Classe : terminale

LYCÉE INTERNATIONAL COURS LUMIÈRE

séquence 2

STRUCTURES DE DONNEES : LES ARBRES

Mr BANANKO K.

Objectif :

Contenus	Capacités attendues	Commentaires
Arbres : structures hiérarchiques. Arbres binaires : nœuds, racines, feuilles, sous-arbres gauches, sous-arbres droits.	Identifier des situations nécessitant une structure de données arborescente. Évaluer quelques mesures des arbres binaires (taille, encadrement de la hauteur, etc.).	On fait le lien avec la rubrique « algorithmique ».

I- NOTION D'ARBRE

Nous allons travailler sur la structure de donnée "arbre". Cette structure n'est pas linéaire mais **hiérarchique**.

Un organisateur de tournoi de rugby recherche la meilleure solution pour afficher les potentiels quarts de final, demi-finales et finale :

Au départ nous avons 4 poules de 4 équipes. Les 4 équipes d'une poule s'affrontent dans un mini championnat (3 matchs par équipe). À l'issue de cette phase de poule, les 2 premières équipes de chaque poule sont qualifiées pour les quarts de finale.

Dans ce qui suit, on désigne les 2 qualifiés par poule par :

- Poule 1 => 1er Eq1 ; 2e Eq8
- Poule 2 => 1er Eq2 ; 2e Eq7
- Poule 3 => 1er Eq3 ; 2e Eq6
- Poule 4 => 1er Eq4 ; 2e Eq5

En quart de final on va avoir :

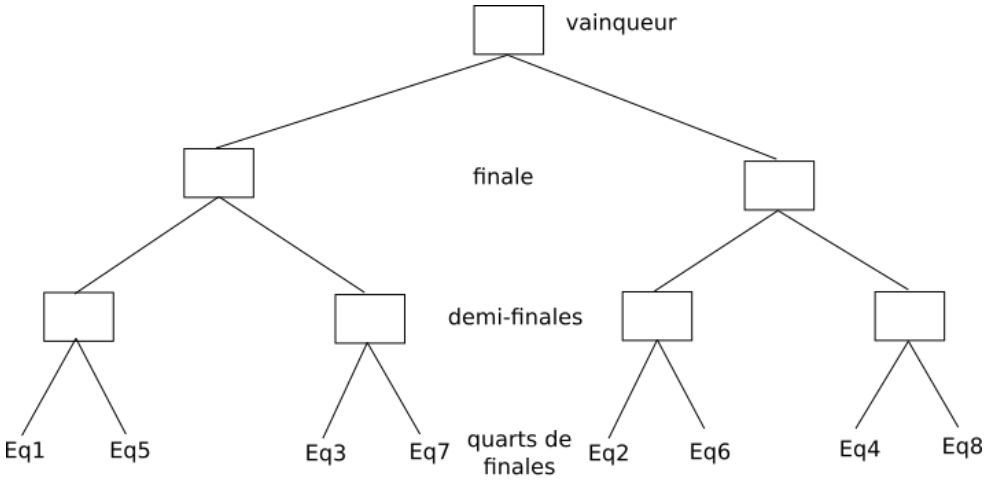
- quart de finale 1 => Eq1 contre Eq5
- quart de finale 2 => Eq2 contre Eq6
- quart de finale 3 => Eq3 contre Eq7
- quart de finale 4 => Eq4 contre Eq8

Pour les demi-finales on aura :

- demi-final 1 => vainqueur quart de finale 1 contre vainqueur quart de finale 3
- demi-final 2 => vainqueur quart de finale 2 contre vainqueur quart de finale 4

L'organisateur du tournoi affiche les informations ci-dessus le jour du tournoi. Malheureusement, la plupart des spectateurs se perdent quand ils cherchent à déterminer les potentielles demi-finales (et ne parlons pas de la finale !)

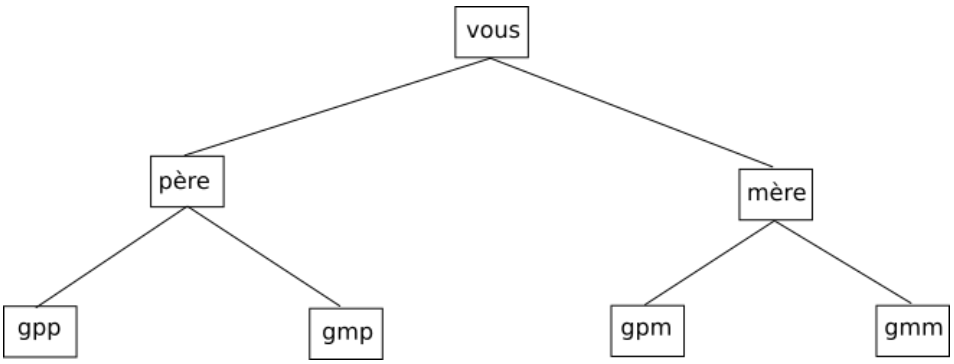
Pourtant, un simple graphique aurait grandement simplifié les choses :



