Chapitre 1
Introduction à la théorie des graphes

Les graphes font partie des structures abstraites les plus fondamentales en mathématique discrète. Ils modélisent les situations où des objets sont en relation, et ont donc de nombreuses applications en ingénierie : informatique (réseaux, automates, intelligence artificielle, ...), sciences physiques (transport, optimisation, liaisons chimiques, ...), sciences de l'information (communication, trafic, ...), biologie (épidémiologie, ...), sciences sociales, linguistique, etc.

Le but de cette premier chapitre est d'introduire le vocabulaire de base de la théorie des graphes et de présenter des algorithmes de résolution de problèmes classiques liés aux graphes.

1 Définitions et vocabulaire

1.1 Graphes

Définition 1

*Un **graphe** $G = (S, A)$ est la donnée*

- *d'un ensemble fini S dont les éléments sont appelés les **sommets** de G*
- *et d'un ensemble de paires de sommets, c'est-à-dire*

$$A \subset \Big\{ \{s, s'\} \;\Big|\; (s, s') \in S^2 \text{ et } s \neq s' \Big\} = \Big\{ \alpha \in \mathscr{P}(S) \mid \mathrm{card}(\alpha) = 2 \Big\}$$

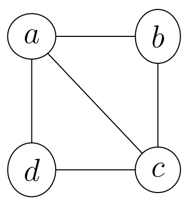
  *dont les éléments sont appelés les **arêtes** de G.*

*On dit que chaque arête $\{s, s'\}$ **relie** les sommets s et s', ou que les sommets s et s' sont les **extrémités** de l'arête $\{s, s'\}$.*

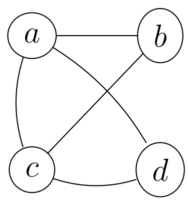
Exemple. Le graphe ayant pour ensemble de sommets $S = \{a, b, c, d\}$ et ensemble d'arêtes

$$A = \Big\{ \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{c, d\} \Big\}$$

peut être représenté schématiquement de la manière suivante :



Dans ce type de schéma, la position des sommets et la forme des arêtes sont sans importance, seule la présence ou non d'une arête entre chaque paire de sommets compte. Ainsi, le schéma suivant représente le même graphe que le précédent :



Remarque. Dans un graphe, on dit aussi **graphe simple**, on n'autorise ni les **boucles** (c'est-à-dire les arêtes reliant un sommet à lui même), ni les **arêtes multiples** (c'est-à-dire plusieurs arêtes différentes reliant la même paire de sommets). Sinon, on parle de **multigraphe** qui est une notion plus générale.

Définition 2

*Le nombre $n = \mathrm{card}(S)$ de sommets d'un graphe $G = (S, A)$ est appelé son **ordre**. Et le nombre $\mathrm{card}(A)$ d'arêtes est appelé sa **taille**.*

Propriété 1

Soit S un ensemble fixé de sommets avec $n = \mathrm{card}(S)$. Il y a exactement $2^{n(n-1)/2}$ façons possibles de construire un ensemble d'arêtes A afin que $G = (S, A)$ forme un graphe. Autrement dit, il y a $2^{n(n-1)/2}$ graphes d'ordre n.

Démonstration. Le nombre total de paires de sommets possibles est égal à :

$$\mathrm{card}\Big(\Big\{\alpha \in \mathscr{P}(S) \mid \mathrm{card}(\alpha) = 2\Big\}\Big) = \binom{n}{2} = \frac{n(n-1)}{2}.$$

Puisque l'ensemble des arêtes A est une partie de l'ensemble des paires de sommets, il y a donc $2^{n(n-1)/2}$ choix possibles pour A.

$\square$

Exemple. Pour $n = 10$ sommets, on obtient $2^{45} \approx 3,5 \times 10^{13}$ graphes d'ordre 10. C'est beaucoup.

Définition 3

Soit $n \in \mathbb{N}^*$.

- *Le **graphe complet**, noté $\boldsymbol{K_n}$, est défini par :*

$$S = \{1, 2, \dots, n\} \quad \textit{et} \quad A = \Big\{\alpha \in \mathscr{P}(S) \mid \mathrm{card}(\alpha) = 2\Big\}.$$

  Sa taille est égale à $\frac{n(n-1)}{2}$.

