

# Option Informatique en Spé MP et MP\*

## Devoir à rendre après les vacances de Noël

Polyominos; mots et chemins de Dyck et de Motzkin

### Résumé

Polymères ramifiés, changements de phase, viscosité, percolation, traitement d'images, hydrographie, structure secondaire des ARN, sédimentation, diffusion à travers une membrane, élasticité, jeu de Tetris : quelques uns des domaines d'intervention des polyominos.

De multiples questions de programmation seront pour vous l'occasion de faire preuve de votre compétence en Caml.

*Veillez rédiger chaque partie sur une copie séparée.*

## Table des matières

1	Polyominos	2
2	Mots de Dyck	2
3	Chemins de Dyck	3
4	Codage d'un mot de Dyck par ses pics et ses creux	3
5	Codage des polyominos parallélogrammes	4
6	Chemins et mots de Motzkin	4
7	Chemins et mots de Motzkin colorés	4

# 1 Polyominos

► La figure ci-dessous présente trois polyominos :

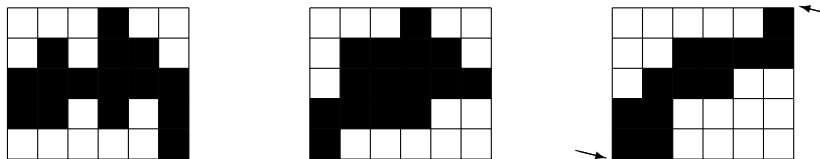


Figure 1: trois polyominos

Un peu de vocabulaire : une *dalle* est un carré  $[j, j + 1] \times [k, k + 1]$  où  $j, k \in \mathbb{Z}$ . Deux dalles sont *adjacentes* si elles ont un côté commun. Un polyomino est un ensemble fini de dalles, *connexe* au sens suivant : deux dalles quelconques peuvent être reliées par un chemin formé de dalles adjacentes.

Le polyomino de gauche est *connexe par colonnes* : dans chacune des colonnes, il n'y a aucun trou. La *largeur* de ce polyomino est 6, sa *hauteur* est 5 ; son *aire* est le nombre de dalles qui le constituent, soit 15 dans notre cas ; son *périmètre* est le nombre de côtés séparant un carré blanc d'un carré noir, soit 28 dans notre cas.

Le polyomino du milieu est convexe par lignes et par colonnes.

Le polyomino de droite est un polyomino *parallélogramme* : il est convexe par lignes et par colonnes, mais, de plus, il contient les «coins» inférieur gauche et supérieur droit (indiqués par des flèches). Nous mettons en évidence deux chemins reliant le coin inférieur gauche et le coin supérieur droit, constitués d'arêtes bordant le polyomino. En codant les déplacements vers la droite (resp. vers le haut) par la lettre h (resp. la lettre v), le chemin supérieur sera codé par le mot  $S = vvhvvhvhvh$  et le chemin inférieur sera codé par le mot  $I = hhvvhvhvhv$ .

► Un polyomino de hauteur  $n$  et de largeur  $p$  sera représenté par une matrice à  $n$  lignes et  $p$  colonnes, à coefficients dans l'ensemble  $\{0,1\}$  : un carré blanc sera codé 0, un carré noir sera codé 1.

**Question 1** • Rédigez en Caml une fonction :

```
aire : int matrix -> int
```

spécifiée comme suit : `aire m` calcule l'aire du polyomino décrit par la matrice  $m$ .

**Question 2** • Rédigez en Caml une fonction :

```
périmètre : int matrix -> int
```

spécifiée comme suit : `périmètre m` calcule le périmètre du polyomino décrit par la matrice  $m$ .

**Question 3** • Rédigez en Caml une fonction :

```
est_lignes_convexe : int matrix -> bool
```

spécifiée comme suit : `est_lignes_convexe m` décide si  $m$  représente un polyomino convexe par lignes.

**Question 4** • Rédigez en Caml une fonction :

```
est_parallélogramme : int matrix -> bool
```

spécifiée comme suit : `est_parallélogramme m` décide si  $m$  représente un polyomino parallélogramme.

## 2 Mots de Dyck

► Soit  $s = s_1 s_2 \dots s_{2n}$  un mot de longueur  $2n$  sur l'alphabet  $\{u, d\}$ . Les lettres  $u$  et  $d$  ont un *poids* :  $\pi(u) = +1$  et  $\pi(d) = -1$ . Pour  $p \in \llbracket 0, 2n \rrbracket$ , nous noterons  $h_p = \sum_{1 \leq k \leq p} \pi(s_k)$ . Nous dirons que  $s$  est un mot de DYCK s'il vérifie

les deux contraintes suivantes :  $h_p \geq 0$  pour tout  $p \in \llbracket 1, 2n - 1 \rrbracket$  ; et  $h_{2n} = 0$ .

