Très Grands Nombres

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

Durée de l'épreuve: 3 heures 30 minutes

Juin/Juillet 2017

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 $\widetilde{u_0}$ 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 $\widetilde{u_0}$ au lieu de u_0. Vous indiquerez vos réponses (correspondant à votre u_0) sur la seconde et vous la remettrez à l'examinateur 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.

Si $a \geqslant 0$ et $b > 0$ sont deux entiers, on note $a \bmod b$ le reste de la division euclidienne de a par b, autrement dit l'unique entier r avec $0 \leqslant r < b$ tel qu'il existe un entier q satisfaisant $a = bq + r$.

On fixe pour tout le sujet

$$m = 2^{31} - 1 = 2\,147\,483\,647$$

et l'on définit une suite $(u(n))_{n\in[0,1\,000]}$ par

$$u(0) = u_0, \qquad u(n+1) = \big(16\,807 \times u(n) + 17\big) \bmod m.$$

avec u_0 l'entier qui vous a été donné (à reporter sur votre fiche réponse).

Question 1 *Calculer* $u(n) \bmod 101$ *pour*
a) $n = 5$ ***b)*** $n = 100$ ***c)*** $n = 997$

On définit $(v(k,n))_{k\in[0,61],n\in[0,1\,000]}$ par

$$v(k,n) = u(n) \bmod 2^k + 2^k$$

Question 2 *Calculer* $v(k, 97k \bmod 997) \bmod 101$ *pour*
a) $k = 5$ ***b)*** $k = 30$ ***c)*** $k = 61$

1 Représentation trichotomique

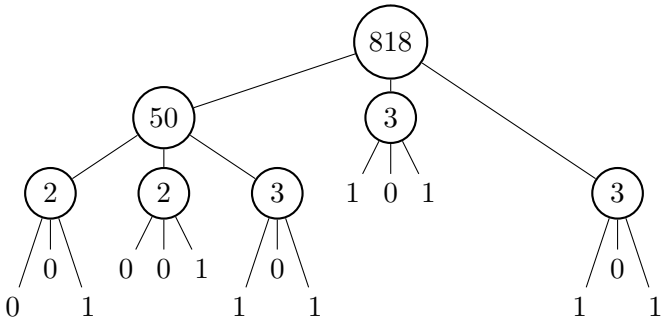
Dans ce sujet, on s'intéresse à la manipulation informatique de très grands nombres, de l'ordre du gogolplex ($10^{10^{100}}$), qu'il est physiquement impossible de représenter dans le système décimal. Afin d'exploiter au mieux les capacités micro-architecturales des ordinateurs, nous étudions des variations du système de numération binaire. On note ainsi $\mathbf{x}(p) = 2^{2^p}$.

Question à développer pendant l'oral 1 *Soit $n \in \mathbb{N}$. On appelle longueur binaire, notée $\mathbf{l}(n)$, la longueur du tableau de bits représentant le nombre n dans le système binaire. Donner une borne supérieure asymptotique ("en grand O") de $\mathbf{l}(n)$. Pour $p \in \mathbb{N}$, quelle est la complexité spatiale du nombre $\mathbf{x}(p)$ dans le système binaire ?*

Question à développer pendant l'oral 2 *Prouver que tout entier naturel $n \geqslant 2$ se décompose de façon unique en un triplet $(g,p,d) \in \mathbb{N}^3$ tel que $n = g + \mathbf{x}(p) \times d$ avec $0 \leqslant g < \mathbf{x}(p)$ et $0 < d < \mathbf{x}(p)$.*

Partant de cette observation, nous adoptons une représentation *trichotomique* des entiers naturels sous forme d'arbres ternaires. Un nœud contient un sous-arbre gauche représentant l'entier g, un sous-arbre central représentant l'entier p et un sous-arbre droit représentant l'entier d. Par abus de langage, on écrira "$g + \mathbf{x}(p) \times d$" pour dénoter aussi bien un tel nœud que le nombre qu'il représente. Les feuilles permettent de représenter les nombres strictement inférieurs à $\mathbf{x}(0) = 2$, soit 0 et 1.

