Algorithmique numérique

L3

Devoir écrit du 26 mai 2009

Document autorisé : un feuillet personnel blanc manuscrit recto-verso uniquement

Durée : deux heures

Eléments de correction

Exercice 1 :

Représentation des réels en machine

Nous considérons le petit programme expérimental ci-contre écrit en C et le résultat de son exécution (page suivante). La boucle «while» calcule le terme de la suite $u_k=2^{-k}$ tant qu'il reste strictement supérieur à zéro.

```
#include <stdio.h>
#include <math.h>
float u=1, v=1;
int Tu=sizeof(u), n, k=0;

main()
{
printf("Taille d'un flottant simple=");
printf("%d",8*Tu); printf("bits\n\n");
printf("-----------------------\n");
printf("Vers l'infiniment petit:\n");
while (u>0)
      {
      printf("k=%d\t u=%1.20e \n",k,u);
      k=k+1; u=u/2;
      }
getchar();
}
```

- Rappeler la représentation des réels selon la norme IEEE754 simple précision en précisant le décalage de la puissance, la taille de la pseudo-mantisse et de l'exposant.

- Quelle est, dans cette représentation, **l'expression binaire** de u_k pour k=0, pour k=1, pour k=126 et pour k=127 ?

- Que représente u_k pour k=149 ? Quelle est sa représentation binaire pour la machine ?

