8	0.003	7	2.85
RAND30.9	5	5	5	0.005	5	0
RAND40.1	29	31	30.2	0.003	31	2.58
RAND40.3	16	17	16.6	0.004	17	2.35
RAND40.5	11	11	11	0.003	11	0
RAND40.7	6	6	6	0.003	8	25
RAND40.9	4	5	4.4	0.003	5	12
RAND50.1	31	33	32.2	0.003	34	5.29
RAND50.3	16	17	16.4	0.004	18	8.88
RAND50.5	10	11	10.6	0.004	13	18.46
RAND50.7	7	7	7	0.003	8	12.5
RAND50.9	5	7	6	0.003	7	14.28
RAND60.1	38	39	38.6	0.007	40	3.5
RAND60.3	18	20	19	0.006	20	5
RAND60.5	10	12	11.4	0.005	13	12.3
RAND60.7	7	9	8.2	0.004	9	8.88
RAND60.9	5	5	5	0.005	6	16.66
RAND70.1	40	42	41.4	0.005	44	5.9
RAND70.3	19	20	19.4	0.006	22	11.81
RAND70.5	12	14	12.8	0.003	14	8.57
RAND70.7	7	8	7.6	0.005	9	15.55
RAND70.9	5	5	5	0.005	6	16.66

Tableau 5.6 . Résultats obtenus avec le critère 6

Les résultats détaillés des 6 tableaux précédents peuvent être résumés dans le tableau suivant qui contient les écarts moyens (en %) par rapport à la valeur optimale ainsi que le nombre d’instances pour lesquelles chaque algorithme a réussi à produire une solution optimale.

Critère	1	2	3	4	5	6
Écart moyen	22.01	12.63	9.38	7.73	8.77	8.92
Nombre d’optima	1	6	12	14	13	13

Tableau 5.7 . Résumé des résultats pour les 25 instances

D’après les courbes, où l’axe des abscisses représente les graphes en fonction de l’ordre et la densité (nous considérons les points 30.1, 30.3...), l’axe des ordonnées représente les moyennes (voir figure 5.1) et le tableau (5.7) les meilleurs résultats sont ceux obtenus en utilisant le critère 4, mais globalement la différence n’est pas très grande entre les critères 4, 5 et 6. En fin de classement nous avons les critères 1 et 2, ce qui n’est

pas étonnant puisqu'il s'agit des critères les plus basiques dans lesquels nous avons mis un premier jet.

Améliorations

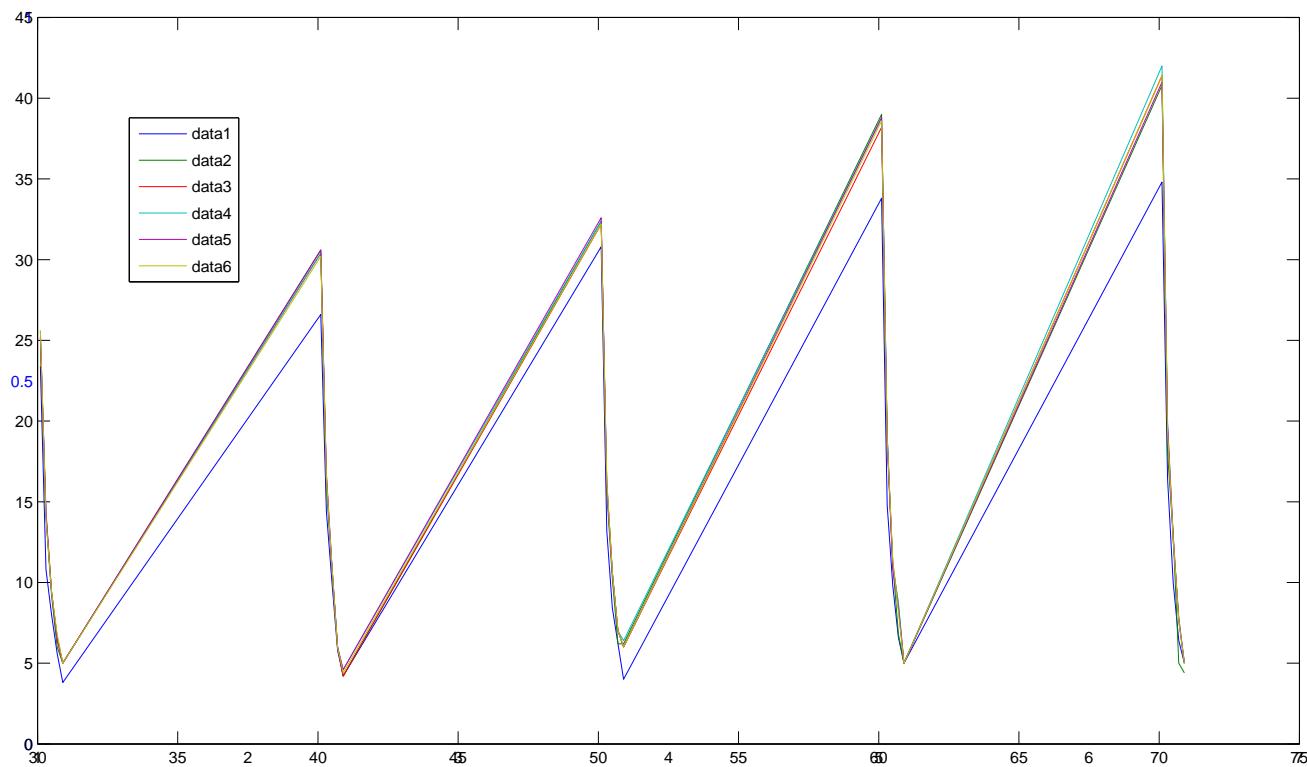
Dans le but de tenter d'améliorer un peu les résultats de la section précédente, sans toutefois trop augmenter les temps de calcul, nous avons tenté de combiner plusieurs critères entre eux. Pour ce faire, nous appliquons tout d'abord l'algorithme constructif avec un critère A; puis nous ôtons aléatoirement un sous-ensemble k de sommets de la forêt ainsi obtenue. Nous tentons ensuite à nouveau d'agrandir la forêt, mais cette fois-ci en utilisant un critère $B \neq A$. Après avoir effectué quelques essais préliminaires, nous avons fixé k égal à $\lfloor \frac{F}{2} \rfloor$, où F est la cardinalité de la forêt courante. Ceci peut être résumé comme suit.

