

# Introduction à la théorie des graphes (résumé-memento)

J.C. Fournier

## 1. Introduction à la notion de graphe

Des points et des lignes reliant certains de ces points sur le plan, voilà une image « naïve » d'un « graphe ». Elle suffit pour imaginer déjà un peu tout ce que cette notion peut représenter dans des domaines très divers. Un premier exemple s'impose à l'esprit : la représentation d'un réseau de communication, les points représentant les centres, ou les noeuds, du réseau et les lignes représentant les liaisons. A partir de ce « modèle » du réseau, on peut poser beaucoup de questions, par exemple : ce réseau permet-il à tous les centres de communiquer entre eux, c'est-à-dire d'un centre donné quelconque peut-on aller à n'importe quel autre, soit directement soit en passant par des centres intermédiaires ? Ou encore : étant donnés deux centres trouver les chemins qui les relient, en particulier trouver un « plus court chemin », plus court du point de vue du nombre de liaisons intermédiaires. A propos de cette dernière question, on peut élargir la problématique en considérant des valeurs associées aux lignes, représentant par exemple des distances entre les points reliés, ou des temps, des coûts, etc. La longueur d'un chemin sera alors la somme des valeurs des lignes qui le composent, et un plus court chemin sera bien entendu un chemin de plus petite longueur parmi ceux reliant les deux points considérés. On a là une représentation de ce qu'on appellera un « graphe valué ».

Nous allons maintenant formaliser cette notion de graphe, les points vont s'appeler les « sommets » et les lignes les « arêtes ». Cette première notion sera complétée ensuite par celle de « graphe orienté », les lignes arêtes étant cette fois orientées dans un sens. A vrai dire, beaucoup d'applications sont concernées par le modèle de graphe orienté. Ainsi dans certains réseaux de communication les liaisons ne peuvent être utilisées que dans un sens défini *a priori*, on peut penser par exemple à un plan de ville avec des rues à sens unique. Un autre exemple particulièrement intéressant, évoqué plus loin, est la modélisation de la gestion d'un projet composé de tâches qui ont entre elles des dépendances dans le temps, représentées dans ce qu'on appelle le graphe potentiel-tâches

## 2. Généralités sur les graphes

### Définition

Un *graphe (non orienté)*  $G$  est défini par deux ensembles finis : un ensemble  $X$ , non vide, d'éléments appelés *sommets*, un ensemble  $E$  (qui peut être vide) d'éléments appelés *arêtes* avec associés à chaque arête  $e$  deux sommets  $x$  et  $y$ , distincts ou non, appelés les *extrémités* de  $e$ .

### Notations

On écrit  $G = (X, E)$ . Les ensembles  $X$  et  $E$  peuvent être notés  $X(G)$  et  $E(G)$ . On note habituellement par  $n$  ou  $n_G$  le cardinal de  $X$ , c'est-à-dire le nombre de sommets du graphe, et par  $m$  ou  $m_G$  le cardinal de  $E$ , c'est-à-dire le nombre d'arêtes du graphe. La paire des extrémités  $x$  et  $y$  d'une arête  $e$  est notée ici simplement  $xy$  (ou  $yx$ ), au lieu de la notation mathématique habituelle  $\{x, y\}$  lorsque  $x \neq y$ .

### Représentation

Il est habituel, et bien pratique, de dessiner les graphes sur un plan de la façon suivante : les sommets sont représentés par des points, les arêtes par des lignes simples (qu'on peut définir mathématiquement de façon précise) qui relient les points extrémités concernés.

### Terminologie

Lorsque  $x$  et  $y$  sont les extrémités d'une arête  $e$ , on dit que  $e$  *relie* les sommets  $x$  et  $y$ , que les sommets  $x$  et  $y$  sont *voisins*, que l'arête  $e$  est *incidente* au sommet  $x$  et au sommet  $y$ . On peut avoir  $x = y$ , dans ce cas l'arête  $e$  est appelée une *boucle*. Deux arêtes  $e$  et  $e'$ , ou plus, peuvent avoir mêmes extrémités  $x$  et  $y$ , on dit alors qu'elles sont *parallèles* ou qu'on a une *arête multiple* (*double*, *triple*, etc., suivant le nombre d'arêtes) entre  $x$  et  $y$ .

Un graphe est dit *simple* s'il n'a ni boucles ni arêtes multiples. Dans ce cas, souvent considéré, chaque arête s'identifie à la paire de ses extrémités, qui sont distinctes, et on écrit par exemple  $e = xy$ . Un graphe simple peut être défini, plus simplement que de la façon générale donnée plus haut, comme un couple d'ensembles finis  $(X, E)$  où  $X$  est non vide et  $E$  est un ensemble de parties à deux éléments de  $X$ .

### Isomorphismes et graphes non étiquetés

On définit de façon naturelle un *isomorphisme* d'un graphe  $G = (X, E)$  sur un graphe  $H = (Y, F)$  par deux bijections,  $\phi$  de  $X$  sur  $Y$  et  $\psi$  de  $E$  sur  $F$ , telles que, pour  $e \in E$  et  $x, y \in X$ , l'arête  $\psi(e)$  a pour extrémités  $\phi(x)$  et  $\phi(y)$  dans  $H$  si et seulement si l'arête  $e$  a pour extrémités  $x$  et  $y$  dans  $G$ . C'est-à-dire, ces bijections préservent la relation d'incidence des arêtes sur les sommets. Les graphes  $G$  et  $H$  sont dits *isomorphes*.