- Déduire de cette expérience la valeur du **plus petit réel normalisé** représentable en simple précision.

```
C:\temp\DE-L3-2009\FloatMin.exe
Taille d'un flottant simple=32bits

-------------------------
Vers l'infiniment petit:
k=0       u=1.00000000000000000000e+000
k=1       u=5.00000000000000000000e-001
k=2       u=2.50000000000000000000e-001
k=3       u=1.25000000000000000000e-001
k=4       u=6.25000000000000000000e-002
k=5       u=3.12500000000000000000e-002
k=6       u=1.56250000000000000000e-002
k=7       u=7.81250000000000000000e-003
k=8       u=3.90625000000000000000e-003
k=9       u=1.95312500000000000000e-003
k=10      u=9.76562500000000000000e-004
k=11      u=4.88281250000000000000e-004
k=12      u=2.44140625000000000000e-004
k=13      u=1.22070312500000000000e-004
k=14      u=6.10351562500000000000e-005
k=15      u=3.05175781250000000000e-005
k=16      u=1.52587890625000000000e-005
k=17      u=7.62939453125000000000e-006
k=18      u=3.81469726562500000000e-006
k=19      u=1.90734863281250000000e-006
k=20      u=9.53674316406250000000e-007
k=21      u=4.76837158203125000000e-007
k=22      u=2.38418579101562500000e-007
k=23      u=1.19209289550781250000e-007
k=24      u=5.96046447753906250000e-008
k=25      u=2.98023223876953130000e-008
k=26      u=1.49011611938476560000e-008
k=27      u=7.45058059692382810000e-009
k=28      u=3.72529029846191410000e-009
k=29      u=1.86264514923095700000e-009
k=30      u=9.31322574615478520000e-010
k=31      u=4.65661287307739260000e-010
k=32      u=2.32830643653869630000e-010
k=33      u=1.16415321826934810000e-010
k=34      u=5.82076609134674070000e-011
k=35      u=2.91038304567337040000e-011
k=36      u=1.45519152283668520000e-011
k=37      u=7.27595761418342590000e-012
k=38      u=3.63797880709171300000e-012
k=39      u=1.81898940354585650000e-012
k=40      u=9.09494701772928240000e-013
k=41      u=4.54747350886464120000e-013
k=42      u=2.27373675443232060000e-013
k=43      u=1.13686837721616030000e-013
k=44      u=5.68434188608080150000e-014
k=45      u=2.84217094304040070000e-014
k=46      u=1.42108547152020040000e-014
k=47      u=7.10542735760100190000e-015
k=48      u=3.55271367880050090000e-015
k=49      u=1.77635683940025050000e-015
k=50      u=8.88178419700125230000e-016
k=51      u=4.44089209850062620000e-016
k=52      u=2.22044604925031310000e-016
k=53      u=1.11022302462515650000e-016
k=54      u=5.55111512312578270000e-017
k=55      u=2.77555756156289140000e-017
k=56      u=1.38777878078144570000e-017
k=57      u=6.93889390390722840000e-018
k=58      u=3.46944695195361420000e-018
k=59      u=1.73472347597680710000e-018
k=60      u=8.67361737988403550000e-019
k=61      u=4.33680868994201770000e-019
k=62      u=2.16840434497100890000e-019
k=63      u=1.08420217248550440000e-019
k=64      u=5.42101086242752220000e-020
k=65      u=2.71050543121376110000e-020
k=66      u=1.35525271560688050000e-020
k=67      u=6.77626357803440270000e-021
k=68      u=3.38813178901720140000e-021
k=69      u=1.69406589450860070000e-021
k=70      u=8.47032947254300340000e-022
k=71      u=4.23516473627150170000e-022
k=72      u=2.11758236813575080000e-022
k=73      u=1.05879118406787540000e-022
```

```
C:\temp\DE-L3-2009\FloatMin.exe
k=73      u=1.05879118406787540000e-022
k=74      u=5.29395592033937710000e-023
k=75      u=2.64697796016968860000e-023
k=76      u=1.32348898008484430000e-023
k=77      u=6.61744490042422140000e-024
k=78      u=3.30872245021211070000e-024
k=79      u=1.65436122510605530000e-024
k=80      u=8.27180612553027670000e-025
k=81      u=4.13590306276513840000e-025
k=82      u=2.06795153138256920000e-025
k=83      u=1.03397576569128460000e-025
k=84      u=5.16987882845642300000e-026
k=85      u=2.58493941422821150000e-026
k=86      u=1.29246970711410570000e-026
k=87      u=6.46234853557052870000e-027
k=88      u=3.23117426778526440000e-027
k=89      u=1.61558713389263220000e-027
k=90      u=8.07793566946316090000e-028
k=91      u=4.03896783473158040000e-028
k=92      u=2.01948391736579020000e-028
k=93      u=1.00974195868289510000e-028
k=94      u=5.04870979341447560000e-029
k=95      u=2.52435489670723780000e-029
k=96      u=1.26217744835361890000e-029
k=97      u=6.31088724176809440000e-030
k=98      u=3.15544362088404720000e-030
k=99      u=1.57772181044202360000e-030
k=100     u=7.88860905221011810000e-031
k=101     u=3.94430452610505900000e-031
k=102     u=1.97215226305252950000e-031
k=103     u=9.86076131526264760000e-032
k=104     u=4.93038065763132380000e-032
k=105     u=2.46519032881566190000e-032
k=106     u=1.23259516440783090000e-032
k=107     u=6.16297582203915470000e-033
k=108     u=3.08148791101957740000e-033
k=109     u=1.54074395550978870000e-033
k=110     u=7.70371977754894340000e-034
k=111     u=3.85185988877447170000e-034
k=112     u=1.92592994438723590000e-034
k=113     u=9.62964972193617930000e-035
k=114     u=4.81482486096808960000e-035
k=115     u=2.40741243048404480000e-035
k=116     u=1.20370621524202240000e-035
k=117     u=6.01853107621011200000e-036
k=118     u=3.00926553810505600000e-036
k=119     u=1.50463276905252800000e-036
k=120     u=7.52316384526264010000e-037
k=121     u=3.76158192263132000000e-037
k=122     u=1.88079096131566000000e-037
k=123     u=9.40395480657830010000e-038
k=124     u=4.70197740328915000000e-038
k=125     u=2.35098870164457500000e-038
k=126     u=1.17549435082228750000e-038
k=127     u=5.87747175411143750000e-039
k=128     u=2.93873587705571880000e-039
k=129     u=1.46936793852785940000e-039
k=130     u=7.34683969263929690000e-040
k=131     u=3.67341984631964850000e-040
k=132     u=1.83670992315982420000e-040
k=133     u=9.18354961579912120000e-041
k=134     u=4.59177480789956060000e-041
k=135     u=2.29588740394978030000e-041
k=136     u=1.14794370197489010000e-041
k=137     u=5.73971850987445070000e-042
k=138     u=2.86985925493722540000e-042
k=139     u=1.43492962746861270000e-042
k=140     u=7.17464813734306340000e-043
k=141     u=3.58732406867153170000e-043
k=142     u=1.79366203433576590000e-043
k=143     u=8.96831017167882930000e-044
k=144     u=4.48415508583941460000e-044
k=145     u=2.24207754291970730000e-044
k=146     u=1.12103877145985370000e-044
k=147     u=5.60519385729926830000e-045
k=148     u=2.80259692864963410000e-045
k=149     u=1.40129846432481710000e-045
```

Eléments de correction :

- Mantisse : 23b – Exposant : 8b – décalage = 127

$$\boxed{u = (-1)^S . 2^{E-d} . \left[\sum_{i=0}^{M} 2^{-i} . m_i \right] \quad \text{avec} \quad m_0 = 1}$$