Algorithme hybride pour la plus grande forêt

1. Déterminer une forêt F à l'aide d'un algorithme constructif en partant d'un ensemble vide de sommets et en utilisant un critère $A \in \{1, \dots, 6\}$. Poser $\text{Meilleur} = |F|$.
2. Ôter $\lfloor \frac{F}{2} \rfloor$ sommets de F , en choisissant les sommets à éliminer de manière aléatoire. Soit F' la nouvelle forêt.
3. Déterminer une forêt F'' à l'aide d'un algorithme constructif en partant de la forêt F' et en utilisant un critère $B \in \{1, \dots, 6\}$ différent de A. Poser $\text{Meilleur} = \max\{|F|, |F''|\}$.

Nous avons réalisé quelques tests avec trois variantes: Variante 1, avec $A = 3$, $B = 4$; variante 2 avec $A = 3$, $B = 5$; et variante 3 avec $A = 1$ et $B = 4$.

Figure 5.1 . Courbes représentatives des résultats obtenus par les six critères



Nous présentons ci-dessous le tableau récapitulatif des résultats obtenus après cinq essais. La description des données par colonne est la suivante: $|E|$ dénote le nombre d'arêtes; Pi, la pire solution; Mo, la moyenne des solutions; Me, la meilleure solution et Ts, le temps en secondes.

Résultats obtenus en utilisant les variantes 1, 2 et 3:

Graphes	$ E $	Variante 1				Variante 2				Variante 3				Algorithme exacte	
		Pi	Mo	Me	Ts	Pi	Mo	Me	Ts	Pi	Mo	Me	Ts	Algorithme exacte	Temps(s)
RAND30.1	42	25	25	25	0.008	26	26	26	0.008	24	24	24	0.004	26	0.83
RAND30.3	125	15	15	15	0.01	12	12	12	0.01	14	14	14	0.004	15	3.7
RAND30.5	214	10	10	10	0.01	9	9.8	10	0.01	7	7	7	0.008	10	1.06
RAND30.7	312	6	6.8	7	0.02	7	7	7	0.02	5	5	5	0.01	7	0.6
RAND30.9	391	5	5	5	0.02	4	4	4	0.03	4	4.8	5	0.01	5	0.49
RAND40.1	86	29	29.6	30	0.01	29	30.2	31	0.01	28	28.8	29	0.008	31	130
RAND40.3	239	17	17	17	0.02	16	16.6	17	0.04	13	14.2	15	0.01	17	121
RAND40.5	399	10	10.8	11	0.03	10	10.8	11	0.03	7	7.2	9	0.02	11	5.49
RAND40.7	558	8	8	8	0.03	5	5	5	0.05	5	5.4	6	0.03	8	1.07
RAND40.9	709	5	5	5	0.03	4	4.4	5	0.04	3	4	5	0.03	5	0.78
RAND50.1	127	30	31.2	32	0.03	31	31	31	0.04	28	29.4	30	0.02	34	>7200
RAND50.3	374	15	16.2	17	0.06	17	17.2	18	0.07	15	15.2	16	0.02	18	2560
RAND50.5	615	10	10.6	11	0.06	9	9.6	10	0.07	8	8.6	11	0.04	13	27.2
RAND50.7	861	7	7.4	8	0.06	7	7	7	0.07	5	6.2	7	0.04	8	2.27
RAND50.9	1107	6	6.8	7	0.06	6	6	6	0.08	4	4.8	5	0.3	7	1.16
RAND60.1	171	39	39.6	40	0.05	38	38.8	39	0.06	35	35.8	36	0.02	40	> 7200
RAND60.3	512	19	19.8	20	0.08	18	19	20	0.1	16	16.8	17	0.05	20	> 7200
RAND60.5	882	10	11.8	13	0.08	10	10.6	11	0.1	10	10	10	0.08	13	127
RAND60.7	1237	7	8.6	9	0.1	9	9	9	0.1	7	7.8	8	0.09	9	5.23
RAND60.9	1600	5	5.8	6	0.1	5	5	6	0.1	4	4.4	5	0.09	6	1.62
RAND70.1	235	43	43	43	0.06	41	41.4	42	0.1	31	36.6	37	0.04	44	> 7200
RAND70.3	704	20	21.2	22	0.1	19	19.6	20	0.1	15	17.2	19	0.08	22	> 7200
RAND70.5	1210	11	12.6	13	0.1	12	12.8	14	0.1	9	10	11	0.1	14	490
RAND70.7	1698	7	7	7	0.3	6	7.6	8	0.1	7	7.8	8	0.1	9	10.5
RAND70.9	2189	5	5	5	0.1	5	5	5	0.2	4	4.8	6	0.1	6	2.28

