Recherche Opérationnelle - Cours 4

Olivier Baudon

Université de Bordeaux

5 octobre 2023

Parcours

Introduction

Il existe principalement deux algorithmes permettant de parcourir un graphe (sommets et arêtes ou arcs) en temps linéaire, c'est à dire en $O(n+m)$:

- le parcours en largeur, que l'on appellera également BFS pour "Breadth-First-Search", basé sur l'utilisation d'un file ;
- le parcours en profondeur, que l'on appellera également DFS pour "Depth-First-Search", basé sur l'utilisation d'un pile.

Parcours en largeur

Principe

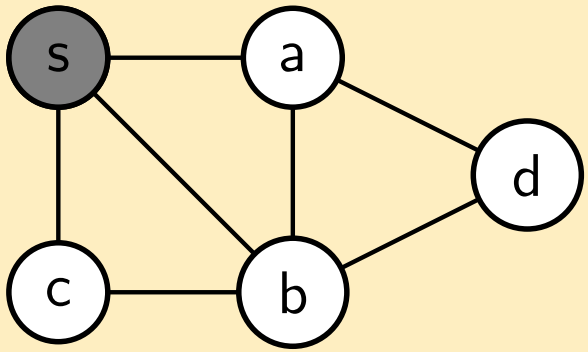
On examine les sommets du graphe à partir d'un sommet source s. Tous les sommets accessibles depuis s par une chaîne ou un chemin seront visités, et seulement ceux-là.
Les sommets sont examinés un par un. Ils sont blancs tant qu'ils n'ont pas été visités, puis gris quand ils sont présents dans la file d'attente, puis noirs après avoir été sortis de la file d'attente.
Quand on examine un sommet, on ajoute tous ses voisins ou successeurs qui sont encore blancs dans la file d'attente et l'on met à jour leur distance par rapport au sommet source.

Énoncé

Voir le document "Algorithmes de graphes" page 8.

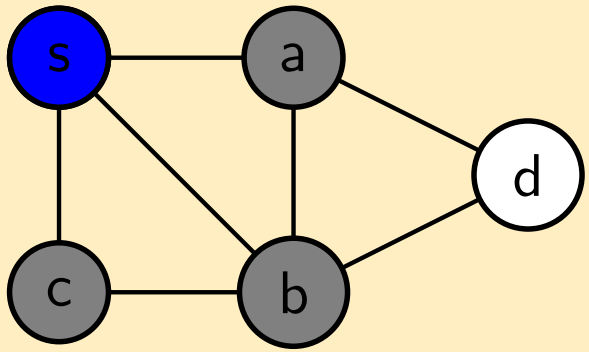
Exemple dans le cas non orienté

La couleur "NOIR" sera affichée avec du bleu afin de laisser l'étiquette lisible.



