Rapport : “NODE- AND EDGE-DELETION NP-COMPLETE PROBLEMS”

N. Perrin

5 mai 2007

1 Introduction

Cet article de 1978, écrit par Mihalis Yannakakis, introduit plusieurs méthodes générales permettant de démontrer la NP-complétude d’une large classe de problèmes sur les graphes, et ouvre ainsi les portes d’une approche alternative à l’habituel traitement au “cas par cas” et tous les gadgets complexes qui vont avec. Cette classe de problèmes (les problèmes de suppression de noeuds et de suppression d’arêtes) est par ailleurs suffisamment naturelle pour englober de nombreux problèmes classiques, dont je donnerai quelques exemples par la suite.

Dans ce rapport je tenterai de suivre *grosso modo* le plan de l’article, de présenter clairement les résultats importants ainsi que la preuve principale. Par contre, dans un soucis de concision et afin de présenter l’essentiel, je ne mentionnerai pas certains résultats qui me semblent secondaires.

2 Définition des problèmes de suppression de noeuds et d’arêtes

2.1 Le problème de suppression de noeuds

A partir d’une propriété de graphes π, le problème de suppression des noeuds est défini de la manière suivante : il s’agit, à partir d’un graphe donné, de supprimer un minimum de noeuds afin de trouver un sous-graphe vérifiant π. On notera $\nu_\pi(G)$ le nombre minimum de noeuds supprimés.

On considèrera parfois le problème de suppression restreint à certains ensembles de graphes, comme les graphes planaires (qui ont la bonne idée de le rester après des suppressions de noeuds) ou bipartis (eux aussi), ou connexes (pour eux c'est différent, et une section entière de l'article est dédiée à ce cas).

En outre, nous parleront également du problème de suppression de noeuds pour les graphes *orientés*.

2.2 Le problème de suppression d'arêtes

Idem, mais on supprime des arêtes et non des noeuds.

2.3 Des exemples de problèmes classiques vus comme des problèmes de suppression

Si on choisit π="être une clique", on s'aperçoit que **MAX-CLIQUE se ramène à un problème de suppression de noeuds**.

De même, choisissons π="être un ensemble de noeuds indépendants".
Soit G un graphe, et soit V une couverture par noeuds de G (chaque arête a au moins une de ses extrémités dans V). $G - V$ est nécessairement un ensemble indépendant de noeuds. Réciproquement, si V est un ensemble de noeuds tel que $G - V$ est un ensemble de noeuds indépendants, alors V est une couverture par noeuds de G. Donc, une solution du problème de suppression de noeuds associé à π conduit à une couverture par noeuds de taille minimale (ce sont les noeuds supprimés) : **NODE COVER se ramène également à un problème de suppression de noeuds**.

Par la suite, on utilisera NODE COVER (qui est un problème NP-complet) pour démontrer la NP-complétude d'autres problèmes, et **on notera $\alpha_0(G)$ la taille minimale d'une couverture par noeuds de G**.

2.4 Restrictions sur π

Pour l'instant on ne considère que le problème de suppression de noeuds.

Voici les hypothèses que l'on supposera vérifiées par π par la suite :

- hérédité : si π est vraie pour un graphe G, alors elle l'est également pour tous ses sous-graphes induits par la suppression de certains noeuds.
- non triviale : π est vraie pour un noeud isolé, mais pas vraie pour tous les graphes.
- savoir si un graphe vérifie π ou non est dans P.
- intéressante : la taille des graphes vérifiant π n'est pas bornée (sinon les problèmes de suppression de noeuds et d'arêtes sont dans P).

3 Le premier et principal théorème

3.1 Sans restriction sur le domaine de départ

Théorème 1 :

Le problème de suppression de noeuds pour une propriété π héréditaire, non triviale, dans P et intéressante, est NP-complet.

Lemme 1 :
Soit toutes les cliques, soit toutes les anticliques vérifient π.

Démonstration : on se base sur le théorème de Ramsey, qui affirme que pour tout couple d'entier (m, n), il existe un nombre $r(m, n)$ tel que tout graphe ayant au moins $r(m, n)$ noeuds contient soit la clique de taille m soit l'anticlique de taille n.