Par exemple, le nombre 818 sera représenté par



où les nombres contenus dans les noeuds (818, 50, 3 et 2) sont donnés à titre d'information : c'est la représentation décimale du nombre correspondant au noeud. On observe que $818_{(10)}$ s'écrit $_{(2)}\mathbf{01}00\mathbf{11}00\mathbf{11}$ en notation binaire avec le bit de poids fort à droite.

Question à développer pendant l'oral 3 *Soit $n \in \mathbb{N}$. On appelle profondeur, notée $\mathbf{ll}(n)$, la hauteur de l'arbre ternaire représentant le nombre n dans le système trichotomique. Donner une borne supérieure asymptotique ("en grand O") de $\mathbf{ll}(n)$. Démontrer, par le moyen d'une suite de nombres $(a_k)_{k\in\mathbb{N}}$ à déterminer, que cette borne est asymptotiquement atteinte.*

Une implémentation naïve des opérations présentées dans ce sujet peut potentiellement consommer une quantité astronomique de mémoire. On fera donc preuve de la plus grande prudence en sauvegardant très régulièrement ses fichiers de travail.

On définit la signature d'un arbre ternaire par la formule

$$\begin{aligned}
\mathbf{Sig}(0) &= 0 \\
\mathbf{Sig}(1) &= u(10) \bmod 97 \\
\mathbf{Sig}(g + \mathbf{x}(p) \times d) &= (\mathbf{Sig}(g) + u(20) \times \mathbf{Sig}(d)) \bmod 97 && \text{si } p \text{ est impair} \\
\mathbf{Sig}(g + \mathbf{x}(p) \times d) &= (\mathbf{Sig}(g) + u(30) \times \mathbf{Sig}(d)) \bmod 97 && \text{si } p \text{ est pair}
\end{aligned}$$

Question 3 *Construire l'arbre ternaire t représentant le nombre $v(k, n)$ et calculer $\mathbf{Sig}(t)$*
a) $(k, n) = (1, 10)$ ***b)*** $(k, n) = (2, 20)$ ***c)*** $(k, n) = (32, 30)$ ***d)*** $(k, n) = (61, 40)$

On définit une suite $(\mathbf{h}(n))_{n\in\mathbb{N}}$ de très (très !) grands nombres de façon récursive :

$$\begin{aligned}
\mathbf{h}(0) &= 1 \\
\mathbf{h}(n+1) &= \mathbf{h}(n) + \mathbf{x}(\mathbf{h}(n)) \times \mathbf{h}(n)
\end{aligned}$$

Question 4 *Calculer $\mathbf{Sig}(\mathbf{h}(v(k, 7k)))$ pour*
a) $k = 0$ ***b)*** $k = 2$ ***c)*** $k = 4$ ***d)*** $k = 8$

Afin d'avoir une intuition des nombres manipulés par la suite, nous utilisons la notation des puissances itérées, définie par :

$$\begin{aligned}
2 \uparrow\uparrow 0 &= 1 \\
2 \uparrow\uparrow (n+1) &= 2^{2\uparrow\uparrow n}
\end{aligned}$$

Question à développer pendant l'oral 4 *Proposer un polynôme $p(X)$ pour lequel $2 \uparrow\uparrow p(n)$ fournit une borne inférieure asymptotique (non triviale) de $\mathbf{h}(n)$. En particulier, on cherchera une approximation vérifiant $2 \uparrow\uparrow p(n) \leqslant \mathbf{h}(n)$ pour tout $n \in \mathbb{N}$.*

Montrer que

$$\mathbf{l}(\mathbf{h}(n)) = 1 + \sum_{0 \leqslant k < n} 2^{\mathbf{h}(k)}$$

pour tout $n > 0$. À partir de quelle valeur de n est-ce que $\mathbf{h}(n)$ n'est plus représentable sous la forme d'un tableau de bits sur un ordinateur "standard" ?