$F = \{s\}$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	0	∞	∞	∞	∞

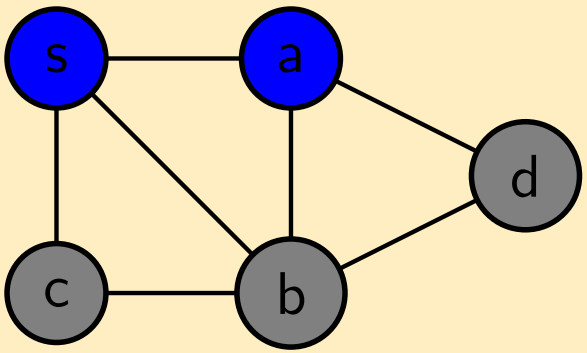


$F = \{a, b, c\}$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	0	1	1	1	∞

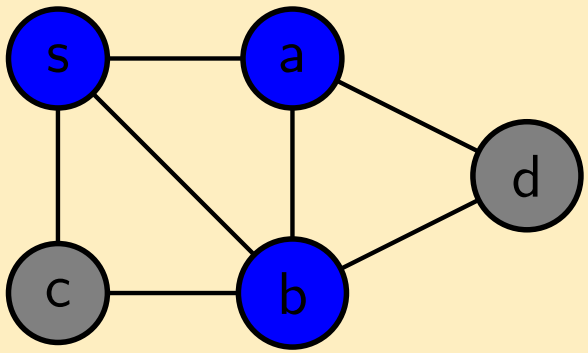
Parcours en largeur

Exemple dans le cas non orienté



$F = \{b, c, d\}$

sommet v	s	a	b	c	d
$d(v)$	0	1	1	1	2

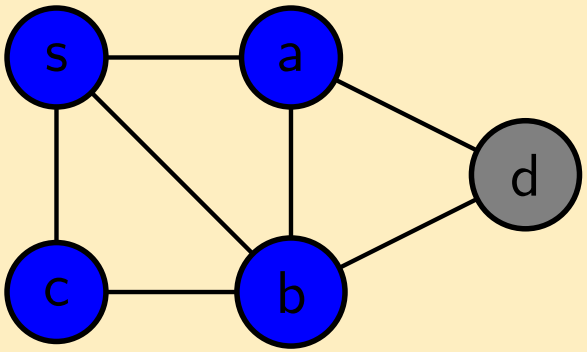


$F = \{c, d\}$

sommet v	s	a	b	c	d
$d(v)$	0	1	1	1	2

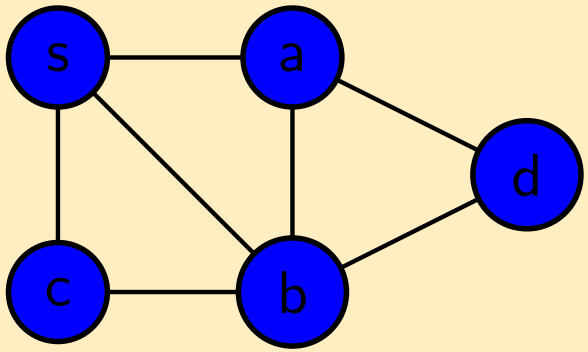
Parcours en largeur

Exemple dans le cas non orienté



$F = \{d\}$

sommet v	s	a	b	c	d
$d(v)$	0	1	1	1	2

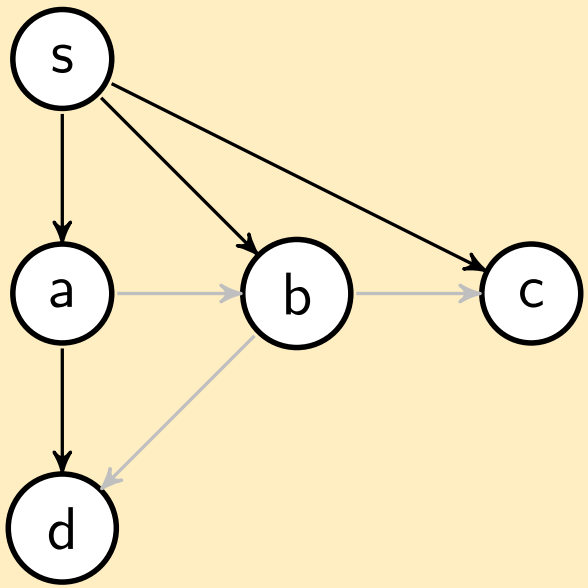


$F = \{\}$

sommet v	s	a	b	c	d
$d(v)$	0	1	1	1	2

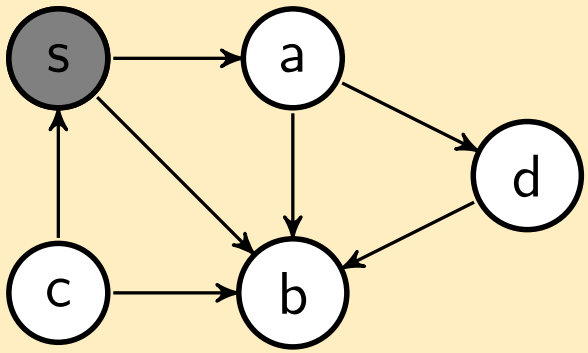
Parcours en largeur

Arborescence du parcours en largeur



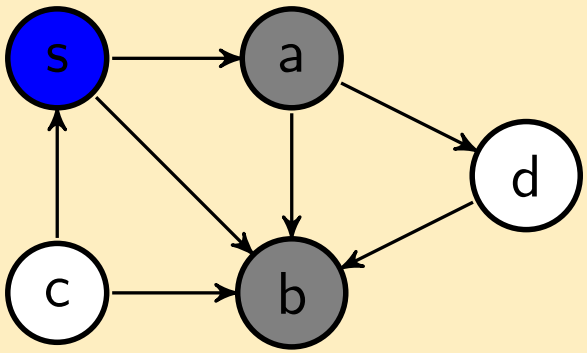
Parcours en largeur

Exemple dans le cas orienté



$F = \{s\}$

sommet v	s	a	b	c	d
$d(v)$	0	∞	∞	∞	∞

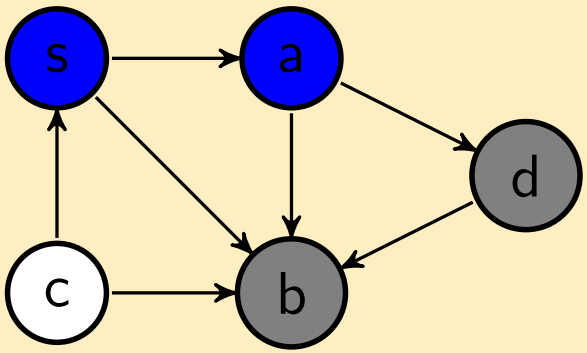


$F = \{a, b\}$

sommet v	s	a	b	c	d
$d(v)$	0	1	1	∞	∞

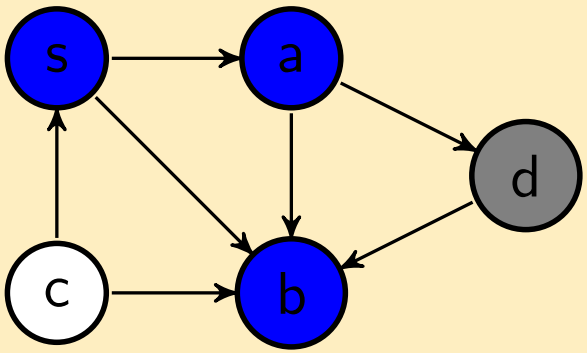
Parcours en largeur

Exemple dans le cas orienté



$F = \{b, d\}$

sommet v	s	a	b	c	d
$d(v)$	0	1	1	∞	2

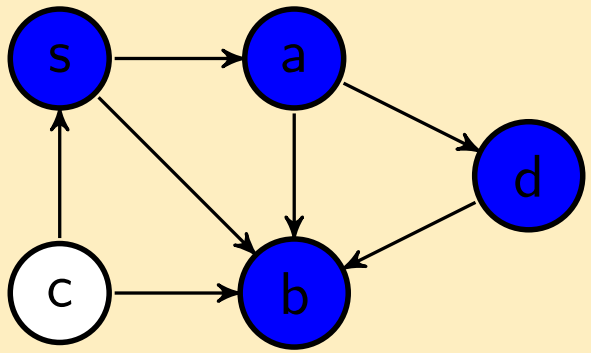


$F = \{d\}$

sommet v	s	a	b	c	d
$d(v)$	0	1	1	∞	2

Parcours en largeur

Exemple dans le cas orienté

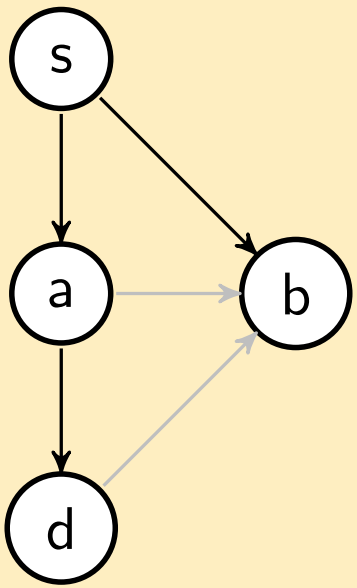


$F = \{\}$

sommet v	s	a	b	c	d
$d(v)$	0	1	1	∞	2

Parcours en largeur

Arborescence du parcours en largeur



Parcours en largeur

Complexité

L'algorithme de parcours en largeur a une complexité en $O(n+m)$. En effet, chaque sommet est introduit au plus une fois dans la file (tous uniquement si tous les sommets sont accessibles depuis s par une chaîne ou un chemin) et pour chaque sommet extrait de la file, on parcourt toutes ses arêtes incidentes (ou arc sortant).

Parcours en largeur

Validité

L'objectif du parcours en largeur est de calculer $d(v)$ qui contiendra en sortie $d(s, v)$: la distance dans le graphe G entre s et tout sommet v du graphe (∞ s'il n'existe pas de chaîne ou de chemin de s à v).

