

Calculs et tableurs

Épreuve pratique d'algorithmique et de programmation
Concours commun des Écoles normales supérieures

Durée de l'épreuve : 3 heures 30 minutes

Juin 2024

ATTENTION !

N'oubliez en aucun cas de recopier votre u_0
à l'emplacement prévu sur votre fiche réponse

Important.

Il vous a été donné un numéro u_0 qui servira d'entrée à vos programmes. Les réponses attendues sont généralement courtes et doivent être données sur la fiche réponse fournie à la fin du sujet. À la fin du sujet, vous trouverez en fait deux fiches réponses. La première est un exemple des réponses attendues pour un \tilde{u}_0 particulier (précisé sur cette même fiche et que nous notons avec un tilde pour éviter toute confusion !). Cette fiche est destinée à vous aider à vérifier le résultat de vos programmes en les testant avec \tilde{u}_0 au lieu de u_0 . Vous indiquerez vos réponses (correspondant à votre u_0) sur la seconde et vous la remettrez à l'examineur à la fin de l'épreuve.

En ce qui concerne la partie orale de l'examen, lorsque la description d'un algorithme est demandée, vous devez présenter son fonctionnement de façon schématique, courte et précise. Vous ne devez en aucun cas recopier le code de vos procédures !

Quand on demande la complexité en temps ou en mémoire d'un algorithme en fonction d'un paramètre n , on demande l'ordre de grandeur en fonction du paramètre, par exemple : $O(n^2)$, $O(n \log n)$,...

Il est recommandé de commencer par lancer vos programmes sur de petites valeurs des paramètres et de *tester vos programmes sur des petits exemples que vous aurez résolus préalablement à la main ou bien à l'aide de la fiche réponse type fournie en annexe*. Enfin, il est recommandé de lire l'intégralité du sujet avant de commencer afin d'effectuer les bons choix de structures de données dès le début.

1 Préliminaires

Les sections 3, 4, 5 et 6 peuvent être traitées de manière indépendante mais il est néanmoins recommandé de les traiter dans l'ordre.

1.1 Notations

Pour deux entiers naturels a et b avec $b \neq 0$, on notera $a \bmod b$ le reste de la division euclidienne de a par b , c'est-à-dire l'unique entier r avec $0 \leq r < b$ tel que $a = k \times b + r$ et $k \in \mathbb{N}$.

Attention : En OCaml, le résultat de la fonction `mod` peut être négatif si a est lui-même négatif.

On rappelle que $\mathbb{Z}/k\mathbb{Z}$ est l'anneau des entiers modulo k . Un élément de $\mathbb{Z}/k\mathbb{Z}$ sera en général représenté par un entier $0 \leq x < k$.

On parlera de booléen pour indiquer le type de données contenant vrai et faux.

On note $A \wedge B$ la conjonction des deux formules A et B , et $A \vee B$ leur disjonction.

1.2 Génération de nombres pseudo-aléatoires

Étant donné u_0 , on considère la suite u générée par la relation de récurrence suivante :

$$u_{n+1} = (178\,481 \times u_n + 131\,071) \bmod 869\,753$$

L'entier u_0 vous est donné, et doit être recopié sur votre fiche réponse avec vos résultats. Une fiche réponse type vous est donnée en exemple, et contient tous les résultats attendus pour une valeur de u_0 différente de la vôtre (notée \widehat{u}_0). Il vous est conseillé de tester vos algorithmes avec cet \widehat{u}_0 et de comparer avec la fiche de résultats fournie. Pour chaque calcul demandé, avec le bon choix d'algorithme le calcul ne devrait demander qu'au plus de l'ordre de quelques secondes, jamais plus d'une minute.

On aura souvent besoin de nombreuses valeurs consécutives de la suite u . Il est donc conseillé que votre implémentation calcule le tableau de tous les u_n jusqu'à un certain rang choisi soigneusement.