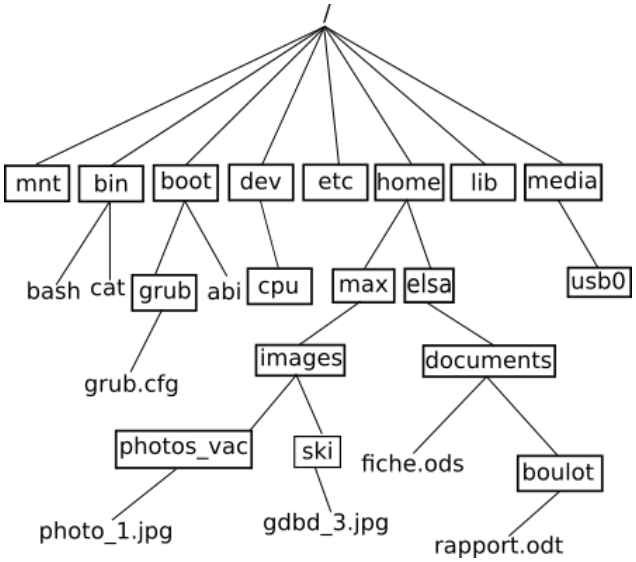
Poule 1 => 1er Eq1 ; 2e Eq8
Poule 2 => 1er Eq2 ; 2e Eq7
Poule 3 => 1er Eq3 ; 2e Eq6
Poule 4 => 1er Eq4 ; 2e Eq5

Les spectateurs peuvent alors recopier sur un bout de papier ce schéma et ensuite se livrer au jeu des pronostics.

Nous avons ci-dessous ce que l'on appelle une structure en arbre. On peut aussi retrouver cette même structure dans un arbre généalogique :



Dernier exemple, les systèmes de fichiers dans les systèmes de type UNIX ont aussi une structure en arbre (notion vue l'année dernière)



Système de fichiers de type UNIX

Définition :

Un arbre est une structure de données constituée de **nœuds**.

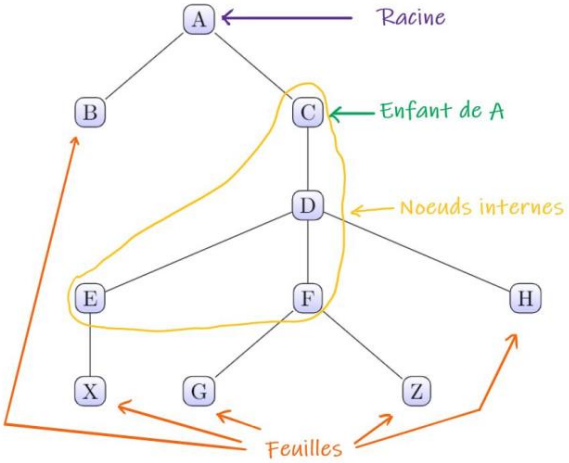
Le sommet de l'arbre s'appelle la **racine**.

Le nœud B situé sous un nœud A est appelé **enfant** du nœud A

Un nœud qui ne possède pas d'enfant est appelé **feuille**.

Les nœuds autre que la racine et les feuilles sont appelés **nœuds internes**.

Une **branche** est une suite de nœuds consécutifs de la racine vers une feuille.



Dans cet arbre:

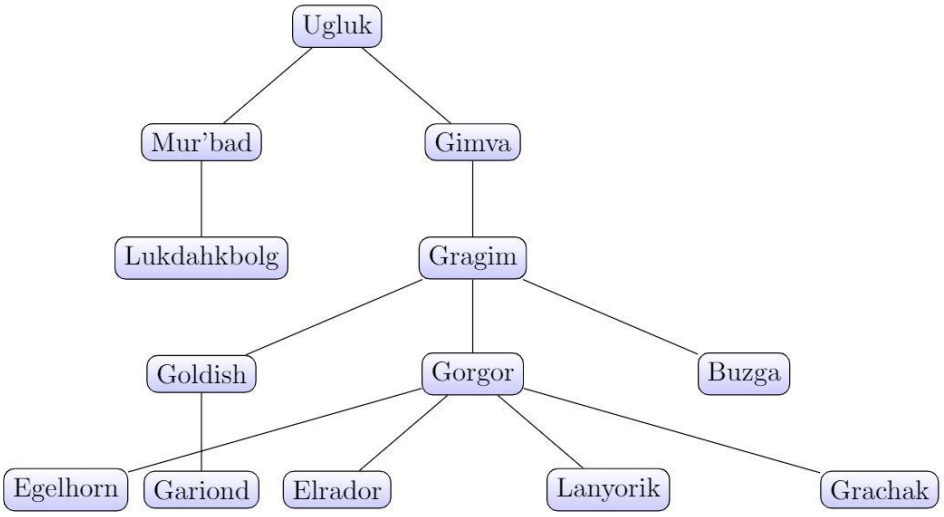
- il y a 10 nœuds.
- Le nœud A est la racine.
- Il y a 5 feuilles donc 5 branches.
- Le nœud D est l'enfant du nœud C.
- Il y a 4 nœuds internes.

