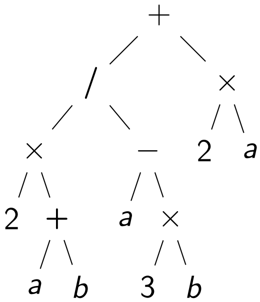
Les Arbres

Sandrine Vial
sandrine.vial@uvsq.fr

Novembre 2020

Les arbres

Une des structures les plus importantes et les plus utilisées en informatique



- Arbres généalogiques
- Arbres de classification
- Arbres d'expression

Représentation de l'expression

$$(2 \times (a + b))/(a - 3 \times b) + 2 \times a$$

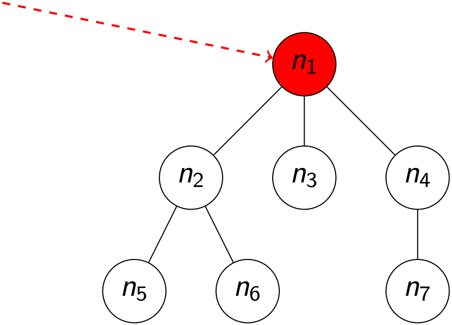
Terminologie

- Un **arbre** : un ensemble de nœuds reliés entre eux par des arêtes.
- Trois propriétés pour les arbres **enracinés** :
  1. Il existe un nœud particulier nommé **racine**. Tout nœud c autre que la racine est relié par une arête à un nœud p appelé **père** de c.
  2. Un arbre est **connexe**.
  3. Un arbre est **sans cycle.**

Terminologie

- Un nœud peut avoir 0 ou plusieurs fils.
- Un nœud a exactement un père.

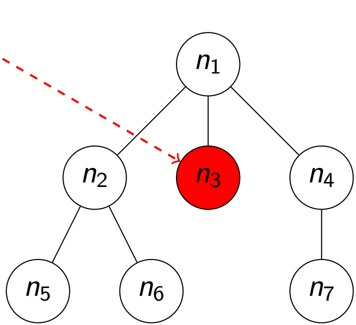
La racine



Terminologie

- Un nœud peut avoir 0 ou plusieurs fils.
- Un nœud a exactement un père.

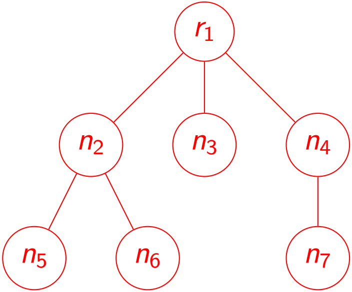
La racine



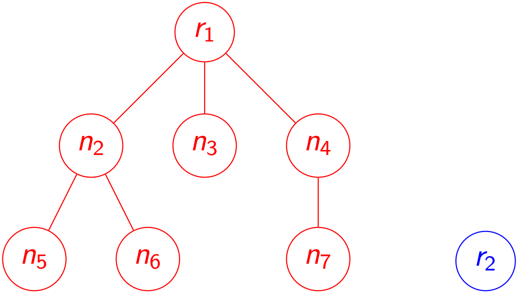
Définition récursive

- **Base :**
  - Un nœud unique n est un arbre
  - n est la racine de cet arbre.
- **Récurrence :**
  - Soit r un nouveau nœud
  - $T_1, T_2, \ldots, t_k$ sont des arbres ayant pour racine $r_1, r_2, \ldots, r_k$.
  - Création d'un nouvel arbre ayant pour racine r et on ajoute une arête entre r et r_1 r et r_2, …, r et r_k.