► Nous noterons  $\mathcal{D}_n$  le langage constitué des mots de DYCK de longueur  $2n$ , et  $\mathcal{D} = \bigcup_{n \in \mathbb{N}} \mathcal{D}_n$ .

**Question 5** • Soient  $s$  et  $t$  deux mots de DYCK. Montrez que  $st$  est lui aussi un mot de DYCK.

**Question 6** • Énumérez les mots de DYCK de longueur 8.

**Question 7** • Le langage  $\mathcal{D}$  est-il rationnel ?

### 3 Chemins de Dyck

► Au mot de DYCK  $s$  de longueur  $2n$ , nous associons un chemin dans le plan, partant de l'origine et constitué d'une suite de vecteurs  $(1, 1)$  ou  $(1, -1)$ . Le chemin va donc toujours vers la droite, mais il peut monter ou descendre, sans toutefois passer sous l'axe des abscisses ; il se termine en  $(2n, 0)$ . Il est naturel de coder  $u$  le vecteur  $(1, 1)$  et  $d$  le vecteur  $(1, -1)$ . La figure 3 représente deux chemins de DYCK : celui de gauche est codé par le mot  $uuduudd$ , celui de droite est codé par le mot  $uududdduudd$ .

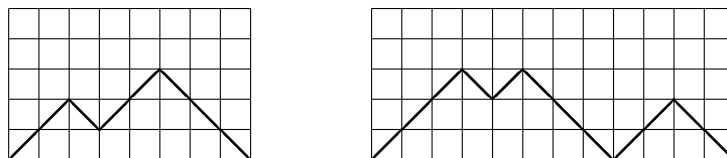


Figure 2: deux chemins de DYCK

Remarquons que, pour  $p \in \llbracket 0, 2n \rrbracket$ , le nombre  $h_p$  (défini à la partie 2) est l'ordonnée du point d'abscisse  $p$  de ce chemin.

► Soit  $s$  un mot de DYCK non vide. Deux situations peuvent se présenter :

1. dans le chemin associé, il n'y a que deux points d'ordonnée nulle : celui d'abscisse 0 et celui d'abscisse  $2n$  ;
2. le chemin associé passe par au moins un point d'ordonnée nulle et d'abscisse comprise entre 1 et  $2n - 1$ .

**Question 8** • Montrez que l'on est dans le premier cas ssi  $s = uvd$ , où  $v$  est un mot de DYCK.

► Plaçons-nous maintenant dans le deuxième cas. Notons  $q$  l'abscisse du premier point d'ordonnée nulle par lequel repasse le chemin.

**Question 9** • Montrez que  $q$  est pair, et que  $s[1..q]$  et  $s[q+1..2n]$  sont des mots de DYCK.

**Question 10** • Montrez que  $\mathcal{D}$  vérifie l'équation  $\mathcal{D} = \varepsilon + u\mathcal{D}d\mathcal{D}$ .

► Notons  $C_n$  le nombre de mots de DYCK de longueur  $2n$ .

**Question 11** • Utilisez les résultats précédents pour établir la formule  $C_{n+1} = \sum_{0 \leq k \leq n} C_k C_{n-k}$ .

► Les  $C_n$  sont les nombres de CATALAN ; ils sont parfaitement définis par la donnée de  $C_0 = 1$  et la relation de récurrence complète que nous venons d'établir. Nous admettrons la formule  $C_n = \frac{1}{n+1} \binom{2n}{n}$ .

► Soit  $s$  un mot de DYCK ; notons  $\pi(s)$  le nombre de points à coordonnées entières situés entre l'axe des abscisses et le chemin (inclus). Par exemple,  $\pi(ud) = 4$ .

**Question 12** • Prouvez la formule  $\sum_{s \in \mathcal{D}_n} \pi(s) = 4^n$ .

### 4 Codage d'un mot de Dyck par ses pics et ses creux

► Soit  $s$  un mot de DYCK de longueur  $2n$  ( $n \geq 1$ ). Un *pic* de  $s$  est un entier  $k \in \llbracket 1, 2n - 1 \rrbracket$  tel que  $h_k > h_{k-1}$  et  $h_k > h_{k+1}$ . Définissons de même un *creux* de  $s$  : c'est un entier  $k \in \llbracket 1, 2n - 1 \rrbracket$  tel que  $h_k < h_{k-1}$  et  $h_k < h_{k+1}$ .