Les arbres sont des types abstraits très utilisés en informatique. On les utilise notamment quand on a besoin d'une structure hiérarchique des données : dans l'exemple ci-dessous le fichier grub.cfg ne se trouve pas au même niveau que le fichier rapport.odt (le fichier grub.cfg se trouve "plus proche" de la racine / que le fichier rapport.odt). On ne pourrait pas avec une simple liste qui contiendrait les noms des fichiers et des répertoires, rendre compte de cette hiérarchie (plus ou moins "proche" de la racine).

On trouve souvent dans cette hiérarchie une notion de temps (les quarts de finale sont avant les demi-finales ou encore votre grand-mère paternelle est née avant votre père), mais ce n'est pas une obligation (voir l'arborescence du système de fichiers).

Activité 1

On donne l'arbre suivant :



1. Quelle est la racine de cet arbre ?
2. Combien y-a-t-il de nœuds ?
3. Combien y-a-t-il de feuilles ?
4. Combien y-a-t-il de branches ?
5. L'ensemble des nœuds internes compte combien d'éléments ?
6. Quels sont les enfants de Gragim ?

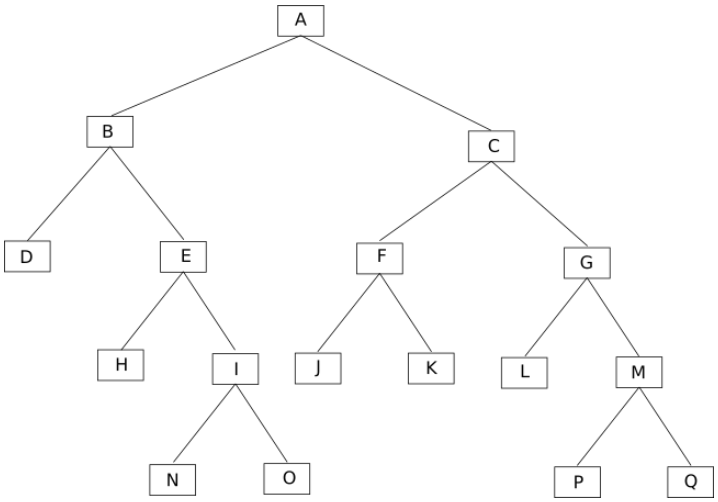
II- LES ARBRES BINAIRES

a) INTRODUCTION

Les arbres binaires sont des cas particuliers d'arbre : l'arbre du tournoi de rugby et l'arbre "père, mère..." sont des arbres binaires, en revanche, l'arbre représentant la structure du système de fichier n'est pas un arbre binaire. Dans un arbre binaire, on a au maximum 2 **arête** qui partent d'un élément (pour le système de fichiers, on a 7 branches qui partent de la racine : ce n'est donc pas un arbre binaire).

Dans la suite nous allons uniquement travailler sur les arbres binaires.

Soit l'arbre binaire suivant :



b) UN PEU DE VOCABULAIRE

- chaque élément de l'arbre est appelé **nœud** (par exemple : A, B, C, D,...,P et Q sont des nœuds)
- le nœud initial (A) est appelé **nœud racine** ou plus simplement **racine**
- On dira que le nœud E et le nœud D sont **les fils** du nœud B. On dira que le nœud B est le père des nœuds E et D
- Dans un arbre binaire, un nœud possède au plus 2 **fils**
- Un nœud n'ayant aucun **fils** est appelé **feuille** (exemples : D, H, N, O, J, K, L, P et Q sont des feuilles)
- À partir d'un nœud (qui n'est pas une feuille), on peut définir un **sous-arbre gauche** et un **sous-arbre droite** (exemple : à partir de C on va trouver un sous-arbre gauche composé des nœuds F, J et K et un sous-arbre droit composé des nœuds G, L, M, P et Q)
- On appelle **arête** le segment qui relie 2 nœuds.
- On appelle **taille** d'un arbre le nombre de nœuds présents dans cet arbre
- On appelle **profondeur** d'un nœud ou d'une feuille dans un arbre binaire le nombre de nœuds du chemin qui va de la racine à ce nœud. La racine d'un arbre est à une profondeur 1, et la profondeur d'un nœud est égale à la profondeur de son prédécesseur plus 1. Si un nœud est à une profondeur p, tous ses successeurs sont à une profondeur p+1.