On pourra par ailleurs remarquer que la suite u est cyclique de longueur 869 752.

Question 1 Calculez les valeurs suivantes :

a) $u_1 \bmod 1000$

b) $u_{24} \bmod 1000$

c) $u_{7852} \bmod 1000$

d) $u_{869\,751} \bmod 1000$

1.3 Tableaux

Les tableaux sont des outils logiciels qui contiennent des feuilles de calcul et qui ont des utilisations variées telles que la comptabilité, les statistiques, etc.

Dans ce sujet, on considère une version simplifiée des tableaux. Un tableau va être vu comme une matrice composée de cellules contenant chacune soit un scalaire pris dans l'anneau $\mathbb{Z}/p\mathbb{Z}$ pour un certain p , soit une formule faisant référence à d'autres cellules du tableau. Nous donnons un exemple de tel tableau avec $p = 11$ en Figure 1.

Ici le tableau est de taille 5×3 avec des colonnes numérotées de 0 à 4 et des lignes de 0 à 2. La notation (i, j) fait référence à la cellule située colonne i et ligne j . Par exemple, le contenu

	0	1	2	3	4
0	3	8	PLUS((0, 0), (2, 1))	1	PROD((2, 1), (4, 2))
1	2	MOINS((0, 2))	9	5	PROD((3, 2), (0, 1))
2	10	PLUS((1, 1), (0, 2))	0	MOINS((2, 0))	7

FIGURE 1 – Exemple de tableur

de la cellule (3, 2) est MOINS((2, 0)) qui est une formule faisant référence à la cellule (2, 0). Pour évaluer cette formule, il faut donc évaluer la cellule (2, 0) et prendre son opposé modulo p . La cellule contient elle-même une autre formule : PLUS((0, 0), (2, 1)). Celle-ci fait référence aux cellules (0, 0) et (2, 1) qui contiennent cette fois-ci les valeurs 3 et 9 respectivement, ce qui donne comme résultat $3 + 9 = 12 = 1$ pour la cellule (2, 0) et donc $-1 = 10$ pour (3, 2). De même, on évalue une formule PROD(c_1, c_2) en faisant le produit (modulo p) des valeurs correspondant à c_1 et c_2 .

Évaluer toutes les cases d'un tableur – quand c'est possible – donne lieu à une vue qui est une matrice de même taille ne contenant que des scalaires. La vue associée au tableur de la Figure 1 est donnée en Figure 2.

	0	1	2	3	4
0	3	8	1	1	8
1	2	1	9	5	9
2	10	0	0	10	7

FIGURE 2 – Vue d'un tableur

1.4 Génération de formules et cellules

Dans ce sujet on va considérer que les formules ne sont jamais imbriquées et sont donc toujours de la forme PLUS(c_1, c_2), PROD(c_1, c_2) ou MOINS(c) où c, c_1 et c_2 font référence à des cellules. Ces références sont donc de la forme (i, j). Le contenu d'une cellule est soit une formule, soit un entier.

Afin de générer formules et cellules, on définit deux fonctions, v et w par :

$$v(m, i, j) = u_{3(im+j)+1} \bmod m \quad w(m, i, j) = u_{3(im+j)+2} \bmod m$$

On définit une fonction \mathfrak{c} qui renvoie le contenu d'une cellule (donc soit une formule soit un entier). \mathfrak{c} prend pour arguments un booléen b et cinq entiers n, m, p, i, j et est définie comme suit :

$$\mathfrak{c}(b, n, m, p, i, j) = \begin{cases} \text{PLUS}((i-1, v(m, i, j)), (i-1, w(m, i, j))) & \text{si } r = 0 \\ \text{PROD}((i-1, v(m, i, j)), (i-1, w(m, i, j))) & \text{si } r = 1 \\ \text{MOINS}((i-1, v(m, i, j))) & \text{si } (r = 2) \vee ((r = 3) \wedge \neg b') \\ \text{MOINS}((n-1, v(m, i, j))) & \text{si } (r = 3) \wedge b' \\ u_{3(im+j)+1} \bmod p & \text{sinon} \end{cases}$$

où $r = u_{3(im+j)} \bmod 6$ et $b' = b \wedge (u_{3(im+j)+2} \bmod 17 = 0)$.