Procédons par l'absurde et supposons alors qu'il existe une clique de taille m et une anticlique de taille n ne vérifiant pas π. Comme π est intéressante, il existe un graphe possédant plus de $r(m, n)$ noeuds et vérifiant π. Donc, comme π est héréditaire, soit la clique de taille m, soit l'anticlique de taille n vérifie π : contradiction.

Définition : on note $\overline{\pi}$ la propriété telle que : G vérifie $\overline{\pi}$ si et seulement si le complémentaire de G vérifie π.

Clairement, le problème de suppression de noeuds associé π se ramène en temps polynomial au problème de suppression de noeuds associé à $\overline{\pi}$, et réciproquement.

On peut donc désormais, quitte à choisir $\overline{\pi}$ à la place de π, supposer que toutes les anticliques (les ensembles de noeuds indépendants) vérifient π.

Quelques constructions préliminaires :

Soit G un graphe, et soient $G_1, \dots, G_t$ ses composantes connexes.

Pour tout G_i, et tout noeud c_i séparateur de G_i (i.e. après suppression de c_i il y a au moins 2 composantes connexes), on considère la séquence $< n_{i1}, \dots, n_{ij_i} >$ constituée des tailles *décroissantes* des composantes connexes obtenues. On choisit un point c_i minimisant cette séquence au sens lexicographique ; on note alors α_i la séquence obtenue. Si il n'existe aucun noeud séparateur de G_i, alors on pose $\alpha_i = |G_i|$, et le noeud séparateur c_i est un noeud quelconque de G_i.

En classant les α_i par ordre lexicographique *décroissant*, on obtient :

$$\beta_G =< \alpha_{i1}, \dots, \alpha_{it} >$$

En comparant lexicographiquement les séquences β_G, on obtient un *ordre total* parmi les graphes, que l'on note $<_R$.

On note J le plus petit graphe selon cet ordre ne pouvant pas être répété un nombre arbitrairement grand de fois sans invalider π, i.e. il existe un graphe constitué de copies indépendantes de J ne validant pas π. **On note $k \geq 1$ le nombre minimum de copies de J qu'il faut faire pour invalider π.**

On peut remarquer que J a au moins une composante de taille ≥ 2.

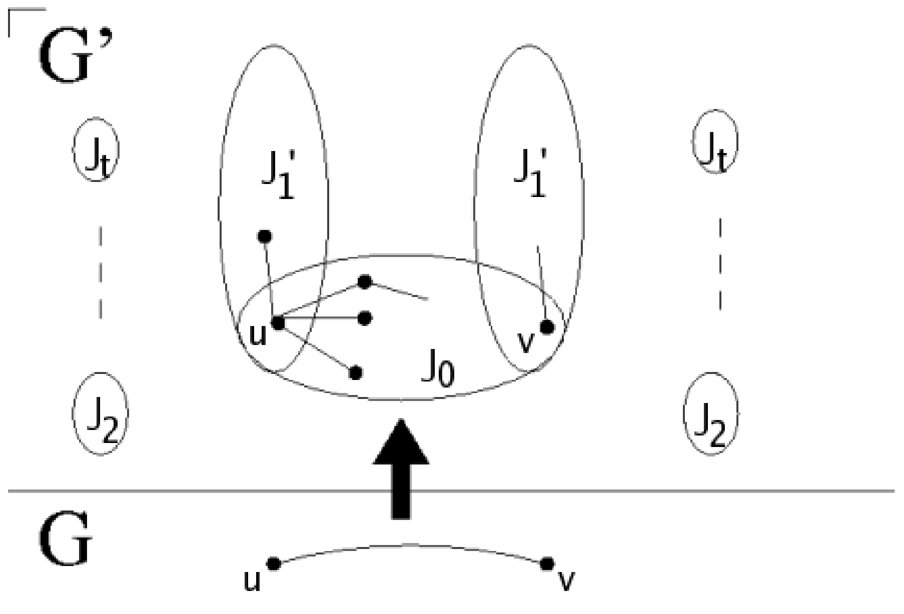
Désignons par $J_1, \ldots, J_t$ les composantes connexes de J triées selon leurs α_i, et soit c_1 le noeud séparateur de J_1 (celui qui conduit à α_1). Notons alors J^0 une plus grande composante de J_1 après suppression de c_1, et J_0 le sous-graphe obtenu à partir de J_1 en conservant seulement les noeuds de J^0 et c_1.

