Probabilités, Statistiques, Combinatoire : TD3

Combinatoire : codage des arbres binaires

Tous les arbres mentionnés dans les exercices sont des arbres binaires complets (ABC)/Tous les mots sont sur l'alphabet $\{a, b\}$.

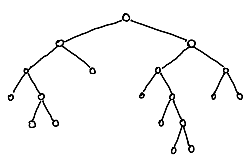
Exercice 3.1

Pour un arbre binaire complet t, son *mot de codage*, sur l'alphabet $\{a, b\}$, est défini de la manière suivante (ce codage est le même que celui qui a été décrit en cours par parcours préfixe, même si ce n'est pas évident) :

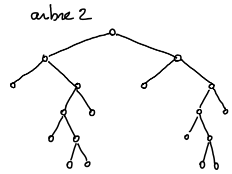
- l'arbre réduit à une feuille, est codé par le mot vide ε ;
- un arbre t non réduit à une feuille, dont les sous-arbres gauche et droit sont respectivement t_g et t_d, est codé par le mot $a.w_g.b.w_d$, où w_g est le mot de codage de t_g, et w_d, le mot de codage de t_d.

Calculer les mots codant les arbres ci-dessous.

arbre 1



arbre 2



Exercice 3.2

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.

1. Dessiner les peignes gauches et peignes droits à 4, 5 et 6 nœuds internes, et calculer leurs mots de codage. Proposer une formule pour les mots de codage d'un peigne droit et d'un peigne gauche, en général (c'est-à-dire une formule donnant, en fonction de n, le mot de codage de l'arbre peigne gauche à n nœuds internes, et une autre pour le mot de codage de l'arbre peigne droit à n nœuds internes).

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

Exercice 3.3

La **bijection de Rémy** (vu en cours), permet de passer bijectivement d'un couple formé d'un arbre binaire complet à n nœuds internes avec un nœud marqué (quelconque : interne ou feuille) et d'un choix parmi $\{\text{Gauche}, \text{Droite}\}$, à un arbre binaire complet à $n+1$ nœuds internes avec une feuille marquée.

Pour cet exercice, organisez-vous en binômes. Chacun choisit, sans le montrer à son partenaire, un arbre binaire à au moins 8 nœuds interne, un de ses nœuds et une direction, et calcule l'arbre à feuille marquée résultat de la bijection. Il ou elle transmet uniquement ce résultat au partenaire ; le partenaire doit alors retrouver l'arbre marqué et la direction de départ. Si ce n'est pas correct, l'un ou l'autre a fait une erreur : concertez-vous jusqu'à tomber d'accord.

Exercice 3.4

Parmi les arbres binaires de taille $n+1$, avec une feuille marquée, que l'on peut obtenir en appliquant la bijection de Rémy, il y a les *peignes droits* ; un peigne de taille $n+1$ ayant $n+2$ feuilles, identifier les $n+2$ façons d'obtenir le peigne droit de taille $n+1$ via la bijection de Rémy (*i.e.* les $n+2$ couples de départ, formés chacun d'un arbre de taille n avec un nœud marqué et d'une direction).

Vous pouvez commencer par prendre un petit exemple, par exemple $n = 4$: le peigne droit à 5 noeuds internes a 6 feuilles, il y a donc 6 façons différentes de l'obtenir avec la bijection.

Exercice 3.5

Un mot de Dyck non vide commence forcément par la lettre a, et après un certain nombre de a, il y aura forcément un b (pourquoi ?). Appelons *longueur de la première montée* d'un mot de Dyck w, le nombre de a consécutifs qui précèdent le premier b dans le mot w. Par exemple, la longueur de la première montée du mot $\underline{aaab}abbabbabaabb$ est 3.

