Matroïdes et algorithme de Prim

Yves Lemaire

8 mai 2003

Résumé

L'exposé de Pierre Bejian hier[1] faisait apparaître l'algorithme de Kruskal comme un cas particulier de l'algorithme glouton standard sur les matroïdes; mais pas l'algorithme de Prim pour lequel des structures plus générales sont nécessaires (gloutonoïdes). On donne ici un algorithme glouton′ sur les matroïdes qui, sans sortir de la théorie des matroïdes, généralise l'algorithme de Prim et d'autres (au moins un autre ...) algorithmes gloutons classiques. Il est probable que cet algorithme est connu car sa preuve n'est pas très longue, mais il n'apparaît dans aucun des exposés sur les matroïdes que nous avons utilisé Pierre et moi.

1 Rappels et compléments sur les matroïdes

1.1 Indépendants

Dans toute la suite, on fixe un *matroïde* $\mathcal{M} = (E, \mathcal{I})$: E est un ensemble fini et $\mathcal{I}$ un ensemble de parties de E vérifiant les axiomes suivants:

(i) $\emptyset \in \mathcal{I}$;

(ii) $\forall I \in \mathcal{I}, \forall I',\ I' \subset I \Rightarrow I' \in \mathcal{I}$;

(iii) $\forall I, I' \in \mathcal{I},\ |I'| > |I| \Rightarrow \exists e \in I' \setminus I :\ I + e \in \mathcal{I}$.[2]

Les éléments de $\mathcal{I}$ sont dits *indépendants* ou *libres.*

1.2 Rang, bases, circuits

On déduit immédiatement des axiomes que, pour toute partie F de E, le cardinal de I, partie de F indépendante maximale, ne dépend pas du choix de I. Le cardinal commun des sous ensembles indépendants maximaux de F est appelé le *rang* de F et noté $\mathrm{rg}(F)$. En particulier le rang de E est appelé le rang du matroïde $\mathcal{M}$ et les parties de E indépendantes maximales sont appelées des *bases* de $\mathcal{M}$. On note $\mathcal{B}$ l'ensemble des bases de $\mathcal{M}$.

Un *circuit* est un sous ensemble de E dépendant (non indépendant) minimal. L'ensemble des circuits est noté $\mathcal{C}$.

1. Rencontre « Algorithmique et Programmation », CIRM (Luminy) 5 - 9 mai 2003

2. On adopte des notations simplificatrices telles que $A + e - e'$ pour $(A \cup \{e\}) \setminus \{e'\}$.

Proposition 1.1 *Si $I \in \mathcal{I}$ et $I+e \notin \mathcal{I}$ alors $I+e$ contient un circuit et un seul.*

Preuve Soit I un ensemble indépendant de cardinal minimum pour lequel la proposition est fausse. $I+e$ contient deux circuits distincts C_1 et C_2. On a $e \in C_1 \cap C_2$. Par minimalité de C_2 et C_1 respectivement, il existe $e_1 \in C_1 \setminus C_2$ et $e_2 \in C_2 \setminus C_1$.

$J = I + e - e_1 - e_2$ est indépendant. En effet, si J est dépendant, $I' = I - e_1$ est indépendant (car contenu dans I) et $I' + e$ contient deux circuits distincts: C_2 et un circuit contenu dans l'ensemble dépendant J; ceci contredit la minimalité de $|I|$.

Comme $|J| < |I|$, on a $J + e_i \in \mathcal{I}$ pour un $i \in \{1, 2\}$, ce qui est impossible car $J + e_i$ contient l'ensemble dépendant C_i. ■

1.3 Dual d'un matroïde

On pose $\mathcal{M}^* = (E, \mathcal{I}^*)$ avec $\mathcal{I}^* = \{J \subset E \mid \exists B \in \mathcal{B},\, J \subset E \setminus B\}$.

Proposition 1.2 *$\mathcal{M}^*$ est un matroïde.*

Preuve Les axiomes (i) et (ii) se vérifient sans peine. Avant de démontrer (iii), on montre d'abord que, pour toute partie F de E, les sous-ensembles de F éléments de $\mathcal{I}^*$ et maximaux pour ces deux propriétés ont tous le même cardinal. Soit J un tel ensemble. Comme J est dans $\mathcal{I}^*$, il existe une base B telle que $J \subset E \setminus B$. Par maximalité, on a $J = F \cap (E \setminus B)$. Posons $I = B \cap (E \setminus F)$. On a $I \in \mathcal{I}$ (car $I \subset B \in \mathcal{I}$) et $I \subset E \setminus F$. Montrons que I est maximal pour ces deux propriétés: Si ce n'est pas le cas, on peut trouver un $e \in E \setminus (F \cup B)$ vérifiant $I + e \in \mathcal{I}$. Si $p = |B \cap F| = 1$ alors $|I + e| = |B| = \operatorname{rg} E$ donc $B' = I + e$ est une base de $\mathcal{M}$ et $J \subsetneqq J' = F \subset E \subset B'$ ce qui contredit la maximalité de J. Si $p \geqslant 2$, d'après l'axiome (iii), il existe $e' \in B \setminus I$ tel que $I + e + e' \in \mathcal{I}$. Si $p = 2$, on trouve une contradiction en posant cette fois $B' = I + e + e'$ et $J' = F - e'$. Si $p \geqslant 3$, etc. On a donc montré que I était un sous-ensemble de $E \setminus F$ indépendant maximal. Par suite $|I| = \operatorname{rg}(E \setminus F)$ et $|J| = |F| - |B| + |I| = |F| - \operatorname{rg} E + \operatorname{rg}(E \setminus F)$ est bien indépendant du choix de J.