N bits										
1b	(N-M-1) bits				M bits					
S	...	E_2	E_1	E_0	m_1	m_2	m_3	m_4	m_5	…

k=0	S	E7	E6	E5	E4	E3	E2	E1	E0	M1	M2	M3	M4	M5	M6																	M23
$u=2^{-0}$	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

E=127	3	F	8	0	0	0	0	0

k=1	S	E7	E6	E5	E4	E3	E2	E1	E0	M1	M2	M3	M4	M5	M6																	M23
u=2^{-1}	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E=126	3				F				0				0				0				0				0				0			

k=2	S	E7	E6	E5	E4	E3	E2	E1	E0	M1	M2	M3	M4	M5	M6																	M23
u=2^{-2}	0	0	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E=125	3				E				8				0				0				0				0				0			

Le plus petit réel normalisé représentable :

k=126	S	E7	E6	E5	E4	E3	E2	E1	E0	M1	M2	M3	M4	M5	M6																	M23
u=2^{-126}	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E=1	0				0				8				0				0				0				0				0			

Pour k=127 il faut dénormaliser

k=127	S	E7	E6	E5	E4	E3	E2	E1	E0	M1	M2	M3	M4	M5	M6																	M23
u=2^{-127}	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E=0	0				0				4				0				0				0				0				0			

Le plus petit réel dénormalisé représentable :

k=149	S	E7	E6	E5	E4	E3	E2	E1	E0	M1	M2	M3	M4	M5	M6																	M23
u=2^{-149}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E=0	0				0				0				0				0				0				0				1			

Exercice 2 :

Résolution d'un système linéaire

Partie 1 :

Soit le système suivant :

$$\left(\begin{array}{ccc|c} 0 & 2 & 3 & 5 \\ 1 & 0 & -2 & 4 \\ 2 & 1 & 0 & 8 \end{array}\right)$$

dont les solutions sont : $x_1 = 2$, $x_2 = 4$ et $x_3 = -1$

remarque : Ce système a donc des solutions, ce qui implique que le déterminant de la matrice n'est pas nul !

Citer deux méthodes de résolution qu'il n'est pas judicieux d'utiliser pour résoudre ce système. **Expliquer pourquoi**.

Eléments de correction :

3 méthodes ne sont pas judicieuses :

***Le pivot de gauss « simple »**, car les pivots sont pris sur la diagonale, et sont tous nuls, on ne peut pas diviser par le pivot.*

***La méthode de Jacobi**, car cette méthode assure la convergence si la matrice est à diagonale dominante, c'est à dire si $|Aii| > \sum_{j=1, j \neq i}^{n} |Aij|$, ce qui n'est pas le cas pour ce système. La convergence n'est donc pas garantie.*

Remarque : ce n'est pas une équivalence, car des systèmes à diagonale non dominante peuvent être résolus par cette méthode !

***La méthode de Gauss-Seidel**, car cette méthode est une variante de la méthode de Jacobi et nécessite les mêmes hypothèses de diagonale dominante pour assurer la convergence vers un résultat.*

Remarque : les méthodes de Newton, de Lagrange n'ont absolument aucune raison d'être invoquées ici.

Citer une méthode de résolution que l'on peut utiliser pour résoudre ce système.

On peut utiliser la méthode du pivot de Gauss partiel ou total (en expliquant qu'on ne peut utiliser le pivot de Gauss simple)

Partie 2 :

Résoudre le système suivant **en utilisant la méthode du pivot partiel de Gauss** (note : on vous indique que les valeurs x_1 à x_4 des inconnues sont tous des entiers)

$$\left(\begin{array}{cccc|c} 3 & 1 & 1 & 2 & 12 \\ -4 & 2 & 1 & -1 & -16 \\ 1 & 0 & -2 & 1 & 5 \\ 1 & 3 & 2 & -2 & -4 \end{array}\right)$$

Eléments de correction :

Remarque : il faut utiliser les combinaisons linéaires définies par la méthode du pivot de gauss et on celles qui font uniquement apparaître des entiers. Par exemple, après le choix du pivot pour la première itération, on calcule : L2 ← L2 +3/4 L1 et non pas L2 ← 4L2+3L1. Il faut également intégrer le second membre(Bi) aux calculs.

choix du pivot :

-4	2	1	-1	-16
3	1	1	2	12
1	0	-2	1	5
1	3	2	-2	-4

apres pivotage de la ligne :

-4	2	1	-1	-16
0	5/2	7/4	5/4	0
0	1/2	-7/4	3/4	1
0	7/2	9/4	-9/4	-8

choix du pivot :

-4	2	1	-1	-16
0	7/2	9/4	-9/4	-8
0	1/2	-7/4	3/4	1
0	5/2	7/4	5/4	0

apres pivotage de la ligne :

