

Mini-Projet

Problème de Couplage (Matching)

Nathanael Bayard — L3 DLMI — Algorithmique des Graphes

12 décembre 2018

1 Présentation du Problème et Théorème de Hall

On peut définir le problème de façon abstraite, de la manière suivante :

Δ Définition. On appelle *problème de couplage* la donnée d'un triplet $P = (X, Y, \phi)$ tel que :

- (i) X et Y soient des ensembles finis quelconques non vides.
- (ii) ϕ est une application de X vers $\mathcal{P}(Y)$

Un élément $y \in Y$ est dit compatible avec un élément $x \in X$ si $y \in \phi(x)$. L'ensemble des paires d'éléments de $X \times Y$ liés par cette relation de compatibilité sera noté E , et on a alors $(x, y) \in E$ si et seulement si y est compatible avec x par rapport au problème P .

On définit alors une solution au problème de la manière suivante :

Δ Définition. On appelle *couplage* pour le problème de couplage $P = (X, Y, \phi)$ toute application $f : A \subseteq X \rightarrow Y$ qui soit injective, et qui respecte ϕ dans le sens que :

$$\forall x \in A, \quad f(x) \in \phi(x)$$

De façon équivalente, on peut dire que le graphe de f (au sens d'un graphe d'application) est inclus dans E .

On appelle *couplage parfait* un couplage f dont l'ensemble de définition soit X tout entier. L'existence d'un couplage parfait entraîne évidemment que nécessairement $|X| \leq |Y|$. En particulier, si l'on suppose dès le départ que, dans P , $|X| \geq |Y|$, alors un couplage parfait pour P est simplement bijection entre X et Y dont le graphe d'application soit inclus dans E .

On appelle *couplage maximal* pour le problème P , tout couplage f de P tel que pour tout autre couplage g , $|\text{Im } f| \geq |\text{Im } g|$. Dit autrement, puisque f et g sont injectives, f est un couplage maximal s'il n'existe pas de couplage g de P dont l'ensemble de définition soit plus grand (en cardinal) que l'ensemble de définition de f . On peut aussi le formuler en disant que f est maximal au sens du cardinal de son graphe d'application $\Gamma_f = \{(x, f(x)) \mid x \in A\}$.

On appellera *ordre* d'un couplage $f : A \subseteq X \rightarrow Y$, l'entier $|A|$. Un couplage maximal est donc en particulier un couplage maximal au sens de l'ordre.

On dira qu'un problème de couplage P est *régulier* lorsque $|X| = |Y|$.

Tout couplage parfait est bien sûr maximal, mais la réciproque n'est à priori pas vraie. Le Théorème de Hall donne une condition nécessaire et suffisante sur P pour qu'il admette un couplage parfait. Avant de l'énoncer et de le prouver, on aura besoin de la définition suivante.

Δ Définition. Soit $\phi : X \rightarrow \mathcal{P}(Y)$ une application quelconque, X et Y étant des ensembles quelconques. On dira que $\gamma : X \rightarrow \mathcal{P}(Y)$ est incluse dans ϕ si

$$\forall x \in X, \gamma(x) \subseteq \phi(x)$$

On notera alors $\gamma \subseteq \phi$.

⊤ Théorème. (Hall) Soit $\phi : X \rightarrow \mathcal{P}(Y)$, et $P = (X, Y, \phi)$ le problème de couplage associé. Alors P admet un couplage parfait si et seulement si ϕ respecte la condition suivante, dite *condition de couplage* :

$$\forall A \subseteq X, \quad |A| \leq \left| \bigcup_{x \in A} \phi(x) \right| \quad (1)$$

Preuve.

Prouvons la nécessité de la condition. Remarquons en premier lieu que si $f : A \rightarrow Y$ est un couplage pour P , alors comme $f(x) \in \phi(x)$ pour tout $x \in A$, on a l'inclusion :

$$\forall B \subseteq A, \quad f(B) \subseteq \bigcup_{x \in B} \phi(x)$$

En particulier si P admet un couplage parfait $f : A = X \rightarrow Y$, l'inclusion ci-dessus entraîne bien la condition énoncée par le Théorème de Hall :

$$\forall B \subseteq X, \quad |B| = |f(B)| \leq \left| \bigcup_{x \in B} \phi(x) \right|$$

La première égalité venant de l'injectivité de f .