On est maintenant en mesure de démontrer (iii). Soient $J, J' \in \mathcal{I}^*$ avec $|J'| > |J|$. Posons $F = J \cup J'$. Soit $J'' \in \mathcal{I}^*$ maximal tel que $J' \subset J'' \subset F$. S'il n'existe aucun $e \in J' \setminus J$ pour lequel $J + e \in \mathcal{I}^*$, alors $J \subset F$, $J \in \mathcal{I}^*$ et J est maximal pour ces deux propriétés. D'après ce qui précède, $|J| = |J''|$, ce qui est faux. ■

Un ensemble indépendant (resp. une base, un circuit) du matroïde $\mathcal{M}^*$ est dit *coindépendant* (resp. une *cobase*, un *cocircuit*) pour $\mathcal{M}$. Le rang d'un ensemble $F \subset E$ au sens de $\mathcal{M}^*$ est appelé le *corang* de F et noté $\operatorname{rg}^* F$. Au passage, la démonstration de la propriété 1.2 montre que

$$\operatorname{rg}^* F = |F| - \operatorname{rg} E + \operatorname{rg}(E \setminus F)$$

et, en particulier $\operatorname{rg}^* E = |E| - \operatorname{rg} E$.

Proposition 1.3 *Si C est un circuit et C' un cocircuit, alors $|C \cap C'| \neq 1$.*

Preuve Supposons $C \cap C' = \{e\}$. $I' = C' - e$ est coindépendant donc $\mathrm{rg}(E \setminus I') = \mathrm{rg}\, E$; ainsi, comme $I = C - e$ est indépendant et contenu dans $E \setminus I'$, il existe une base B telle que $I \subset B \subset E \setminus I'$. Si $e \in B$ alors $C \subset B$ et sinon $C' \subset E \setminus B$; dans le premier cas, C est indépendant et, dans le deuxième, C' est coindépendant: on a une contradiction dans les deux cas. ■

2 L'algorithme glouton$'$

2.1 Cas général

On suppose que chaque élément e de E est affecté d'un poids $w(e) \in \mathbf{R}$ et on cherche une base de poids minimum; le poids $w(F)$ d'une partie de E étant la somme des poids de ses éléments.

Proposition 2.1 *Si I est une partie de E contenue strictement dans une base B de poids minimum, alors*
1) *il existe au moins un cocircuit disjoint de I;*
2) *si D est une réunion de cocircuits disjoints de I et si e est un élément de D de poids minimum, alors $I + e$ est contenu dans une base B' de poids minimum.*

Preuve 1) Comme l'ensemble $E \setminus I$ contient strictement la cobase $E \setminus B$, il est codépendant et il contient donc un cocircuit.
2) Soit D_0 un cocircuit contenu dans D et contenant e. Si $e \notin B$ (prendre $B' = B$ dans le cas $e \in B$), soit C le circuit contenu dans $B + e$ (proposition 1.1). Comme $e \in C \cap D_0$ et d'après la proposition 1.3, il existe $e' \in C \cap D_0 - e$. Alors $I + e \subset B' = B + e - e'$ qui est une base (car $B + e$ ne contient qu'un circuit C, donc B' n'en contient pas) de poids minimum (car le poids de e est $\leqslant$ à celui de e'). ■

Cette proposition prouve la validité de l'algorithme suivant

Algorithme 1 Algorithme glouton$'$ pour les matroïdes

soit *GLOUTON*$'(\mathcal{M}, w) =$
  soit $I \leftarrow \emptyset$ **dans**
  tant que I n'est pas une base **faire**
    soit D une réunion de cocircuits disjointe de I **dans**
    soit $e \in D$ de poids minimum **dans**
    $I \leftarrow I + e$
  renvoyer I

2.2 Cas des graphes

Soit $G = (S, A)$ un graphe (non orienté) connexe. On suppose à partir de maintenant que $E = A$ et $\mathcal{I} = \{I \subset A \,|\, I \text{ acyclique}\}$. D'après l'exposé de Pierre (exemple 3 p. 9), il s'agit bien d'un matroïde.

On adopte les notations suivantes.

Si $T, T' \subset S$, $\omega(T, T')$ désigne l'ensemble des arêtes de G dont une extrémité est dans T et l'autre dans T'. $\omega(T) = \omega(T, S \setminus T)$ est appelé un *cocycle* de G.