Deux graphes isomorphes sont en fait identiques quant à leurs structures de graphe, ils ont exactement les mêmes propriétés et ne se distinguent que par leurs ensembles d'éléments, sommets et arêtes, concrètement par les noms ou *étiquettes* donnés à ces éléments. Dès lors qu'on ne s'intéresse qu'aux propriétés des graphes en tant que graphes, il est naturel de ne pas considérer comme différents deux graphes isomorphes. C'est ce qu'on fait ici, considérant toujours ce qu'on appelle des *graphes non étiquetés*. Autrement dit, en langage algébrique, un graphe non étiqueté est une classe d'équivalence suivant la relation d'équivalence définie sur l'ensemble des graphes par la relation d'isomorphisme.

## Graphes planaires

Un cas particulier remarquable de graphe est celui où on peut imposer à une représentation plane du graphe la condition que deux lignes arêtes ne se coupent pas, en dehors d'extrémités communes (dans le cas d'arêtes ayant une extrémité commune). Ceci définit les graphes *planaires*, lesquels jouent un rôle important dans la théorie du fait de leurs propriétés remarquables.

## Graphes complets

Les *graphes complets* sont les graphes simples tels que deux sommets distincts quelconques sont reliés par une arête. Comme graphe non étiqueté, un graphe complet est simplement déterminé par son nombre  $n$  de sommets, et on le note d'une façon générale  $K_n$ .

Notons que le nombre d'arêtes  $m$  de  $K_n$  est égal au coefficient binomial  $\binom{n}{2}$ , c'est-à-dire :

$$m = \frac{n(n-1)}{2}$$

## Sous-graphes

Soit  $G = (X, E)$  un graphe. Un *sous-graphe* de  $G$  est un graphe de la forme  $H = (Y, F)$  où  $Y \subseteq X$  et  $F \subseteq E$  sont tels que toute arête de  $F$  a ses extrémités dans  $Y$ . Notons que le fait qu'un sous-graphe est un *graphe* implique cette propriété que toute arête de  $F$  a ses extrémités dans  $Y$ .

Un sous-graphe  $H$  de  $G$  est dit *engendré* (ou *induit*), et on peut préciser par un ensemble de sommets  $Y \subseteq X$ , s'il est de la forme  $H = (Y, F)$  où  $F$  est l'ensemble des arêtes de  $E$  qui ont leurs extrémités dans  $Y$ . On note ce sous-graphe  $G_Y$ . En particulier  $G_X = G$ .

Un sous-graphe  $H = (Y, F)$  de  $G$  est dit *couvrant* si  $Y = X$ . On dit aussi dans ce cas que  $H$  est un *graphe partiel* de  $G$ . On peut préciser graphe partiel *engendré par*  $F$ , c'est le graphe  $(X, F)$  et on le note  $G(F)$ .

Les notations suivantes sont courantes :

$G - Y$  où  $Y \subseteq X$  : sous-graphe de  $G$  engendré par  $X \setminus Y$  (sous-graphe obtenu en enlevant de  $G$  les sommets de  $Y$ , avec leurs arêtes incidentes).

$G - F$  où  $F \subseteq E$  : graphe partiel de  $G$  engendré par  $E \setminus F$  (graphe partiel de  $G$  obtenu en enlevant les arêtes de  $F$ ).

On écrira en particulier :  $G - x$  à la place de  $G - \{x\}$  pour  $x \in X$  et  $G - e$  à la place de  $G - \{e\}$  pour  $e \in E$ .

## Degrés

On appelle *degré* d'un sommet  $x$  d'un graphe  $G$  le nombre d'arêtes de  $G$  incidentes à  $x$ , c'est-à-dire admettant  $x$  comme extrémité. Précisons que chaque boucle compte deux fois. On note  $d(x)$  ou  $d_G(x)$  cet entier.

Un sommet est dit *isolé* si son degré est nul.

PROPOSITION 1 - Dans un graphe  $G$  quelconque, on a :

$$\sum_{x \in X} d_G(x) = 2m$$

COROLLAIRE 2 - Dans un graphe (quelconque) le nombre de sommets de degrés impairs est pair.

On appelle *degré minimum* d'un graphe  $G$  le minimum des degrés, quantité notée  $\delta_G$  ou simplement  $\delta$ . Remarque que  $\delta_G$  est le degré d'au moins un sommet du graphe. On définit de même le *degré maximum* de  $G$ , maximum des degrés, quantité notée  $\Delta_G$  ou simplement  $\Delta$ . C'est aussi le degré d'au moins un sommet du graphe.

REMARQUE : Il résulte de la proposition 1 précédente les inégalités :

$$n\delta_G \leq 2m \leq n\Delta_G$$

## Graphes réguliers

Un graphe  $G$  est dit *régulier* lorsque les degrés de ses sommets sont tous égaux. On peut préciser le degré commun  $k$  des sommets en disant *k-régulier*. Les graphes 3-réguliers sont dits *cubiques*.

REMARQUE : On a pour un graphe  $G$   $k$ -régulier la relation suivante entre  $k$  et les nombres de sommets  $n$  et d'arêtes  $m$  :

$$nk = 2m$$

## Chaînes

Une *chaîne* d'un graphe  $G = (X, E)$  est une suite de la forme :

$$(x_0, e_1, x_1, \dots, e_k, x_k)$$

où  $k$  est un entier  $\geq 0$ , les  $x_i$  sont des sommets de  $G$  et les  $e_i$  sont des arêtes de  $G$  tels que pour  $i = 0, \dots, k-1$ ,  $x_i$  et  $x_{i+1}$  sont les extrémités de  $e_{i+1}$ . L'entier  $k$  est la *longueur*. Une chaîne peut être de longueur nulle, elle se réduit alors à une suite en comprenant que le sommet  $x_0$ . Les sommets  $x_0$  et  $x_k$  sont les *extrémités* de la chaîne.