► Il est clair qu'un mot  $s$  de longueur  $2n$  comporte au moins un pic, et au plus  $n$  ; le nombre de creux est compris entre 0 et  $n - 1$  ; deux pics successifs sont séparés par un creux, et inversement.

► Sur la figure 3, le mot de droite comporte trois pics :  $p_1 = 3$ ,  $p_2 = 5$  et  $p_3 = 10$  ; et deux creux  $c_1 = 4$  et  $c_2 = 8$ . Nous pouvons coder le mot  $s$  par la liste «entrelacée»  $\lambda(s) = (3, 2, 3, 0, 2)$  des hauteurs de ses pics et de ses creux.

**Question 13** • Soient :  $n \geq 1$ ,  $j \in \llbracket 1, 2n - 1 \rrbracket$ , et  $r = (r_i)_{1 \leq i \leq j}$  une liste d'entiers. Énoncez et démontrez une condition nécessaire et suffisante pour que  $r$  soit la liste des pics et des creux d'un mot de DYCK de longueur  $2n$ .

**Question 14** • Montrez que la fonction  $\lambda$  est injective.

**Question 15** • Rédigez en Caml une fonction :

```
pics_et_creux : int list -> int list
```

spécifiée comme suit : `pics_et_creux` u rend la liste  $(p_1, c_1, p_2, c_2, \dots, p_{j-1}, c_{j-1}, p_j)$  où  $j$  est le nombre de pics.

## 5 Codage des polyominos parallélogrammes

► Un polyomino parallélogramme de largeur  $p$  peut être codé efficacement, au moyen de deux suites d'entiers :

1. la suite  $(h_k)_{1 \leq k \leq p}$  où  $h_k$  est la hauteur de la  $k$ -ième colonne ;
2. la suite  $(\delta_k)_{1 \leq k < p}$  où  $\delta_k$  est le nombre de contacts entre cases des colonnes  $k$  et  $k + 1$ .

Par exemple, dans la figure 1, le polyomino de droite sera codé par les suites  $h = (2, 3, 2, 2, 1, 2)$  et  $\delta = (2, 1, 2, 1, 1)$ .

**Question 16** • Rédigez en Caml une fonction :

```
codage_poly_par : int matrix -> (int list) * (int list)
```

spécifiée comme suit : `codage_poly_par` m rend le couple  $(h, \delta)$  de listes associé au polyomino décrit par la matrice  $m$ .

**Question 17** • Rédigez en Caml une fonction :

```
décodage_poly_par : (int list) * (int list) -> int matrix
```

spécifiée comme suit : `décodage_poly_par` (h,d) construit la matrice  $m$  décrivant le polyomino codé par le couple  $(h, \delta)$ .

**Question 18** • Montrez qu'il existe un et un seul mot de DYCK dont  $(h_k)_{1 \leq k \leq p}$  est la suite des hauteurs de pics, et  $(\delta_k - 1)_{1 \leq k < p}$  est la suite des hauteurs des creux.

## 6 Chemins et mots de Motzkin

► Rappel : le nombre de chemins de DYCK de longueur  $2n$  est  $C_n = \frac{1}{n+1} \binom{2n}{n}$ .

► Notons  $u$  le vecteur  $(1, 1)$ ,  $d$  le vecteur  $(1, -1)$  et  $h$  le vecteur  $(1, 0)$ . Un chemin de MOTZKIN de longueur  $n$  est une suite de pas  $u$ ,  $d$  et  $h$ , commençant en  $(0, 0)$ , finissant en  $(n, 0)$  et ne passant jamais sous l'axe des abscisses. Remarquons qu'un chemin de Dyck est un chemin de Motzkin ne comportant aucun pas  $h$ .

**Question 19** • Montrez que le nombre de chemins de Motzkin ayant  $n - 2k$  pas  $h$  est  $C_k \binom{n}{2k}$ .

**Question 20** • Dessinez tous les chemins de MOTZKIN de longueur 3.

► Notons  $\Delta_n$  l'ensemble des mots de DYCK de longueur  $2n$  dont tous les creux ont une abscisse paire.

**Question 21** • Soit  $\mathbf{m} \in \Delta_n$ . Montrez que le facteur  $du$  ne peut apparaître qu'à des positions paires dans  $\mathbf{m}$ .

**Question 22** • Exhibez une bijection de  $\Delta_n$  sur l'ensemble des mots de MOTZKIN de longueur  $n$ .

## 7 Chemins et mots de Motzkin colorés

► Dans un chemin de MOTZKIN coloré, chaque pas  $h$  est coloré en rouge ou en bleu ; nous noterons  $b$  un pas «bleu» vers la droite et  $r$  un pas «rouge» vers la droite.