Prouvons maintenant que la condition est suffisante pour obtenir l'existence d'un couplage parfait. Soit $\gamma \subseteq \phi$ respectant la condition (1), et minimale au sens que pour tout $\xi \subset \gamma$, ξ ne respecte pas la condition (1).

Soit $z \in X$. Puisque γ respecte la condition de couplage, on a l'inégalité $|\gamma(z)| \geq |\{z\}| = 1$. Montrons que $|\gamma(z)| = 1$ par l'absurde.

Supposons que $\{u, v\} \subseteq \gamma(z)$, avec $u \neq v$. On définit $\lambda_u \subset \gamma$ égal à γ sur tout $X \setminus \{z\}$, et tel que $\lambda_u(z) = \gamma(z) \setminus \{u\}$. On définit symétriquement $\lambda_v \subset \gamma$ tel que $\lambda_v(z) = \gamma(z) \setminus \{v\}$ et $\lambda_v = \gamma$ partout ailleurs. Comme γ est minimal au sens de l'inclusion par rapport à la condition de couplage, ni λ_u ni λ_v ne peuvent vérifier celle-ci, et alors :

$$\begin{aligned} \exists B_u \subseteq X, \quad |B_u| > |W_u| \quad \text{avec} \quad W_u &:= \bigcup_{x \in B_u} \lambda_u(x) \\ \exists B_v \subseteq X, \quad |B_v| > |W_v| \quad \text{avec} \quad W_v &:= \bigcup_{x \in B_v} \lambda_v(x) \end{aligned}$$

Comme λ_u et λ_v sont égales à γ sur tout $X \setminus \{z\}$, et comme γ respecte elle la condition de couplage, λ_u et λ_v doivent respecter la condition de couplage pour tout $E \subseteq X \setminus \{z\}$. On en déduit que nécessairement z appartient à B_u et aussi à B_v . Par suite, $v \in \lambda_u(z) \subseteq W_u$ et de même $u \in W_v$. Définissons $B := (B_u \cap B_v) \setminus \{z\}$, $N := \bigcup_{x \in B} \gamma(x)$, et $W := W_u \cap W_v$.

Soit $a \in B$. Alors $a \in B_u$ et comme $\gamma(a) = \lambda_u(a)$ (car $a \neq z$),

$$\gamma(a) \subseteq \bigcup_{x \in B_u} \lambda_u(x) = W_u$$

D'où $N \subseteq W_u$, et de même $N \subseteq W_v$ pour les mêmes raisons : finalement, $N \subseteq W$.

Nous aurons aussi besoin du résultat suivant.

$$\begin{aligned} W_u \cup W_v &= \left(\bigcup_{x \in B_u} \lambda_u(x) \right) \cup \left(\bigcup_{x \in B_v} \lambda_v(x) \right) \\ &= \left(\bigcup_{x \in (B_u \cup B_v) \setminus \{z\}} \gamma(x) \right) \cup (\lambda_u(z) \cup \lambda_v(z)) = \bigcup_{x \in B_u \cup B_v} \gamma(x) \end{aligned}$$

La deuxième égalité se justifie car $\lambda_u(x) = \lambda_v(x) = \gamma(x)$ pour tout $x \neq z$; la troisième vient du fait que $\lambda_u(z) \cup \lambda_v(z) = \gamma(z)$. Comme γ respecte la condition de couplage, on en déduit que

$$|W_u \cup W_v| = \left| \bigcup_{x \in B_u \cup B_v} \gamma(x) \right| \geq |B_u \cup B_v| \quad (2)$$

Finalement :

$$\begin{aligned} |N| &= \left| \bigcup_{x \in B} \gamma(x) \right| \leq |W| = |W_u \cap W_v| \quad \text{car} \quad N \subseteq W \\ &= |W_u| + |W_v| - |W_u \cup W_v| \\ &\leq (|B_u| - 1) + (|B_v| - 1) - |B_u \cup B_v| \\ &\quad \text{puisque} \quad |W_u| < |B_u| \quad \text{et} \quad |W_v| < |B_v| \quad \text{et} \quad (2) \\ &= |B_u \cap B_v| - 2 \\ &= |(B_u \cap B_v) \setminus \{z\}| - 1 \\ &= |B| - 1 \end{aligned}$$

D'où finalement $\left| \bigcup_{x \in B} \gamma(x) \right| < |B|$: absurde puisque γ respecte la condition de couplage.