1. En calculant, pour divers exemples de mots de Dyck, à chaque fois l'arbre binaire codé par le mot, essayez de proposer une *interprétation* : comment peut-on “lire” sur un arbre, sans calculer son mot de codage, la longueur de la première montée de son mot de codage ? (optionnel : prouver votre interprétation).
2. Inversement : la *taille* (nombre de nœuds internes) du sous-arbre gauche d'un arbre, peut être “lue” sur le mot qui code cet arbre. Expliquez quel paramètre calculé directement sur le mot de Dyck correspond à la taille du sous-arbre gauche.
3. Un autre exemple de paramètre qu'on peut calculer sur un mot (ou un chemin) de Dyck est le *nombre de retours à* 0 : le nombre de points du chemin (en comptant le point d'arrivée, mais pas celui de départ : donc ce nombre vaut au moins 1) qui se trouvent sur l'axe horizontal. Là encore, proposer une interprétation de ce paramètre sur l'arbre binaire correspondant.
4. Parmi les 1978261657756160653623774456 mots de Dyck de longueur 100, y en a-t-il plus qui commencent par $aaaaaaaaaaaab$ (12 a suivis d'un b), ou qui présentent exactement 12 retours à 0 ? (on ne demande pas de dire combien il y en a dans chaque cas !)

Exercice 3.6

Principe de réflexion étendu

On considère dans cet exercice des chemins faisant des pas Nord et Est, codés respectivement par les lettres a et b. Soit un entier $k \geq 0$ fixé.Le cas $k = 0$ a déjà été vu en cours.

1. Quelle est, en fonction de n, la longueur des chemins qui vont du point $(0, k)$ au point $(n, n + k)$? Combien y a-t-il de tels chemins ?
2. Parmi les chemins de la question précédene, on appelle **positifs** les chemins qui ne passent jamais (strictement) **sous** la diagonale (droite d'équation $y = x$).
   Si n est suffisamment petit, **tous** les chemins de $(0, k)$ à $(n, n + k)$ sont positifs. À partir de quelle valeur de n (en fonction de k) est-ce que ce n'est plus le cas ?
3. En adaptant le principe de réflexion, chercher une formule donnant, en fonction de n et k, le nombre de chemins de $(0, k)$ à $(n, n + k)$ qui ne sont **pas** des chemins positifs (**Indication :** il est donné sous la forme d'un coefficient binomial). En déduire une formule donnant, toujours en fonction de n et k, le nombre de chemins positifs de $(0, k)$ à $(n, n + k)$.
4. En exploitant votre formule, montrer que, pour $k = 1$, le nombre de chemins positifs de $(0, 1)$ à $(n, n + 1)$ est C_{n+1} (la suite des nombres de Catalan, décalée d'un indice ; cela nécessite de manipuler un peu des fractions avec des factorielles). En d'autres termes, qu'il y a autant de chemins positifs de $(0, 1)$ à $(n, n + 1)$, que de chemins de Dyck de longueur $2n + 2$. Chercher une preuve "directe" de cette relation, en raisonnant directement sur les chemins et sans passer par le principe de réflexion.

Exercice 3.7

Arbres ternaires complets

Un **arbre ternaire complet** est un arbre dans lequel chaque noeud interne (noeud autre qu'une feuille) a exactement trois enfants ; comme pour les arbres binaires, on distingue un ordre sur les enfants. La **taille** d'un arbre ternaire complet est son nombre de noeuds internes.

1. Dessiner tous les arbres ternaires complets à 0, 1, 2 et 3 noeuds internes (cela prend un peu de temps, pour 3 noeuds internes). Déterminer ainsi le début de la suite de comptage des arbres ternaires complets.
2. Proposer une formule donnant le nombre de feuilles d'un arbre ternaire complet, en fonction de son nombre de noeuds internes. (Il n'est pas demandé de prouver votre formule ; mais ce n'est pas très compliqué)
3. En vous inspirant du codage des arbres binaires complets par des mots de Dyck (en prenant la définition par parcours préfixe), proposer un codage des arbres **ternaires** complets par des mots sur l'alphabet $\{a, b\}$. Il n'est pas demandé de prouver que votre codage soit injectif (deux arbres différents sont toujours codés par des mots différents), mais c'est un objectif.
4. En interprétant les mots comme des chemins à pas Nord et Est, comme pour les mots de Dyck qui codent les arbres binaires complets, dessiner les chemins correspondant au codage des arbres que vous avez dessinés à la question 1.
5. Au vu de vos dessins de la question précédente, proposer une description de l'ensemble des chemins correspondant aux mots de codage de tous les arbres ternaires complets (de toutes les tailles).