Tableau 5.8 . Résultats obtenus par les variantes 1, 2 et 3

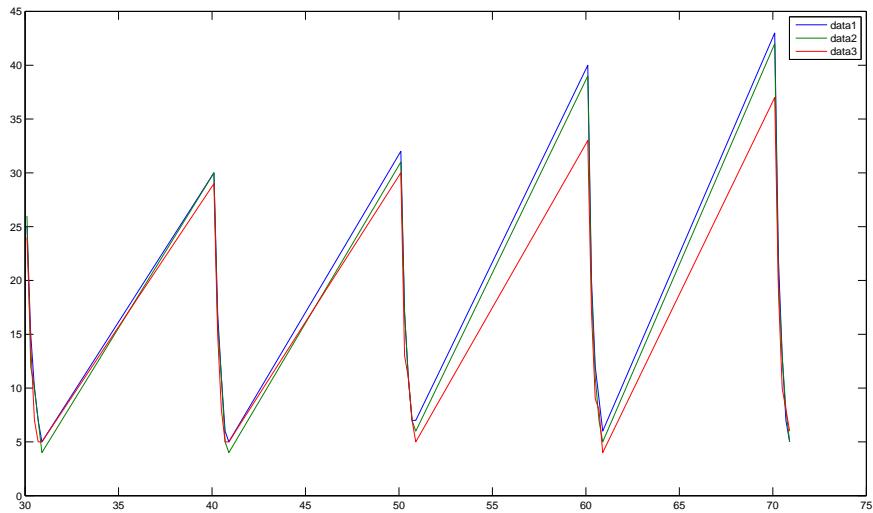


Figure 5.2 . Courbes représentatives des trois variantes

Variante	1	2	3
écart moyen	5.3	8.7	20.2
nombre d'optima	16	12	3

Tableau 5.9 . Résumé des résultats

Remarques:

Les tests et les courbes (figure 5.2) montrent clairement que les résultats obtenus par les variantes 1 et 2 sont meilleurs que les solutions trouvées par la variante 3.

Un point mis en évidence par ces résultats, est que le choix des sommets par des critères qui prennent en considération les degrés des sommets améliore considérablement les valeurs trouvées. Par ailleurs, en appliquant des essais randomisés durant une minute (ce qui est nettement inférieur aux temps mis par l'algorithme exact dans la majorité des cas) les variantes 1 et 2 aboutissent à trouver tous les optima (sauf pour l'instance 70.9 pour laquelle nous avons dû réaliser des essais durant une minute et

demie pour que les deux variantes puissent trouver l'optimum), ce qui n'est pas le cas pour la variante 3.

Nous exposons dans la suite les résultats obtenus par les variantes 1 et 2 sur des graphes aléatoires d'ordre 250, 500 et 1000 sommets. Nous avons effectué 5 tests sur chaque graphe.

Résultats obtenus en utilisant les variantes 1 et 2:

Graphes	E	Variante 1				Variante 2			
		Pire cas	Moyenne	Meilleur cas	Temps	Pire cas	Moyenne	Meilleur	Temps
RAND250.1	9090	19	20.4	23	4.8	21	22.6	22	7.1
RAND250.3	9278	20	22.8	24	5.1	22	22.4	24	7.3
RAND250.5	15760	13	13.4	15	6.05	14	14.8	16	7.1
RAND250.7	21768	8	8.6	9	6.67	7	8.4	9	8.09
RAND250.9	28021	7	7	7	6.8	6	6	6	10.03
RAND500.1	37264	25	25.4	26	39.96	25	25.2	26	58
RAND500.3	37422	29	29	29	42.1	28	28.6	29	57
RAND500.5	62547	16	16	16	49.2	14	14	14	65.4
RAND500.7	87226	10	10.8	11	48.5	9	9.8	10	70.3
RAND500.9	112319	6	6	6	295.29	6	6	6	60.3
RAND1000.1	50148	83	83.6	84	368.4	79	79	79	381.5
RAND1000.3	149516	27	27.8	29	414	28	28	28	441
RAND1000.5	249703	16	16	16	430	16	16	16	443
RAND1000.7	349722	6	6	6	440.3	5	5	5	500.8
RAND1000.9	449710	7	7	7	460	5	6	7	543

Tableau 5.10 . Résultats obtenus par les variantes 1 et 2

Remarques:

Nous constatons que la variante 1 fournit les meilleurs résultats. Dès lors nous pouvons supposer que la combinaison des critères 3 et 4 était un bon choix pour mesurer la

qualité des sommets à insérer. Cependant, nous remarquons qu'il y a un accroissement des temps de calcul. Pour améliorer cela, nous pouvons mémoriser pour chaque sommet v n'appartenant pas à la forêt, le nombre $n_i(v)$ de sommets adjacents à v dans une composante connexe C_i . Ainsi, le rajout d'un sommet v est possible si et seulement si, pour tout i , $n_i(v) \leq 1$.

Hypercubes

Nous avons également testé nos algorithmes sur une classe de graphes bipartis appelés hypercubes. En effet, plusieurs chercheurs ont cherché à déterminer le transversal de cycles d'un hypercube [11].