On vient donc de prouver que $\forall z \in X, |\gamma(z)| = 1$.

On peut définir une application $f : X \rightarrow Y$ qui à $x \in X$ lui associe l'unique $y \in Y$ tel que $\gamma(x) = \{y\}$. Reste à démontrer l'injectivité de f : si $f(x) = f(x') = y$, c'est que $\gamma(x) = \gamma(x') = \{y\}$, et alors comme γ respecte la condition de couplage :

$$|\{x, x'\}| \leq |\gamma(x) \cup \gamma(x')| = |\{y\}| = 1$$

D'où $x = x'$, et f est bien injective, c'est donc bien un couplage parfait pour P , ce qui achève cette démonstration. \square

De nombreux problèmes concrets peuvent être modélisés par un problème de couplage. L'objectif est souvent de construire une bijection $f : X \rightarrow Y$ avec des contraintes sur la valeur que peut prendre $f(x)$, contraintes dépendant de x . Il est à remarquer que le problème de couplage parfait est en un certain sens symétrique par rapport à X et Y .

Δ Définition. Soit $P = (X, Y, \phi)$ un problème de couplage. Alors on définit son *symétrique* P^{-1} comme le problème de couplage $P^{-1} := (Y, X, \gamma)$, avec :

$$\forall y \in Y, \quad \gamma(y) := \{x \in X ; y \in \phi(x)\}$$

Lorsque $|X| = |Y|$, c'est-à-dire lorsque P est régulier alors P^{-1} l'est bien sûr aussi, et la recherche d'un couplage parfait pour $P = (X, Y, \phi)$ se résume dans ce cas à trouver une bijection entre X et Y respectant les contraintes exercées par ϕ . On en déduit que cela est équivalent à la recherche d'un couplage parfait pour $P^{-1} = (Y, X, \gamma)$, puisqu'en effet ϕ et γ exercent les mêmes contraintes sur une éventuelle bijection entre X et Y . Les problèmes P et P^{-1} peuvent alors être vus comme équivalents de ce point de vue-là : f est un couplage parfait pour P si et seulement si f^{-1} est un couplage parfait pour P^{-1} .

Pour des cardinaux de X et Y pas forcément égaux, tout ce que l'on peut dire c'est qu'à tout couplage f de P , on peut associer un couplage f' de P^{-1} de même ordre, en vertu qu'une injection $f : A \subseteq X \rightarrow Y$ est aussi une bijection $g : A \rightarrow \text{Im } f \subseteq Y$, et donc $g^{-1} = f'$ est une injection de $\text{Im } f \subseteq Y$ dans X , donc un couplage pour P^{-1} , et l'ordre de f' est $|\text{Im } f| = |A|$, qui est aussi l'ordre de f .

Par suite, on en déduit tout de même le fait important que tout couplage maximal de P entraîne l'existence d'un couplage maximal de P^{-1} de même ordre, et ce, que P soit régulier ou non.

2 Modélisation par un Graphe

Ce type de problème se prête naturellement bien à la modélisation par un graphe (au sens de la théorie des graphes).

Δ Définition. On appelle *graphe du problème de couplage* $P = (X, Y, \phi)$, le graphe non orienté, simple, biparti $G := (V, E)$, avec $V := X \cup Y$ (on supposera ici X et Y disjoints) et

$$E := \{\{x, y\} \in X \times Y \mid \text{tel que } y \in \phi(x)\}$$

Un *couplage* du graphe G correspondra à un sous-ensemble $M \subseteq E$ tel qu'il existe un couplage $f : A \subseteq X \rightarrow Y$ de P , tel que le graphe d'application Γ_f de f soit M (en identifiant 2-uplets et doubletons), c'est-à-dire :

$$M = \Gamma_f = \{\{x, f(x)\} \mid \text{pour } x \in A\}$$

En termes plus graphiques, du fait de l'injection de f , M correspond à un ensemble d'arêtes de G deux à deux disjoints par sommets. En particulier, chaque sommet de G admet au plus un voisin dans le sous-graphe (V, M) .