On définit un générateur pseudo-aléatoire $(\mathbf{gen}(k,n))_{k\in[0,61],n\in[0,996]}$ produisant des nombres sous forme d'arbres ternaires :

$$\begin{aligned}
&\mathbf{gen}(0,n) = 0 && \text{si } u(n) \bmod 7 = 0\\
&\mathbf{gen}(0,n) = 1 && \text{si } u(n) \bmod 7 \neq 0\\
&\mathbf{gen}(k,n) = g && \text{si } k > 0 \text{ et } d = 0\\
&\mathbf{gen}(k,n) = g + \mathbf{x}(p) \times d && \text{si } k > 0 \text{ et } d > 0\\
&\quad \text{avec } k' = \max(0, k - 1 - (u(n) \bmod 2))\\
&\qquad g = \mathbf{gen}(k', (n+1) \bmod 997)\\
&\qquad p = v(k', n)\\
&\qquad d = \mathbf{gen}(k', (n+2) \bmod 997)
\end{aligned}$$

Question 5 *Calculer* $\mathbf{Sig}(\mathbf{gen}(k,n))$ *pour*
a) $(k,n) = (6,35)$ ***b)*** $(k,n) = (8,45)$ ***c)*** $(k,n) = (10,55)$ ***d)*** $(k,n) = (12,65)$ ***e)*** $(k,n) = (14,75)$

Dans le système trichotomique, la décrémentation $\mathbf{dec}(n) = n - 1$ (pour $n > 0$) et l'opérateur $\mathbf{x}'(p) = \mathbf{x}(p) - 1$ (pour $p \in \mathbb{N}$) sont définis de façon mutuellement récursive par :

$$\begin{aligned}
&\mathbf{dec}(1) = 0\\
&\mathbf{dec}(g + \mathbf{x}(p) \times d) = \mathbf{dec}(g) + \mathbf{x}(p) \times d && \text{si } g \neq 0\\
&\mathbf{dec}(g + \mathbf{x}(p) \times d) = \mathbf{x}'(p) && \text{si } g = 0 \text{ et } d = 1\\
&\mathbf{dec}(g + \mathbf{x}(p) \times d) = \mathbf{x}'(p) + \mathbf{x}(p) \times \mathbf{dec}(d) && \text{si } g = 0 \text{ et } d \neq 1
\end{aligned}$$

$$\begin{aligned}
&\mathbf{x}'(0) = 1\\
&\mathbf{x}'(p) = \mathbf{x}'(q) + \mathbf{x}(q) \times \mathbf{x}'(q) && \text{pour } p > 0\\
&\text{avec } q = \mathbf{dec}(p)
\end{aligned}$$

Question 6 *Calculer* $\mathbf{Sig}(\mathbf{dec}(\mathbf{gen}(k, 19k \bmod 997)))$ *pour*
a) $k = 6$ ***b)*** $k = 16$ ***c)*** $k = 26$

Pour plus de sûreté, on ne se contentera pas de tester les opérations arithmétiques avec les résultats fournis pour $\widetilde{u_0}$, la fonction signature injectant parfois des nombres distincts vers des signatures égales. On pourra cependant exploiter les entiers machines et leurs opérations arithmétiques primitives à des fins de test.

Question à développer pendant l'oral 5 *Décrire un algorithme calculant l'incrémentation* $\mathbf{inc}(n) = n + 1$ *(pour $n \in \mathbb{N}$) dans le système trichotomique. Quelle est sa complexité temporelle pour $n \in \mathbb{N}$?*

Question 7 *Calculer* $\mathbf{Sig}(\mathbf{inc}(\mathbf{gen}(k, 17k \bmod 997)))$ *pour*
a) $k = 6$ ***b)*** $k = 7$ ***c)*** $k = 16$

Pour $m, n \in \mathbb{N}^2$, on note $\mathbf{compare}(m, n)$ la fonction de comparaison qui renvoie 0 si $m = n$, -1 si $m < n$ et 1 si $m > n$.