Lemme

Pour toute arête ou arc uv, $d(s, v) \leq d(s, u) + 1$.
En effet, si u est accessible à partir de s, alors le plus court chemin de s à v ne peut être plus long que le plus court chemin de s à u prolongé de uv.
Si u n'est pas accessible, alors $d(s, u) = \infty$ et dans tous les cas $d(s, v) \leq \infty + 1$.

Parcours en largeur

Validité : pour tout sommet $v, d(v) \geq d(s, v)$

On utilise une récurrence sur le nombre d'insertion de sommets dans F.
Au départ, seul s est introduit dans F et $d(s) = d(s, s) = 0$.
Soit w un sommet blanc, voisin d'un sommet u. Lorsque les voisins de u sont étudiés, par hypothèse, $d(u) \geq d(s, u)$.
Donc, quand w va être inséré dans F,
$d(w) = d(u) + 1 \geq d(s, u) + 1$. Or d'après le lemme précédent, on sait que $d(s, w) \leq d(s, u) + 1$. Donc $d(w) \geq d(s, w)$.

Parcours en largeur

Validité : Si $F = \{v_1, \ldots, v_r\}$, alors pour tout i, $d(v_i) \leq d(v_{i+1})$ et $d(v_r) \leq d(v_1) + 1$

On utilise une récurrence sur les opérations sur la file F.

- Au départ, $F = \{s\}$ et l'hypothèse est vérifiée.
- Quand on retire le sommet v_1 en tête de file, v_2 devient la tête de file. Par hypothèse,
  - $d(v_1) \leq d(v_2) \leq \ldots \leq d(v_r)$
  - $d(v_r) \leq d(v_1) + 1$

  donc
  - $d(v_r) \leq d(v_2) + 1$ car $d(v_r) \leq d(v_1) + 1$ et $d(v_1) \leq d(v_2)$
  - $d(v_2) \leq \ldots \leq d(v_r)$.

Parcours en largeur

Validité : Si $F = \{v_1, \ldots, v_r\}$, alors pour tout i, $d(v_i) \leq d(v_{i+1})$ et $d(v_r) \leq d(v_1) + 1$ - suite

- Quand on rajoute un sommet v_{r+1} à F
  - v_{r+1} est inséré à partir d'un sommet v_0 qui se trouvait dans F juste avant v_1 et qui vient d'être retiré.
  - Donc $d(v_0) \leq d(v_1) \ldots \leq d(v_r)$ et $d(v_r) \leq d(v_0) + 1$
  - $d(v_{r+1}) = d(v_0) + 1$ et comme $d(v_r) \leq d(v_0) + 1$ on a bien $d(v_1) \leq d(v_2) \ldots \leq d(v_r) \leq d(v_{r+1})$.
  - De plus, comme $d(v_1) \geq d(v_0)$ et $d(v_{r+1}) = d(v_0) + 1$, on a bien $d(v_{r+1}) \leq d(v_1) + 1$.

Parcours en largeur

Validité : Si deux sommets u et v sont enfilés dans F, et u est enfilé avant v, alors $d(u) \leq d(v)$.

Immédiat d'après le lemme précédent et en remarquant que la valeur $d(x)$ n'est jamais modifiée pour tout sommet x une fois qu'elle a été initialisée à une valeur finie.

Parcours en largeur

Validité : Après l'exécution de PL(G, s), tout sommet accessible à partir de s est visité, et pour tout sommet v, $d(v) = d(s, v)$. De plus, pour tout sommet v accessible à partir de s, alors un des plus courts chemins de s à v est un plus court chemin de s à $pere(v)$ complété par l'arête ou l'arc $pere(v)v$.

Supposons qu'il existe un sommet v tel que $d(v) \neq d(s, v)$.
Considérons le sommet v dans ce cas avec la plus petite valeur $d(v)$ possible.
Il est clair que $v \neq s$.
D'après un lemme précédent, on sait que $d(v) \geq d(s, v)$. Donc $d(v) > d(s, v)$. De plus, v est accessible depuis s, sinon $d(v) = \infty = d(s, v)$.
Soit u un sommet précédent v sur un plus court chemin de s à v.
Comme $d(s, u) = d(s, v) - 1$, on a bien $d(u) = d(s, u)$.
Donc $d(v) > d(s, v) = d(s, u) + 1 = d(u) + 1$.

Suite de la preuve

Quand u devient tête de file de F, si v est blanc, on aura $d(v) = d(u) + 1$ ce qui contredit l'hypothèse $d(v) > d(u) + 1$.
Si v est noir, alors $d(v) \leq d(u)$ ce qui contredit $d(v) > d(u) + 1$.
Si v est gris, alors il a été inséré à partir d'un sommet w inséré avant u et donc $d(w) \leq d(u)$ et $d(v) = d(w) + 1$. Donc $d(v) \leq d(u) + 1$ ce qui contredit $d(v) > d(u) + 1$.
Conclusion, v n'existe pas et donc pour tout sommet, $d(v) = d(s, v)$.
Tout sommet v accessible sera découvert, sinon on aurait $d(v) = \infty > d(s, v)$. De plus, si $pere(v) = u$ alors $d(v) = d(u) + 1$ et on aura bien un plus court chemin de s à v en prenant le plus court chemin de s à u complété de l'arête ou de l'arc uv.

Parcours en largeur

Arbre du parcours en largeur

Soit T le sous-graphe partiel de G avec :

- $V(T) = \{v \in V(G) | pere(v) \neq NIL\} \cup \{s\}$
- $E(T) = \{uv | u = pere(v), v \in V(T), v \neq s\}$

T est l'arbre de parcours en largeur de G.

Arêtes du co-arbre

Si uv est une arête de G n'appartenant pas à l'arbre du parcours en largeur, alors $d(u) - 1 \leq d(v) \leq d(u) + 1\}$.

Graphe biparti

Graphe biparti

Un graphe G non orienté est dit biparti si et seulement si il existe une partition de $V(G)$ en deux sous-ensembles A et B tels que les sous-graphes induits G_A et G_B ne contiennent aucune arête.
En d'autres termes, pour toute arête de G
$uv \in E(G) \Rightarrow (u \in A \Leftrightarrow v \in B)$

Parcours en largeur d'un graphe biparti

Si G est biparti, alors pour tout arête uv de G, $d(u) \neq d(v)$. En d'autres termes, les arêtes du co-arbre de l'arbre de parcours en largeur reliront toujours deux sommets de niveau consécutifs.