  Exemples : profondeur de B = 2 ; profondeur de I = 4 ; profondeur de P = 5

  ATTENTION : on trouve aussi dans certains livres la profondeur de la racine égale à 0 (on trouve alors : profondeur de B = 1 ; profondeur de I = 3 ; profondeur de P = 4). Les 2 définitions sont valables, il faut juste préciser si vous considérez que la profondeur de la racine est de 1 ou de 0.

- On appelle **hauteur** d'un arbre la profondeur maximale des nœuds de l'arbre. Exemple : la profondeur de P = 5, c'est un des nœuds les plus profonds, donc la hauteur de l'arbre est de 5.

ATTENTION : comme on trouve 2 définitions pour la profondeur, on peut trouver 2 résultats différents pour la hauteur : si on considère la profondeur de la racine égale à 1, on aura bien une hauteur de 5, mais si l'on considère que la profondeur de la racine est de 0, on aura alors une hauteur de 4.

!!! Ainsi un *arbre vide* a pour hauteur 0.

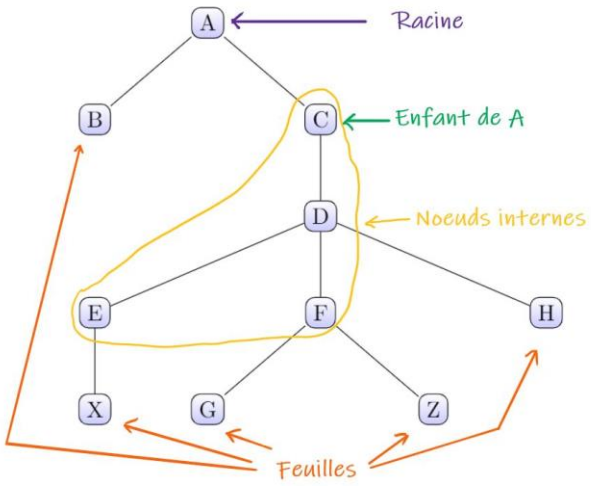
Définition	**Hauteur de l'arbre avec un nœud**	**Hauteur de l'arbre vide**
Par le nombre d'arêtes	0	−1
Par le nombre de nœuds	1	0

On peut caractériser un arbre par différentes caractéristiques :

- Son **arité** : le nombre maximal d'enfants qu'un nœud peut avoir.

Exemple :

On reprend l'arbre de la définition d'un arbre :



- L'arité de cet arbre est 3.
- La taille de cet arbre est 10.
- La hauteur de cet arbre est 4.

Remarque : On appelle **arbre binaire** un arbre d'arité inférieure ou égale à 2.

c) STRUCTURE RECURSIVE

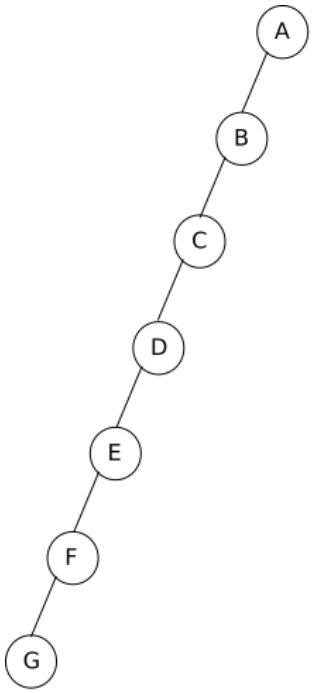
Il est aussi important de bien noter que l'on peut aussi voir les arbres comme des structures récursives : les fils d'un nœud sont des arbres (sous-arbre gauche et un sous-arbre droite dans le cas d'un arbre binaire), ces arbres sont eux-mêmes constitués d'arbres.