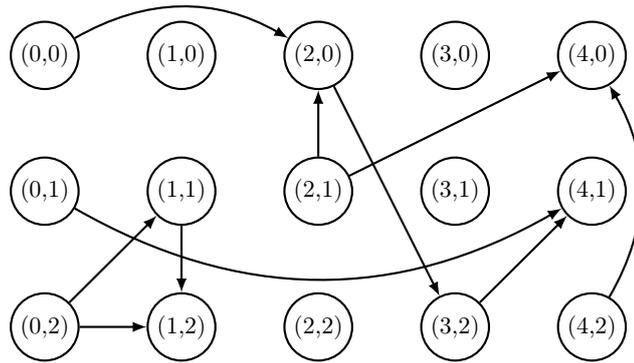


FIGURE 3 – Graphe de dépendances

Question 4 Calculez le nombre, modulo 100 000, de nœuds avec des arcs sortants dans les graphes de dépendances associés aux tableurs suivants :

- | | |
|--|--|
| a) $T(\text{vrai}, 10, 10, 97)$ | b) $T(\text{vrai}, 10, 100, 547)$ |
| c) $T(\text{vrai}, 30, 1000, 1091)$ | d) $T(\text{vrai}, 1000, 2000, 5099)$ |

Question à développer pendant l'oral 2 Détaillez quelle structure de données vous avez choisie pour représenter les graphes de dépendances et justifiez votre choix.

3 Tri topologique et évaluation de tableur

Comme déjà suggéré, le graphe de dépendances, lorsqu'il est acyclique, permet de déterminer un ordre d'évaluation, c'est-à-dire un ordre total sur les cellules qui respecte les dépendances : si une cellule c_2 dépend d'une autre cellule c_1 alors c_1 est "plus petite" que c_2 et doit être évaluée avant. Un tel ordre s'appelle ordre topologique, et il peut en exister plusieurs. Par exemple, dans le graphe en Figure 3, les cellules $(1, 0)$ et $(3, 0)$ peuvent être évaluées dans n'importe quel ordre. On va donc s'intéresser au fait de trouver un ordre topologique sans se préoccuper duquel il s'agit. Pour ce sujet, l'ordre choisi n'a pas d'influence du moment que c'est bien un ordre topologique.

Un tri topologique consiste à trouver un ordre topologique. Il peut se faire en effectuant un parcours en profondeur du graphe de dépendances, tout en prenant soin de distinguer les nœuds déjà visités et intégrés à l'ordre, les nœuds en train d'être visités, et les nœuds pas encore visités.

Un tel algorithme va donc détecter les cycles du graphe. On donne un exemple de tableur avec une référence circulaire en Figure 4 et le graphe de dépendances associé en Figure 5. Remarquez la présence d'un cycle. Celui-ci empêche d'obtenir une vue du tableur car il est impossible d'évaluer la cellule $(0, 0)$ par exemple.

Si un ordre topologique peut être établi (donc s'il n'y a pas de cycles), on peut alors calculer une vue du tableur associé en évaluant les formules dans cet ordre.

