7. Les arbres : parcours préfixé, postfixé et infixé

Définition

Un *arbre* est une structure qui est :

- soit vide[1],
- soit composée d'un *nœud* chaîné à zéro un ou plusieurs sous-arbres ordonnés de gauche à droite.

Un sous-arbre est donc un nœud,
la *racine* est le nœud qui **n'est pas** un sous-arbre,
une *feuille* est un nœud qui **n'a pas** de sous-arbre.

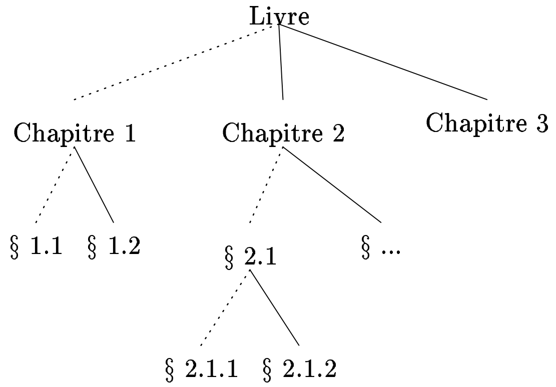

FIG. 7.1 – Une table des matières est un arbre

Étiquettes et expressions

L'arbre *étiqueté* de la figure 7.2 représente l'expression arithmétique $(a+b)*(a+c)$. Les noms attribués aux nœuds sont $n_1, n_2, \ldots, n_7$ et les étiquettes apparaissent, comme c'est l'usage, à côté des nœuds. Les règles imposées à un arbre pour qu'il représente une expression sont les suivantes :

1. Chaque feuille est étiquetée par un opérande et est constituée de cet opérande uniquement. Ainsi la feuille n_4 représente l'expression a.
2. Chaque nœud interne n est étiqueté par un opérateur. Si n est étiqueté par un opérateur binaire θ, comme $+$ ou $*$, si son fils gauche représente l'expression E_1 et son fils droit l'expression E_2, alors n représente l'expression $(E_1)\theta(E_2)$. Les parenthèses peuvent être retirées si aucune ambiguïté n'est à craindre.

Par exemple, l'opérateur $+$ est associé au nœud n_2 et ses fils gauche et droit représentent a et b respectivement. Ainsi e_2 représente $(a) + (b)$, ou simplement $a + b$. Le nœud e_1 représente $(a+b)*(a+c)$, puisque $*$ est l'étiquette associée à e_1 et puisque $a+b$ et $a+c$ sont les expressions représentées par e_2 et e_3 respectivement.

Parcours d'arbres

Il existe plusieurs moyens de parcourir les nœuds d'un arbre. Les trois parcours les plus importants sont les parcours *préfixé*, *postfixé* et *infixé* ; ces parcours peuvent être définis récursivement.

Il existe un moyen pratique pour simuler les trois parcours d'arbre : imaginons que l'on parcourt l'arbre depuis sa racine, dans le sens trigonométrique, en en restant toujours le plus près possible.

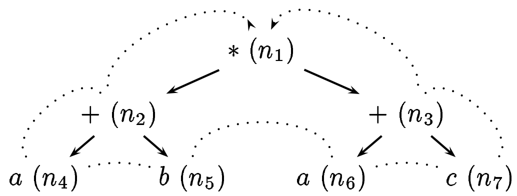

FIG. 7.2 – *L'arbre d'une expression*

Dans un parcours préfixé, on ne considère que le *premier* passage par un nœud donné $(* + ab + ac)$; dans un parcours postfixé, on ne prend en compte que le *dernier* passage par un nœud, lors de la remontée vers son père $(ab + ac + *)$. Pour un parcours infixé, on liste une feuille la première fois qu'on la rencontre, mais on ne liste un nœud non terminal qu'à la deuxième rencontre : $(a + b) * (a + c)$.

1. que nous noterons Λ. Concrètement, Λ pourra être 0, -1 ou tout autre valeur significative dans un contexte particulier.