Question 8 *Calculer* $\mathbf{compare}(\mathbf{gen}(k, 29k \bmod 997), \mathbf{gen}(k, 31k \bmod 997))$ *pour*
a) $k = 6$ ***b)*** $k = 8$ ***c)*** $k = 16$

Question à développer pendant l'oral 6 *Décrire l'algorithme calculant l'addition* $\mathbf{add}(m, n) = m + n$ *dans le système trichotomique.*

Indication : on pourra s'aider de l'opérateur de normalisation $\mathbf{C}(g, p, d)$ *transformant un triplet* (g, p, d) *dénormalisé (c.-à-d. ne vérifiant pas nécessairement* $\max(g, d) < \mathbf{x}(p)$*) vers sa représentation trichotomique et de l'opérateur de normalisation à droite* $\mathbf{C}_1(g, p, d)$ *transformant un triplet* (g, p, d) *dénormalisé à droite (vérifiant* $0 \leqslant g < \mathbf{x}(p)$*) vers sa représentation trichotomique.*

Ces opérateurs sont définis mutuellement avec l'addition par

$$\begin{aligned}
\mathbf{C}(gg + \mathbf{x}(gp) \times gd, p, d) &= \mathbf{C}_1(gg + \mathbf{x}(gp) \times gd, p, d) && \text{si } gp < p \\
\mathbf{C}(gg + \mathbf{x}(gp) \times gd, p, d) &= \mathbf{C}_1(gg, p, \mathbf{add}(gd, d)) && \text{si } gp = p \\
\mathbf{C}(gg + \mathbf{x}(gp) \times gd, p, d) &= \mathbf{C}(\mathbf{C}(gg, p, d), gp, gd) && \text{si } gp > p
\end{aligned}$$

$$\begin{aligned}
\mathbf{C}_1(g, p, dg + \mathbf{x}(dp) \times dd) &= g + \mathbf{x}(p) \times (dg + \mathbf{x}(dp) \times dd) && \text{si } dp < p \\
\mathbf{C}_1(g, p, dg + \mathbf{x}(dp) \times dd) &= (g + \mathbf{x}(p) \times dg) + \mathbf{x}(\mathbf{inc}(p)) \times dd && \text{si } dp = p \\
\mathbf{C}_1(g, p, dg + \mathbf{x}(dp) \times dd) &= \mathbf{C}(\mathbf{C}_1(g, p, dg), dp, \mathbf{C}_1(0, p, dd)) && \text{si } dp > p
\end{aligned}$$

Question 9 *Calculer* $\mathbf{Sig}(\mathbf{add}(\mathbf{gen}(k, 41k \bmod 997), \mathbf{gen}(k, 43k \bmod 997)))$ *pour*
a) $k = 6$ ***b)*** $k = 12$ ***c)*** $k = 16$

Question à développer pendant l'oral 7 *Décrire l'algorithme calculant la multiplication* $\mathbf{mul}(m, n) = m \times n$ *dans le système trichotomique.*

Question 10 *Calculer* $\mathbf{Sig}(\mathbf{mul}(\mathbf{gen}(k, 41k \bmod 997), \mathbf{gen}(k, 43k \bmod 997)))$ *pour*
a) $k = 5$ ***b)*** $k = 8$ ***c)*** $k = 10$

2 Représentation compacte & calcul efficace

La représentation trichotomique manipulée jusqu'à présent reste cependant naïve. Afin de traiter la soustraction (Question 13 et Question 14) sur des instances non triviales, il nous faut améliorer notre représentation ainsi que la façon de calculer sur celle-ci.

On note $\mathbf{t}(n)$ le nombre de noeuds (en excluant les feuilles) de la représentation trichotomique de n. On a ainsi $\mathbf{t}(0) = \mathbf{t}(1) = 0$ tandis que, par exemple, $\mathbf{t}(818) = 7$. On définit les parties d'un entier par

$$\begin{aligned}
\mathcal{S}(0) = \mathcal{S}(1) &= \emptyset \\
\mathcal{S}(g + \mathbf{x}(p) \times d) &= \{g + \mathbf{x}(p) \times d\} \cup \mathcal{S}(g) \cup \mathcal{S}(p) \cup \mathcal{S}(d)
\end{aligned}$$

et la taille des parties par $\mathbf{s}(n) = |\mathcal{S}(n)|$. On a ainsi $\mathbf{s}(818) = 4$.