Nous définissons, tout d'abord, le produit cartésien noté \times entre deux graphes G et H . Le graphe $G \times H$ est défini comme suit: $V(G \times H) = V(G) \times V(H)$ et

$$E(G \times H) = \{\{(x, y), (x', y')\} : [x = x' \wedge yy' \in E(H)] \vee [y = y' \wedge xx' \in E(G)]\}$$

Les hypercubes sont définis récursivement de la manière suivante:

$$Q_1 = K_2 \text{ et } Q_n = K_2 \times Q_{n-1}.$$

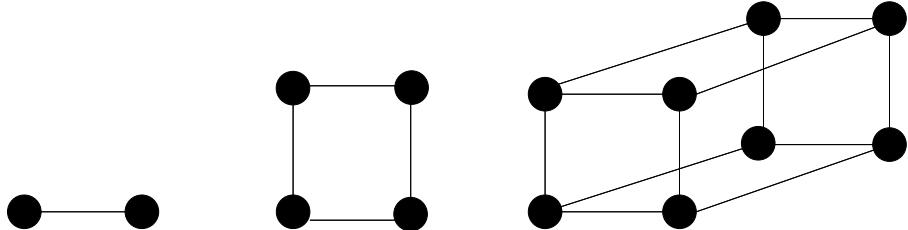


Figure 5.3 . Illustration des hypercubes: Q_1 , Q_2 et Q_3

Dans cet article [11], les auteurs prouvent plusieurs lemmes leur permettant de trou-

ver la valeur exacte de $\nabla(Q_n)$, le transversal de cycles de Q_n , pour $n \leq 8$. Pour les hypercubes Q_9, Q_{10}, Q_{11} et Q_{12} , ils donnent des bornes inférieures et supérieures.

Nous avons implanté un programme qui génère des hypercubes et nous avons testé nos variantes sur ces graphes. Dans les tableaux suivants nous donnons les résultats que nous avons obtenu. Dans ces tableaux, F_{opt} dénote la cardinalité de la forêt optimale. B_i et B_s dénotent les bornes inférieures et supérieures.

Hypercubes	Ordre	E	$\nabla(Q_n)$	F_{opt}	Variante 1				Variante 2			
					Pi	Mo	Me	Ts	Pi	Mo	Me	Ts
Q_4	16	32	6	10	9	9.6	10	0.004	8	8.8	9	0.008
Q_5	32	80	14	18	16	17	18	0.008	17	17	17	0.01
Q_6	64	192	28	36	34	34.8	35	0.04	31	31.8	32	0.08
Q_7	128	448	56	72	64	66.4	67	0.4	63	63.8	64	0.5
Q_8	256	1024	112	144	128	129.6	130	2.4	119	119.6	120	3.9

Hypercubes	Ordre	E	Bi	Bs	Variante 1				Variante 2			
					Pi	Mo	Me	Ts	Pi	Mo	Me	Ts
Q_9	512	2304	200	288	252	253.2	254	28.2	242	242.8	243	29.8
Q_{10}	1024	5120	418	576	510	510	510	155.7	515	515	515	190
Q_{11}	2048	11264	864	1152	1027	1027.6	1028	1043	1007	1007.8	1008	1077
Q_{12}	4096	24576	1872	2304	2048	2049.2	2050	2623	2023	2023	2023	2650

Tableau 5.11 . Résultats obtenus pour les hypercubes

Conclusion

Nous avons réussi à obtenir des solutions de bonne qualité dans tous les tests effectués avec les variantes 1 et 2, en dépit de l'accroissement du temps d'exécution. Les objectifs obtenus par ces deux procédures respectent les bornes supérieures et inférieures. Notons que les valeurs trouvées par la troisième variante (que nous n'avons pas données dans le

tableau 5.11) ne sont pas satisfaisantes car elles ne cadrent pas dans les intervalles, par exemple pour Q_{10} la procédure retourne la valeur $399 < B_i$.

5.3 Algorithme constructif pour la détermination de la plus grande forêt linéaire

Afin de déterminer de grandes forêts linéaires, nous utilisons le critère 5 défini dans la section précédente. Nous ajoutons la contrainte qu'un sommet ne peut être insérable que si le nombre de ses voisins dans la forêt courante est inférieur ou égal à deux. De plus le degré maximum du sous-graphe induit par les sommets de la forêt doit demeurer inférieur ou égal à deux.

Étant donné que nous ne connaissons aucun algorithme exact pour la détermination de la plus grande forêt linéaire, nous comparons les résultats fournis par nos algorithmes à la taille de la plus grande forêt induite qui constitue une borne supérieure sur la valeur de la solution optimale. Nous avons effectué cinq essais sur 40 instances. Voici les résultats obtenus.

Résultats

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Borne sup	Écart moyen
<i>RAND30.1</i>	14	14	14	0.004	26	46.1
<i>RAND30.3</i>	5	6	5.8	0.005	15	61.3
<i>RAND30.5</i>	4	4	4	0.01	10	60
<i>RAND30.7</i>	4	4	4	0.01	7	42.8
<i>RAND30.9</i>	3	3	3	0.06	5	40
<i>RAND40.1</i>	15	17	16.2	0.03	31	47.7
<i>RAND40.3</i>	6	8	7.6	0.05	17	55.2
<i>RAND40.5</i>	4	4	4	0.03	11	63.6
<i>RAND40.7</i>	3	3	3	0.03	8	62.5
<i>RAND40.9</i>	3	3	3	0.03	5	40

Tableau 5.12 . Résultats obtenus par l'algorithme de recherche de la plus grande forêt linéaire induite