Lorsque le graphe est simple, une chaîne peut être définie simplement par la suite  $(x_0, x_1, \dots, x_k)$  de ses sommets. Une *sous-chaîne* d'une chaîne est une chaîne définie comme sous-suite, entre deux sommets, de la suite définissant la chaîne considérée.

Une chaîne est dite *simple* si ses arêtes  $e_i$ , pour  $i = 1, \dots, k$ , sont deux à deux distinctes. On dit que la chaîne *ne passe pas* deux fois par une même arête.

Une chaîne est dite *élémentaire* si ses sommets  $x_i$ , pour  $i = 0, 1, \dots, k$ , sont deux à deux distincts. On dit que la chaîne *ne passe pas* deux fois par un même sommet.

Remarque que *élémentaire* entraîne *simple*.

Une chaîne est dite *fermée* si ses extrémités  $x_0$  et  $x_k$  coïncident.

LEMME 3 - Si dans un graphe deux sommets sont reliés par une chaîne alors ils sont reliés par une chaîne élémentaire.

## Cycles

Un *cycle* est une chaîne de longueur  $\geq 1$  simple et fermée. C'est donc une chaîne de la forme :

$$(x_0, e_1, x_1, \dots, e_k, x_0)$$

où  $k \geq 1$  et les  $e_i$  sont distincts. L'entier  $k$  est la *longueur* du cycle. Notons que, contrairement à une chaîne, un cycle n'est jamais de longueur nulle; comme cas limite il peut être de longueur 1, il est alors constitué d'un sommet avec une boucle. Lorsque le graphe  $G$  est simple, un cycle peut être défini par la suite  $(x_0, x_1, \dots, x_0)$  de ses sommets.

Un cycle est *élémentaire* pour l'expression *cycle élémentaire* si ses sommets  $x_i$ , pour  $i = 0, \dots, k-1$ , sont deux à deux distincts. Dans ce cas, la longueur est le nombre de sommets du cycle (sauf le dernier).

Un cycle est dit *pair* ou *impair* suivant que sa longueur est paire ou impaire.

REMARQUE : On ne distingue pas des cycles qui ne diffèrent que par leurs « paramétrages », c'est-à-dire par les suites qui les définissent. Par exemple, dans un graphe simple, les trois cycles élémentaires suivants, définis par leurs suites de sommets, sont considérés comme un seul et même cycle :

$$(x_0, x_1, x_2, x_3, x_4, x_0)$$

$$(x_3, x_4, x_0, x_1, x_2, x_3)$$

$$(x_0, x_4, x_3, x_2, x_1, x_0)$$

De fait, c'est bien dans le graphe un même cycle mais parcouru de façons différentes.

## Chaînes et cycles comme graphes

Par extension des définitions précédentes, on appelle également *chaîne (élémentaire)* tout graphe qui peut être paramétré comme une chaîne (élémentaire), c'est-à-dire qui est de la forme  $G = (X, E)$  avec  $X = \{x_0, x_1, \dots, x_k\}$  et  $E = \{e_1, \dots, e_k\}$  tels que :

$$(x_0, e_1, x_1, \dots, e_k, x_k)$$

soit une chaîne élémentaire de  $G$ . C'est dans ce sens par exemple qu'on dit qu'une chaîne est un arbre.

De même, on parle d'un *cycle (élémentaire)* pour un graphe qui peut être paramétré comme un cycle élémentaire. Compte tenu de la remarque précédente, et compte tenu aussi qu'on considère toujours ici des graphes non étiquetés, un tel cycle est unique dès lors que sa longueur est fixée. C'est ainsi par exemple qu'on parle du cycle de longueur 5.

## Connexité

Un graphe  $G$  est *connexe* si quels que soient deux sommets de ce graphe ils sont reliés par une chaîne.

Les *composantes connexes* d'un graphe  $G$  sont les sous-graphes engendrés connexes *maximaux* de  $G$ . *Maximal* veut dire ici que le sous-graphe en question n'est pas lui-même sous-graphe *propre*, c'est-à-dire ayant strictement moins de sommets, d'un sous-graphe connexe de  $G$ . Evidemment, un graphe est connexe si et seulement s'il n'a qu'une seule composante connexe.

On vérifie que les composantes connexes d'un graphe sont des sous-graphes deux à deux disjoints, c'est-à-dire n'ayant deux à deux ni sommets ni arêtes en commun. Elles définissent la *décomposition en composantes connexes* du graphe. Cette décomposition est unique.

On peut aussi définir en langage algébrique les composantes connexes d'un graphe  $G = (X, E)$  comme les sous-graphes engendrés par les classes d'équivalence sur  $X$  définies par la relation : les sommets  $x$  et  $y$  sont reliés par une chaîne. Cette relation binaire est en effet une relation d'équivalence sur l'ensemble  $X$  (réflexive, symétrique et transitive).

Pour finir sur la connexité, signalons juste la proposition suivante : *Si un graphe possède un graphe partiel connexe, il est lui-même connexe*. Elle fait partie de ces petites propositions qui sont considérées comme évidentes, ou presque, et qu'on ne se donne pas toujours la peine de démontrer, voire même d'énoncer !