Question 11 *Calculer* $(\mathbf{s}(x), \mathbf{t}(x))$ *avec* $x = \mathbf{gen}(k, 23k \bmod 997)$ *pour*
a) $k = 8$ ***b)*** $k = 16$ ***c)*** $k = 24$ ***d)*** $k = 32$

Question à développer pendant l'oral 8 *En extrapolant à partir des résultats de la Question 11, proposez une représentation compacte des entiers produits par le générateur pseudo-aléatoire. (On pourra tester d'autres exemples !)*

Pour deux nombres m et n dans le système trichotomique, quelle est la complexité temporelle du test d'égalité $m = n$ avec votre représentation ?

Indication : la taille des parties de tous les nombres manipulés dans ce sujet est strictement inférieure à 10^6.

On définit la population d'un entier $n \in \mathbb{N}$ de façon récursive par

$$\begin{aligned} \mathbf{pop}(0) &= 0 \\ \mathbf{pop}(1) &= 1 \\ \mathbf{pop}(g + \mathbf{x}(p) \times d) &= \mathbf{pop}(g) + \mathbf{pop}(d) \end{aligned}$$

Question à développer pendant l'oral 9 *Exprimer, en fonction de $n \in \mathbb{N}$, le nombre d'appels récursifs effectués par la fonction* $\mathbf{pop}(\text{-})$ *lors du calcul de* $\mathbf{pop}(k)$ *pour* $k = \mathbf{h}(n)$. *Commenter.*

Question 12 *Calculer* $\mathbf{pop}(\mathbf{gen}(k, 37k \bmod 997)) \bmod 97$ *pour*
a) $k = 48$ ***b)*** $k = 55$ ***c)*** $k = 61$

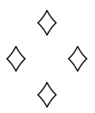
Question 13 *Calculer* $\mathbf{Sig}(\mathbf{dec}(\mathbf{gen}(k, 19k \bmod 997)))$ *pour*
a) $k = 48$ ***b)*** $k = 55$ ***c)*** $k = 61$

On définit la soustraction $\mathbf{sub}(m, n)$ (pour $m > n$) suivant la technique du complément à deux : pour $m = g + \mathbf{x}(p) \times d > n$, on définit $-n$ comme étant l'unique entier vérifiant

$$n + (-n) = 0 \bmod \mathbf{x}(p).$$

Question 14 *Calculer* $\mathbf{Sig}(\mathbf{sub}(\mathbf{gen}(k + 2, 41k \bmod 997), \mathbf{gen}(k, 43k \bmod 997)))$ *pour*
a) $k = 2$ ***b)*** $k = 6$ ***c)*** $k = 8$

Question à développer pendant l'oral 10 *Proposer des améliorations de votre représentation en termes d'efficacité ou d'opérations supportées.*



Fiche réponse type: Très Grands Nombres

$\widetilde{u_0}$: 42

Question 1

a) 33

b) 9

c) 86

Question 2

a) 44

b) 34

c) 36

Question 3

a) 57

b) 65

c) 82

d) 54

Question 4

a) 65

b) 11

c) 4

d) 88

Question 5

a) 9

b) 62

c) 15

d) 66

e) 2

Question 6

a) 63

b) 93

c) 37

Question 7

a) 19

b) 96

c) 30

Question 8

a) 1

b) -1

c) -1

Question 9

a) 38

b) 45

c) 59

Question 10

a) 33

b) 56

c) 57

Question 11

a) (28, 102)

b) (137, 4687)

c) (334, 323015)

d) (710, 11251450)

Question 12

a) 28

b) 11

c) 43

Question 13

a) 60

b) 9

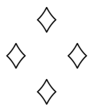
c) 76

Question 14

a) 22

b) 52

c) 82



Fiche réponse: Très Grands Nombres

Nom, prénom, u_0: .....

Question 1

a)

b)

c)

Question 2

a)

b)

c)

Question 3

a)

b)

c)

d)

Question 4

a)

b)

c)

d)

Question 5

a)

b)

c)

d)

e)

Question 6

a)

b)

c)

Question 7

a)

b)

c)

Question 8

a)

b)

c)

Question 9

a)

b)

c)

Question 10

a)

b)

c)

Question 11

a)

b)

c)

d)

Question 12

a)

b)

c)

Question 13

a)

b)

c)

Question 14

a)

b)

c)