Parcours en profondeur

Principe

On examine les sommets du graphe en partant d'un sommet s. Tant que c'est possible, on "descend" dans le graphe de voisin en voisin. Sinon, on remonte jusqu'à être sur un sommet ayant encore un voisin non encore visité, et on redescend à nouveau. Si on remonte jusqu'au premier sommet et qu'aucun de ses voisins (ou successeur) n'a pas été visité, alors s'il reste encore un sommet s' non encore visité, on repart de s'.
Ce processus est répété tant qu'il existe des sommets non visités.
Pour cela, on utilise une boucle et une pile. Le plus simple est d'utiliser la pile d'appels générée par des appels récursifs.
Une variable *temps*, initialisée à 0, sert à définir pour chaque sommet u deux valeurs : $d(u)$ l'heure (ou la date) de début de visite de u et $f(u)$ l'heure (ou la date) de fin de visite de u. *temps* est incrémentée avant chaque mise à jour des tableaux d et f.

Parcours en profondeur

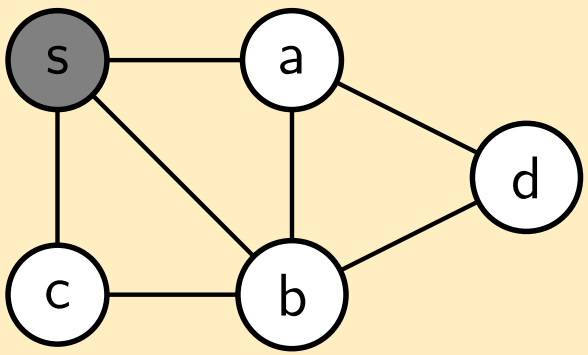
Énoncé

Voir le document "Algorithmes de graphes" page 9.

Parcours en profondeur

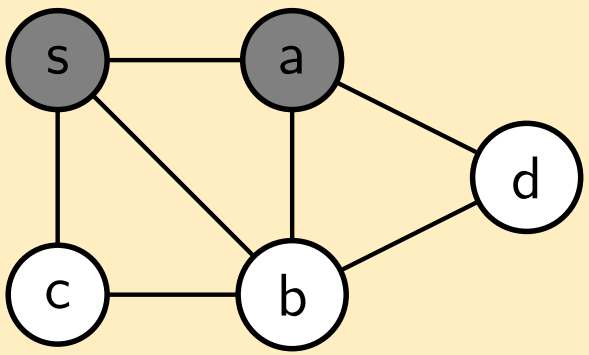
Exemple dans le cas non orienté

La couleur "NOIR" sera affichée avec du bleu afin de laisser l'étiquette lisible.



$Pile = [s]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1				
$f(v)$					

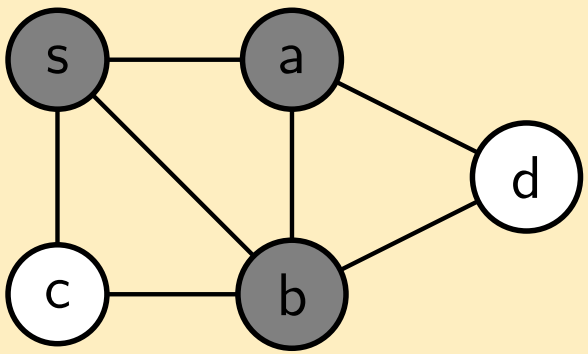


$Pile = [a, s]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2			
$f(v)$					

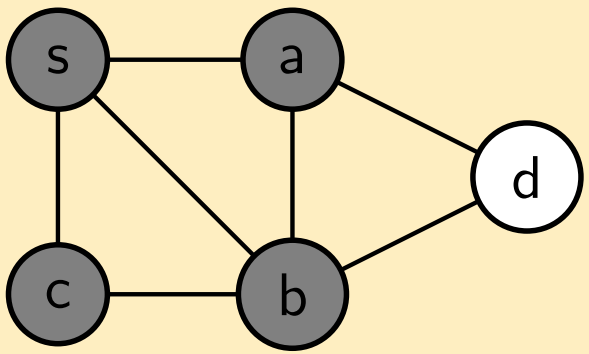
Parcours en profondeur

Exemple dans le cas non orienté



$Pile = [b, a, s]$

sommet v	s	a	b	c	d
$d(v)$	1	2	3		
$f(v)$					

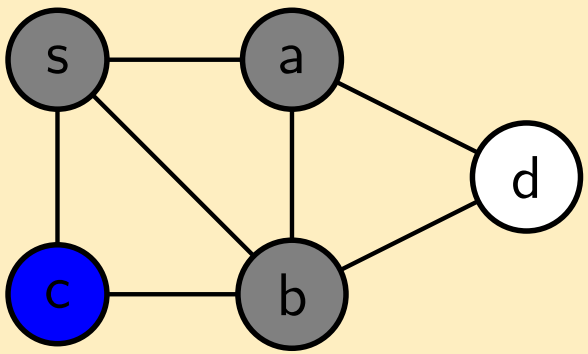


$Pile = [c, b, a, s]$

sommet v	s	a	b	c	d
$d(v)$	1	2	3	4	
$f(v)$					

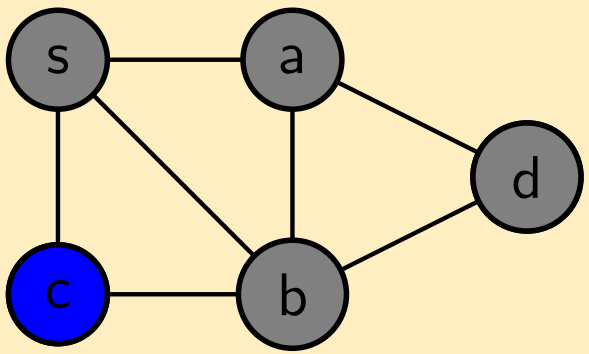
Parcours en profondeur

Exemple dans le cas non orienté



$Pile = [b, a, s]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3	4	
$f(v)$				5	

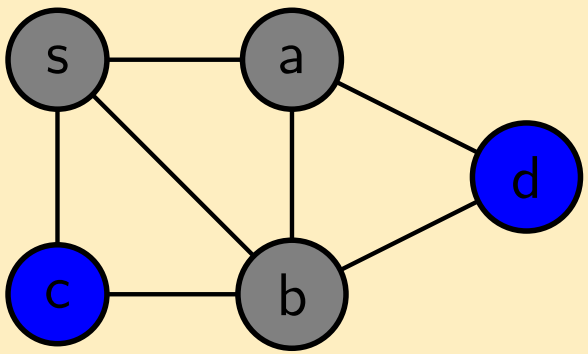


$Pile = [d, b, a, s]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3	4	6
$f(v)$				5	