On peut alors définir un **arbre binaire** de façon *récursive* :

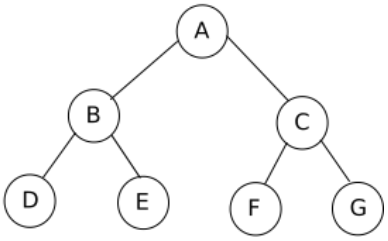
- soit c'est **l'arbre vide** s'il ne contient aucun noeud
- soit il est constitué d'un noeud spécial appelé **racine** de l'arbre dont le *fils droit* et le *fils gauche* sont des **arbres binaires**

d) ENCADREMENT DE LA HAUTEUR D'UN ARBRE

Il est possible d'avoir des arbres binaires de même taille mais de "forme" très différente :



Arbre 1



Arbre 2

Sur le schéma ci-dessus l'arbre 1 est dit " **filiforme**" alors que l'arbre 2 est dit "**complet**" (on dira qu'un arbre binaire est complet si tous les nœuds possèdent 2 fils et que toutes les feuilles se situent à la même profondeur). On pourra aussi dire que l'arbre 1 est **déséquilibré** alors que l'arbre 2 est équilibré.

Si on prend un arbre filiforme de taille n, on peut dire que la hauteur de cet arbre est égale à $n-1$ (si on prend la définition de la hauteur d'un arbre où la racine a une profondeur 0)
Si on prend un arbre complet de taille n, on peut démontrer que la hauteur de cet arbre est égale à la partie entière de $log_2(n)$ (on arrondit à l'entier immédiatement inférieur le $log_2(n)$)). Dans le cas de l'arbre 2,
Nous avons $log_2(7) = 2{,}8$ donc en prenant la partie entière on a bien la hauteur de l'arbre 2 égale à 2.

Un arbre filiforme et un arbre complet étant deux cas extrêmes, on peut affirmer que pour un arbre binaire quelconque

$$\lfloor log_2(n) \rfloor \leq h \leq n-1$$

Avec n la taille de l'arbre et h la hauteur de l'arbre ($\lfloor log_2(n) \rfloor$ permet de prendre la partie entière du logarithme base 2 de n)

Démonstration plus bas 

e) LES ARBRES BINAIRES EN PYTHON

Python ne propose pas de façon native l'implémentation des arbres binaires. Mais nous aurons, plus tard dans l'année, l'occasion d'implémenter des arbres binaires en Python en utilisant la programmation orientée objet.

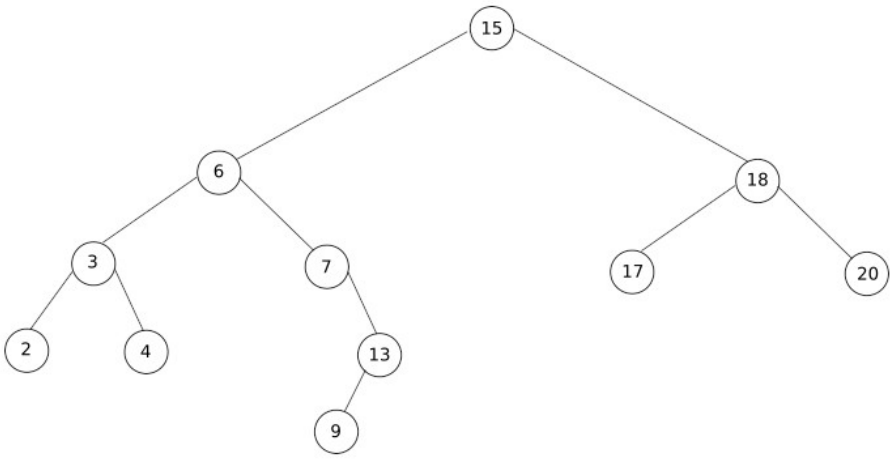
III- les arbres binaires de recherche

Un arbre binaire de recherche est un cas particulier d'arbre binaire.

Un **arbre binaire de recherche** est un arbre binaire dans lequel l'étiquette d'un nœud est appelé **clé** et est un entier. L'arbre binaire vérifie deux propriétés :

- Les clés de tous les nœuds du sous-arbre gauche d'un nœud x sont inférieures ou égales à la clé de x.
- il faut que les clés de nœuds composant l'arbre soient ordonnables (on doit pouvoir classer les nœuds, par exemple, de la plus petite clé à la plus grande)
- soit x un nœud d'un arbre binaire de recherche. Si y est un nœud du sous-arbre gauche de x, alors il faut que y.clé $\leqslant$ x.clé. Si y est un nœud du sous-arbre droit de x, il faut alors que x.clé $\leqslant$ y.clé
- Les clés de tous les nœuds du sous-arbre droit d'un nœud x sont strictement supérieures à la clé de x.