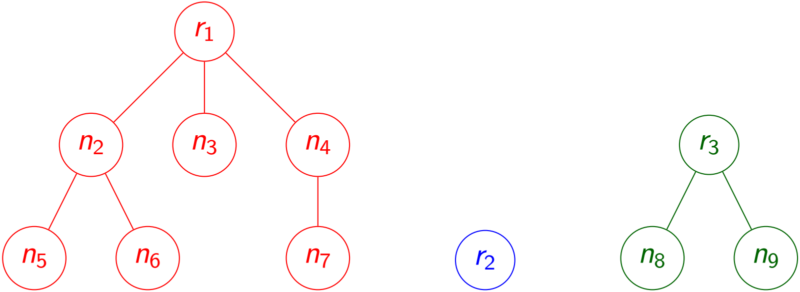
Définition récursive : un exemple



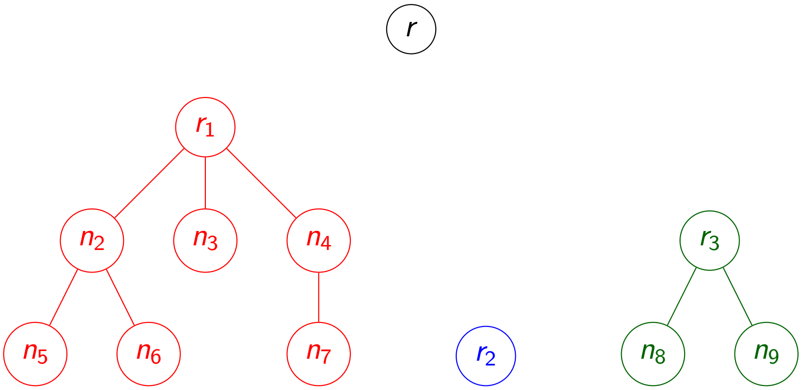
Définition récursive : un exemple



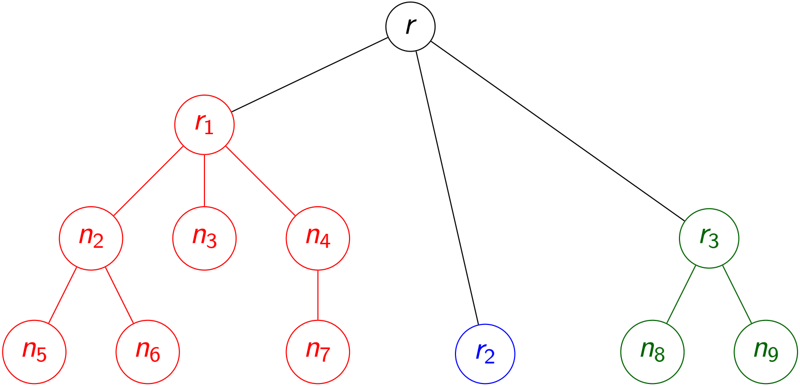
Définition récursive : un exemple



Définition récursive : un exemple



Définition récursive : un exemple



Propriétés d'un arbre

Pour un arbre T a n sommets il y a équivalence entre les propriétés :

- T est un arbre
- T est connexe, et la suppression de toute arête le rend non connexe
- T est acyclique à $n-1$ arêtes
- T est acyclique et l'ajout de toute arête le rend cyclique.

Chemins, ancêtres, descendants, . . .

- Les **ancêtres** d un nœud : *Nœuds trouvés sur le **chemin unique** entre ce nœud et la racine*.
- Le nœud d est un **descendant** de a si et seulement si a est un ancêtre de d.
- **Longueur** d'un chemin = nombre d'arêtes parcourues.

Généalogie

- La racine est un ancêtre de tous les nœuds.
- Chaque nœud est un descendant de la racine.
- Les nœuds ayant le même parent = **enfants**.
- Un nœud n et tous ses descendants = **sous-arbre**

Feuilles et nœuds intérieurs

- Une `feuille` est un nœud qui n'a pas d'enfants
- Un nœud intérieur est un nœud qui a au moins 1 enfant.
- Tout nœud de l'arbre est :
  - Soit une feuille
  - Soit un nœud intérieur

Mesures sur les arbres

- **Taille** de l'arbre T, notée *taille(T)* = nombre de nœuds.
- **Nombre de feuilles** noté *nf(T)*.
- **Longueur de cheminement** de l'arbre T, notée *LC(T)* = somme des longueurs de tous les chemins issus de la racine.

$$LC(T) = \sum_{x \text{ nœud de } T} h(x).$$

- **Longueur de cheminement externe** de l'arbre T, notée *LCE(T)* = somme des longueurs de tous les chemins aboutissant à une feuille issus de la racine.

$$LCE(T) = \sum_{x \text{ feuille de } T} h(x).$$

Hauteur

- La **hauteur d'un nœud** n, notée $h(n)$, est la longueur du chemin depuis la racine jusqu'à n.
- La **hauteur de l'arbre** T, notée $h(T)$:

$$h(T) = \max_{x \text{ nœud de l'arbre}} h(x)$$

Mesures

- **Hauteur moyenne** de l'arbre T, notée *HM(T)*= moyenne des hauteurs de tous les nœuds.

$$HM(T) = \frac{LC(T)}{taille(T)}$$

- **Hauteur moyenne externe** de l'arbre T, notée *HME(T)* = moyenne des longueurs de tous les chemins issus de la racine et se terminant par une feuille.

$$HME(T) = \frac{LCE(T)}{nf(T)}$$