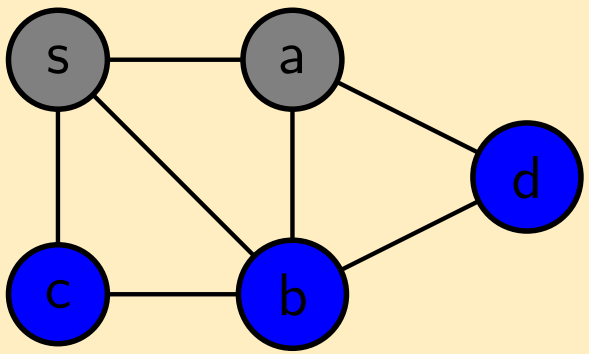
Parcours en profondeur

Exemple dans le cas non orienté



$Pile = [b, a, s]$

sommet v	s	a	b	c	d
$d(v)$	1	2	3	4	6
$f(v)$				5	7

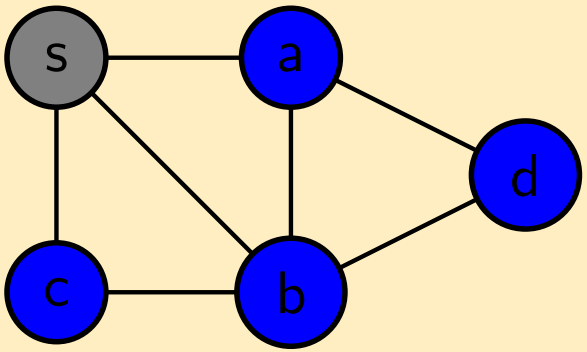


$Pile = [a, s]$

sommet v	s	a	b	c	d
$d(v)$	1	2	3	4	6
$f(v)$			8	5	7

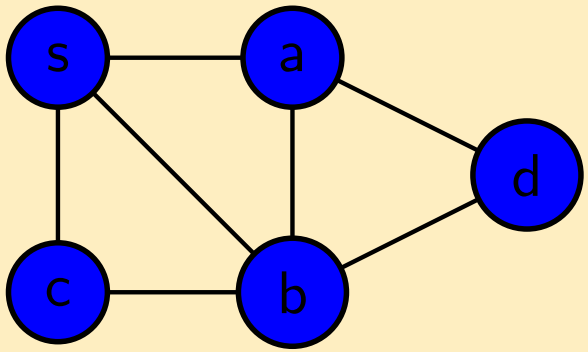
Parcours en profondeur

Exemple dans le cas non orienté



$Pile = [s]$

sommet v	s	a	b	c	d
$d(v)$	1	2	3	4	6
$f(v)$		9	8	5	7

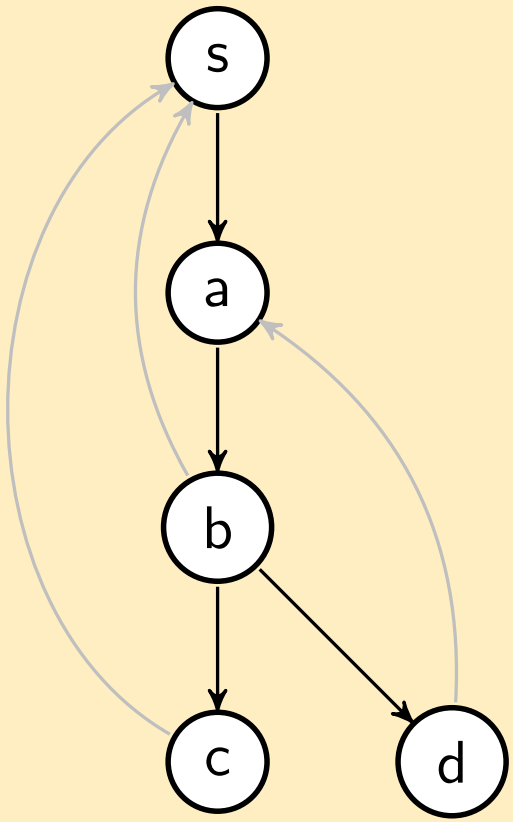


$Pile = []$

sommet v	s	a	b	c	d
$d(v)$	1	2	3	4	6
$f(v)$	10	9	8	5	7

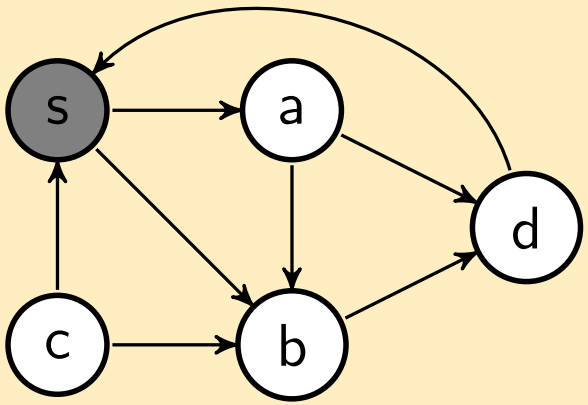
Parcours en profondeur

Arborescence du parcours en profondeur



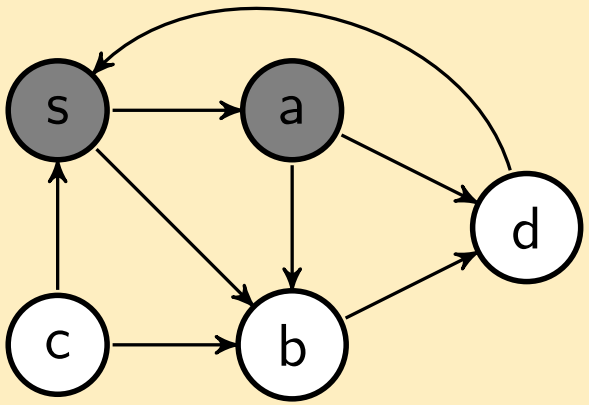
Parcours en profondeur

Exemple dans le cas orienté



$Pile = [s]$

sommet v	s	a	b	c	d
$d(v)$	1				
$f(v)$					

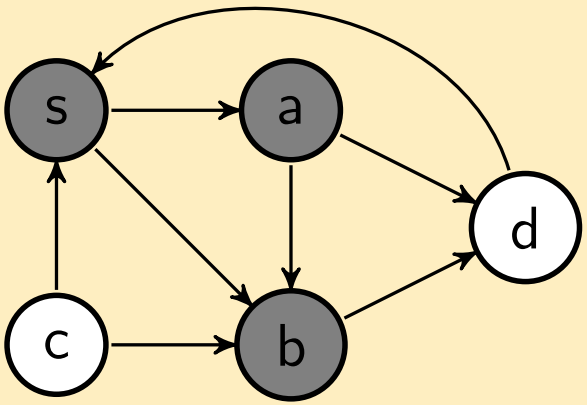


$Pile = [a, s]$

sommet v	s	a	b	c	d
$d(v)$	1	2			
$f(v)$					

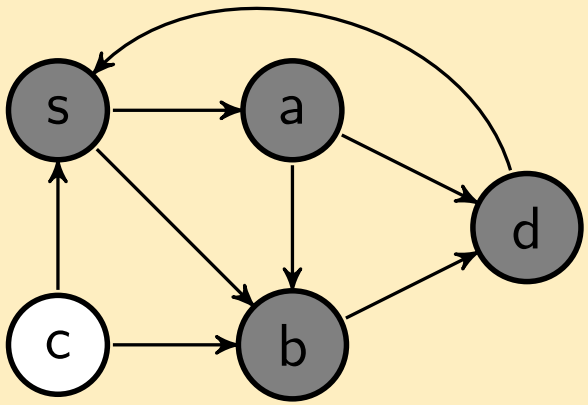
Parcours en profondeur

Exemple dans le cas orienté