Un *couplage parfait* de G correspondra similairement à un ensemble $M = \Gamma_f \subseteq E$ avec $f : X \rightarrow Y$ un couplage parfait pour P , avec la condition supplémentaire que tout sommet de G soit recouvert par M (et pas seulement ceux dans X) : cela implique que f soit bijective, donc aussi que $|X| = |Y|$.

Cette condition supplémentaire se justifie du fait que les graphes des problèmes P et P^{-1} sont exactement égaux, donc tout couplage parfait de l'un doit logiquement être un couplage parfait pour l'autre, ce qui force à ce que tout couplage parfait de P soit un couplage parfait de P^{-1} , d'où l'égalité nécessaire des cardinaux de X et de Y .

Un *couplage maximal* pour G correspondra à un ensemble $M \subseteq E$ avec $f : X \rightarrow Y$ un couplage maximal pour P . Cela correspond comme dit précédemment automatiquement à un couplage maximal pour P^{-1} de même ordre, donc ici il n'y a pas d'ambiguïté par rapport à la régularité de P . Cela correspond aussi à un couplage M pour G de cardinal maximal, puisque le cardinal de M est exactement l'ordre du couplage f de P qui lui correspond.

3 Algorithme de Hopcroft-Karp

Le problème qui consiste à trouver un couplage maximal pour un graphe biparti simple $G = (V = X \cup Y, E)$ peut être résolu en temps polynômial grâce à l'algorithme de Hopcroft-Karp. Plus exactement, la complexité dans le pire cas est en $O(m\sqrt{n})$ avec $n := |V|$ et $m := |E|$.

3.1 L'idée

Δ Définition. Soit $G = (V = X \cup Y, E)$ le graphe d'un problème de couplage, et $M \subseteq E$ un couplage de G . On dit que $v \in V$ est libre par rapport à M s'il est isolé dans (V, M) , c'est-à-dire s'il n'est pas l'extrémité d'une arête dans M .

On dira que C est un *chemin d'augmentation* de M dans G , lorsque C est un chemin de G , simple et élémentaire, de longueur non nulle, et non cyclique, dont les deux sommets d'extrémité sont libres par rapport à M , et dont la suite des arêtes traversées est en alternance dans M et dans $E \setminus M$.

Tout sommet faisant partie d'un chemin d'augmentation de M est, ou bien une des extrémités libres du chemin, ou bien n'est pas libre, puisqu'en effet il est dans ce cas l'extrémité d'exactlyement une arête dans M . Comme exemple trivial, un chemin d'augmentation "minimaliste" peut très bien être une unique arête dans $E \setminus M$ reliant deux sommets libres.

Du fait qu'un chemin d'augmentation C alterne la traversée d'arêtes dans M et dans $E \setminus M$, et comme il commence et se termine par des arêtes dans $E \setminus M$ (puisque ses extrémités sont libres pour M), il est alors nécessaire que la longueur du chemin C soit impaire. Par bipartisme de G , on en déduit que les extrémités de C sont nécessairement un élément de X et un élément de Y .

Si C est un chemin d'augmentation pour M dans G , en identifiant C avec l'ensemble des arêtes qu'il visite, on peut définir l'ensemble $M' := M \oplus C$ (où \oplus désigne la différence symétrique entre ensembles), et remarquer qu'alors M' est encore un couplage de G : tout sommet de (V, M') est en effet encore de degré inférieur ou égal à 1.

Par ailleurs, M' contient une arête de plus que le couplage M . À partir du couplage trivial $M_0 = \emptyset$, on peut ainsi construire une suite finie (M_0, \dots, M_s) de couplages de G qui soit strictement croissante pour le cardinal : l'élément M_{i+1} est défini comme la différence symétrique de M_i avec un chemin d'augmentation pour M_i dans G . La suite s'arrête lorsqu'il ne reste plus de chemins d'augmentation à trouver pour le dernier élément de la suite.