- *Le **graphe linéaire d'ordre $\boldsymbol{n}$**, noté $\boldsymbol{L_n}$, est défini par :*

$$S = \{1, 2, \dots, n\} \quad \textit{et} \quad A = \Big\{\{k, k+1\} \mid k \in [\![1, n-1]\!]\Big\}.$$

  Sa taille est égale à $n - 1$.

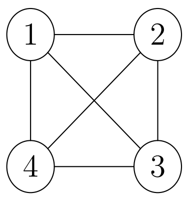
- *Le **graphe cyclique d'ordre $\boldsymbol{n}$**, noté $\boldsymbol{C_n}$, est défini par :*

$$S = \{1, 2, \dots, n\} \quad \textit{et} \quad A = \Big\{\{k, k+1\} \mid k \in [\![1, n-1]\!]\Big\} \cup \Big\{\{1, n\}\Big\}.$$

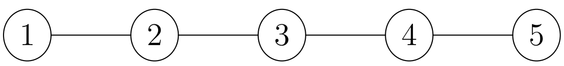
  Sa taille est égale à n.

Exemple.

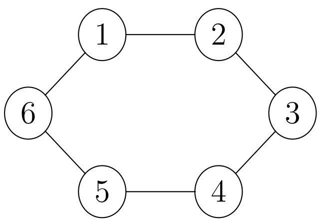
- Le graphe complet K_4 :



- Le graphe linéaire L_5 :



- Le graphe cyclique C_6 :



1.2 Graphes enrichis

Définition 4

*Un **graphe orienté** $G = (S, A)$ est la donnée*

- *d'un ensemble fini S dont les éléments sont appelés les **sommets** de G*
- *et d'un ensemble de couples de sommets, c'est-à-dire*

$$A \subset \left\{ (s, s') \,\middle|\, (s, s') \in S^2 \text{ et } s \neq s' \right\}$$

*dont les éléments sont appelés les **arcs** de G.*

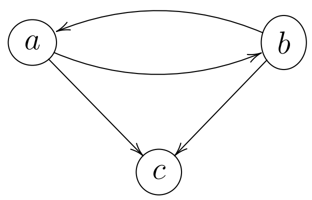
*On dit que chaque arc (s, s') **part** du sommet s et **arrive** au sommet s', ou que les sommets s et s' sont respectivement le **départ** et **l'arrivée** de l'arc (s, s').*

Rappel. Les couples, et donc les arcs, sont ordonnées (car $(s, s') \neq (s', s)$ si $s \neq s'$) alors que les paires, et donc les arêtes, ne le sont pas (car $\{s, s'\} = \{s', s\}$).

Exemple. Le graphe orienté ayant pour ensemble de sommets $S = \{a, b, c\}$ et ensemble d'arcs

$$A = \left\{ (a, b), (a, c), (b, a), (b, c) \right\}$$

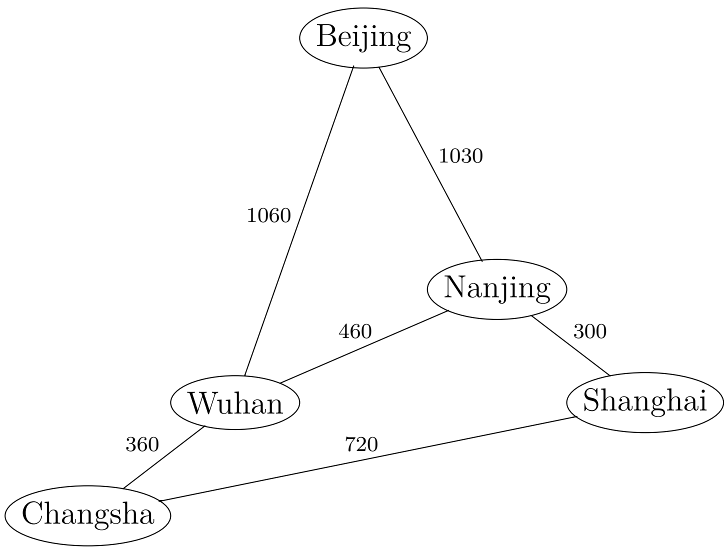
peut être représenté schématiquement de la manière suivante :