La figure 3 donne un exemple de chemin de MOTZKIN coloré ; parmi les quatre pas horizontaux, les deux pas «rouges» ont été tracés en trait **gras**. Le chemin de MOTZKIN coloré est codé par le mot de MOTZKIN coloré `uurduubrddbd`.

► Définissons  $\varphi : \{u, d, b, r\} \mapsto \{u, d\}$  comme suit :  $\varphi(u) = uu$ ,  $\varphi(d) = dd$ ,  $\varphi(b) = ud$  et  $\varphi(r) = du$ .

► Soit  $\mathbf{m} = \mathbf{m}_1 \mathbf{m}_2 \dots \mathbf{m}_n$  un mot de MOTZKIN coloré de longueur  $n$ . Définissons la fonction  $\Phi$  par la formule  $\Phi(\mathbf{m}) = u\varphi(\mathbf{m}_1)\varphi(\mathbf{m}_2) \dots \varphi(\mathbf{m}_n)d$ .

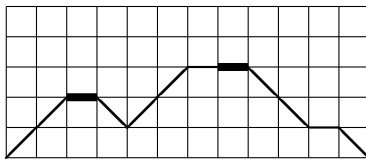


Figure 3: un chemin de MOTZKIN coloré

**Question 23** • Montrez que  $\Phi$  est une bijection de l'ensemble des chemins de MOTZKIN colorés de longueur  $n$  sur l'ensemble des chemins de DYCK de longueur  $2n + 2$ .

**Question 24** • Prouvez la formule de TOUCHARD :  $C_{n+1} = \sum_{0 \leq 2k \leq n} 2^{n-2k} C_k \binom{n}{2k}$ .

► Nous construisons maintenant une bijection de l'ensemble  $\mathcal{P}_{2n+4}$  des polyominos parallélogrammes de périmètre  $2n + 4$  sur l'ensemble  $\mathcal{M}_n$  des chemins de MOTZKIN colorés de longueur  $n$ .

► Un polyomino parallélogramme est parfaitement défini par les mots  $S$  et  $I$  qui lui ont été associés dans la partie 1. Ces deux chemins ont la même longueur  $p + q$ , où  $p$  (resp.  $q$ ) est la largeur (resp. la hauteur) du polyomino. Le mot  $S$  (resp.  $I$ ) commence toujours par  $v$  (resp.  $h$ ) et se termine par  $h$  (resp.  $v$ ). Définissons deux mots  $s$  et  $t$  par les relations  $S = vsh$  et  $I = htv$ ;  $s$  et  $t$  ont même longueur  $n = p + q - 2$ . Nous codons le couple  $(S, I)$  par le mot  $\mathbf{m}$ , de longueur  $n$ , défini comme suit : pour tout  $i \in \llbracket 1, n \rrbracket$ ,  $\mathbf{m}_i = u$  si  $s_{i+1} = t_{i+1} = v$ ;  $\mathbf{m}_i = d$  si  $s_{i+1} = t_{i+1} = h$ ;  $\mathbf{m}_i = b$  si  $s_{i+1} = h$  et  $t_{i+1} = v$ ;  $\mathbf{m}_i = r$  si  $s_{i+1} = v$  et  $t_{i+1} = h$ .

**Question 25** • Notons  $\mathcal{K}$  la fonction qui, au couple  $(S, I)$ , associe le mot de MOTZKIN coloré  $\mathbf{m}$ . Montrez que  $\mathcal{K}$  est une bijection de  $\mathcal{P}_{2n+4}$  sur  $\mathcal{M}_n$ .

**Question 26** • Rédigez en Caml une fonction :

```
vérifie_chemins : string * string -> bool
```

spécifiée comme suit : `vérifie_chemins (s,t)` vérifie que les mots  $S = vsh$  et  $I = htv$  codent bien les chemins supérieur et inférieur d'un polyomino parallélogramme.

**Question 27** • Rédigez en Caml une fonction :

```
codage_SI_vers_motzkin : string * string -> string
```

spécifiée comme suit : `codage_SI_vers_motzkin (s,t)` calcule le mot  $\mathbf{m}$ , image du couple  $(S, I)$  par la bijection définie à la question 25.

**Question 28** • Rédigez en Caml une fonction :

```
codage_motzkin_vers_SI : string -> string * string
```

spécifiée comme suit : `codage_motzkin_vers_SI m` calcule l'antécédent du mot de MOTZKIN  $\mathbf{m}$  par la bijection  $\mathcal{K}$  définie plus haut.

FIN