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Borne sup	Écart moyen
RAND50.1	9	13	10.2	0.04	34	70
RAND50.3	5	7	5.4	0.04	18	70
RAND50.5	3	3	3	0.04	13	76.9
RAND50.7	3	4	3.4	0.03	8	57.4
RAND50.9	3	3	3	0.03	7	57.1
RAND60.1	16	16	16	0.07	40	54.5
RAND60.3	5	9	6.6	0.06	20	67
RAND60.5	3	6	4.2	0.05	13	67.6
RAND60.7	3	5	3.8	0.04	9	57.7
RAND60.9	3	3	3	0.1	6	33.3
RAND70.1	14	20	17.2	0.09	44	60.9
RAND70.3	8	9	8.4	0.07	22	61.8
RAND70.5	4	5	4.4	0.09	14	68.5
RAND70.7	3	4	3.2	0.1	9	64.4
RAND70.9	3	3	3	0.14	6	50
RAND250.1	11	13	11.8	3.5	23	48.6
RAND250.3	8	11	9	4.8	24	62.5
RAND250.5	5	6	5.2	4.5	16	76.5
RAND250.7	4	4	4	5.4	9	55.5
RAND250.9	3	3	3	5.8	7	57.1
RAND500.1	12	12	12	30.8	26	53.8
RAND500.3	10	14	12.6	28.5	29	56.5
RAND500.5	7	7	7	35.5	16	56.2
RAND500.7	5	5	5	39.2	11	54.5
RAND500.9	3	3	3	48.7	6	50
RAND1000.1	38	46	44.2	163.7	84	47.3
RAND1000.3	14	14	14	210	29	51.7
RAND1000.5	8	8	8	250.8	16	50
RAND1000.7	3	3	3	302	6	50
RAND1000.9	3	3	3	378.1	7	57.1

Tableau 5.13 . Résultats obtenus par l'algorithme de recherche de la plus grande forêt linéaire induite

Remarques

Nous constatons, à partir de ces résultats, que l'écart moyen entre la cardinalité de la forêt linéaire et de la plus grande forêt est plus petit pour les graphes de densité 0.1. Cet écart augmente pour les graphes de densité 0.3 et 0.5, pour diminuer par la suite pour les instances de densité 0.7 et 0.9.

5.4 Algorithme constructif pour la détermination du plus grand arbre

Pour déterminer un arbre induit d'un graphe, un sommet s est insérable s'il ne forme pas de cycle avec les sommets de la solution courante. Et de plus, ce sommet doit avoir au moins un voisin dans l'arbre courant afin d'assurer la connexité.

De la même manière que précédemment, nous comparons les résultats fournis par notre algorithme à la taille de la plus grande forêt induite qui constitue une borne supérieure sur la valeur de la solution optimale. Nous avons effectué cinq essais sur 40 instances. Voici les résultats obtenus.

Résultats

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Borne sup	Écart moyen
RAND30.1	13	17	16.4	0.008	26	36.9
RAND30.3	7	9	7.8	0.012	15	48
RAND30.5	5	5	5	0.008	10	50
RAND30.7	4	4	4	0.01	7	42.8
RAND30.9	3	3	3	0.02	5	40
RAND40.1	14	14	14	0.01	31	54.8
RAND40.3	8	9	8.6	0.05	17	49.4
RAND40.5	5	5	5	0.03	11	54.5
RAND40.7	3	5	4.2	0.02	8	47.5
RAND40.9	3	3	3	0.02	5	40

Tableau 5.14 . Résultats obtenus par l'algorithme de recherche du plus grand arbre induit

Graphes	Pire cas	Meilleur cas	Moyenne	Temps (sec)	Borne sup	Écart moyen
<i>RAND</i> 50.1	16	19	17.2	0.03	34	49.4
<i>RAND</i> 50.3	6	10	7.6	0.04	18	57.7
<i>RAND</i> 50.5	4	4	4	0.05	13	69.2
<i>RAND</i> 50.7	4	4	4	0.06	8	50
<i>RAND</i> 50.9	3	3	3	0.06	7	57.1
<i>RAND</i> 60.1	17	21	18.6	0.07	40	53.5
<i>RAND</i> 60.3	7	7	7	0.06	20	65
<i>RAND</i> 60.5	4	6	4.8	0.08	13	63
<i>RAND</i> 60.7	4	4	4	0.1	9	55.5
<i>RAND</i> 60.9	3	3	3	0.1	6	50
<i>RAND</i> 70.1	11	13	12.2	0.1	44	72.2
<i>RAND</i> 70.3	8	13	10	0.1	22	54.5
<i>RAND</i> 70.5	8	8	8	0.1	14	42.8
<i>RAND</i> 70.7	3	3	3	0.1	9	66.6
<i>RAND</i> 70.9	4	4	4	0.1	6	33.3
<i>RAND</i> 250.1	9	11	10.4	4	23	54.7
<i>RAND</i> 250.3	10	10	10	3.8	24	58.3
<i>RAND</i> 250.5	6	8	7.4	6.1	16	53.7
<i>RAND</i> 250.7	3	4	3.8	6.4	9	57.7
<i>RAND</i> 250.9	3	3	3	5.1	7	57.1
<i>RAND</i> 500.1	13	13	13	33	26	50
<i>RAND</i> 500.3	14	14	14	28.2	29	51.7
<i>RAND</i> 500.5	7	7	7	41.1	16	56.2
<i>RAND</i> 500.7	5	5	5	43.7	11	54.5
<i>RAND</i> 500.9	4	4	4	44.8	6	33.3
<i>RAND</i> 1000.1	56	56	56	149	84	33.3
<i>RAND</i> 1000.3	14	15	14.6	216	29	49.6
<i>RAND</i> 1000.5	8	8	8	267	16	50
<i>RAND</i> 1000.7	3	5	4.2	302	6	29.9
<i>RAND</i> 1000.9	3	3	3	402	7	57.1

Tableau 5.15 . Résultats obtenus par l'algorithme de recherche du plus grand arbre induit

Conclusion

Le but de ce chapitre n'était pas de développer une métaheuristique puissante, mais plutôt d'essayer de trouver de bons résultats au moyen d'algorithmes simples. Toutefois, il est nécessaire de poursuivre la recherche pour développer des heuristiques performantes pour résoudre ces problèmes.