Question 5 Calculez la somme modulo p des valeurs des cellules dans la vue associée au tableur $T(\text{vrai}, n, m, p)$ pour les valeurs de n , m et p suivantes. Si le tableur présente un cycle, indiquez

	0	1	2	3
0	MOINS((2, 0))	4	PROD((3, 0), (3, 1))	10
1	5	PLUS((0, 0), (1, 0))	0	MOINS((1, 1))

FIGURE 4 – Exemple de tableur avec une référence circulaire

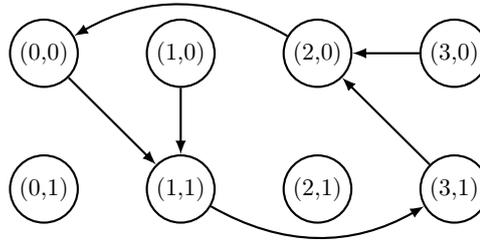


FIGURE 5 – Graphe de dépendances avec cycle

simplement “Cycle” dans la fiche réponse.

- a)** $n = 10, m = 10, p = 97$
- b)** $n = 10, m = 100, p = 547$
- c)** $n = 30, m = 1000, p = 1091$
- d)** $n = 1000, m = 2000, p = 5099$

Question à développer pendant l’oral 3 Décrivez l’algorithme utilisé pour établir l’ordre topologique en insistant sur les structures de données utilisées. Donnez une analyse de la complexité en fonction de n et m , les dimensions du tableur.

Question à développer pendant l’oral 4 Même question pour l’obtention de la somme des valeurs de la vue.

4 Évaluation incrémentale

À partir de maintenant, nous allons considérer que tous les tableurs considérés ne présentent aucun cycle. Attention ! Cela signifie que nous n’allons pas utiliser les mêmes exemples que précédemment.

4.1 Idée générale

Dans cette section nous allons nous intéresser au problème dit du calcul incrémental. Étant donné un tableur et une vue associée, la question est de recalculer efficacement la vue quand seulement une petite partie des données a changé.

Pour ce faire, on va reposer à nouveau sur le graphe de dépendances. Si on revient au tableur Figure 1, changer la valeur de la cellule (0,0) implique de devoir recalculer la cellule (2,0) qui en dépend directement, mais également les cellules (3,2) et (4,1) qui sont dans la fermeture transitive de la relation de dépendance. En revanche, aucune des autres cases n’a besoin d’être recalculée.

Par ailleurs, si on modifie la cellule $(0, 2)$, on peut voir que la cellule $(1, 2)$ a besoin d'être recalculée pour deux raisons : d'abord parce qu'elle dépend de $(0, 2)$, mais aussi parce qu'elle dépend de $(1, 1)$ qui elle-même dépend de $(0, 2)$. Il faut donc prendre soin de ne pas recalculer cette valeur deux fois inutilement.

L'algorithme va donc opérer en deux passes : une première passe qui consiste à invalider les cellules qui dépendent de la cellule changée, suivie d'une deuxième qui va réparer la vue en recalculant uniquement les cellules invalidées.

Note. On ne va considérer que des modifications qui remplacent des scalaires par d'autres scalaires, ainsi le graphe de dépendances n'est pas modifié.

4.2 Passe d'invalidation

Comme indiqué précédemment, cette phase consiste à marquer comme invalidées les cellules qui dépendent de la cellule modifiée, et ce transitivement. Pour éviter les calculs inutiles, on va également calculer, pour chaque cellule c , le nombre de cellules invalidées c' avec un arc vers c . On appellera ce nombre le degré d'une cellule. Pour tenir compte des cellules qui ont subi le changement, on considèrera que celles-ci ont pour degré 1.

Par exemple (toujours Figure 1), après avoir changé la valeur de la cellule $(0, 2)$, la cellule $(1, 1)$ aura pour degré 1, tandis que la cellule $(1, 2)$ aura pour degré 2. La cellule $(0, 2)$ doit elle aussi avoir 1 pour degré. Toutes les autres cellules ont pour degré 0 et n'ont pas besoin d'être changées.

On note $c \leftarrow v$ pour le changement qui remplace le contenu de la cellule c par la valeur v . Par exemple $(0, 2) \leftarrow 6$ remplace le contenu de la cellule $(0, 2)$ par la valeur 6. Appliquer cette modification au tableur Figure 1 donne un nouveau tableur montré en Figure 6. Les cellules invalidées sont surlignées, avec une emphase particulière pour la cellule directement modifiée.