## Graphes bipartis

Un graphe  $G$  est *biparti* si son ensemble de sommets peut être divisé en deux parties, appelées *classes*, disjointes et recouvrant l'ensemble, telles que toute arête a une extrémité dans chaque classe.

On note un graphe biparti  $G = (X, Y, E)$  où  $X$  et  $Y$  sont les deux classes (et par conséquent  $X \cup Y$  est l'ensemble des sommets),  $E$  est l'ensemble des arêtes.

REMARQUES : 1) Il importe de noter que l'un des ensembles  $X$  ou  $Y$  peut être vide. De ce fait, le couple  $(X, Y)$  ne constitue pas une *partition* au sens mathématique strict du terme (les classes d'une partition ne doivent pas être vides). Néanmoins, les termes de « bipartition » et de « classes » sont couramment employés. Notons qu'avec cette définition le graphe réduit à un sommet, et sans arêtes, est biparti.

2) Une bipartition qui définit un graphe comme biparti n'est pas unique en général.

3) Un graphe biparti n'a pas de boucles. En effet une boucle contredirait l'hypothèse qu'une arête a ses extrémités dans des classes différentes. Mais un graphe biparti peut avoir des arêtes multiples.

THÉORÈME 4 - *Un graphe est biparti si et seulement s'il n'a pas de cycles impairs.*

Un graphe biparti  $G = (X, Y, E)$  est *complet* si son ensemble d'arêtes est  $E = \{xy \mid x \in X, y \in Y\}$ , c'est-à-dire si toute paire d'un sommet de  $X$  et d'un sommet de  $Y$  est une arête.

## Aspects algorithmiques

La théorie des graphes est un terrain de prédilection pour l'algorithmique. Beaucoup d'applications mettent en jeu des algorithmes de graphes et il faut donc représenter

en machine les graphes considérés.

Pour classer les différentes possibilités de représentations en machine des graphes, on peut distinguer les trois *principes* suivants :

1. Donner la possibilité de déterminer que deux sommets donnés sont voisins, c'est-à-dire sont reliés par une arête dans le graphe, et même plus précisément dans le cas d'un graphe non simple, donner le nombre d'arêtes reliant les sommets concernés. La façon naturelle de réaliser ceci pour un graphe  $G = (X, E)$  est la *matrice d'adjacence*, définie comme suit : posant  $X = \{x_1, \dots, x_n\}$ , c'est la matrice carrée d'ordre  $n$ ,  $M = (m_{ij})$ , où  $m_{ij}$  est le nombre d'arêtes d'extrémités  $x_i$  et  $x_j$  dans  $G$ .
2. Un deuxième principe de représentation d'un graphe est de donner pour chaque sommet son « voisinage », c'est-à-dire la *liste des sommets voisins*, avec éventuellement les arêtes incidentes correspondantes.
3. Si dans un algorithme « l'entrée » dans le graphe se fait par les arêtes, et non directement par les sommets comme dans les représentations précédentes, un troisième principe de représentation consiste à donner la *liste des arêtes*, avec pour chacune ses extrémités.

### Graphes valués

Dans les applications des graphes, notamment en optimisation, on considère fréquemment des graphes *valués*, aux arêtes, c'est-à-dire des graphes avec des valeurs, entières ou réelles, positives ou non, associées aux arêtes. De façon formelle on a un graphe  $G = (X, E)$  avec une application  $v : E \rightarrow \mathbb{R}$ .

Lorsqu'un graphe valué est simple, en tant que graphe, ce qui est souvent le cas, il est souvent représenté en machine par une matrice, comme la matrice d'adjacence mais avec en entrées les valeurs des arêtes concernées. On convient d'une valeur spéciale, par exemple  $\infty$ , lorsqu'il n'y a pas dans le graphe d'arêtes reliant les sommets correspondants à cette entrée de la matrice. De façon précise, et avec les notations précédentes, c'est la matrice  $M = (v(x_i x_j))$  où  $1 \leq i, j \leq n$ , avec une application  $v$  étendue en posant  $v(x_i x_j) = \infty$  lorsque  $i \neq j$  et  $x_i x_j \notin E$  et  $v(x_i x_j) = 0$  lorsque  $i = j$ . Cette matrice est symétrique.

## 3. Arbres

### Définitions et propriétés

Un *arbre* est un graphe connexe et acyclique, où *acyclique* veut dire sans cycles. Un arbre est un graphe simple (une boucle ou une arête double définirait un cycle).

Sauf notations explicites différentes, on désigne toujours, comme pour les graphes en général, par  $n$  le nombre de sommets d'un arbre et par  $m$  son nombre d'arêtes.

PROPOSITION 5 - Un arbre tel que  $n \geq 2$  possède au moins deux sommets pendants, c'est-à-dire de degré 1.

PROPOSITION 6 - Si  $G$  est un arbre, on a  $m = n - 1$ .

PROPOSITION 7 - Dans un arbre, deux sommets quelconques sont reliés par une chaîne élémentaire unique.

### Forêts

Une *forêt* est un graphe acyclique. Les composantes connexes d'une forêt sont donc des arbres, ce qui explique la terminologie (très naturelle!). Les forêts généralisent les arbres.

PROPOSITION 8 - Dans une forêt  $G$  on a  $m \leq n - 1$ , avec l'égalité si et seulement si  $G$  est un arbre.