Le point important de notre approche est de faire intervenir les degrés des sommets. En effet, intuitivement, le degré d'un sommet est relié au nombre de cycles auxquels appartient ce sommet, d'où l'intérêt de sélectionner les sommets de degré minimum. En vue d'améliorer notre algorithme qui combine deux critères, nous pouvons imposer des règles pour la suppression de sommets dans la deuxième phase de l'algorithme hybride. Là aussi on peut faire intervenir les degrés des sommets.

CONCLUSION

La théorie des graphes va bénéficier grandement du progrès des logiciels. En effet, grâce aux développements récents des systèmes informatiques un grand nombre de conjectures peut être générées et proposées aux chercheurs.

Notre objectif dans cette thèse était de prouver la conjecture suivante, générée par le système Graffiti [30] et ouverte depuis une quinzaine d'années: «Soit b l'ordre du plus grand sous-graphe induit biparti d'un graphe connexe G . Alors la distance moyenne entre les sommets de G est inférieure ou égale à $b/2$.»

Nous avons prouvé ce résultat et nous avons réussi à prouver un résultat encore plus fort, à savoir;

Théorème 5.4.1. *Soit G un graphe connexe de n sommets. Alors*

$$\mu(G) \leq \frac{F(G)}{2}.$$

C'est l'apport important du présent travail.

Dans un deuxième lieu, nous avons essayé de déterminer de nouvelles bornes supérieures sur des différences d'invariants. Dans ce même cadre nous avons tenté de caractériser et d'explorer des graphes extrêmes à l'aide du système AutoGraphiX.

Nous tenons à affirmer que l'utilisation du système AutoGraphiX a constitué une aide considérable pour notre travail. Ce système combine le raisonnement mathématique, la génération et la visualisation des graphes, des fonctions interactives simples et un calcul rapide permettant de déceler d'éventuelles conjectures.

Toutefois, ce travail de génération de conjectures et de caractérisation de graphes

extrêmes peut s'avérer très difficile pour plusieurs raisons. Par exemple, le système peut ne pas réussir à trouver des graphes extrêmes, ou bien l'utilisateur n'arrive pas à décrire les familles de ces graphes.

Voici un tableau récapitulatif des résultats prouvés dans le cadre de ce travail:

Borne inférieure	$Inv_1 - Inv_2$	Borne supérieure
0	$\alpha_2 - F$	$\lfloor \frac{n}{2} \rfloor - 1$
0	$\alpha_2 - LF$	$\lfloor \frac{n}{2} \rfloor$
0	$F - LF$	$\lfloor \frac{n}{4} \rfloor$
0	$\frac{\alpha_2}{2} - \mu$	$\frac{n}{2} - \frac{3n-4}{2(n-1)}$
0	$\frac{F}{2} - \mu$	$\frac{n}{2} + \frac{2}{n} - 2$
?	$\frac{LF}{2} - \mu$	$\frac{n-1}{2} + \frac{2n^2-6n+6}{n(n-1)}$
0	$F - T$?

Tableau 5.16 . Récapitulation des résultats

Nous avons présenté, en dernier lieu, des algorithmes constructifs qui tentent de déterminer la plus grande forêt induite, la plus grande forêt linéaire induite et le plus grand arbre.

Nous posons à la fin de ce travail les deux conjectures suivantes.

Conjecture 5.4.2. *Soit G un graphe connexe d'ordre n , alors on a*

$$F(G) - T(G) \leq n - 2\lfloor \sqrt{n} \rfloor.$$

Conjecture 5.4.3. *Soit G un graphe connexe de n sommets. Alors*

$$\mu(G) \leq \frac{LF(G)}{2}.$$

La découverte scientifique assistée par l'ordinateur est une branche prometteuse en théorie des graphes qui mènera certainement à de nouvelles connaissances et élargira les domaines de recherche. Néanmoins l'intervention de l'humain reste toujours essentielle pour perfectionner et améliorer les résultats de la machine.

Nous concluons par cette citation: «Les ordinateurs sont des outils magnifiques pour la réalisation de nos rêves, mais aucune machine ne peut remplacer l'étincelle humaine de l'esprit, de la compassion, de l'amour, et de la compréhension. » Louis Gerstner [44]

REFERENCES

- [1] Aharoni, R, Berger, E, 2009. Menger's theorem for infinite graphs. *Inventiones Mathematicae*, Vol. 176, p 1-62.
- [2] Alon, N. 1996. Bipartite subgraphs *Combinatorica*, Vol. 16, p 301-311.
- [3] Alon, N, Mubayi, D, Thomas, R. 2001. Large induced forests in sparse graphs. *Journal of Graph Theory*, Vol. 38, p 113-123.
- [4] Aouchiche, M. 2006. Comparaison automatisée d'invariants en théorie des graphes. *Thèse de doctorat, École Polytechnique de Montréal*.
- [5] Aouchiche, M, Brinkmann, G, Hansen, P. 2008. Variable neighborhood search for extremal graphs. 21. Conjectures and results about the independence number. *Discrete Applied Mathematics*, Vol. 156, p 2530-2542.
- [6] Aouchiche, M, Caporossi, G, Hansen, P, Laffay, P. 2005. AutoGraphiX: a survey. *Electronics Notes in Discrete Mathematics*, Vol 22. p 515-520.
- [7] Aouchiche, M, Hansen, P, Caporossi, G, Mélot, H, Stevanović, D. 2003. What Forms Do Interesting Conjectures Have in Graph Theory? *Les Cahiers du Gérard*.
- [8] Aouchiche, M, Hansen, P. 2007. Automated Results and Conjectures on Average Distance in Graphs. *Game Theory in Paris: Special Volume in Memory of Claude Berge*, p 21-36.
- [9] Bafna, V, Berman, P, Fujito, T. 1995. Constant ratio approximations of the weighted feedback vertex set problems for undirected graphs. *Lectures Notes in Computer Science*, Vol. 1004 , p 142-151.