Définition 5

*Un **graphe pondéré** est un graphe $G = (S, A)$ muni d'une **fonction de poids** $f : A \to \mathbb{R}$. Ainsi, à chaque arête $\{s, s'\}$ est associée une valeur réelle $f(\{s, s'\})$ appelée son **poids**.*

Exemple. On peut modéliser le réseau chinois de train à grande vitesse par un graphe pondéré par les distances en km entre les villes. Par exemple pour Beijing, Changsha, Nanjing, Shanghai et Wuhan :



Ainsi l'arête reliant Shanghai à Nanjing modélise l'existence d'une ligne à grande vitesse, et son poids de 300km correspond à la distance entre ces deux villes.

Remarque. Il existe aussi des **graphes orientés pondérés**.

1.3 Adjacence et degré

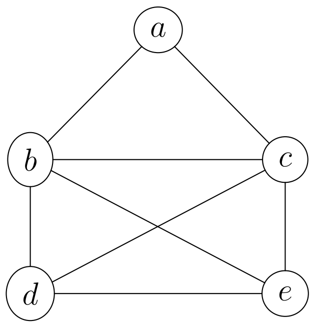
Définition 6

*Deux sommets s et s' d'un graphe sont dits **adjacents** ou **voisins** si $\{s, s'\}$ est une arête.*

Définition 7

*Le **degré** d'un sommet s d'un graphe est le nombre de sommets adjacents à s, donc aussi le nombre d'arêtes dont s est une extrémité. On le note $\mathbf{deg}(\boldsymbol{s})$. Lorsque le degré d'un sommet est nul, on dit que ce sommet est **isolé**.*

Exemple. Dans le graphe suivant, a et d ne sont pas adjacents, alors que b et e le sont.



D'autre part, $\deg(a) = 2$, $\deg(b) = \deg(c) = 4$ et $\deg(d) = \deg(e) = 3$.

Propriété 2

Soit un graphe $G = (S, A)$. Alors :

— $\forall s \in S,\ \deg(s) \leqslant \mathrm{card}(S) - 1$,

— *et* $\displaystyle\sum_{s \in S} \deg(s) = 2\mathrm{card}(A)$.

Démonstration. Le premier point vient du fait qu'un graphe n'a ni boucle ni arête multiple. Pour le deuxième, il suffit de compter de deux façons différentes le nombre total d'extrémités d'arêtes :

— c'est égal à la somme de tous les degrés par définition,

— et c'est aussi égal au double de la taille du graphe puisque chaque arête a deux extrémités.

□

Remarque. On retrouve que $\mathrm{card}(A) = \dfrac{1}{2}\displaystyle\sum_{s \in S} \deg(s) \leqslant \dfrac{n}{2}(n-1) = \binom{n}{2}$ en notant $n = \mathrm{card}(S)$.

Corollaire 1 (*Lemme des poignées de mains*)

Dans un graphe, le nombre de sommets de degré impair est pair. En particulier, lors d'une réunion, il y a toujours un nombre pair de personnes qui serrent un nombre impair de mains.

Démonstration. Puisque $\sum_{s \in S} \deg(s)$ est paire comme étant égale à $2\mathrm{card}(A)$, les contributions impaires dans la somme sont en nombre pair. On peut modéliser la réunion par un graphe $G = (S, A)$ où chaque sommet représente une personne et chaque arrête représente une poignée de mains. D'où le résultat.

□

1.4 Chemin et connexité

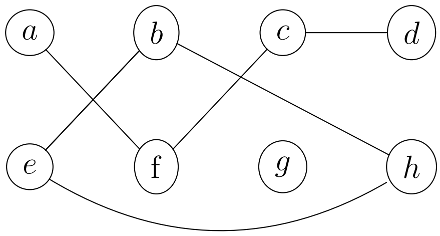
Définition 8

*Dans un graphe, un **chemin** est de façon équivalente :*

- *une suite finie de sommets telle que deux sommets successifs de la suite sont adjacents,*
- *une suite finie d'arêtes telle que deux arêtes successives de la suite ont une extrémité commune.*

*Deux sommets s et s' sont dits **reliés par un chemin** s'il existe au moins un chemin dont le premier sommet est s et le dernier est s'.*

Exemple. Dans le graphe suivant, a et b ne sont pas reliés par un chemin, alors que a et d sont reliés par le chemin (a, f, c, d). g étant isolé, il n'est relié par un chemin à aucun autre sommet.