$Pile = [b, a, s]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3		
$f(v)$					

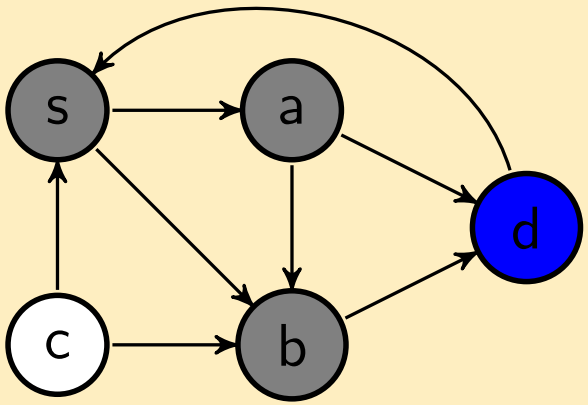


$Pile = [d, b, a, s]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3		4
$f(v)$					

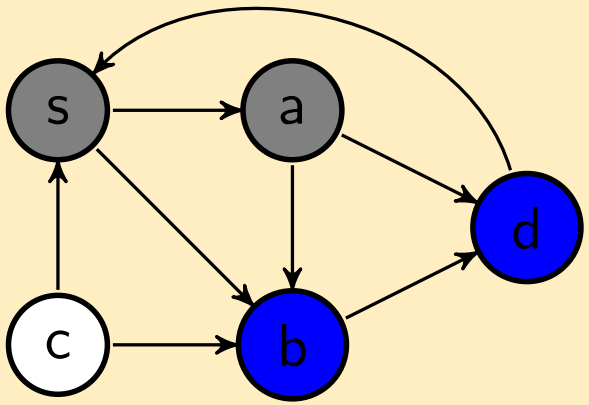
Parcours en profondeur

Exemple dans le cas orienté



$Pile = [b, a, s]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3		4
$f(v)$					5

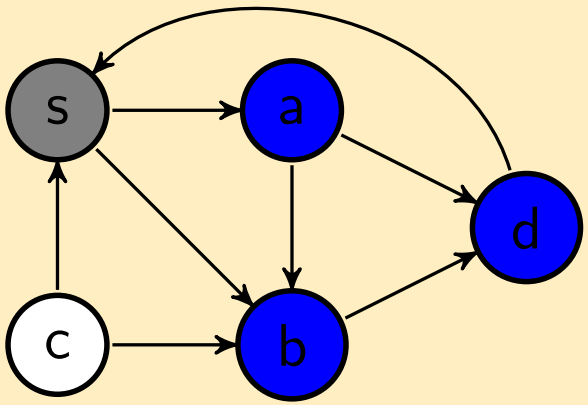


$Pile = [a, s]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3		4
$f(v)$			6		5

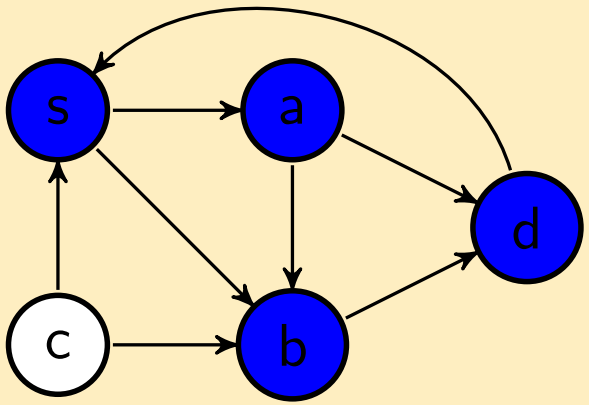
Parcours en profondeur

Exemple dans le cas orienté



$Pile = [s]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3		4
$f(v)$		7	6		5

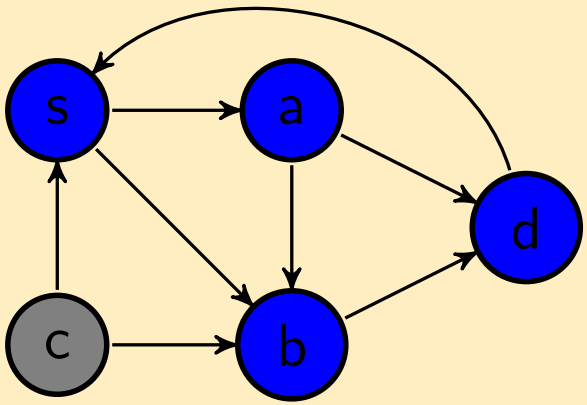


$Pile = []$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3		4
$f(v)$	8	7	6		5

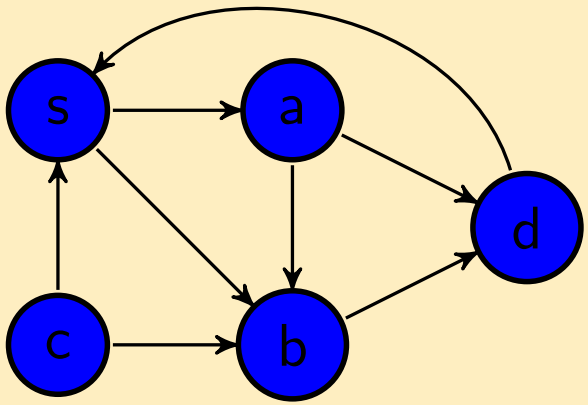
Parcours en profondeur

Exemple dans le cas orienté



$Pile = [c]$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3	9	4
$f(v)$	8	7	6		5