Afin de prouver que le dernier couplage M_s de la suite, obtenu selon cette méthode des chemins d'augmentation, est un couplage maximal de G , (c'est-à-dire que pour tout autre couplage M de G , on aurait $|M_s| \geq |M|$), nous avons le théorème suivant.

⊤ Théorème. (Berge 1957) Un couplage M est maximal si et seulement si G ne contient pas de chemins d'augmentation pour M .

Preuve. Comme montré précédemment, s'il existe un chemin d'augmentation C pour M , alors $C \oplus M$ est un couplage contenant une arête de plus que M , ce qui implique que M ne peut être maximal.

Réciproquement, supposons M non maximal. Soit M^* un couplage maximal dans G . On doit donc avoir $c := |M^*| - |M| > 0$. Définissons $S := M \oplus M^*$. S est composé de toutes les arêtes $S_1 := S \cap M$ de M qui ne sont pas dans M^* , et de toutes les arêtes $S_2 := S \cap M^*$ de M^* qui ne sont pas dans M . On doit avoir $|S_2| = |S_1| + c$.

Soit $x \in V$. Dans le sous-graphe (V, S) , si x est l'extrémité d'une arête de M , alors il ne peut être l'extrémité d'une autre arête de M , puisque M est un couplage. De la même façon, x est l'extrémité d'au plus une arête de M^* . Par suite, tous les sommets de (V, S) sont de degré inférieurs ou égaux à 2.

On en déduit que S contient un ensemble de cycles et de chemins élémentaires maximaux de G (que l'on ne peut pas prolonger dans S) deux à deux sommet-disjoints, et où il y a toujours alternance entre une arête de M et une arête de M^* dans chacun de ces cycles et chemins.

Les cycles sont donc tous de longueur paire, et contiendront aussi autant d'arêtes de M que d'arêtes de M^* . De même, tout chemin maximal de S de longueur paire contiendra autant d'arêtes de M que d'arêtes de M^* .

Comme S_2 est supposé être plus grand de cardinal que S_1 , cela implique forcément qu'il existe dans S au moins un chemin C ayant plus d'arêtes dans M^* que dans M . Un tel chemin est alors de longueur impaire, commençant et se terminant par deux arêtes de $M^* \setminus M$, donc par deux sommets libres par rapport à M , car si les extrémités de C n'étaient pas libre pour M , le chemin C pourrait être prolongé d'au moins une arête, et ne serait donc pas maximal (ou bien ce serait un cycle, ce qui a déjà été exclus), c'est donc précisément un chemin d'augmentation pour M dans G , ce qui achève la preuve de la réciproque.

Précisons un point qui sera utile pour prouver le théorème qui va suivre : il n'existe aucun chemin maximal dans S contenant plus d'arêtes de M^* que de M : si un tel chemin existait, ce serait un chemin d'augmentation pour M^* , ce qui est impossible du fait que M^* est un couplage maximal pour G . \square

⊤ **Théorème.** (Hopcroft-Karp) On reprend les hypothèses et notations du théorème précédent. Alors, il existe, dans $S = M^* \oplus M$, exactement $c = |M^*| - |M| > 0$ chemins d'augmentation pour M qui soient deux-à-deux sommet-disjoints.

Preuve. Il suffit de voir que chaque chemin d'augmentation pour M dans S contient exactement une arête de plus provenant de M^* , que d'arêtes provenant de M . Il existe donc nécessairement c chemins d'augmentation disjoints dans S pour justifier que M^* contienne c arêtes de plus que M , tout cela en tenant compte que, comme montré précédemment, les chemins d'augmentation pour M sont les seuls chemins maximaux de S qui influent sur la différence de taille entre M et M^* . \square

Par application du théorème précédent, M_s étant le dernier couplage de la suite, comme aucun chemin d'augmentation n'existe pour celui-ci, M_s est donc nécessairement un couplage maximal. Cela achève la preuve de la validité de la méthode de résolution du problème de couplage maximal par chemins d'augmentation.

3.2 L'algorithme

On améliore l'algorithme esquissé précédemment en cherchant, à chaque étape, un ensemble maximal K de chemins d'augmentation deux à deux sommet-disjoints et de longueur aussi petite que possible. Cela permet d'augmenter la taille du couplage à chaque étape de manière beaucoup plus efficace, afin d'obtenir la complexité attendue.