On note J_1' le graphe obtenu à partir de J_1 en supprimant tous les noeuds de J_0 sauf c_1.

On note également $J' = J_1' \cup J_2 \cup \cdots \cup J_3$, et soit d un noeud quelconque de J_0 (il existe car J_1 a au moins 2 noeuds).

Soit G un graphe, et G^* $n \times k$ copies indépendantes de G. **On construit** G' de la manière suivante : à tout noeud u de G^* on attache une copie de J' en identifiant c_1 à u. Toute arête (u, v) de G^* est remplaçée par une copie de J_0, attaché à u par c_1 et à v par d(ou bien l'inverse, cela n'a pas d'importance).

La figure suivante résume la construction :



—*figure 1*—

On va montrer le résultat suivant : $\alpha_0(G) \leq l \Leftrightarrow \nu_\pi(G') \leq n \times k \times l$**, ce qui prouvera bien que la NP-complétude de NODE COVER implique celle du problème de suppression de noeuds associé à** π**, et concluera la démonstration du théorème 1.**

1) $\Rightarrow$

Soit V une couverture par noeuds de G, avec $|V| \leq l$. On supprime V dans chaque copie de G dans G'. Ce qui reste de G est un ensemble de noeuds indépendants, donc les composantes connexes du résultat dans G' sont (voir figure précédente et retirer soit v soit u soit les deux) des composantes des graphes suivants :

- (a) soit une composante J_i de J autre que J_1
- (b) soit un graphe formé de J_1' et de plusieurs copies de J_0, en retirant c_1 ou d de chaque copie de J_0.
- (c) soit $J_1' - \{c_1\}$
- (d) soit $J_0 - \{c_1, d\}$

Remarquons que les graphes (c) et (d) sont des sous-graphes d'un graphe (b) (quel que soit le nombre (≥ 1) de copies de J_0).

Le graphe obtenu peut donc être vu comme un *sous-graphe d'une répétition du graphe* J^* composé de J_2, J_3, …, J_t, et de tous les graphes possibles du cas (b).

Etudions le graphe J^* :

considérons un sous-graphe de J^* issu du cas (b), et supposons-le connexe (on le note J_b^*). Pour le noeud séparateur c_1, la taille de la plus grande composante connexe issue de la suppression de c_1 est forcément *strictement* plus petite que celle $|J_0|$ (cf. définition de J_0). Ainsi, $\alpha_{J_b^*} <_{lex} \alpha_{J_1}$ (on ne choisit pas nécessairement le noeud séparateur c_i, mais le noeud choisi ne peut pas conduire à une composante plus grande).

Cette inégalité est également vérifiée pour chaque composante connexe de J_b^* si ce-dernier n'est pas connexe. De plus, les autres composantes connexes de J^* sont J_2, J_3, …, J_t. On en déduit finalement l'inégalité suivante :

$$\beta_{J^*} <_R \beta_J$$

Donc $J*$ vérifie nécessairement π, par définition de J. On a donc obtenu à partir de G' un graphe vérifiant π, en supprimant exactement $n \times k \times |V|$ noeuds.

Donc $\nu_\pi(G') \leq n \times k \times l$.

2) $\Leftarrow$

Supposons $\alpha_0(G) \geq l + 1$, et soit V une solution du problème de suppression de noeuds associé à π pour le graphe G' construit précédemment.

G' est constitué de $n \times k$ répétitions d'un même graphe issu de G. Dans $G' - V$ apparaissent $n \times k$ sous-graphes dont au plus $k-1$ contiennent J. Donc au moins $(n-1) \times k + 1$ ne contiennent pas J.