### Isthmes

Un *isthme* d'un graphe  $G$  est une arête  $e$  telle que  $G - e$  a une composante connexe de plus que  $G$ . Ce qui revient à dire que dans  $G - e$  les extrémités  $x$  et  $y$  de  $e$  ne sont plus reliées par une chaîne. On dit parfois aussi que l'arête  $e$  *sépare* les sommets  $x$  et  $y$ . Lorsque  $G$  est connexe, un isthme est une arête  $e$  telle que  $G - e$  n'est pas connexe.

Les isthmes sont caractérisés comme suit.

LEMME 9 - Une arête d'un graphe  $G$  est un isthme si et seulement si elle n'appartient pas à un cycle de  $G$ .

Les isthmes interviennent dans les propriétés des arbres par le résultat suivant.

PROPOSITION 10 - Dans un arbre toute arête est un isthme.

### Caractérisations des arbres

Les propriétés précédentes conduisent aux caractérisations suivantes des arbres.

THÉORÈME 11 - Les conditions suivantes pour un graphe  $G$  sont équivalentes :

- (1)  $G$  est un arbre.
- (2)  $G$  est connexe et on a  $m = n - 1$ .
- (3)  $G$  est acyclique et on a  $m = n - 1$ .
- (4)  $G$  est connexe et toute arête est un isthme.
- (5) Dans  $G$  deux sommets quelconques sont reliés par une chaîne élémentaire unique.

### Arbres couvrants

Un *arbre couvrant* d'un graphe  $G$  est un graphe partiel de  $G$  qui est un arbre.

PROPOSITION 12 - Un graphe connexe  $G$  a (au moins) un arbre couvrant.

COROLLAIRE 13 - Si  $G$  est connexe alors  $m \geq n - 1$ , avec l'égalité si et seulement si  $G$  est un arbre.

PROPOSITION 14 - Un graphe partiel d'un graphe connexe  $G$  est un arbre couvrant de  $G$  si et seulement s'il est connexe et minimal avec cette propriété relativement à la suppression d'arêtes.

PROPOSITION 15 - Un graphe partiel d'un graphe connexe  $G$  est un arbre couvrant de  $G$  si et seulement s'il est acyclique et maximal avec cette propriété relativement à l'ajout d'arêtes.

PROPOSITION 16 - Etant donné un arbre couvrant  $T$  de  $G$  et une arête  $e$  de  $G$  qui n'appartient pas à  $T$ , le graphe partiel  $T + e$  contient un seul cycle élémentaire.

LEMME 17 (D'ÉCHANGE) - Etant donné un arbre couvrant  $T$  de  $G$ , une arête  $e$  de  $G$  qui n'appartient pas à  $T$  et une arête  $f$  du cycle de  $T + e$ , alors  $T + e - f$  est un arbre couvrant de  $G$ .

LEMME 18 (D'ÉCHANGE FORT) - Etant donné deux arbres couvrants  $T$  et  $T'$  de  $G$ , et une arête  $e \in T' \setminus T$ , il existe une arête  $f \in T \setminus T'$  telle que  $T + e - f$  et  $T' + f - e$  sont des arbres couvrants de  $G$ .

### Application : problème de l'arbre couvrant minimum

C'est, sous une forme simplifiée, un problème de réseau de communication. Par exemple, étant donné des coûts de liaisons entre les paires de centres il s'agit de trouver un réseau d'un coût total qui soit le plus petit possible. En termes de graphes, le problème général peut s'énoncer de la façon suivante :

**Données** : un graphe simple connexe  $G = (X, E)$  valué par l'application  $v$  à valeurs dans les réels strictement positifs, qui peuvent représenter des *coûts*, des *distances*, des *temps*, etc.

**Sortie** : un graphe partiel de  $G$ ,  $T = (X, A)$  où  $A \subset E$ , qui soit connexe et tel que  $v(T) = \sum_{e \in A} v(e)$  soit minimum.

Il est facile de voir d'abord qu'une solution  $T$  est nécessairement un arbre couvrant de  $G$ . Sinon en effet, en application de la proposition 14, il existerait une arête  $e$  de  $T$  telle que  $T - e$  serait toujours connexe, et on aurait  $v(T - e) = v(T) - v(e) < v(T)$  (car  $v(e) > 0$ ) ce qui contredirait la minimalité de  $v(T)$ . Ce qu'on cherche donc est un *arbre couvrant minimum* de  $G$ .

Le très classique algorithme suivant, de Kruskal, résout le problème.

```

procedure Kruskal( $G, v$ ) ;
begin
   $F := E$  ;  $A := \emptyset$  ;
  while  $|A| < n-1$  loop
    trouver  $e \in F$  tel que  $v(e)$  soit minimum ;
     $F := F - \{e\}$  ;
    if  $G(A \cup \{e\})$  acyclique then

```

```

   $A := A \cup \{e\}$  ;
end if ;
end loop ;
  --  $G(A)$  est un arbre couvrant minimum
end Kruskal ;

```

### Connectivité

Le nombre de connectivité  $\kappa(G)$  d'un graphe  $G$  est défini comme le plus petit nombre de sommets dont la suppression dans  $G$  donne un graphe non connexe ou réduit à un seul sommet.

Un graphe  $G$  est dit *k-connecte* si  $\kappa(G) \geq k$ .

Une application directe de ces notions concerne les questions de vulnérabilité des réseaux de communication, vulnérabilité par rapport aux pannes de centres. Par exemple, le problème du réseau de coût minimum traité plus haut peut être prolongé avec une contrainte supplémentaire de résistance minimale aux pannes, de centres de la façon suivante.