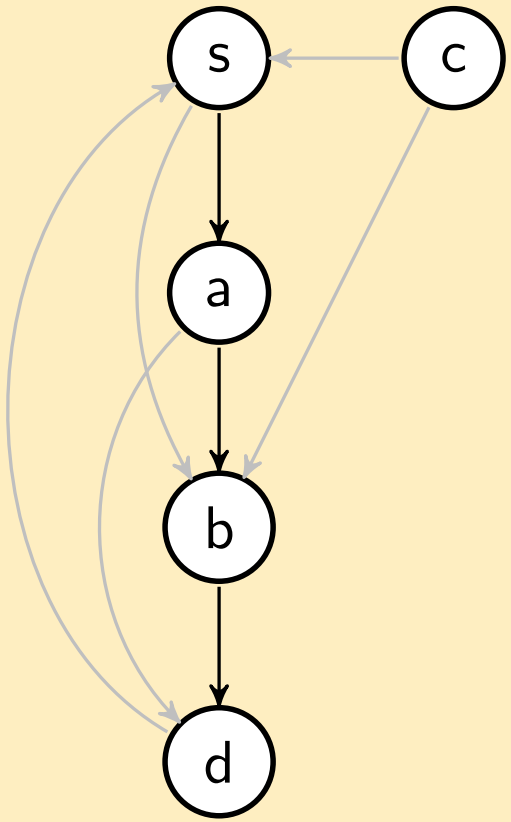


$Pile = []$

sommet v	*s*	*a*	*b*	*c*	*d*
$d(v)$	1	2	3	9	4
$f(v)$	8	7	6	10	5

Parcours en profondeur

Arborescence du parcours en profondeur



Parcours en profondeur

Type des arcs de l'arborescence

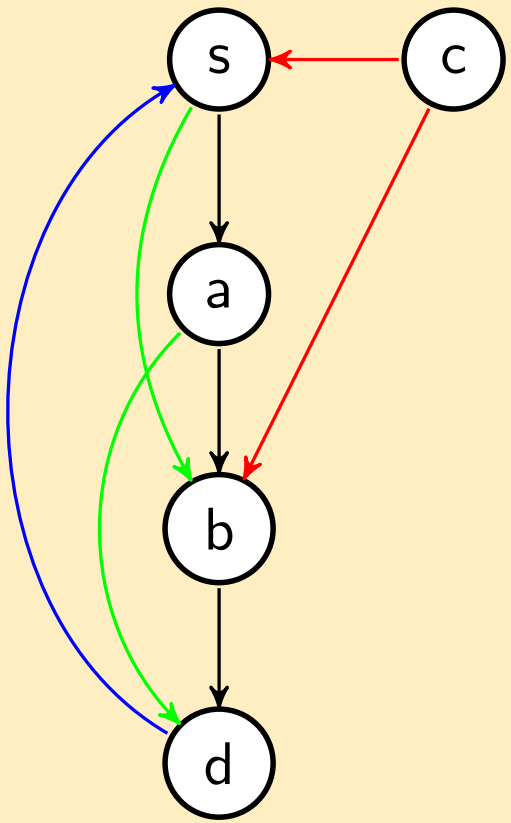
On distingue 4 types d'arcs :

- de liaison : ce sont les arcs reliant un sommet à ses fils ;
- avant : ce sont les arcs reliant un sommet à ses descendants $\neq$ de ses fils ;
- arrière ou de retour : ce sont les arcs reliant un sommet à l'un de ses ancêtres ;
- transverse : ce sont les arcs reliant un sommet aux sommets avec lesquels il n'a aucune relation dans l'arborescence.

Parcours en profondeur

Type des arcs de l'arborescence

Dans le dessin suivant, les arcs de liaison sont en noir, les arcs de retour en bleu, les arcs avant en vert et les arcs transverses en rouge.



Parcours en profondeur

Classe des arcs en fonction des couleurs des sommets

L'idée est de connaître le type d'un arc lors de sa visite en fonction de la couleur du sommet destination (le sommet origine étant toujours GRIS). Soit uv l'arc visité.

- Si v est BLANC, alors l'arc uv sera un arc de liaison.
- Si v est GRIS, alors l'arc uv sera un arc retour (aussi appelé arrière.
- Si v est NOIR, alors l'arc uv sera un arc avant ou transverse selon que u est un ancêtre ou non de v.

Parcours en profondeur

Classe des arcs dans le cas non orienté

Dans le cas du parcours en profondeur d'un graphe G non orienté, alors les arcs sont soit de liaison, soit retour. Il n'y a pas d'arc avant, ni d'arc transverse.
Justification : Lorsque l'on examine pour la première fois une arête uv, disons à partir d'un sommet u, si v n'a pas encore été visité, alors uv sera un arc de liaison.
Sinon, v est forcément un ancêtre de u car on ne peut finir de visiter v sans regarder l'arête uv. Donc v ne peut être NOIR et donc uv est un arc de retour.

Parcours en profondeur

Théorème des parenthèses

Soit u et v deux sommets de G, et $[d(u), f(u)]$, $[d(v), f(v)]$ les intervalles définis par leurs heures de début et fin de visite. On suppose que $d(u) < d(v)$. Deux cas sont possibles :

- Soit $[d(v), f(v)] \subseteq [d(u), f(u)]$. Dans ce cas v est un descendant de u.
- Soit $[d(v), f(v)] \cap [d(u), f(u)] = \emptyset$. Dans ce cas, aucun des deux n'est le descendant de l'autre dans l'arborescence.

Justification :

- dans le cas où $[d(v), f(v)] \subseteq [d(u), f(u)]$, cela signifie que v a été visité quand u était GRIS. Donc v sera un descendant de u et on finira la visite de v avant celle de u.
- Si $[d(v), f(v)] \cap [d(u), f(u)] = \emptyset$, alors comme $d(v) > d(u)$, on a $f(v) > d(v) > f(u) > d(u)$ et donc, on aura fini de visiter u avant de commencer la visite de v.

Parcours en profondeur

Théorème du chemin blanc

Dans une forêt obtenu par un parcours en profondeur d'un graphe $G = (V, E)$, un sommet v est un descendant d'un sommet u si et seulement si lors du début de visite du sommet u, il existe une chaîne ou un chemin reliant u à v composé exclusivement de sommets blancs.

