MINISTÈRE
DE L'ÉDUCATION
NATIONALE,
DE LA JEUNESSE
ET DES SPORTS

Liberté
Égalité
Fraternité

VOIE GÉNÉRALE

2DE | 1RE | TLE

Numérique et sciences informatiques

ENSEIGNEMENT SPÉCIALITÉ

GÉNÉRALITÉS SUR LES ARBRES

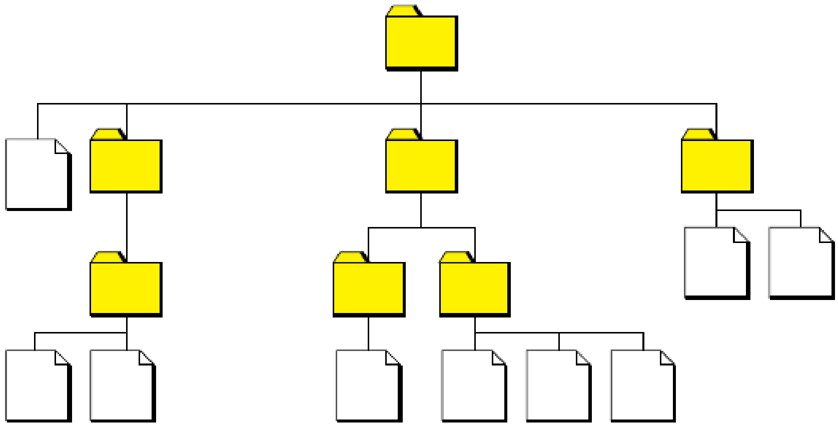
Introduction

Les *arbres* sont des graphes possédant une structure relationnelle particulière appelée *structure hiérarchique*. Ils sont parmi les structures de données les plus importantes et les plus utilisées en informatique, car ils interviennent dans de très nombreux problèmes. On les classe généralement en plusieurs catégories suivant ce qu'ils modélisent :

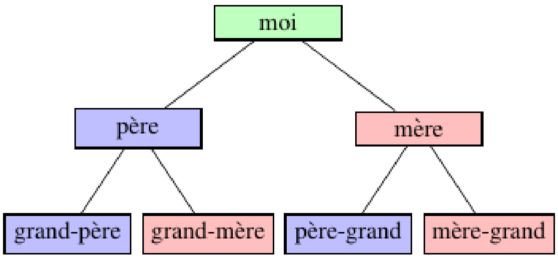
- *arbres de classification* permettant la représentation d'une arborescence de fichiers, du DOM d'une page web… ;
- *arbres généalogiques* descendants ;
- *arbres syntaxiques* permettant la représentation d'expressions arithmétiques, l'analyse d'une phrase… ;
- *arbres de parcours* résultants d'un parcours en largeur ou en profondeur d'un graphe ;
- *arbres couvrants* servant à « capturer » une structure particulière d'un graphe connexe ;
- *arbres lexicographique* permettant de représenter un dictionnaire… ;
- …

Retrouvez éduscol sur

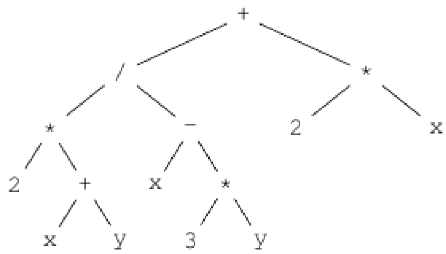
La figure 1 ci-dessous donne trois exemples classiques :



(a) Arbre de classification : arborescence de fichiers



(b) Arbre généalogique



(c) Arbre syntaxique : traduction de l'expression `(2*(x+y))/(x-3*y)+2*x`

FIGURE 1 – Quelques exemples d'utilisation des arbres

Suivant les problèmes à modéliser, les arbres peuvent être de différents types : arbres généraux, arbres binaires, arbres binaires de recherche, B-arbres, arbres rouges et noirs, etc. Dans cette ressource, nous nous intéressons uniquement aux arbres généraux.

Définitions et exemples

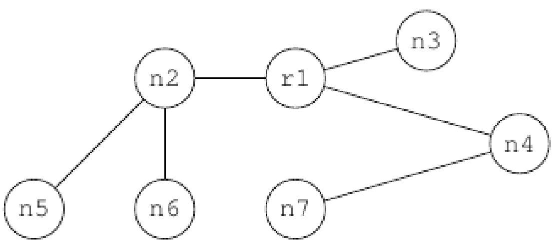
Arbres libres et arbres enracinés

Définition 1 - Arbre libre, nœud et taille

On appelle *arbre libre* un graphe $\mathcal{A}$ non orienté, connexe et acyclique. Les sommets de $\mathcal{A}$ sont communément appelés *nœuds*. Le nombre de nœuds est appelé la *taille* de $\mathcal{A}$ et est noté taille($\mathcal{A}$).