**Données** : un graphe simple connexe valué  $G$  et un entier  $k \geq 1$ .

**Sortie** : un graphe partiel  $H$  de  $G$   $k$ -connecte et dont la somme des valeurs des arêtes soit minimum.

Avec  $k = 1$ , on retrouve comme cas particulier le problème de l'arbre couvrant minimum traité plus haut. Pour  $k > 1$ , le problème est connu pour être difficile.

On a une notion équivalente relative aux pannes de liaisons, lesquelles concernent les arêtes du graphe.

## 4. Graphes orientés

### Définitions et généralités

Un *graphe orienté*  $G$  est défini par deux ensembles finis : un ensemble  $X$  non vide de *sommets*, un ensemble  $A$  d'*arcs*, ou *arêtes orientées*, avec associé à chaque arc  $a$  un *couple*<sup>1</sup>  $(x, y)$  de sommets qui sont les *extrémités* de  $a$ , le sommet  $x$  étant en particulier appelé l'*origine*.

### Notations

On écrit  $G = (X, A)$ . Les ensembles  $X$  et  $A$  peuvent être aussi notés  $X(G)$  et  $A(G)$ . On conserve les notations  $n_G$  ou  $n$  pour le nombre de sommets et  $m_G$  ou  $m$  pour le nombre d'arcs.

<sup>1</sup> Rappelons la distinction, classique en mathématiques, entre *couple* et *paire* : un couple est ordonné, une paire ne l'est pas.

## Représentation

Les graphes orientés se représentent sur le plan comme les graphes non orientés avec simplement en plus sur chaque ligne arc une flèche qui indique son orientation, flèche dans le sens de  $x$  à  $y$  si le couple  $(x, y)$  est associé à l'arc.

Indiquons, pour finir sur la définition des graphes orientés, que les graphes orientés considérés sont des graphes *non étiquetés*, c'est-à-dire considérés à un isomorphisme près. Nous ne revenons pas plus sur cette notion qui se définit exactement comme pour les graphes non orientés et qui ne pose pas de problème pratiquement.

## Terminologie

Lorsque  $(x, y)$  est le couple associé des extrémités de l'arc  $a$ , on dit : que l'arc  $a$  va du sommet  $x$  au sommet  $y$ , que l'arc  $a$  est *incident* aux sommets  $x$  et  $y$ , ou encore, plus précisément, que l'arc  $a$  *sort* du sommet  $x$  et qu'il *entre* dans le sommet  $y$ . Le sommet  $y$  est appelé un *successeur* de  $x$  et le sommet  $x$  un *prédécesseur* de  $y$ .

Comme en non orienté, on a une *boucle* dans le cas  $x = y$  et un arc *multiple* dans le cas d'arcs ayant un même couple  $(x, y)$  associé. On peut préciser, suivant le nombre d'arcs concernés : *arc double*, *triple*, etc. Deux arcs ayant pour couples associés respectivement  $(x, y)$  et  $(y, x)$  sont dits *opposés*.

Un graphe orienté est dit *strict* s'il n'a ni boucles ni arcs multiples (il peut avoir des arcs opposés). Dans ce cas, souvent considéré, chaque arc s'identifie au couple de ses extrémités, qui sont alors distinctes, et on écrit par exemple  $a = (x, y)$ . Pour un graphe orienté strict  $G = (X, A)$  on peut définir l'ensemble des arcs  $A$  directement comme partie du produit cartésien  $X \times X$ .

Un graphe orienté strict est dit *symétrique* si pour tout arc  $(x, y)$  il existe également l'arc opposé  $(y, x)$ . Cette notion est très proche de celle d'un graphe non orienté.

## Représentations en machine des graphes orientés

On retrouve en orienté les mêmes principes de représentations en machine qu'en non orienté, avec quelques précisions supplémentaires dues à l'orientation.

## Graphe non orienté associé et notions « non orientées »

Etant donné un graphe orienté  $G$ , on considère le *graphe non orienté associé* défini de façon évidente en « oubliant » l'orientation des arcs de  $G$ , c'est-à-dire chaque arc dont le couple associé est  $(x, y)$  est remplacé par une arête dont la paire de sommets associée est  $xy$ .

REMARQUE : Etant donné un graphe non orienté  $G$  il y a  $2^m$  graphes orientés dont le graphe non orienté associé est  $G$  ( $m$  étant le nombre d'arêtes de  $G$ ).

Toutes les notions définies sur les graphes non orientés s'appliquent également aux graphes orientés par le biais du graphe non orienté associé : *degré* d'un sommet, *connexité* et *composantes connexes*, *chaînes* et *cycles*, etc. Par exemple, on dit qu'un

graphe orienté est *connexe* si son graphe non orienté associé l'est. Ces notions sont donc indépendantes de l'orientation des arcs.

Il existe par ailleurs des notions qui prennent en compte l'orientation des arcs et que nous définissons maintenant.

## Notions « orientées »

Certaines notions se transposent directement du cas non orienté au cas orienté, en remplaçant le mot *arête* par le mot *arc*. Ce sont : les *sous-graphes*, les *sous-graphes engendrés*, les *graphes partiels*, les *chemins*, *directed paths* ou *chaînes orientées*, les *circuits* ou *cycles orientés*<sup>2</sup>, avec les cas *simple* et *élémentaire* définis comme en non orienté. Notons la différence terminologique, un peu arbitraire mais commode, entre *chaîne* pour le cas non orienté et *chemin* pour le cas orienté, de même entre *cycle* et *circuit*. Un *chemin* est donc une suite d'éléments alternativement sommets et arcs, commençant et finissant par un sommet, et telle que chaque arc a pour origine le sommet précédent dans la suite et pour autre extrémité le sommet suivant. Le premier et le dernier sommet sont les *extrémités* du chemin. On dit qu'un chemin, ou un circuit, *passé par* un arc ou un sommet s'il contient cet arc ou ce sommet. Un chemin est *simple* s'il ne passe pas deux fois par un même arc, *élémentaire* s'il ne passe pas deux fois par un même sommet. Un *circuit* est un chemin *simple* et *fermé* (ses extrémités coïncident), et un circuit est *élémentaire* s'il est élémentaire en tant que chemin, à l'exception de ses extrémités qui sont identiques. La *longueur* d'un chemin ou d'un circuit est toujours son nombre d'arcs. Comme en non orienté, un chemin peut être de longueur nulle, par contre un circuit doit être de longueur  $\geq 1$ .