Exemple d'arbre binaire de recherche :

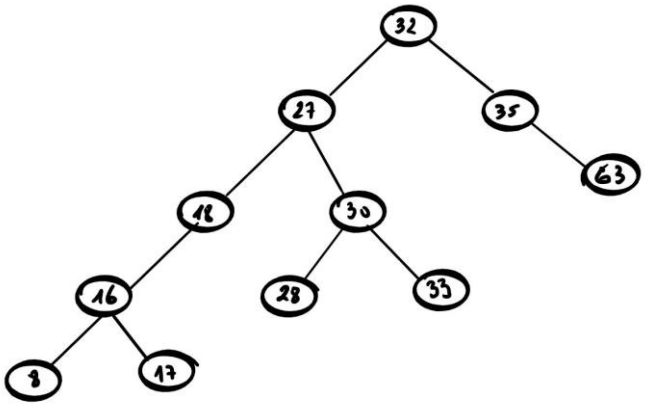


Vous pouvez vérifier que le fils gauche d'un nœud a une valeur plus petite que son père (par exemple 3 < 6) et que le fils droit d'un nœud a une valeur plus grande que son père (par exemple 7 > 6)

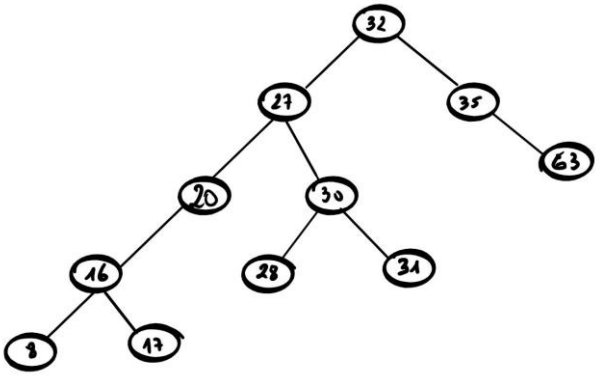
Attention : pour un nœud donné A, tous les nœuds de l'arbre gauche de A auront des valeurs plus petites que la valeur du nœud A et tous les nœuds de l'arbre droit de A auront des valeurs plus grandes que la valeur du nœud A.

Exemple :

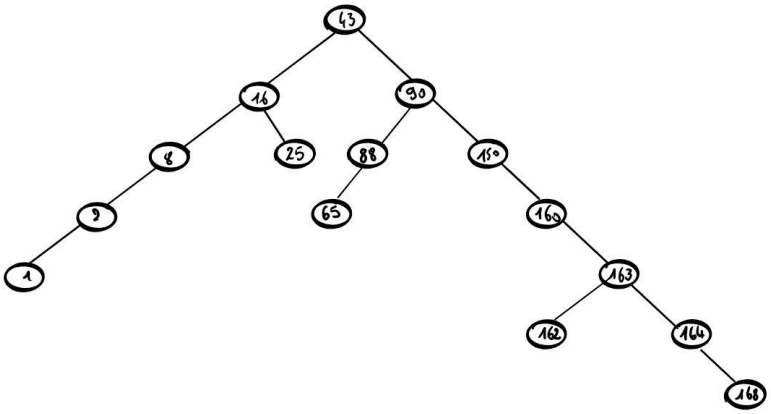
Quels sont parmi les arbres proposés ci-dessous les arbres binaires de recherche?



1



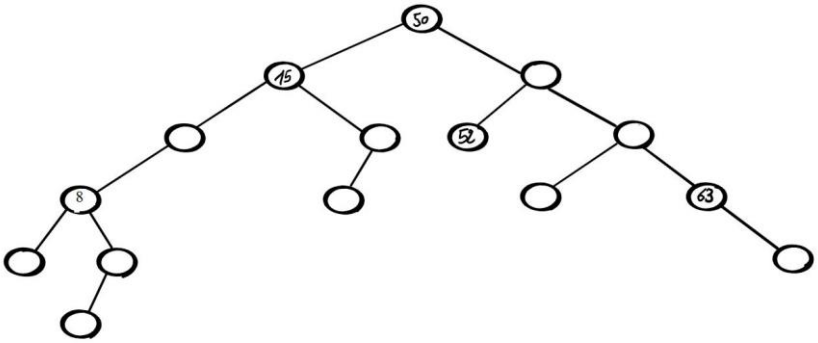
2



3

Exercice

Compléter cet arbre pour que ce soit un arbre binaire de recherche :

