

Probabilités, Statistiques, Combinatoire : TD2

Combinatoire : codage des arbres binaires

Tous les arbres mentionnés dans les exercices de cette feuille sont des arbres binaires complets, même si l'énoncé ne le précise pas. Tous les mots sont sur l'alphabet $\{a, b\}$, même si l'énoncé ne le précise pas.

2.1 Définitions du cours et révisions

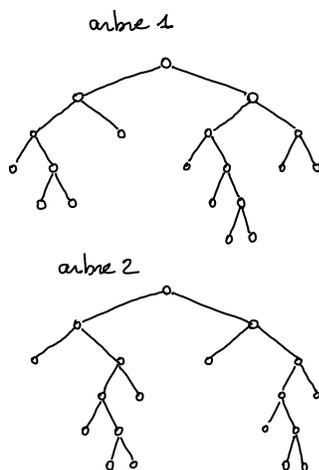
Les exercices de cette section ont pour but de faire pratiquer les définitions et leur application. Ils ne sont censés présenter aucune difficulté si les définitions sont connues.

Exercice 2.1

On rappelle la définition du codage par un mot d'un arbre binaire complet, vue en cours :

- l'arbre réduit à une feuille, est codé par le mot vide ;
- un arbre t non réduit à une feuille, dont les sous-arbres gauche et droit t_1 et t_2 sont codés respectivement par les mots w_1 et w_2 , est codé par le mot $a.w_1.b.w_2$.

Calculer les mots codant les arbres suivants.



Exercice 2.2

Parmi les arbres binaires complets, on distingue certains arbres particuliers :

- un arbre est un *peigne droit* si chacun de ses nœuds internes a pour fils gauche une feuille ;
- un arbre est un *peigne gauche* si chacun de ses nœuds internes a pour fils droit une feuille ;
- un arbre est *parfait* si toutes ses feuilles sont à la même hauteur.

Dessiner les peignes gauches et peignes droits à 4, 5 et 6 nœuds internes, et calculer leurs mots de codage. Proposer une formule pour le mot de codage d'un peigne droit et d'un peigne gauche, en général.

Dessiner les arbres parfaits de hauteurs 2 et 3, et calculer leurs mots de codage.

2.2 Compléments du cours

Les exercices de cette section sont l'occasion de refaire une partie des calculs vus en cours à l'identique, ou d'en faire d'autres selon le même principe.

Exercice 2.3

Prouver, par récurrence sur le nombre de nœuds internes, les formules proposées à l'exercice précédent pour le codage des arbres peignes.

Exercice 2.4

Un mot binaire w sur l'alphabet $\{a, b\}$ est dit *équilibré* s'il contient autant d'occurrences de a que de b : $|w|_a = |w|_b$.

Une propriété énoncée en cours, mais non démontrée, est que le codage d'un arbre binaire complet est un mot équilibré. Prouver cette propriété, par récurrence (forte) sur la taille des arbres.

Exercice 2.5

On convient de dire qu'un mot w sur l'alphabet $\{a, b\}$ est *positif* si *chacun de ses préfixes contient au moins autant d'occurrences de a que de b* : pour tous mots u et v tels que $w = u.v$, on a $|u|_a \geq |u|_b$.

Une propriété énoncée en cours, mais non démontrée, est que le codage d'un arbre binaire complet est un mot équilibré. Prouver cette propriété, par récurrence (forte) sur la taille des arbres.

(Les deux exercices précédents, pris ensemble, montrent que le mot de codage d'un arbre binaire est toujours un *mot de Dyck*)

Exercice 2.6

Une propriété énoncée en cours, mais non démontrée, est que le codage des arbres par les mots de Dyck est *injectif* : deux arbres différents ont toujours des mots de codages différents.

1. Prouver (par récurrence sur n) la propriété suivante : pour tout entier n , le mot de codage de tout arbre binaire complet à n nœuds internes est un mot de longueur $2n$. Ceci implique que deux hypothétiques arbres ayant le même mot de codage, auraient forcément la même taille.

2. On se propose de prouver la propriété par l'absurde. En supposant la propriété fausse, soit n le plus petit entier pour lequel elle échoue, et soient t et t' deux arbres de taille n , ayant le même mot de codage w . Soient t_1 et t_2 les sous-arbres gauche et droit de t , et t'_1 et t'_2 ceux de t' ; et w_1, w_2, w'_1 et w'_2 leurs mots de codage respectifs.

Montrer que w_1 et w'_1 sont de même longueur (**indication** : si on suppose que w_1 est strictement plus court que w'_1 , montrer que $w_1.b$ est un préfixe de w'_1 , et que cela n'est pas compatible avec le fait que w'_1 soit un mot positif), ainsi que w_2 et w'_2 . En déduire les égalités $w_1 = w'_1, w_2 = w'_2$, puis $t_1 = t'_1$ et $t_2 = t'_2$. Conclure.

2.3 Exercices et problèmes

Les exercices de cette section sont prévus pour demander plus de réflexion et d'effort.

Exercice 2.7

Décodage d'un mot de Dyck en arbre binaire. L'objectif de cet exercice est de décrire, pour tout mot de Dyck w , la construction d'un arbre binaire complet dont w soit le mot de codage, ce qui (avec les exercices précédents) justifiera le fait que le codage des arbres établit bien une bijection entre les arbres binaires complets et les mots de Dyck.

1. Soit w un mot de Dyck non vide. Justifier l'existence d'un *plus court préfixe non vide de w qui soit un mot de Dyck*. (**Indication** : le mot w est un de ses propres préfixes)
2. Soit w' le plus court préfixe non vide de w qui soit un mot de Dyck. Quelles sont les première et dernière lettres de w' ? Justifier qu'il existe un mot w'' tel que l'on ait $w' = a.w''.b$, et que w'' est un mot de Dyck.
3. Soit v le mot tel que l'on ait $w = w'.v$. Justifier que v est forcément un mot de Dyck. Montrer que, s'il existe un arbre t_1 dont le mot de codage est w'' et un arbre t_2 dont le mot de codage est v , alors il existe un arbre (le décrire!) dont w est le mot de codage.
4. Terminer la preuve (**indication** : on peut procéder par l'absurde, en supposant la propriété fausse et en prenant un mot de Dyck de longueur minimale qui ne soit pas le codage d'un arbre).

Exercice 2.8

Description alternative du codage d'un arbre binaire

Reprendre les différents exemples d'arbres binaires et de leurs mots de codage. Pour chaque arbre, parcourir la liste de ses nœuds **en ordre préfixe**, en notant pour chaque nœud s'il s'agit d'un nœud interne (I) ou d'une feuille (F).

Proposer une description alternative du mot de codage d'un arbre quelconque. En s'inspirant des preuves des exercices précédent, prouver que ce codage alternatif est bien le même que celui vu en cours.