Il est à noter que l'ensemble K est “maximal” au sens qu'il n'existe pas d'autre chemins d'augmentation à lui ajouter qui soit “sommet-disjoint” des chemins déjà présents dans K . Il n'est pas nécessaire que K soit *maximum* au sens d'être l'ensemble de chemins d'augmentation deux à deux sommet-disjoints de cardinal le plus grand que l'on puisse trouver, ce qui serait autrement plus long et difficile à construire.

Algorithme Hopcroft-Karp.

Entrée : Graphe biparti simple non orienté $G = (V, E) = (X \cup Y, E)$.

Sortie : Couplage maximal M de G .

```

M ← ∅
K ← {C1, ..., Cn}
// K étant défini comme ci-dessus comme un ensemble maximal
// de chemins d'augmentation pour M les plus courts possibles et 2 à 2 disjoints
Faire {
    K ← {C1, ..., Cn}
    M ← M ⊕ ⋃C ∈ K C
} Tant que K ≠ ∅.
Retourner M.
```

3.3 Construction de K

À chaque étape de l'algorithme, on a besoin de créer un ensemble K de chemins d'augmentation qui soit sommet-disjoints, maximal, et aussi court que possible.

On commence par parcourir en largeur, et sans revisites, le graphe G , en commençant par traiter tous les sommets libres par rapport à M dans X . Le parcours en largeur est ensuite effectué de telle sorte que, pour un sommet visité quelconque, si l'arête empruntée pour le visiter est dans M , alors toutes les arêtes à traverser à partir de ce sommet devront être dans $E \setminus M$, et vice-versa. Comme les arêtes du niveau de profondeur de départ sont toutes dans $E \setminus M$ (puisque les sommets de départ du parcours en largeur sont tous des sommets libres par rapport à M), chaque niveau de profondeur est donc composé exclusivement ou bien d'arêtes de M , ou bien d'arêtes de $E \setminus M$.

On arrête prématurément le parcours en largeur au premier niveau de profondeur qui visite un sommet libre de Y . En particulier, du fait de la spécificité de ce parcours en largeur, cela équivaut à dire que l'on s'arrête à la profondeur où l'on rencontre le plus court chemin d'augmentation pour M qui existe dans G . Il est important de préciser que l'on arrête le parcours à un certain niveau de profondeur, et non pas immédiatement après avoir visité le sommet libre de Y en question : s'il reste des sommets à visiter au niveau en question, on termine de les visiter normalement. En particulier, il est possible, et même souhaitable, de pouvoir trouver plusieurs sommets libres de Y au dernier niveau de profondeur du parcours. On notera pour la suite k le niveau de profondeur en question : la longueur de la plus longue branche dans la forêt de visite de ce parcours en largeur, c'est-à-dire le niveau où l'on a trouvé le premier sommet libre de Y .

Une fois ce parcours en largeur (prématurément) arrêté, on lance séquentiellement plusieurs parcours, cette fois en profondeur, à partir de chaque sommet libre de X , et ce en étant guidé par le résultat du

parcours en largeur : les arêtes traversées doivent alterner entre M et $E \setminus M$, et chaque nouveau sommet visité doit augmenter le niveau de profondeur atteint, au sens du parcours en largeur. Un tel parcours en profondeur depuis un sommet libre $x \in X$ spécifique se termine dès que l'on dépasse la profondeur k limite, ou bien dès que l'on tombe sur un sommet libre de Y , c'est-à-dire dès que l'on parvient à trouver un chemin d'augmentation de M qui part du sommet $x \in X$ en question.