	0	1	2	3	4
0	3	8	PLUS((0, 0), (2, 1))	1	PROD((2, 1), (4, 2))
1	2	MOINS((0, 2))	9	5	PROD((3, 2), (0, 1))
2	6	PLUS((1, 1), (0, 2))	0	MOINS((2, 0))	7

FIGURE 6 – Tableur modifié

Note. Si l'on venait à modifier à nouveau la cellule $(0, 2)$ — par exemple avec le changement $(0, 2) \leftarrow 4$ — les degrés restent néanmoins inchangés (ils ne se cumulent pas).

Nous allons maintenant supposer que quelques changements ont été effectués dans le première colonne d'un tableur T. On considère la séquence $C(k, p, m)$ de 10 changements suivante :

$$(0, u_{k+2l} \bmod m) \leftarrow u_{k+2l+1} \bmod p$$

pour $0 \leq l < 10$.

Dans les deux questions qui suivent, on considère uniquement la première passe qui invalide les cellules sans effectuer la passe de réparation.

Question 6 Calculez la somme d des degrés ainsi que le nombre N de cellules invalidées après avoir effectué la séquence de changements $C(79\,730, p, m)$ sur le tableur $T(\text{faux}, n, m, p)$ pour les

valeurs de n , m et p suivantes. Écrivez le résultat sous la forme du couple (d, N) .

a) $n = 10, m = 10, p = 97$

b) $n = 10, m = 100, p = 547$

c) $n = 30, m = 1000, p = 1091$

d) $n = 1000, m = 2000, p = 5099$

Question à développer pendant l'oral 5 Quelle est la complexité en temps dans le pire cas de cette première passe d'invalidation en fonction des dimensions d'un tableur n et m ? Qu'observe-t-on empiriquement au regard des valeurs demandées ci-dessus? On pourra séparer la phase d'initialisation dans l'analyse de complexité.

4.3 Passe de réparation

La deuxième passe effectue la réparation de la vue en partant des cellules dépendant directement des cellules modifiées. Pour chaque cellule visitée, on décrémente le nombre de dépendances invalides, et si ce nombre atteint 0, alors on peut évaluer la cellule à nouveau sans risque (toutes ses dépendances ont été réparées).

En partant du tableur modifié en Figure 6, on commence par remettre le degré de $(0, 2)$ à 0 avant de visiter les cellules $(1, 2)$ et $(1, 1)$ qui en dépendent. $(1, 2)$ a pour degré 2, nombre que l'on décrémente à 1, signifiant qu'une de ses dépendances n'a toujours pas été réparée. On visite alors $(1, 1)$ qui est de degré 1, on peut donc descendre ce degré à 0 et effectuer le calcul, ce qui donne 5. On visite maintenant les dépendances de $(1, 1)$, c'est-à-dire $(1, 2)$ à nouveau. On descend son degré à 0 et on calcule sa valeur : 0. Ceci donne ainsi la vue réparée en Figure 7.

	0	1	2	3	4
0	3	8	1	1	8
1	2	5	9	5	9
2	6	0	0	10	7

FIGURE 7 – Vue réparée

Nous allons maintenant effectuer la réparation des vues pour les quatre tableurs que nous avons modifiés ci-dessus. Votre algorithme devra commencer par propager les changements à partir des dépendances des dix cellules modifiées.

Question 7 Calculez la somme modulo p des valeurs des cellules dans la vue associée au tableur $T(\text{faux}, n, m, p)$ après l'avoir réparée pour les valeurs de n , m et p suivantes.

a) $n = 10, m = 10, p = 97$

b) $n = 10, m = 100, p = 547$

c) $n = 30, m = 1000, p = 1091$

d) $n = 1000, m = 2000, p = 5099$

Question à développer pendant l'oral 6 Commentez la complexité des deux passes en fonction des dimensions, mais aussi en fonction des données du graphe de dépendances.