Propriété 3

La propriété d'être relié par un chemin est une relation d'équivalence, c'est-à-dire qu'elle est :

- *réflexive (un sommet est toujours relié à lui-même par un chemin sans arête),*
- *symétrique (si s est relié à s' par un chemin alors s' est relié à s par le chemin décrit en sens inverse),*
- *et transitive (si s est relié à s' par un chemin et s' est relié à s'' par un chemin alors s est relié à s'' par le chemin obtenu en concaténant les deux chemins précédents).*

Définition 9

*La **composante connexe** d'un sommet s d'un graphe est l'ensemble des sommets reliés par un chemin à s (autrement dit sa classe d'équivalence pour la propriété d'être relié par un chemin). Un graphe est dit **connexe** s'il admet une unique composante connexe, c'est-à-dire lorsque toute paire de sommets peut être reliée par au moins un chemin.*

Exemple. Le graphe de l'exemple précédent n'est pas connexe, ces composantes connexes sont $\{a, c, d, \mathrm{f}\}$, $\{b, e, h\}$ et $\{g\}$. D'autre part, les graphes K_n, L_n et C_n (voir la définition 3) sont connexes.

Remarque. La composante connexe d'un sommet isolé est réduite à un singleton. En particulier, tout graphe connexe contenant au moins deux sommets n'a pas de sommet isolé.

Propriété 4

Si un graphe $G = (S, A)$ est connexe alors $\mathrm{card}(A) \geqslant \mathrm{card}(S) - 1$.

Démonstration. On raisonne par récurrence sur l'ordre $n = \mathrm{card}(S)$.

Initialisation. Le résultat est évident pour $n = 1$ (et même $n = 2$).

Hérédité. On suppose que la propriété est vérifiée pour tous les graphes connexes d'ordre $n \in \mathbb{N}^*$. Soit $G = (S, A)$ un graphe connexe d'ordre $n + 1$. La connexité assure qu'il n'y a pas de sommets isolés, donc que $\forall s \in S,\ \deg(s) \geqslant 1$. Il y a alors deux cas :

- 1[er] cas : $\exists s \in S,\ \deg(s) = 1$. Alors le graphe $G' = (S', A')$ obtenu en supprimant s (c'est-à-dire $S' = S \setminus \{s\}$) et l'arête dont il est l'extrémité (donc $\mathrm{card}(A') = \mathrm{card}(A) - 1$) est connexe d'ordre n. D'après l'hypothèse de récurrence, $\mathrm{card}(A') \geqslant n - 1$ donc $\mathrm{card}(A) \geqslant n$.
- 2[e] cas : $\forall s \in S,\ \deg(s) \geqslant 2$. Alors on a d'après la propriété 2 :

$$2\mathrm{card}(A) = \sum_{s \in S} \deg(s) \geqslant \sum_{s \in S} 2 = 2\mathrm{card}(S) \quad \text{donc} \quad \mathrm{card}(A) \geqslant n.$$

□

Remarque. La réciproque est fausse. Par exemple, si on ajoute un sommet isolé à K_3, on obtient un graphe non connexe d'ordre 4 et de taille $3 = 4 - 1$. Le calcul de l'ordre et de la taille ne suffit donc pas à caractériser la connexité.

Définition 10

Dans un graphe, un chemin est dit :
- ***élémentaire*** *s'il ne passe pas deux fois par un même sommet,*
- ***simple*** *s'il ne passe pas deux fois par une même arête.*

La ***longueur*** *d'un chemin est le nombre de ses arêtes.*

Remarque. Un chemin élémentaire est nécessairement simple mais la réciproque est fausse. Par exemple dans K_4, le chemin $(1, 2, 3, 1, 4)$ est simple mais pas élémentaire.