Théorème du chemin blanc - Justification

si v est un descendant de u, les sommets reliant u à v dans l'arborescence étaient tous BLANC au moment du début de la visite de u.
Inversement, supposons qu'il existe un chemin composé de sommets blancs entre u et v à l'instant $d(u)$, et que v ne soit pas un descendant de u. Choisissons v de telle sorte qu'il soit le sommet de ce chemin blanc le plus proche de u qui ne soit pas un de ses descendants. Et soit w son prédécesseur sur le chemin blanc. w sera bien un descendant de u. Or si v n'est pas un descendant de u et donc de w, quand on finit de visiter w, v sera encore blanc et devrait donc être visité via l'arête wv. Donc v sera un fils de w et donc un descendant de u.

Parcours en profondeur

Applications

Il existe de nombreuse applications issues du parcours en profondeur. Parmi elles, citons :

1. le tri topologique ;
2. le calcul des composantes fortement connexes ;
3. la recherche de points d'articulation, un point d'articulation étant un sommet dont la suppression augmente le nombre de composantes connexes ;
4. le test de planarité d'un graphe (Hopcroft et Tarjan).

A noter que dans tous les cas cités ci-dessus, les algorithmes ainsi obtenus sont en temps linéaire, i.e. en $O(n + m)$.

Tri topologique

Définition

Le tri topologique d'un graphe G orienté est un ordonnancement de ses sommets de telle sorte que si $v_1, \ldots, v_n$ est l'ordre obtenu, alors :
$\forall v_i, v_j \in V(G), v_i v_j \in E(G) \Rightarrow i < j$

Théorème

Un graphe G orienté admet un tri topologique si et seulement si il est sans circuit.
Justification :
$\Rightarrow$ (par contradiction)
soit $v_1 v_2 \ldots v_{k-1} v_1$ un circuit de G. Si G admet un tri topologique, alors v_1 doit être rangé avant v_2 car il existe un arc $v_1 v_2$, v_2 avant $v_3, \ldots v_{k-1}$ avant v_1, ce qui est impossible.
Donc "G graphe orienté admet un tri topologique" $\Rightarrow$ "G est sans circuit".

Tri topologique

Justification (suite)

$\Leftarrow$

A l'inverse, si G est sans circuit, montrons que G admet un tri topologique.

On peut monter qu'il existe un sommet de G de degré entrant 0.

En effet, si l'on considère un chemin P de longueur maximale dans G, disons $P = v_0, v_1, \ldots v_k$, alors v_0 est de degré entant 0 dans G. Sinon, soit il a un prédécesseur hors de $\{v_1, \ldots v_k\}$ et dans ce cas, P n'est pas un chemin de longueur maximale, soit il a un prédécesseur parmi $\{v_1, \ldots v_k\}$ et dans ce cas, G n'est pas sans circuit.

On prend le sommet v de degré entrant 0 comme le premier sommet dans le tri topologique. $G - v$ est forcément sans circuit, donc possède également un sommet de degré entrant 0, que l'on prendra en deuxième position. On répète cette opération jusqu'à ce que tous les sommets de G aient été placés.

Tri topologique

Algorithme (basé sur le parcours en profondeur)

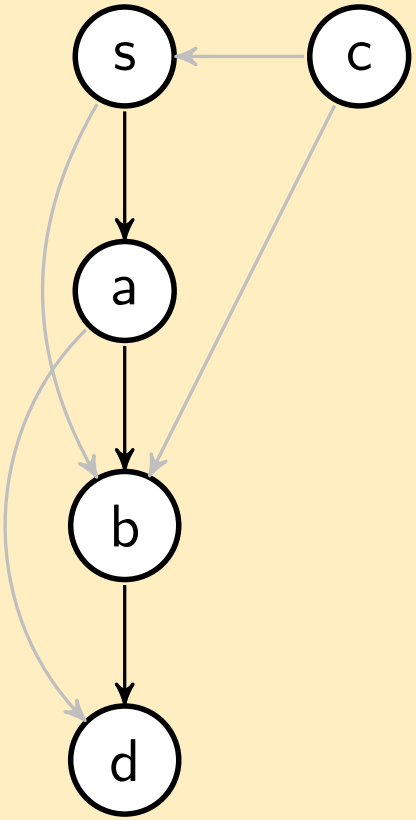
Voir le document "Algorithmes de graphes" page 4.

Complexité

Il est facile de constater que la complexité du tri topologique basé sur le parcours en profondeur est identique à celle du parcours en profondeur.
Donc le tri topologique est en $O(n + m)$

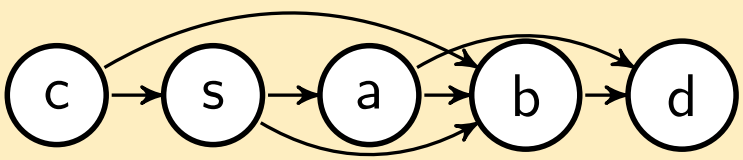
Tri topologique

Exemple



sommet v	s	a	b	c	d
$d(v)$	1	2	3	9	4
$f(v)$	8	7	6	10	5

Tri topologique : c, s, a, b, d



Tri topologique

Justification de l'algorithme

Théorème : un graphe G orienté est sans circuit si et seulement si le parcours en profondeur de G ne produit aucun arc de retour.
Preuve : on va en fait montrer que G possède un circuit si et seulement si son parcours en profondeur produit un arc retour.
Si lors du parcours, on a un arc retour uv, alors le graphe G possède un circuit constitué des arcs de liaisons reliant v à u complété de l'arc uv.
Si G possède un circuit $C = v_1, \dots v_k, v_1$, alors supposons (sans perte de généralité) que v_1 soit le premier sommet de C visité.
D'après le théorème du chemin blanc, v_k sera un descendant de v_1 et donc $v_k v_1$ un arc retour.

Tri topologique

Justification de l'algorithme

Montrons maintenant que l'ordre inverse de fin de visite de $PP(G)$ est bien un tri topologique de G. Il faut donc montrer que $uv \in E(G) \Rightarrow f(u) > f(v)$

En effet, soit uv un arc de G.

Si uv est un arc de liaison ou avant, alors v est un descendant de u et donc $f(v) < f(u)$ car $[d(v), f(v)] \subset [d(u), f(u)]$.

Si uv est un arc transverse, alors $d(v) < f(v) < d(u) < f(u)$.

Donc dans tous les cas, on aura bien $uv \in E(G) \Rightarrow f(u) > f(v)$.