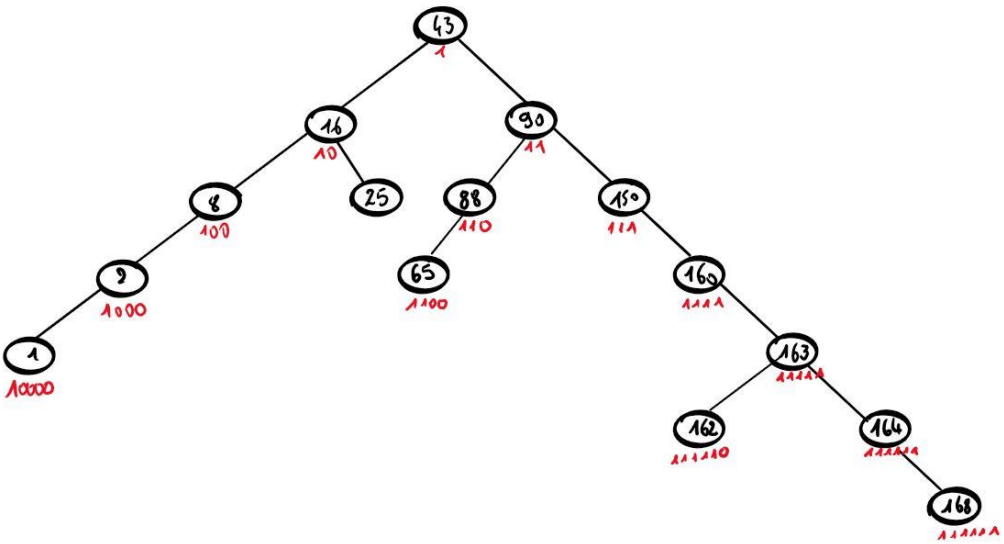


Propriété :

Un **étiquetage intéressant** d'un arbre de recherche est le suivant :

1. La racine est étiquetée 1.
2. La premier nœud du sous arbre gauche prend l'étiquette de son père auquel on ajoute un 0.
3. La premier nœud du sous arbre droit prend l'étiquette de son père auquel on ajoute un 1.

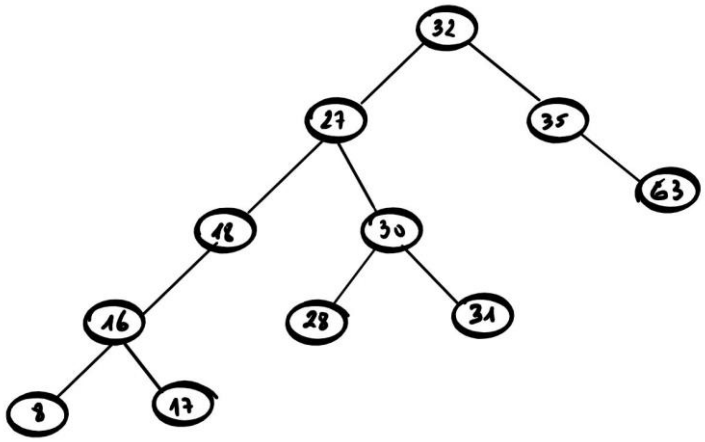
Exemple :



Quelle est l'étiquette de 25 dans l'arbre précédent ?

Exercice :

Etiqueter comme l'exemple précédent l'arbre suivant :



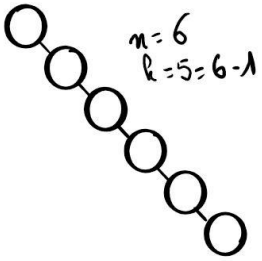
Propriété :

Pour un arbre binaire de recherche de hauteur h et de taille n les inégalités : $\lfloor log_2(n) \rfloor \leq h \leq n-1$

Pour démontrer ce résultat nous allons observer le cas où l'arbre est le plus profond et le cas où l'arbre est le moins profond.

- Cas où l'arbre est le plus profond :

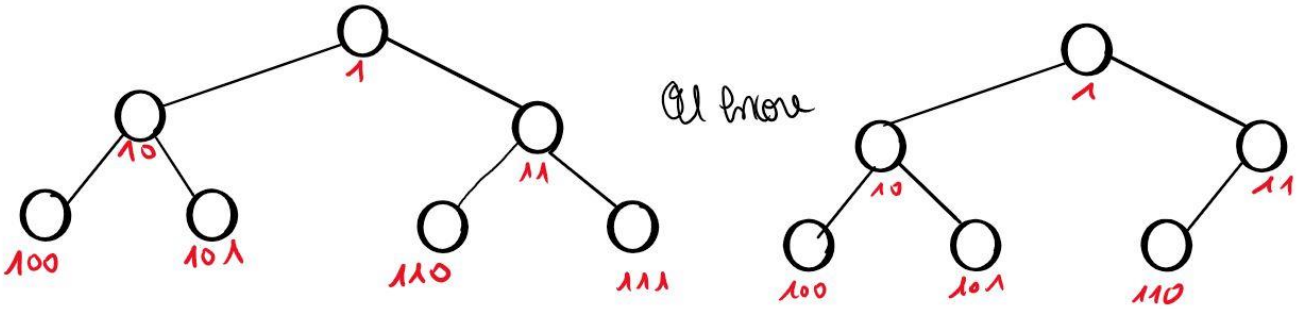
Pour une taille n fixée, la hauteur maximale d'un arbre binaire est $h = n - 1$, qu'on obtient avec des arbres « filiformes » comme cet arbre :