D'autres notions, les suivantes, ont une définition qui est spécifique au cas orienté.

## Degrés intérieurs et extérieurs

Le *degré intérieur*, respectivement le *degré extérieur*, d'un sommet  $x$  d'un graphe orienté  $G$  est le nombre d'arcs entrant dans  $x$ , respectivement le nombre d'arcs sortant de  $x$ . On note respectivement  $d_G^-(x)$ , ou  $d^-(x)$ , et  $d_G^+(x)$ , ou  $d^+(x)$ , ces entiers.

On a facilement :

$$\sum_{x \in X} d_G^-(x) = \sum_{x \in X} d_G^+(x) = m$$

Les boucles ne sont pas souvent considérées dans les graphes orientés, néanmoins cette formule les prend en compte correctement, sachant que chaque boucle dans un graphe orienté compte pour une unité de degré extérieur et une unité de degré intérieur au sommet concerné.

<sup>2</sup> Le qualificatif « orienté » est à comprendre comme une orientation cohérente, c'est-à-dire avec tous les arcs orientés dans un même sens de parcours du « cycle ». Cette précision vaut également pour les « chaînes orientées ».

On retrouve la formule sur la somme des degrés donnée à la proposition 1, en observant que pour tout sommet  $x$  on a  $d_G(x) = d_G^-(x) + d_G^+(x)$ , d'où :

$$\sum_{x \in X} d_G(x) = \sum_{x \in X} (d_G^-(x) + d_G^+(x)) = \sum_{x \in X} d_G^-(x) + \sum_{x \in X} d_G^+(x) = 2m$$

### Composantes fortement connexes

Un graphe orienté  $G$  est *fortement connexe* si quels que soient deux sommets distincts  $x$  et  $y$  il existe un chemin allant de  $x$  à  $y$ . Remarque que cette définition est symétrique en  $x$  et  $y$  : il existe aussi un chemin de  $y$  à  $x$ .

Une *composante fortement connexe* de  $G$  est un sous-graphe engendré de  $G$  fortement connexe et maximal. *Maximal* veut dire qu'il n'existe pas de sous-graphe engendré fortement connexe qui contient strictement (pour les sommets) ce sous-graphe. On peut aussi définir ces composantes comme étant les sous-graphes engendrés par les classes d'équivalence suivant la relation d'équivalence sur les sommets : il existe un chemin de  $x$  à  $y$  et un chemin de  $y$  à  $x$ . Ces composantes définissent une partition de l'ensemble des sommets (mais pas des arcs), et constituent la *décomposition en composantes fortement connexes* de  $G$ . Cette décomposition est unique.

Les composantes fortement connexes sont moins simples à déterminer que les composantes connexes, mais il existe néanmoins de bons algorithmes pour les trouver.

### Graphes orientés sans circuits

Les graphes orientés sans circuits ont d'importantes applications. Nous allons en donner une propriété caractéristique très utile.

D'une façon générale, un sommet *source* d'un graphe orienté est un sommet dont le degré intérieur est nul, c'est-à-dire un sommet sans arcs entrants. De même un sommet *puits* est un sommet de degré extérieur nul, c'est-à-dire sans arcs sortants.

LEMME 19 - *Dans un graphe orienté sans circuits il existe un sommet source et un sommet puits.*

On appelle *tri topologique*, ou *numérotation acyclique*, des sommets d'un graphe orienté  $G = (X, A)$  une bijection  $f$  de  $X$  sur l'intervalle d'entiers de 1 à  $n$  ( $n$  est le nombre de sommets de  $G$ ), telle que si le couple  $(x, y)$  est associé à un arc de  $G$  alors  $f(x) < f(y)$ . Pour cette notion on peut supposer que le graphe est connexe (au sens non orienté) et également qu'il est strict.

On a la caractérisation suivante des graphes sans circuits.

PROPOSITION 20 - *Un graphe orienté  $G$  est sans circuits si et seulement s'il admet un tri topologique de ses sommets.*

Les graphes orientés sans circuits ont d'importantes applications, notamment en gestion de projet. Soit un projet décomposé en tâches élémentaires. Pour chaque tâche il est donné sa durée (dans une unité de temps qui peut être très différente suivant la nature du projet, de la fraction de seconde à la journée), et les tâches dites *antérieures* c'est-à-dire celles qui doivent être achevées pour que cette tâche puisse

commencer (ce sont les contraintes *d'antériorité*). Le problème est la planification dans le temps des différentes tâches, ce qu'on appelle un *ordonnement*.

Ce problème se modélise par un graphe dont les sommets sont les tâches et les arcs les contraintes d'antériorité entre tâches. En outre ce graphe est valué, les valeurs des arcs étant les durées des tâches concernées. Ce graphe, appelé le graphe *potentiel-tâches*, est sans circuits, comme il est facile de le voir (il y aurait sinon une impossibilité de réaliser le projet), et le tri topologique est fondamental pour calculer un ordonnement du projet.

