Proposition de sujet d'alternance 1A
2023-24

Laboratoire : LIS, UMR 7020

Titre du sujet : Énumération lexicographique d'ensembles indépendants maximaux

Encadrant * (s) : Oscar Defrain et Arnaud Labourel

Nom : Defrain
Prénom : Oscar
Qualité ** : Maître de conférences
Localisation : LIS, Luminy, Marseille
Coordonnées e-mail : oscar.defrain@lis-lab.fr

Nom : Labourel
Prénom : Arnaud
Qualité ** : Maître de conférences
Localisation : LIS, Luminy, Marseille
Coordonnées e-mail : arnaud.labourel@lis-lab.fr

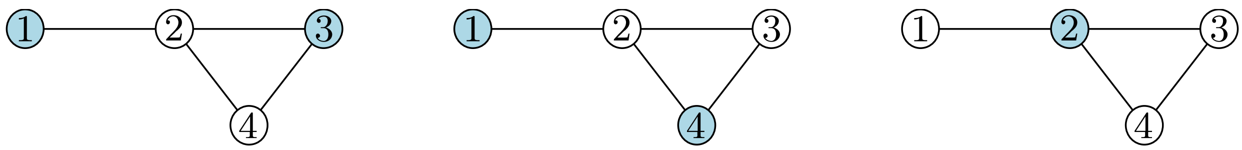
Descriptif du sujet et de la mission (au moins sur la 1er année) :

Un graphe (simple) est une manière de représenter une relation entre des éléments appelés *sommets* en indiquant pour chaque paire d'éléments si ces deux éléments sont reliés par une *arête* (on dit alors qu'ils sont adjacents) ou non. Les graphes sont un outil de modélisation extrêmement utile qui permettent entre autres de représenter des réseaux sociaux (deux personnes sont reliées si elles sont amies), des réseaux de machine (deux machines sont reliées si elles sont directement connectées via un câble), des réseaux routiers (deux carrefours sont reliés si un tronçon de rue les connecte), ...

Un ensemble indépendant est un ensemble de sommets dans un graphe tel que deux d'entre eux ne sont pas adjacents. La recherche d'un ensemble indépendant maximum, c'est-à-dire d'un ensemble indépendant de taille maximale, est un problème fondamental en informatique et a été l'un des premiers problèmes prouvés comme étant NP-complet et donc généralement considéré comme étant coûteux à résoudre. Dans certains cas, il peut être utile de produire des ensembles indépendants qui ne sont que maximaux, c'est-à-dire des ensembles indépendants pour lesquels il n'existe aucun sommet extérieur à l'ensemble qui puisse le rejoindre. Par définition, les Ensembles Indépendants Maximaux (EIM) sont des ensembles dominants (ensemble pour lequel chaque sommet du graphe est dans l'ensemble ou adjacent à un sommet de l'ensemble). Les EIM sont une notion importante en théorie des graphes ainsi que dans d'autres domaines comme l'analyse des réseaux en sciences sociales. Dans certaines applications de ce domaine, un EMI peut correspondre à un échantillon représentatif d'une population. Pour ce contexte, la production d'un seul EIM n'est généralement pas totalement satisfaisante, car elle peut ne pas tenir compte de certaines contraintes qui ne sont pas entièrement formalisées. Dans ce cas, l'une solution consiste à produire tous les EIM et à sélectionner les plus intéressants.

L'énumération des EIM est un problème classique d'énumération qui remonte aux années 1970 avec l'article de Tsukiyama et al. [3]. Cet article est le premier à donner un algorithme énumérant tous les ensembles indépendants maximaux avec un délai polynomial, c'est-à-dire avec une complexité en temps entre la production de deux EIM consécutifs bornée par un polynôme en la taille du graphe. La notion de délai polynomial est utile pour étudier la complexité de problèmes d'énumération car, dans la plupart des cas, le nombre d'objets à énumérer est exponentiel en la taille du graphe. La complexité en temps de l'algorithme d'énumération doit être dans ce cas au moins exponentielle afin juste de produire les EIM. C'est le cas pour les EIM, puisqu'un graphe à n sommets peut contenir jusqu'à 3^(n/3) EIM (borne atteinte pour une union disjointe de triangles).

Un autre article important pour les problèmes d'énumération est celui de Johnson, Yannakakis et Papadimitriou [2] qui décrit un algorithme d'énumération avec un délai polynomial pour lister tous les ensembles indépendants dans l'ordre lexicographique. Intuitivement, imposer un ordre lexicographique signifie que, étant donné un ordre total sur tous les sommets, on énumère d'abord les EMI contenant les plus petits sommets de l'ordre. Plus formellement, on attribue des nombres distincts de 1 à n aux sommets, ce qui nous donne un ensemble totalement ordonné, et à chaque EMI, on associe le tuple contenant tous les numéros de ses sommets dans l'ordre croissant. L'ordre lexicographique des EMI correspond à l'ordre lexicographique des tuples ainsi construits. La figure ci-dessous illustre les EIM (sommets en bleu) d'un graphe dans l'ordre lexicographique (de gauche à droite).



$\{1,3\}$ $\{1,4\}$ $\{2\}$

Le but du stage est d'étudier les questions laissées ouvertes par l'article de Johnson et al. [2]. Leur algorithme utilise un espace mémoire exponentiel dans la taille du graphe et, à ce jour, le problème est ouvert de savoir si cela est vraiment nécessaire. Cette question est restée ouverte pendant trente-cinq ans et pourrait donc s'avérer difficile à résoudre. Cependant, des cas spécifiques tels que la restriction à une famille spécifique de graphes pourrait être résolus. Ceci est déjà connu pour les arbres, pour lesquels il est possible de concevoir un algorithme d'énumération stable avec un délai polynomial et un espace linéaire en taille d'arbre [1]. On peut se demander s'il en va de même pour d'autres familles de graphes tels que les graphes chordaux, les graphes planaires, etc. Un autre problème ouvert à considérer est celui de l'énumération des EMI dans l'ordre lexicographique inverse pour des familles spécifiques de graphes (commencer par énumérer le dernier, puis l'avant-dernier et ainsi de suite). De manière quelque peu surprenante, il a été démontré que ce problème est difficile à résoudre pour les graphes généraux (pas d'algorithme à délai polynomial dans les hypothèses de complexité classiques). Par exemple, nous pourrions étudier si, pour certaines classes de graphes (tels que les arbres, les graphes chordaux, les graphes planaires, ...), il pourrait y avoir un algorithme d'énumération EMI à délai polynomial pour l'ordre lexicographique inverse.

[1] Y.H. Chang, J.S. Wang, and R.C.T. Lee. Generating all maximal independent sets on trees in lexicographic order. *Information Sciences*, 76(3):279–296, 1994.

[2] David S Johnson, Mihalis Yannakakis, and Christos H Papadimitriou. On generating all maximal independent sets. *Information Processing Letters*, 27(3):119–123, 1988.

[3] Shuji Tsukiyama, Mikio Ide, Hiromu Ariyoshi, and Isao Shirakawa. A new algorithm for generating all the maximal independent sets. *SIAM Journal on Computing*, 6(3):505–517, 1977.

Validation pour mise en ligne ECM : DM.