Matroïde et algorithme glouton

Référence : A guide to algorithm design, Yves Robert

Définition 1 (Matroïde, Whitney, 1935). *Le couple (S, I) est un matroïde si S est un ensemble à n éléments, $I \subset P(S)$ et si :*

1. *$X \in I \Rightarrow (\forall Y \subset X, X \in I)$ (hérédité) ;*
2. *$(A, B \in I, |A| < |B|) \Rightarrow \exists x \in B \backslash A, A \cup \{x\} \in I$ (propriété d'échange).*

*Tout ensemble $X \in I$ est dit **indépendant**.*

Exemple de matroïde. cet exemple nous permettra de montrer que l'algorithme de Kruskal est optimal.

Théorème 1. *Soit $G = (V, E)$ un graphe, $S = E$ et $I = \{A \subset E | A \text{ ne contient pas de cycle}\}$. Le couple (S, I) est un matroïde.*

Démonstration. Un ensemble $X \in I$ est indépendant ss'il forme une forêt du graphe G.

1. Un sous-ensemble d'une forêt est encore une forêt. En effet, si $X \subset I$ et il existe $Y \subset X$ contenant un cycle, alors X contient aussi ce cycle, ce qui est absurde.
2. Soient A et B deux forêts de G avec $|A| < |B|$. Chaque sommet de G est contenu dans un arbre de A (si le sommet est isolé, alors il est dans un arbre à un sommet). Ainsi, A contient $|V| - |A|$ arbres (resp. $|V| - |B|$ pour B). En effet, dès qu'on ajoute une arête, on fusionne deux arbres et le nombre total d'arbre diminue de un.

   Ainsi, B contient moins d'arbre que A, et donc il existe un arbre T de B qui n'est pas inclut dans un arbre de A (si ce n'était pas le cas, on aurait $|V| - |B| \geqslant |V| - |A|$ car chaque arbre de B est dans un arbre de A). Autrement dit, il existe deux sommets $u, v \in V$ tels que u et v sont dans T mais pas dans le même arbre de A. Il existe donc une arête (x, y) sur le chemin de T allant de u à v qui n'est pas dans A. On a alors $A \cup \{(x, y)\}$ qui est toujours une forêt.

□

Ensemble indépendant maximal.

Définition 2. *Soit $F \in I$. $x \notin F$ est une **extension** de F si $F \cup \{x\} \in I$. Un ensemble indépendant est **maximal** s'il n'a aucune extension.*

Lemme 1. *Tous les ensembles indépendants maximaux d'un matroïde ont le même cardinal.*

Démonstration. Par l'absurde, on utilise la propriété d'échange et on contredit la maximalité du plus petit des deux ensembles indépendants. □

Définition 3 (Matroïde pondéré). *Un matroïde pondéré est un matroïde (S, I) muni d'une fonction de poids $w : S \to \mathbb{N}$. Le poids d'un ensemble $X \subset S$ est défini par $w(X) = \sum_{x \in X} w(x)$.*

Algorithme 1 : GloutonMatroide(S,I, w)

Données : matroïde (S, I) avec $S = \{s_1, ..., s_n\}$ trié par poids décroissants
Résultat : Ensemble indépendant A de poids maximum.
$A \leftarrow \varnothing$;
pour $i = 1...n$ **faire**
  si $A \cup \{s_i\} \in I$ **alors**
    $A \leftarrow A \cup \{s_i\}$;

Algorithme glouton sur un matroïde pondéré. On cherche à trouver un ensemble indépendant de poids maximum. On peut montrer que dans le cas d'un matroïde, l'algorithme glouton est optimal.

Théorème 2. *L'algorithme GloutonMatroïde renvoie une solution optimale.*

Démonstration. Soit $X = \{x_1, ..., x_m\}$ la solution retourné par l'algorithme glouton. On a alors $w(x_1) \geqslant ... \geqslant w(x_m)$. Soit $Y = \{y_1, ..., y_m\}$ une solution optimale avec $w(y_1) \geqslant ...w(y_m)$. On montre que pour tout $1 \leqslant i \leqslant m$, $w(x_i) \geqslant w(y_i)$. L'optimalité de la solution en découle directement.

Si ce n'était pas le cas, prenons k le plus petit indice tel que $w(x_k) < w(y_k)$. On remarque que $k > 1$ car $\{x_1\}$ est le singleton indépendant de poids maximum. On regarde alors $A = \{x_1, ..., x_{k-1}\}$ et $B = \{y_1, ..., y_k\}$. puisque $|B| = |A| + 1$, on peut appliquer la propriété d'échange et il existe $1 \leqslant i \leqslant k$ tel que $A \cup \{y_i\}$ est indépendant et $y_i \notin A$. On a alors $w(y_i) \geqslant w(y_k) > w(x_k)$. L'algorithme glouton aurait alors choisit y_i avant x_k (il suffit de prendre le sous-ensemble de $A \cup \{y_i\}$ des éléments ayant un poids inférieur à $w(y_j)$, qui est indépendant par hérédité). $\square$

Optimalité de l'algorithme de Kruskal. Le théorème précédent nous permet de montrer que l'algorithme glouton qui construit un arbre couvrant de poids minimum est optimal. En effet, les arêtes sont triés par poids croissant et on choisit de manière gloutonne la prochaine arête qui ne crée pas de cycle.