Exemple 1

Le graphe suivant est un arbre libre de taille 7 :



Le résultat suivant donne plusieurs caractérisations importantes d'un arbre libre.

Retrouvez éduscol sur

Théorème 1
Soit $\mathcal{A}$ un graphe non orienté. Les propriétés suivantes sont équivalentes :

1. $\mathcal{A}$ est un arbre libre.
2. Deux sommets quelconques de $\mathcal{A}$ sont reliés par un unique chemin simple.
3. $\mathcal{A}$ est connexe, mais ne l'est plus si on enlève n'importe laquelle de ses arêtes.
4. $\mathcal{A}$ est connexe et son nombre d'arêtes est égal à son nombre de sommets moins 1.
5. $\mathcal{A}$ est acyclique, mais ne l'est plus si on ajoute n'importe quelle arête.
6. $\mathcal{A}$ est acyclique et son nombre d'arêtes est égal à son nombre de sommets moins 1.

Pour prendre en compte toute la spécificité des arbres par rapport aux graphes généraux, et notamment pour stocker plus facilement des informations, il est souvent

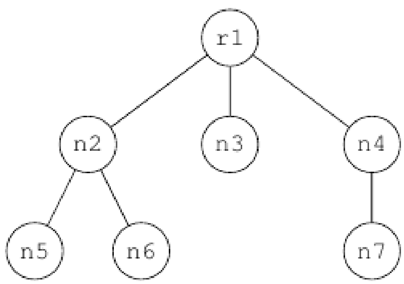
> **Remarques**
> Lorsqu'un graphe est non orienté et acyclique, mais non connexe, chacune de ses composantes connexes est un arbre libre. On dit alors que le graphe est une *forêt*. Ce type de situation intervient souvent lors d'un parcours en profondeur.
> Contrairement aux graphes, la taille d'un arbre désigne son nombre de sommets et non son nombre d'arêtes. Compte tenu du théorème 2.3, le nombre d'arêtes d'un arbre $\mathcal{A}$ est égal à taille($\mathcal{A}$) - 1.

utile (et c'est ce que nous ferons dans la suite) de particulariser un sommet qui va servir "d'ancre" à l'arbre. On a alors la définition suivante :

Définition 2 - Arbre enraciné, racine
On appelle *arbre enraciné*, ou simplement *arbre*, un arbre libre dans lequel l'un des sommets se distingue des autres. Ce sommet particulier est appelé la *racine* de l'arbre.

Pour dessiner un arbre, on procède de la façon suivante : on place sa racine en haut, puis tous les nœuds situés à distance i de la racine à la ligne numéro i. Autrement dit, en informatique, les arbres poussent de haut en bas...

Exemple 2
En prenant comme racine le sommet `r1` dans l'arbre libre de l'exemple 1, on obtient l'arbre suivant :



Exemple 3
- L'arborescence des fichiers d'un système `Unix` est un arbre dont la racine est le dossier `Root`.
- Le DOM d'une page web est un arbre dont la racine est l'élément `html`.
- Un arbre lexicographique permettant de construire un dictionnaire de mots commençant par la lettre A est un arbre de racine A.

Retrouvez éduscol sur

Exercice 1
Quelle est la racine d'un arbre de parcours en largeur ?

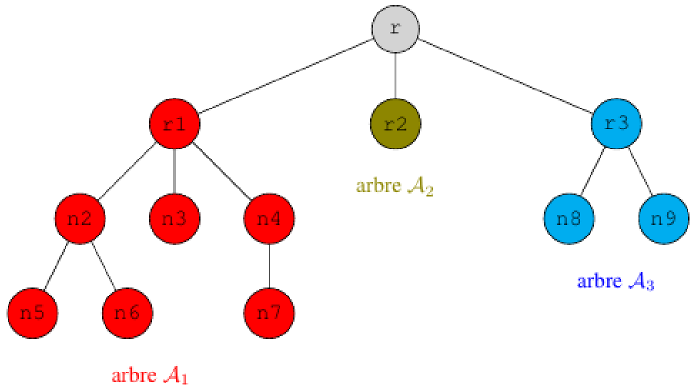
Définition récursive

Il est souvent utile de voir un arbre (enraciné) comme une structure de données récursive :

- *cas de base* : un nœud unique `r` est un arbre de racine `r` ;
- *récurrence* : un arbre avec au moins deux nœuds est constitué d'une racine `r` et d'une suite de sous-arbres $\mathcal{A}1$,… $\mathcal{A}k$ dont les racines `r1`,…, `rk` sont les successeurs de `r`.