Propriété 5

Si deux sommets d'un graphe sont dans la même composante connexe, alors ils peuvent être reliés par un chemin élémentaire. Plus précisément, tout chemin de longueur minimale qui les relie est un chemin élémentaire.

Démonstration. Soit $(s_0 = s, s_1, \ldots, s_{\ell-1}, s_\ell = s')$ un chemin de longueur ℓ minimale qui relie deux sommets s et s'. Par l'absurde, on suppose que ce chemin n'est pas élémentaire. Alors il existe deux indices $1 \leqslant i < j \leqslant \ell$ tels que $s_i = s_j$. On remarque que :

$$\left(s_0 = s, s_1, \ldots, s_{i-1}, s_i = s_j, s_{j+1}, \ldots, s_\ell = s'\right)$$

est un chemin reliant s et s' de longueur $\ell - (j - i) < k$ ce qui contredit la minimalité de ℓ.

□

Définition 11

La ***distance*** *entre deux sommets de la même composante connexe d'un graphe est la longueur minimale des chemins qui les relie.*

1.5 Cycles

Définition 12

Dans un graphe, un ***cycle*** *est un chemin simple de longueur non nulle reliant un sommet à lui-même. Un graphe est dit* ***acyclique*** *s'il ne contient aucun cycle.*

Remarque. Par définition, tout cycle a au minimum 3 arêtes et est donc de longueur supérieure ou égale à 3.

Exemple. L_n est acyclique. Mais K_n et C_n contiennent des cycles dès que $n \geqslant 3$. De plus, K_n contient des cycles de toute longueur comprise entre 3 et n alors que les cycles de C_n sont tous de longueur n.

Propriété 6

Si un graphe a tous ses sommets de degré supérieur ou égal à 2 alors il contient au moins un cycle.

Démonstration. On fixe arbitrairement un sommet noté s_0 puis on construit par récurrence une liste de sommets $(s_0, s_1, s_2, \ldots, s_k)$ deux à deux différents, en choisissant pour chaque s_{i+1} n'importe quel sommet adjacent à s_i qui n'est pas encore dans la liste. La construction s'arrête lorsque tous les voisins de s_k sont déjà dans la liste. Puisque $\deg(s_k) \geqslant 2$, la liste contient nécessairement un autre voisin de s_k que s_{k-1}. Autrement dit, il existe un indice $j < k - 1$ tel que s_j est adjacent à s_k. On en déduit que $(s_k, s_j, s_{j+1}, \ldots, s_{k-1}, s_k)$ est un chemin simple d'au moins 3 arêtes reliant s_k à lui-même, donc un cycle.

□

Corollaire 2

Si un graphe $G = (S, A)$ est acyclique alors $\text{card}(A) \leqslant \text{card}(S) - 1$.

Démonstration. On raisonne par récurrence sur l'ordre $n = \text{card}(S)$.

Initialisation. Le résultat est évident pour $n = 1$ (et même $n = 2$).

Hérédité. On suppose que la propriété est vérifiée pour tous les graphes acycliques d'ordre $n \in \mathbb{N}^*$. Soit $G = (S, A)$ un graphe acyclique d'ordre $n + 1$. D'après la propriété 6, il existe au moins un sommet s de degré 0 ou 1. Alors le graphe $G' = (S', A')$ obtenu en supprimant s (c'est-à-dire $S' = S \setminus \{s\}$) et l'éventuelle arête dont il est l'extrémité (donc $\text{card}(A') \geqslant \text{card}(A) - 1$) est acyclique d'ordre n. D'après l'hypothèse de récurrence, $\text{card}(A') \leqslant n - 1$ donc $\text{card}(A) \leqslant n$.

□

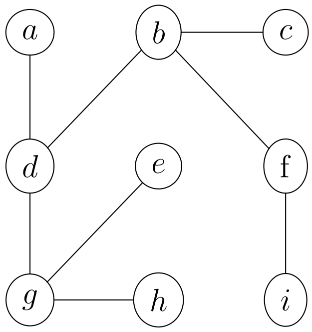
Remarque. La réciproque est fausse. Par exemple, si on ajoute un sommet isolé à K_3, on obtient un graphe d'ordre 4 et de taille $3 = 4 - 1$ qui contient un cycle de longueur 3. Le calcul de l'ordre et de la taille ne suffit donc pas à caractériser l'acyclicité.