- [10] Bar-Yehuda, R, Geiger, D, Naor, J, Roth, R.M. 1998. Approximation algorithms for the vertex feedback set problem with applications to constraint satisfaction and Bayesian inference. *SIAM Journal. Comput*, Vol. 27, p 942-959.
- [11] Beineke, L. W, Vandell, R.C. 1996. Decycling Graphs. *Journal of Graph Theory*, Vol. 25, p 59-77.
- [12] Berge, C. 1970. Graphes et Hypergraphes. *Dunod*.
- [13] Biggs, N, L, Lloyd, E, K, Wilson, R, J. 1998 Graph theory, p 1736-1936. *Oxford University Press* .
- [14] Bollobás, B, Riordan, O. 1998 On some conjectures of Graffiti. *Discrete Mathematics*, Vol. 179, p 223-230.
- [15] Brigham, C, Dutton, R.D. 1991 A compilation of relations between graph invariants. *Networks*, Vol. 21, p 421-455.
- [16] Brigham, C, Dutton, R. D, Gomez, F. 1989 INGRID: A graph invariant manipulator. *J. Symbolic Computation*, Vol. 7, p 163-177.
- [17] Brandstädt, A, Kratsh, D. 1985. On the restriction of some NP-complete graph problems to permutation graphs. *Lectures Notes in Comp. Sci*, Vol. 199, p 53-62.
- [18] Caporossi, G, Hansen, P. 2000. Variable neighborhood search for extremal graphs. 1. The AutoGraphiX system. *Discrete Math*, Vol. 212 , p 29-44.
- [19] Caporossi, G, Hansen, P. 2004. Variable neighborhood search for extremal graphs. 5: Three ways to automate finding conjectures. *Discrete Math*, Vol. 276, p 81-94.
- [20] Chang, M. S, Liang, Y.D. 1997. Minimum feedback vertex sets in cocomparability graphs and convex bipartite graphs. 1997 *Acta Informatica*, Vol. 34, p 337-346.

- [21] Cheng, K.T, Agrawal, V. D. 1990. A partial scan method for sequential circuits with feedback. 1990 *IEEE Transactions on Computers*, Vol. 39, No 4 , p 544-548.
- [22] Chung, F. R. K. 1988. The Average Distance and the Independence Number. *Journal of Graph Theory*, Vol. 12, p 229-235.
- [23] Cvetkovic, D, Kraus, L, Simić, S.K. 1981 Discussing graph theory with a computer, Implementation of graph theoretic algorithms. *Univ. Beograd Publ*, Vol. 733, p 100-104.
- [24] Dankelmann, P. 1994. Average Distance and Independence Number. *Discrete Applied Mathematics*, Vol. 51, p 75-83.
- [25] Dankelmann, P. 1997. Average Distance and Domination Number. *Discrete Applied Mathematics*, Vol. 80, p 21-35.
- [26] Dankelmann, P. 2006. Average Distance, Minimum Degree, and Size. *Utilitas Mathematica*, Vol. 69, p 233-243.
- [27] DeLaViña, E, Waller, B, Pepper, R. 2008. A note on dominating sets and average distance. *Discrete Mathematics*, Vol. 309, p 2615-2619.
- [28] DeLaViña, E, Waller. 2008. Some Conjectures Of Graffiti.pc On The Maximum Order Of Induced Subgraphs. *Preprint*.
- [29] DeLaViña, E. 2002. Graffiti.pc: A variant of Graffiti. *Preprint*.
- [30] DeLaViña, E Web site of bibliographical information on conjectures of Graffiti, Web adress: <http://cms.dt.uh.edu/faculty/delavinae/wowref.html>.
- [31]de Werra, D, Hertz, A. 1993. On the stability number of AH-free graphs. *Journal of graph theory*, Vol. 17, -1, p 53-63.

- [32] Erdős, P, Posa, L. 1962. On the maximal number of disjoint circuits of a graph. *Pub. Math. Debrecen*, Vol. 9, p 3-12.
- [33] Erdős, P, Saks, M, Sós, V. 1986. Maximum induced trees in graphs. *Journal of Graph Theory*, Vol. 41, p 61–79.
- [34] Erdős, P. 1961. Graph theory and probability. *Canad J. Math* p 346-352.
- [35] Fajtlowicz, S ,Waller, W. A. 1986. On two conjectures of Graffiti. *Preprint*, Vol. 13.
- [36] Fajtlowicz, S. 1988. A characterization of radius-critical graphs. *Journal of Graph Theory*, Vol. 12, p 529-532.
- [37] Favaron, O, Maheo, M, Sacle, J. 1990. On conjectures of Graffiti-1. *Ars Combinatoria*, Vol. 29C, p 90-106.
- [38] Favaron, O, Maheo, M, Sacle, J. 1993. Some eigenvalue properties in graphs. *Discrete Mathematics*, Vol. 111, p 197-220.
- [39] Feo, T.A, Resende, MGC. 1995. Greedy randomized adaptive search procedures. *Journal of global optimization*, Vol. 6, p 109–133.
- [40] Festa, P, Pardalos, P. M, Resende, G.C. 2000. Feedback set problems. *Encyclopedia of optimization*.
- [41] Fox, J, Po-shen, L, Sudakov, B. 2008 Large induced trees in K_r -free graphs. *Journal of combinatorial theory, Series B* 99, p 494-501.
- [42] Frank, A. 1975 Some polynomial algorithms for certain graphs and hypergraphs. *Proceedings of the fifth British combinatorial conference*, p 211-226.
- [43] Galinier, P. Habib, M. Hao, J. K. 1999. Métaheuristiques pour l'optimisation combinatoire et l'affectation sous contraintes. *Revue d'intelligence artificielle*.