Exemple 4

La figure suivante est un arbre de racine r constitué de trois sous-arbres $\mathcal{A}1$, $\mathcal{A}2$ et $\mathcal{A}3$ dont les racines `r1`, `r2` et `r3` sont les successeurs de `r`. L'arbre $\mathcal{A}2$ correspond au cas de base, car il est constitué d'un nœud unique. L'arbre $\mathcal{A}3$ est constitué de deux sous-arbres possédant chacun un nœud unique (ce qui termine la description de $\mathcal{A}3$).



Exercice 2

Décrire récursivement l'arbre $\mathcal{A}1$ de l'exemple 4.

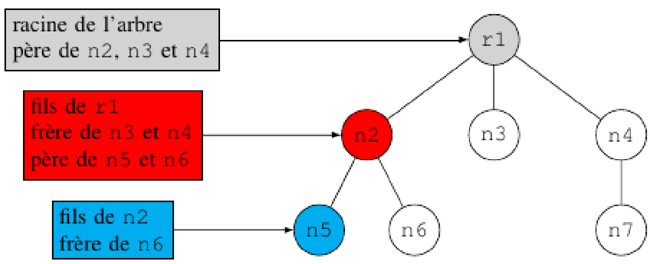
Généalogie

Définition 3 - Père, fils et frères

- Un nœud x est *le père* d'un nœud y si y est un successeur de x.
- Deux nœuds sont *frères* s'ils ont le même père.

Exemple 5

On donne ci-dessous quelques liens de parenté



Remarques

- La racine est le seul nœud de l'arbre n'ayant pas de père.
- Un nœud sans fils est appelé *nœud externe* ou *feuille*. Le nombre de feuilles d'un arbre $\mathcal{A}$ est noté nf($\mathcal{A}$).
- Un nœud qui n'est pas une feuille est appelé *nœud interne*.
- La racine peut être une feuille ou un nœud interne suivant les arbres.

Exercice 3
Donner les feuilles et les nœuds internes de l'arbre de l'exemple 5.

Définition 4 - Ancêtres, descendants et sous-arbres
- On appelle *ancêtres* d'un nœud x tous les nœuds distincts de *x* trouvés sur le chemin simple allant de la racine à *x*.
- Un nœud *y* est appelé *descendant* d'un nœud *x* si *x* est un ancêtre de *y*.
- On appelle *sous-arbre de racine x* l'arbre induit par les descendants de *x*.

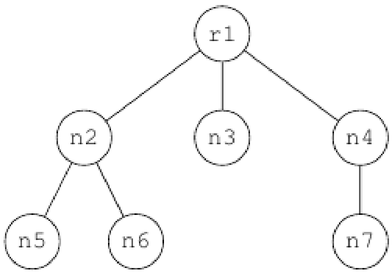
Exemple 6
Dans l'arbre de l'exemple 5 : `r1` et `n2` sont les ancêtres de `n6`, `n5` et `n6` sont les descendants de `n2`. Le sous-arbre de racine `n2` est l'arbre constitué des nœuds `n2`, `n5` et `n6`.

Attributs
Un nœud `n` d'un arbre possède différents attributs permettant de stocker diverses données. On utilise principalement :
- `n.clé` pour désigner la "valeur" contenue dans `n` ;
- `n.père` pour désigner le père de `n` ;
- `n.fils1` pour désigner un fils de `n` ;
- `n.fils2` pour désigner un deuxième fils de `n` ;
- ...

Il peut aussi être utile dans certains cas d'utiliser directement un tableau `n.fils` pour stocker tous les fils de `n`. Bien sûr, suivant ce que l'on désire stocker dans le nœud, d'autres attributs comme des couleurs peuvent être utilisés.

> **Remarque**
> Lorsqu'un attribut est vide, on lui attribue la valeur particulière `null`.

Exercice 4
On considère l'arbre suivant :



Compléter les attributs suivants :

- `r1.père` = ...
- `n2.père` = ...
- `n3.père` = ...
- `r1.fils1` = ...
- `n5.père` = ...
- `n2.fils2` = ...
- `n4.fils` = ...
- `n7.fils` = ...

Implémentation

Puisqu'un arbre est un graphe particulier, on peut utiliser en `Python` une implémentation sous forme de dictionnaire où les clés sont les "valeurs" des différents nœuds et où les valeurs associées sont des listes regroupant les attributs nécessaires (père, fils, etc.).

Pour certains types d'arbres, on utilise une implémentation plus pertinente fondée sur la définition récursive d'un arbre et utilisant la programmation objet (étude ultérieure).

Quelques mesures

Il existe plusieurs mesures classiques sur les arbres. Nous en donnons ici quelques exemples.