Question à développer pendant l'oral 7 À quel point l'efficacité de l'approche présentée ici est dépendante de la génération aléatoire des données?

5 Grandes feuilles de calcul

Dans cette partie, on considère des feuilles de calcul de très grande taille.

L'objectif de cette partie est de calculer la valeur d'un petit nombre de cellules, en l'occurrence les trois premières lignes de la dernière colonne. Par exemple pour un tableur de dimensions 10×27 , cela correspond aux cellules $(9, 0)$, $(9, 1)$ et $(9, 2)$.

Question 8 Calculez la somme modulo p des valeurs des trois premières cellules de la dernière colonne dans la vue associée au tableur $T(\text{faux}, n, m, p)$ pour les valeurs de n , m et p suivantes.

- a) $n = 1000, m = 20\,000, p = 97$
- b) $n = 10\,000, m = 34\,000, p = 127$
- c) $n = 70\,000, m = 567\,000, p = 257$
- d) $n = 234\,000, m = 9\,170\,000, p = 307$

Question à développer pendant l'oral 8 Décrivez l'algorithme ainsi que les structures de données utilisés.

6 Mesure d'impact

On s'intéresse pour finir au nombre de cellules qui se retrouveraient affectées par le changement d'une cellule donnée. Autrement dit, on veut calculer combien de cellules dépendent directement et indirectement d'une cellule donnée. On appellera ce nombre l'impact de la cellule en question.

Si on reprend l'exemple de la Figure 1 on observe que l'impact de la cellule $(0, 0)$ est de 3 car les cellules $(2, 0)$, $(3, 2)$ et $(4, 1)$ en dépendent.

Plus précisément, la question que l'on se pose est de savoir combien de cellules ont un impact strictement supérieur au nombre de colonnes n .

Question 9 Calculez le nombre de cellules à impact strictement supérieur au nombre de colonnes n dans les tableurs suivants.

- | | |
|--|---|
| a) $T(\text{faux}, 10, 100, 97)$ | b) $T(\text{faux}, 200, 300, 151)$ |
| c) $T(\text{faux}, 1000, 1000, 251)$ | d) $T(\text{faux}, 1000, 3000, 373)$ |
| e) $T(\text{faux}, 1000, 10\,000, 7013)$ | f) $T(\text{faux}, 500, 30\,000, 3931)$ |

Question à développer pendant l'oral 9 Expliquez l'algorithme que vous avez utilisé.



Fiche réponse type : Calculs et tableurs

\widetilde{u}_0 : 237

Question 1

a) 924

b) 915

c) 743

d) 944

Question 2

a) 4696

b) 71

c) 216

d) 432

Question 3

a) 4367

b) 4010

c) 9312

d) 4185

Question 4

a) 54

b) 589

c) 18 380

d) 58 704

Question 5

a) 2

b) 481

c) 447

d) Cycle

Question 6

a) (24, 23)

b) (85, 85)

c) (412, 408)

d) (94, 94)

Question 7

a) 11

b) 121

c) 537

d) 2645

Question 8

a) 4

b) 112

c) 27

d) 264

Question 9

a) 146

b) 3093

c) 20 226

d) 71 479

e) 227 260

f) 494 959



Fiche réponse : Calculs et tableurs

Nom, prénom, u₀ :

Question 1

a)

b)

c)

d)

Question 2

a)

b)

c)

d)

Question 3

a)

b)

c)

d)

Question 4

a)

b)

c)

d)

Question 5

a)

b)

c)

d)

Question 6

a)

b)

c)

d)

Question 7

- a)
- b)
- c)
- d)

Question 8

- a)
- b)
- c)

d)

Question 9

- a)
- b)
- c)
- d)
- e)
- f)