- [44] Grestner, L. 2010. <http://www.best-quotes-poems.com/francais/citations-dordinateurs.html>.
- [45] Hajnal, A, 1965. A theorem on k -saturated graphs. *Canad. Journal. Math*, Vol. 17, P 720-724.
- [46] Hansen, P, Mladenovic, N. 2009. Variable neighborhood search methods. *Encyclopedia of Optimization 2nd Edition Springer 2009*, p 3975-3989.
- [47] Hansen, P, Hertz, A, Kilani, R, Marcotte, O, Schindl, D. 2009. Average distance and maximum induced forest. *Journal of Graph Theory*, Vol. 60, p 31-54.
- [48] Hansen, P. Mélot, H. 2002. Computers and discovery in algebraic graph theory. *Linear algebra and its applications*, Vol. 356, p 211-230.
- [49] Hertz, A. 2005 *Notes de cours: Théorie des graphes et des réseaux*, École polytechnique, Montréal, Canada.
- [50] Hertz, A. 1995. Polynomially solvable cases for the maximum stable set problem *Discrete applied mathematics*, Vol. 60, p 195-210.
- [51] Hosoya, H, 1971. Topological Index. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bulletin of the Chemical Society of Japan*, Vol. 44, p 2332-2339.
- [52] Jou, H. M, Lin, J. Y, 2000. On computing the minimum feedback vertex set of a directed graph by contraction operations. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on computer-aided design of integrated circuits and systems*, Vol. 19, -3, p 295-307.
- [53] Karp, R, M. 1972. Reducibility among combinatorial problems *Complexity of computer computations*, p 85-103.

- [54] Levy, H, Lowe, L. 1988. A contraction algorithm for finding small cycle cutsets *Journal of algorithm*, Vol. 9 , p 470-493.
- [55] Lipták, L, Lovász, L. 2001. Critical facets of the stable set polytope. *Combinatorica*, Vol. 21, p 61-88.
- [56] Liang, Y.D. 1994 On the feedback vertex set problem in permutation graphs *Information processing letters*, Vol. 52, -3, p 123-129.
- [57] Lovász, L, Plummer, M. D. 1986. Matching theory. *North-Holland Publishing Co*, Vol. 121.
- [58] Lung, L. C. Yi Tang, C. 1997. A linear time algorithm for the weighted feedback vertex problem on interval graphs. *Information processing letters*, Vol. 61, p 107-111.
- [59] Marathe, M.V, Pandu, C, Ravi, R. 1992. Efficient algorithms for generalized clique covering on interval graphs *Discr. Appl. Math*, Vol. 39, -10, p 1875-1891.
- [60] Melot, H. 2008. Facet defining inequalities among graph invariants: The system GraPHedron *Discr. Appl. Math*, Vol. 156 , p 87-93.
- [61] Monien, B, Schultz, R. 1981. Four approximation algorithms for the feedback vertex set problems. 1981 7th Conference on graph theoretic concepts of computer science, p 315-326.
- [62] Minty, G, J. 1980. On maximal independent sets of vertices in claw-free graphs. *Journal of Combinatorial Theory, Series B*, Vol. 28, p 284-304.
- [63] Pardalos, P. M, Qian, T, Resende, M.G.C. 1999 A greedy randomized adaptive search procedure for feedback vertex set. 1999. *J. Comb. Opt*, Vol. 2 , p 399-412.
- [64] Pier, J, P, 2000. Development of Mathematics. *Centre Universitaire du Luxembourg, Birkhäuser*.

- [65] Plesnik, G, J, 1984. On the sum of all distances in a graph or digraph. *Journal of Graph Theory*, Vol. 8, p 1-24.
- [66] Poljak, A. 1974. A note on stable sets and coloring of graphs. *Comment. Math. Univ. Carolinae*, Vol. 15, p 307-309.
- [67] Sewell, E. Trotter, L.E. 1993. Stability critical graphs and subdivisions of K_4 . *J. Combin. Theory Series B*, Vol. B59, p 74-84.
- [68] Shamir, A. 1979. A linear time algorithm for finding minimum cutsets in reduced graphs. *Acta Informatica*, Vol. 34 , p 337-346.
- [69] Wiener, H, 1947. Structural determination of paraffin boiling points. *Journal of the American Chemical Society*, Vol. 69, p 17–20.
- [70] Yannakakis, M. 1978 Node and edge-deletion NP-complete problems. *10th Annual ACM Symposium on the theory of computing*, p 253-264.
- [71] Yannakakis, M, Gavril, F. 1987 The maximum k-colorable subgraph problem for chordal graphs. *Info Process. Lett*, Vol. 24, p 133-137.
- [72] Zheng, M, Lu, X. 1990. On The Maximum Induced Forests of a Connected Cubic Graph without Triangles. *Discrete Mathematics*, Vol. 85, p 89–96.