Profondeur, niveau et hauteur

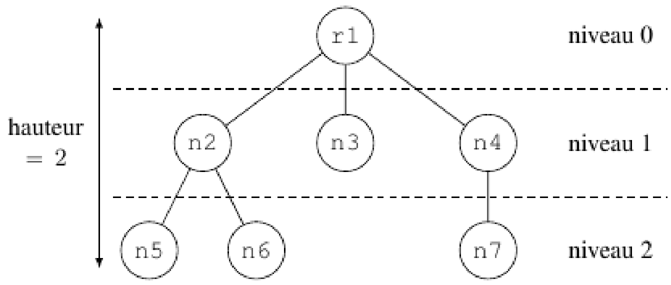
Définition 5

Soit $\mathcal{A}$ un arbre de racine `r`. On appelle :

- *profondeur d'un nœud* `n` la longueur du chemin simple entre `r` et `n` ;
- *niveau* l'ensemble de tous les nœuds de même profondeur ;
- *hauteur d'un nœud* `n` la longueur du chemin simple le plus long qui relie `n` à une feuille dont il est un ancêtre ;
- *hauteur de* $\mathcal{A}$ la hauteur de la racine `r`, c'est-à-dire la longueur du chemin simple le plus long qui relie `r` à une feuille de l'arbre.

Exemple 7

La figure ci-dessus donne les niveaux et la hauteur de l'arbre : `r1` niveau 0 :



Le nœud `n2` a une profondeur de 1 et une hauteur de 1. Le nœud `n3` a une hauteur de 0 puisque c'est une feuille.

Cheminement et profondeur moyenne

Définition 6

Soit $\mathcal{A}$ un arbre de racine `r`. On appelle :

- *longueur de cheminement de* $\mathcal{A}$ la somme LC($\mathcal{A}$) des longueurs de tous les chemins issus de `r` ;
- *longueur de cheminement externe de* A la somme LCE($\mathcal{A}$) des longueurs de tous les chemins issus de `r` et aboutissant à une feuille.

Définition 7

Soit $\mathcal{A}$ un arbre de racine `r`. On appelle :

- *profondeur moyenne d'un nœud de* $\mathcal{A}$ la moyenne PM($\mathcal{A}$) des profondeurs de tous les nœuds de $\mathcal{A}$:

$$PM(A) = \frac{LC(A)}{taille(A)}$$

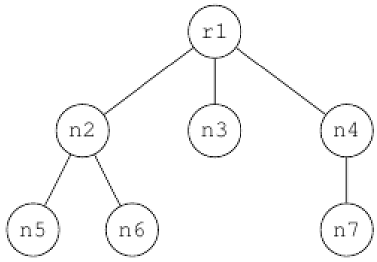
- *profondeur moyenne externe d'une feuille de* $\mathcal{A}$ la moyenne PME($\mathcal{A}$) des profondeurs de toutes les feuilles de $\mathcal{A}$ *:*

$$PME(A) = \frac{LCE(A)}{nf(A)}$$

Exemple 8

Dans l'arbre A suivant, on a les mesures :

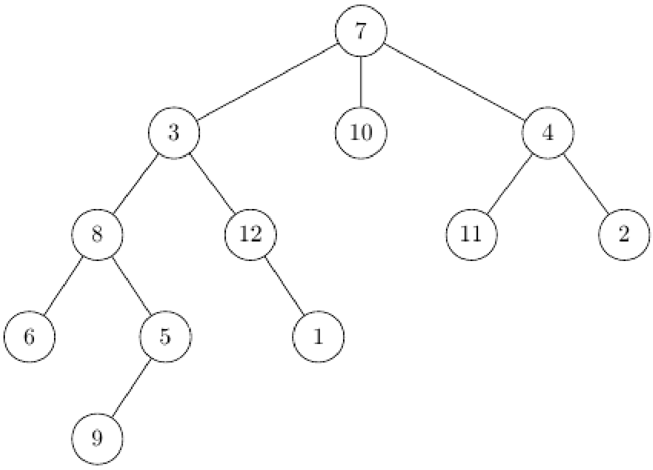
- taille($\mathcal{A}$) = 7
- nf($\mathcal{A}$) = 4
- LC($\mathcal{A}$) = 9
- LCE($\mathcal{A}$) = 7
- PM($\mathcal{A}$) =
- PME($\mathcal{A}$) =



Retrouvez éduscol sur

Exercice 5

Considérons l'arbre A suivant :



1. Déterminer la taille, le nombre de feuilles et le nombre de nœuds internes de $\mathcal{A}$.
2. Quels sont les nœuds de niveau 3 ?
3. Quelles sont la profondeur et la hauteur des nœuds 12, 10 et 3 ?
4. Quelle est la hauteur de l'arbre $\mathcal{A}$?
5. Quelle est la longueur de cheminement de $\mathcal{A}$? En déduire la profondeur moyenne d'un nœud de $\mathcal{A}$.
6. Quelle est la profondeur moyenne externe d'une feuille de $\mathcal{A}$?