Si $S' \subset S$, on note $A_{S'} = \omega(S', S')$ et $G_{S'} = (S', A_{S'})$. On dira que S' est connexe si $G_{S'}$ l'est.

Si $A' \subset A$, on note $S_{A'}$ l'ensemble des sommets qui sont extrémité d'au moins une arête de A', et $G_{A'} = (S_{A'}, A')$.

Rappelons qu'une partie B de A est une base de $\mathcal{M}$ ($B \in \mathcal{B}$) si B est un arbre couvrant de G.

Proposition 2.2 *Pour tout ensemble d'arêtes $D \subset A$, les conditions suivantes sont équivalentes.*

(1) D est un cocircuit.

(2) $(\forall B \in \mathcal{B},\ D \cap B \neq \emptyset)$ et $(\forall e \in D,\ \exists B \in \mathcal{B},\ D \cap B = \{e\})$.

(3) $D = \omega(T)$ où T et $S \setminus T$ sont connexes et non vides.

Preuve L'équivalence (1) $\Leftrightarrow$ (2) provient immédiatement des définitions.

Faisons l'hypothèse (3): $D = \omega(T)$ où T et $S \setminus T$ sont connexes et non vides; et vérifions (2). Comme un arbre couvrant B doit connecter les sommets de T et ceux de $S \setminus T$, il doit contenir au moins une arête de $\omega(T) = D$ et donc $D \cap B \neq \emptyset$. D'autre part, pour toute arête $e \in D$, on peut construire un arbre couvrant B tel que $D \cap B = \{e\}$ en posant $B = B_1 \cup \{e\} \cup B_2$ avec B_1 et B_2 arbres couvrants de G_T et $G_{S \setminus T}$ respectivement.

Réciproquement on vérifie (3) avec l'hypothèse (2): Pour tout $e \in D$, notons B_e un arbre couvrant vérifiant la condition $D \cap B_e = \{e\}$. Fixons une arête $a \in D$ et notons T et $T' = S \setminus T$ les deux composantes connexes du graphe $(S, B_a - a)$. Vérifions que $D = \omega(T)$: $\omega(T) \subset D$ car s'il existe $e \in \omega(T) \setminus D$, alors $(B_a \cap A_T) \cup \{e\} \cup (B_a \cap A_{T'})$ est un arbre couvrant disjoint de D, ce qui est impossible. Réciproquement, si $e \in D$, B_e contient un élément $e' \in \omega(T)$, donc $e' \in D \cap B_e = \{e\}$ et $e = e' \in \omega(T)$. ∎

Proposition 2.3 *Si un ensemble $T \subset S$ de sommets est connexe, alors le cocycle $\omega(T)$ est une réunion de cocircuits.*

Preuve Si T ou $S \setminus T$ est vide, c'est trivial (réunion sur le vide). Sinon $\omega(T) = \omega(T_1) \cup \cdots \cup \omega(T_p)$ où $T_1, \ldots, T_p$ sont les composantes connexes de $S \setminus T$. Pour tout i, T_i est connexe et la connexité de T implique celle de $S \setminus T_i = T \cup \left(\bigcup_{j \neq i} T_j\right)$. D'après la proposition précédente, $\omega(T_i)$ est un cocircuit. ∎

En appliquant l'algorithme glouton$'$ au matroïde graphique associé à G, on obtient l'algorithme 2. Sa correction se déduit de la correction de l'algorithme glouton$'$ en vérifiant que I est toujours connexe, ce qui prouve, à l'aide de la proposition 2.3, que $\omega(S_I)$ est bien une réunion de cocircuits.

On reconnaît l'algorithme de Prim.

Algorithme 2 Algorithme glouton$'$ pour les graphes

soit $GLOUTON\_GRAPH'(G, w) =$
  soit s un sommet fixé de G **dans**
  soit $I \leftarrow \emptyset$ **dans**
  tant que I n'est pas un arbre couvrant **faire**
    soit $D = \omega(S_I)$ **dans** (* on convient ici que $S_\emptyset = \{s\}$ *)
    soit $e \in D$ de poids minimum **dans**
    $I \leftarrow I + e$
  renvoyer I

2.3 Un autre cas

Chercher une base de poids minimum étant équivalent, par complémentation, à trouver une cobase de poids maximum, chaque algorithme de recherche d'une base de poids minimum possède une version duale. Par exemple, une version duale de glouton$'$ appliquée à un matroïde graphique est donnée par l'algorithme 3 (on a posé $K = A \setminus I$).

Algorithme 3 Algorithme dual de glouton$'$ pour les graphes

soit $GLOUTON\_DUAL\_GRAPH'(G, w) =$
  soit $K \leftarrow A$ **dans**
  tant que K n'est pas un arbre couvrant **faire**
    soit D un cycle contenu dans K **dans**
    soit $e \in D$ de poids maximum **dans**
    $K \leftarrow K - e$
  renvoyer K