Notons G'_i l'un de ces sous-graphes ; et soit V_i l'ensemble des noeuds de G'_i appartenant initialement au graphe G (cf. construction de G', fig. 1), tels que : V contient soit un noeud de la copie de J' attachée à u (cf. fig. 1), soit un noeud d'une copie de J_0 qui a remplaçé une arête (v, u), avec $v < u$ (l'ordre sur les noeuds est arbitraire et n'a pas d'importance).

On peut facilement vérifier que $|V_i| \leq |V \cap G'_i|$ (où l'on a confondu G'_i et l'ensemble de ses noeuds). Supposons qu'il existe u et v dans G'_i formant initiallement une arête (u, v) de G, tels que $u \notin V_i$ et $v \notin V_i$; alors en réexaminant la figure 1, on peut s'apercevoir que $J(= J_0 \cup J'_1 \cup J_2 \cup \cdots \cup J_t)$ serait un sous-graphe de G_i, ce qui est exclu. On en déduit donc que V_i constitue une couverture par noeuds de G, et donc que $|V_i| \geq l + 1$.

D'où :

$$|V| \geq ((n-1) \times k + 1) \times (l + 1)$$

Et ainsi $\nu(G') > n \times k \times l$ (car $n > l + 1$), ce qui termine la démonstration.

3.2 Pour les graphes planaires

La démonstration s'adapte bien, et le résultat reste vrai si on restreint l'ensemble de départ aux graphes planaires.

3.3 Pour les graphes bipartis

Comme NODE COVER devient polynomial pour les graphes bipartis, on aurait pu pensé qu'il en serait de même pour les problèmes plus généraux de suppression de noeuds. En fait, le théorème de NP-complétude reste vrai, mais en supposant de plus que tous les ensembles **d'arêtes** indépendants vérifient π.

3.4 Pour les graphes orientés

Encore une fois, le théorème s'applique (sans restriction supplémentaire sur π).

3.5 Pour les graphes connexes

Le théorème principal est toujours vrai : le problème de suppression de noeuds pour une propriété π héréditaire sur les graphes (connexes) induits, non triviale, dans P et intéressante sur les graphes connexes, est NP-complet.

On prouve tout d'abord que π est validée par : soit toutes les cliques, soit toutes les étoiles, soit tous les chemins. La suite de la démonstration est l'étude de ces trois cas ; dans les deux premiers la preuve du théorème 1 s'adapte plus ou moins facilement, et dans le troisième l'auteur nous renvoie à l'un de ses articles antérieurs.

4 Suppression d'arêtes

Le problème de suppression d'arêtes ne se prête pas aussi bien aux généralisations que le problème de suppression de noeuds. L'article démontre que de nombreux problèmes particuliers de suppression d'arêtes sont NP-complets, mais ne parvient pas à extraire une méthode globale.

5 Conclusion

Pour un article de 11 pages, la quantité de théorèmes présents est assez remarquables. La plupart sont des déclinaisons du premier théorème, mais elles sont nombreuses, et je n'ai pas mentionné plusieurs cas étudiés (ex : propriétés déterminées par les blocs d'un graphe, i.e. vérifiées pour le graphe entier dès qu'elles le sont pour une de ses composantes 2-connexes, NP-complétude de problèmes généraux d'approximations, etc.).
En outre, non seulement les propriétés démontrées sont intéressantes, mais leurs démonstrations également.

En particulier, La preuve du théorème 1 est impressionnante car elle utilise une construction tout à fait inattendue, et conduit (presque par magie) à un résultat très puissant, car il unifie de nombreuses démonstrations de NP-complétude, qui pourtant sont souvent non triviales, et dont les ressemblances ne sautent pas aux yeux.

Depuis 1978, les travaux dans le même sens que cet article furent apparemment peu nombreux, on peut citer par exemple "The complexity of some edge deletion problems" de E.S. El-Mallah et C.J. Colbourn en 1997, ou encore d'autres articles de Yannakakis, mais il semble difficile de faire une liste très longue…Et pourtant il reste beaucoup à découvrir dans ce domaine : trouver des méthodes générales pour démontrer la NP-complétude de classes de problèmes aussi larges que possible est un challenge à la fois motivant et probablement très difficile à relever…!