### Arborescences

On appelle *racine* d'un graphe orienté  $G$  un sommet  $r$  tel qu'il existe pour tout sommet  $x$  de  $G$  un chemin de  $r$  à  $x$ .

Une *arborescence* est un graphe orienté qui admet une racine et dont le graphe non orienté associé est un arbre.

REMARQUE : Une arborescence n'a qu'une seule racine (mais d'une façon générale, un graphe orienté peut avoir plusieurs racines).

On représente habituellement une arborescence avec la racine en haut (ce qui fait dire qu'ici les arbres ont leur racine en l'air...), et les autres sommets rangés horizontalement par niveaux égaux de distance à la racine. On omet parfois les flèches qui indiquent les orientations des arcs, celles-ci étant implicitement définies de haut en bas.

REMARQUE : Une arborescence s'identifie à un arbre (non orienté) avec un sommet distingué, appelé la racine ; en effet on retrouve l'orientation des arcs en prenant sur chaque arête le sens de l'éloignement par rapport à cette racine (noter que ce sens est bien défini de façon unique du fait de l'unicité de chemin de la racine à un sommet dans une arborescence, suivant une propriété donnée plus loin). C'est peut-être pourquoi en informatique on appelle couramment « arbre » ce qui est en fait une arborescence au sens de la théorie des graphes.

Une arborescence, comme tout graphe sans circuits, a au moins un sommet puits, c'est-à-dire sans successeurs, et en général il y en a même beaucoup. Un tel sommet s'appelle une *feuille*. Les sommets d'une arborescence en général s'appellent des *noeuds*.

On appelle *profondeur* d'un sommet sa distance à la racine (*distance* au sens de plus petite longueur des chemins). Ce terme se réfère à la représentation habituelle des arborescences, dans laquelle, comme on l'a vu, les sommets de même profondeur sont placés à ce qu'on appelle un même *niveau* de profondeur.

On utilise aussi une terminologie inspirée des arbres généalogiques. On appelle ainsi *fil*, ou *enfant*, d'un sommet tout successeur de ce sommet. Un sommet qui n'a aucun fil est une feuille. Si le sommet  $y$  est fil du sommet  $x$ , on dit tout naturellement que  $x$  est le *père*, ou le *parent*, de  $y$ . Tout sommet d'une arborescence a un père, lequel est unique, sauf la racine qui n'en a pas. Des sommets qui ont même père

sont, tout aussi naturellement, dits *frères*. Cette terminologie, *père*, *fil*, *frère*, est commode pour « naviguer » dans une arborescence. On appelle plus généralement *descendant* d'un sommet  $x$  d'une arborescence tout sommet  $y$  tel qu'il existe dans l'arborescence un chemin de  $x$  à  $y$ , c'est-à-dire  $y$  est un successeur ou, de façon éventuellement répétée un successeur de successeur de  $x$ .

On a les caractérisations suivantes des arborescences.

**THÉORÈME 21** - *Les conditions suivantes pour un graphe orienté  $G$  sont équivalentes :*

- (1)  $G$  est une arborescence.
- (2)  $G$  a une racine et son graphe non orienté associé est acyclique.
- (3)  $G$  a une racine et on a  $m = n - 1$ .
- (4) Il existe un sommet  $r$  tel que pour tout sommet  $x$  de  $G$  il existe un chemin unique de  $r$  à  $x$ .
- (5)  $G$  est connexe et on a  $d^-(x) = 1$  pour tout sommet  $x$  sauf pour un sommet, soit  $r$ , pour lequel on a  $d^-(r) = 0$  (on appelle dans la suite « condition des degrés intérieurs » cette deuxième condition).
- (6) Le graphe non orienté associé de  $G$  est acyclique et on a la condition des degrés intérieurs.
- (7)  $G$  est sans circuits et on a la condition des degrés intérieurs.

REMARQUE : Dans la condition (4) l'unique chemin du sommet  $r$  à lui-même est le chemin de longueur nulle qui ne comporte que le sommet  $r$ .

### Sous-arborescences

Etant donné une arborescence  $T$  et un sommet  $s$  de celle-ci, on définit la *sous-arborescence de racine  $s$*  comme le sous-graphe de  $T$  engendré par  $s$  et tous ses descendants dans  $T$ . Ce sous-graphe est bien une arborescence, dont la racine est le sommet  $s$ .

### Arborescences ordonnées

Il est fréquent dans les applications de considérer ce qu'on peut appeler une *arborescence ordonnée*, à savoir une arborescence dans laquelle est donné sur l'ensemble des fils de chaque sommet un ordre (total). C'est généralement le cas pratiquement, puisqu'on définit en effet le plus souvent une arborescence par la donnée pour chaque sommet de ses fils dans une liste, c'est-à-dire avec un certain ordre. On peut dans le cas d'une arborescence ordonnée parler par exemple du *premier fils* ou du *frère suivant* d'un sommet.

### Forêts orientées

On définit plus généralement une *forêt orientée* comme un graphe orienté dont chaque composante connexe est une arborescence. Le graphe non orienté associé est bien sûr une forêt, et une forêt orientée s'identifie à une forêt dont chaque composante connexe a un sommet distingué qui est sa racine. Les propriétés des forêts orientées se déduisent de celles des arborescences appliquées à ses composantes.

## Référence

J.C. Fournier, *Théorie des graphes et applications*, Hermès (2006)