Ainsi $h \leq n - 1$

- Cas où l'arbre est le moins profond

Les arbres de taille n de hauteur minimale sont les arbres comme ces arbres :



Les clés des feuilles sont supérieures ou égales en binaire à 100...0100...0 où 0 est répété p$\square$ fois

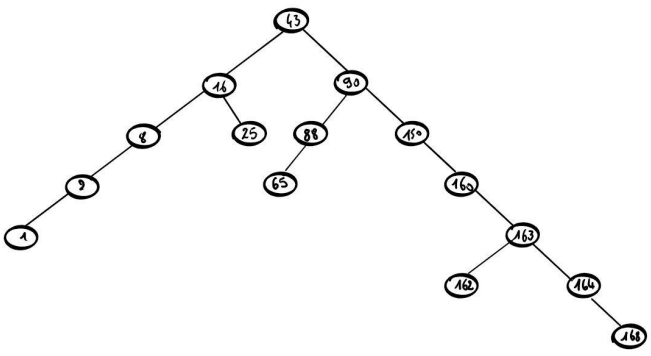
Dans ce cas la hauteur h est égale à p, il y a au moins n=2^p nœuds

Or $log_2(2^p) = p$

Ainsi $log_2(n) \leq p$

Exemple :

Vérifions cette inégalité sur cet arbre :



La taille de cet arbre est 15.

La hauteur est 6. $log_2(15) = 3{,}9$ à 10^{-1} $près$

⌊3,9⌋=3

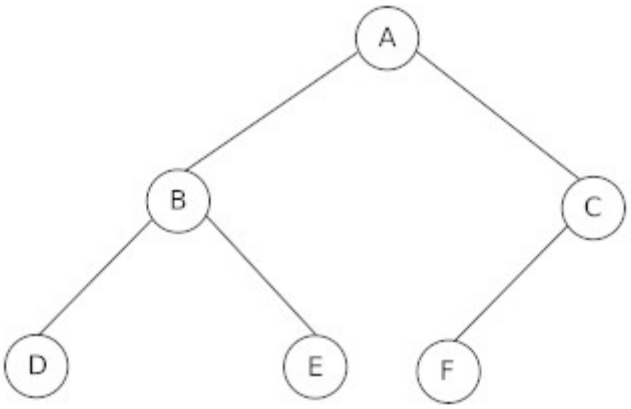
On a bien : 3≤6≤14

Activité 1

Trouvez un autre exemple de données qui peuvent être représentées par un arbre binaire (dans le domaine de votre choix). Dessinez au moins une partie de cet arbre binaire. Déterminez la hauteur et la taille de l'arbre que vous aurez dessiné.

Activité 2

Soit l'arbre binaire suivant :



1) Cet arbre est-il un arbre binaire ? Justifiez votre réponse.

2) Donnez la clé (valeur) de la racine de cet arbre.

3) Quels sont les fils du nœud B.

4) Donnez l'arbre droit du nœud A.

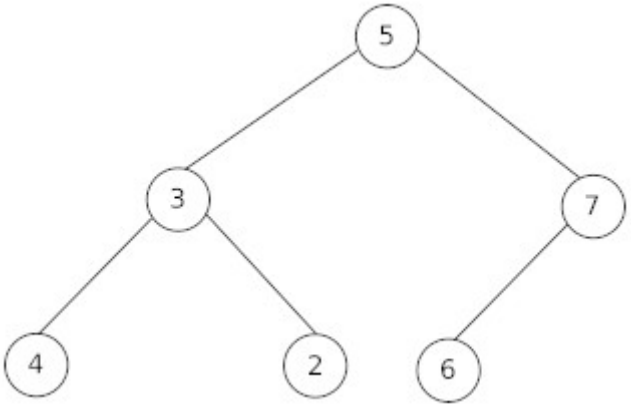
5) Le nœud C est-il une feuille ? Justifiez votre réponse.

6) Donnez la taille de cet arbre.

7) Donnez la profondeur du nœud B (on prendra la profondeur de la racine égale à 0).

8) Donnez la hauteur de cet arbre (on prendra la profondeur de la racine égale à 0).

Activité 3



1) Expliquez pourquoi cet arbre binaire n'est pas un arbre binaire de recherche.

2) Modifiez cet arbre pour le transformer en arbre binaire de recherche.

Activité 4

Soit les valeurs suivantes : 14, 22, 8, 47, 42, 13, 1, 24, 33, 74.

Construisez un arbre binaire de recherche à partir de ces valeurs.