1.6 Arbres

Définition 13

*Un **arbre** est un graphe connexe acyclique. Les sommets de degré 1 d'un arbre sont appelés des **feuilles**, les autres sommets (de degré supérieur ou égal à 2) sont appelés des **nœuds**.*

Exemple. Le graphe suivant est un arbre. a, c, e, h et i sont des feuilles alors que b, d, f et g sont des nœuds.



Théorème 1

Soit $G = (S, A)$ un graphe. Les propositions suivantes sont équivalentes :

(i) *G est un arbre ;*

(ii) *G est connexe et $\text{card}(A) = \text{card}(S) - 1$;*

(iii) *G est connexe minimal : il n'est plus connexe si on lui supprime n'importe quelle arête ;*

(iv) *G est acyclique et $\text{card}(A) = \text{card}(S) - 1$;*

(v) *G est acyclique maximal : il contient au moins un cycle si on lui ajoute n'importe quelle arête ;*

(vi) *pour toute paire de sommets de G, il existe un unique chemin élémentaire qui les relie.*

Démonstration. On a déjà (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) et (i) $\Longrightarrow$ (iv) $\Longrightarrow$ (v) d'après la propriété 4 et le corollaire 2. Pour (iii) $\Longrightarrow$ (i), on raisonne par l'absurde : si G contient un cycle alors il reste connexe si on lui supprime une arête de ce cycle. De même pour (v) $\Longrightarrow$ (i) : si G n'est pas connexe alors il reste acyclique si on lui ajoute une arête reliant deux sommets situés dans des composantes connexes différentes. Enfin (i) $\Longleftrightarrow$ (vi) est évident par définition.

□

Corollaire 3

Tout arbre non réduit à un sommet isolé a au moins deux feuilles.

Démonstration. Soit $G = (S, A)$ un arbre d'ordre $n = \text{card}(S) \geqslant 2$. On note f le nombre de feuilles de G. On sait d'après la propriété 2 et le théorème 1 que :

$$2\text{card}(A) = \sum_{s \in S} \deg(s) \quad \text{et} \quad \text{card}(A) = n - 1.$$

Par conséquent :

$$2(n-1) = \sum_{\substack{s \in S \\ \deg(s)=1}} \deg(s) + \sum_{\substack{s \in S \\ \deg(s)\geqslant 2}} \deg(s) \geqslant \mathrm{f} + 2(n - \mathrm{f}) \quad \text{donc} \quad \mathrm{f} \geqslant 2.$$

□

Définition 14

*Un **arbre enraciné** est un arbre dont un sommet porte le nom de **racine**. Cette notion de racine permet de munir l'arbre d'une fonction $h : S \to \mathbb{N}$ qui associe à chaque sommet s sa **hauteur** $h(s)$ égale à la longueur de l'unique chemin élémentaire reliant s à la racine, c'est-à-dire à la distance entre s et la racine (la hauteur de la racine est nulle). On définit alors :*

— *les **ancêtres** de s : les sommets de l'unique chemin élémentaire reliant s à la racine,*

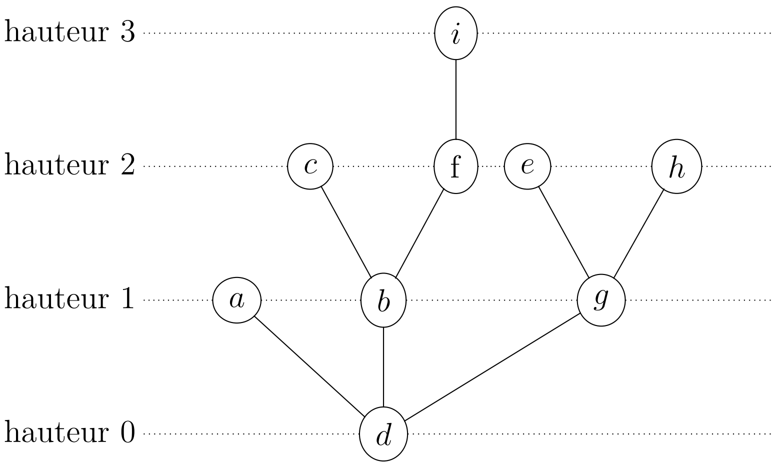
— *le **parent** de s : l'ancêtre de hauteur $h(s) - 1$,*

— *les **descendants** de s : les sommets pour lesquels s est un ancêtre,*

— *les **enfants** de s : les descendants de hauteur $h(s) + 1$.*

Remarque. Enraciner un arbre revient donc à l'orienter de la racine vers ses feuilles.