Tous ces parcours en profondeur effectués pour chaque $x \in X$ libres nous permettent de récupérer un ensemble maximal K de chemins d'augmentation les plus courts possibles (de longueur inférieure ou égale à k) et surtout qui soient deux à deux disjoints : il suffit pour cela de refuser toute revisite d'un sommet, même par un autre parcours en profondeur. En effet, lors d'un parcours profondeur à partir d'un sommet libre $x \in X$, une fois un sommet z visité par ce parcours, ou bien il fera partie d'un chemin d'augmentation démarrant à x et atteignant Y avec une longueur inférieure à k (et alors il ne doit plus être revisité dans la recherche d'un autre chemin d'augmentation sommet-disjoint du premier), ou bien z ne fera pas partie d'un chemin d'augmentation partant de x , ce qui implique qu'il n'existe pas de chemin d'augmentation passant par z ayant une longueur inférieure à k : cela est garanti par le fait que le parcours en profondeur respecte les niveaux de profondeur résultant du parcours largeur précédent.

Du fait que l'on refuse les revisites, la complexité de l'ensemble de ces parcours en profondeur ne dépasse pas celle, standard, d'un parcours complet, en largeur ou profondeur, du graphe G , qui vaut ici $O(n + m)$.

3.4 Complexité

L'algorithme est grossièrement composé d'une unique boucle **faire ... tant que**. Chaque passage dans la boucle sera appelé une *étape* de l'algorithme.

Chaque étape de l'algorithme requiert, pour construire K , un parcours en largeur de G , et un parcours en profondeur "généralisé" (partant de plus d'un sommet) de G . Quant à la différence symétrique entre M et chaque chemin trouvé de K , on peut le réaliser à la volée durant le parcours en profondeur, sans donc modifier la complexité d'une étape.

Une seule de ces étapes peut donc être réalisée en temps $O(2(n + m)) = O(n + m)$. Comme le graphe est biparti toutes les arêtes partent de X et atterrissent dans Y donc on a finalement

$$m \leq |X| \times |Y| \leq \left(\frac{n}{2}\right)^2 = \frac{n^2}{4}$$

puisque $n = |X| + |Y|$ et pour tous $x, y \in \mathbb{R}$,

$$4xy \leq (x + y)^2 = x^2 + y^2 + 2xy \iff 0 \leq x^2 + y^2 - 2xy = (x - y)^2$$

Donc $m = O(n^2)$ et la complexité d'une étape de l'algorithme est donc $O(m)$ (puisque m l'emporte sur n dans la somme $n + m$ dans le pire cas).

Il peut être montré qu'à chaque étape, la longueur minimale k pour laquelle on puisse trouver un chemin d'augmentation, augmente d'au moins 1. Ce résultat est non trivial, on peut en trouver une preuve dans [Meena Mahajan, "Matchings in Graphs", 2010].

Par suite, après l'étape \sqrt{n} , tous les chemins d'augmentation que l'on peut trouver seront de longueur au moins égale à \sqrt{n} . Soit M le couplage obtenu à ce moment-là de l'algorithme, M^* un couplage maximal pour G , et $S := M \oplus M^*$. S contient entre autres un certain nombre de chemins d'augmentation de M deux-à-deux sommet-disjoints (*c.f.* le théorème de Berge), donc nécessairement tous de longueur supérieure ou égale à \sqrt{n} . Comme ils sont sommet-disjoints, chaque chemin d'augmentation passe par un peu plus de \sqrt{n} sommets qu'il ne partage pas avec les autres chemins dans S , et alors il ne peut donc y avoir plus de \sqrt{n} chemins d'augmentation dans S , puisque le nombre total de sommets disponibles est limité à n .

D'après le théorème de Hopcroft-Karp, le nombre de chemins d'augmentation dans S est égal à la différence $|M^*| - |M|$. D'où ici, $|M^*| - |M| \leq \sqrt{n}$. Comme chaque étape de l'algorithme trouve au moins un chemin d'augmentation, à chaque étape M croît en cardinal par au moins 1. Par suite, à l'étape $2\sqrt{n}$ (ou avant si l'algorithme termine plus tôt), on a forcément $|M| = |M^*|$, et M est donc le couplage maximal recherché.

Par suite, l'algorithme de Hopcroft-Karp se termine donc toujours après $O(2\sqrt{n}) = O(\sqrt{n})$ étapes, chaque étape ayant une complexité de $O(m)$, d'où la complexité finale de Hopcroft-Karp qui est bien $O(m\sqrt{n}) = O(n^2\sqrt{n}) = O(n^{2.5})$.

3.5 brouillon

— remplacer maximal par maximum