-4	2	1	-1	-16
0	7/2	9/4	-9/4	-8
0	0	-29/14	15/14	15/7
0	0	1/7	20/7	40/7

choix du pivot :

-4	2	1	-1	-16
0	7/2	9/4	-9/4	-8
0	0	-29/14	15/14	15/7
0	0	1/7	20/7	40/7

apres pivotage de la ligne :

-4	2	1	-1	-16
0	7/2	9/4	-9/4	-8
0	0	-29/14	15/14	15/7
0	0	0	85/29	170/29

RESULTATS : x1= 3 : x2= -1 : x3= 0 : x4= 2 :

Exercice 3 :

On souhaite déterminer la densité relative de l'air ρ à l'altitude h = 10.5 km. On dispose des mesures suivantes, réalisées à différentes altitudes h :

h(km)	0	1.525	3.050	4.575	6.10	7.625	9.150
$\boldsymbol{\rho}$	1	0.8617	0.7385	0.6292	0.5328	0.4481	0.3741

Quels types de méthodes numériques pourrait-on préconiser ? Pourquoi la méthode de régression linéaire convient ? Utiliser cette méthode pour calculer la densité de l'air à 10.5 km d'altitude (**détailler et expliquer les calculs**).

Eléments de correction :

Il s'agit de calculer une valeur approximative de la densité de l'air à une altitude pour laquelle on ne dispose pas de mesure. D'après le cours, on sait que les méthodes d'interpolation polynômiale ou les méthodes d'approximation pourraient être utilisées.

Mais ici la valeur à calculer se situe à l'extérieur de l'intervalle de mesure. Il s'agit donc d'une extrapolation. Etant donné le nombre de points de mesure, il est peu judicieux d'employer l'interpolation polynômiale qui conduit à des oscillations donc des extrapolations dangereuses.

En l'occurrence, on s'aperçoit en dessinant ces quelques points que la courbe est proche d'une droite (on le saurait déjà avec quelques notions de physique…). La régression linéaire, qui consiste à calculer l'équation d'une droite approximant au mieux un nuage de points, semble donc appropriée.

Il s'agit ensuite d'une application directe de la méthode de régression linéaire:

1. On calcule les coefficients a_0 et a_1 de la droite d'équation $\boldsymbol{\rho} = a_1 \boldsymbol{h} + a_0$ grâce au système d'équation simple donné dans le formulaire (qui permet également de retrouver rapidement les formules du cours, PPT page 127).
2. On en déduit $\boldsymbol{\rho}$=0,251 environ pour $\boldsymbol{h}$ = 10,5.

Exercice 4 :

Calculer une expression du polynôme d'interpolation passant par les points suivants (**expliquer et détailler les calculs**) :

x	-3	2	-1	3	1
y	0	5	-4	12	0

Peut-on trouver d'autres polynômes passant par ces mêmes points ?

Eléments de correction :

Il s'agit de trouver l'expression du polynôme d'interpolation de degré inférieur ou égal à 4 (pour 5 points) qui passe par les points donnés. On peut utiliser la méthode d'interpolation de Lagrange (déconseillée car fastidieuse avec 5 points) ou celle de Newton (plus rapide) grâce à la table des différences divisées (utiliser les formules fournies en annexes du sujet).

D'après le théorème d'unisolvance, il n'existe qu'un seul polynôme de dégré inférieur ou égal à 4 passant par ces 5 points (cf. cours PPT page 84).

Formulaire

- **Expression du polynôme d'interpolation de Newton sur *n* points :**

$$P_{n-1}(x) = a_1+(x-x_1)a_2+(x-x_1)(x-x_2)a_3+\ldots+(x-x_1)(x-x_2)\ldots(x-x_{n-1})a_n$$

où les coefficients a_i vérifient :

$$a_1 = y_1 \qquad a_2 = \nabla y_2 \qquad a_3 = \nabla^2 y_3 \qquad \ldots \qquad a_n = \nabla^{n-1} y_n$$

avec

$$\nabla^k y_i = \frac{\nabla^{k-1} y_i - \nabla^{k-1} y_k}{x_i - x_k}, \; i = k+1, k+2, \ldots n$$

- **Les coefficients a_0 et a_1 de la droite issue d'une régression linéaire sur un ensemble de points (x_i, y_i) sont donnés par :**

$$\begin{bmatrix} 1 & \overline{x} \\ \overline{x} & \overline{x^2} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \overline{y} \\ \overline{xy} \end{bmatrix}$$

où $\overline{x}$ désigne la moyenne des x_i.