Exemple. Dans l'exemple précédent, si on choisit d comme racine, alors on peut représenter schématiquement l'arbre enraciné de la manière suivante :



Définition 15

*Une **forêt** est un graphe acyclique, autrement dit un graphe dont toutes les composantes connexes sont des arbres.*

Propriété 7

Soit $G = (S, A)$ un graphe. On note p son nombre de composantes connexes. Alors G est une forêt si et seulement si $\text{card}(A) = \text{card}(S) - p$.

Démonstration. Les composantes connexes de G définissent des graphes notés $\Big(G_i = (S_i, A_i)\Big)_{i \in [\![1,p]\!]}$. Puisque chaque composante connexe est connexe, on a d'après la propriété 4 :

$$\text{card}(A) = \sum_{i=1}^{p} \text{card}(A_i) \geqslant \sum_{i=1}^{p} \Big(\text{card}(S_i) - 1\Big) = \text{card}(S) - p.$$

De plus, il y a égalité si et seulement si $\forall i \in [\![1,p]\!],\ \text{card}(A_i) = \text{card}(S_i) - 1$, c'est-à-dire si et seulement si chaque composante connexe G_i est acyclique d'après le théorème 1. D'où le résultat.

□

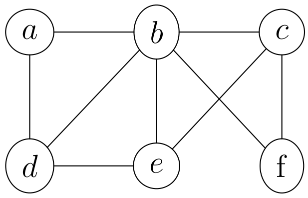
1.7 Sous-graphes

Définition 16

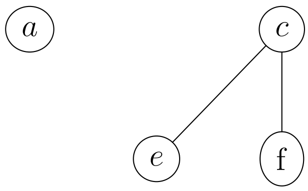
*Un **sous-graphe** d'un graphe $G = (S, A)$ est un graphe $G' = (S', A')$ tel que $S' \subset S$ et $A' \subset A$. Un sous-graphe est dit :*

- ***induit** si $A' = \left\{\{s, s'\} \in A \mid (s, s') \in S'^2\right\}$, c'est-à-dire s'il est obtenu en supprimant des sommets de G et seulement les arêtes dont les sommets supprimés sont des extrémités,*
- ***couvrant** si $V = V'$, c'est-à-dire s'il est obtenu en supprimant seulement des arêtes de G.*

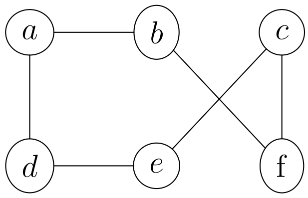
Exemple. On considère le graphe suivant :



Le sous-graphe induit par les sommets a, c, e et f est :



Ce n'est pas un graphe couvrant. Par contre, le sous-graphe suivant est couvrant (mais pas induit) :



Remarque. Tout graphe d'ordre n peut être vu comme un sous-graphe couvrant du graphe complet K_n.

Propriété 8

*Un graphe est connexe si et seulement s'il admet un **arbre couvrant**, c'est-à-dire un sous-graphe couvrant qui est un arbre.*

Démonstration. Si un graphe admet un arbre couvrant alors il est évidemment connexe par définition. Réciproquement, on considère un graphe connexe pour lequel on va construire un arbre couvrant par récurrence. Si c'est un arbre, c'est fini. Sinon, il contient des cycles. On choisit arbitrairement un cycle auquel on supprime une arrête. On obtient un sous-graphe couvrant connexe contenant un cycle de moins. En itérant cette opération, la construction s'arrête lorsqu'il n'y a plus de cycles et donc qu'on obtient un arbre couvrant.

□