

Algorithmes et calcul scientifique

Première session 2018

Durée : 2 heures. Aucun document ni machine autorisé.

Barème indicatif sur 21. Exercices : 5 / 5 / 3 / 7; qualité de la rédaction = ± 1 .

Exercice 1.

1. Rappeler les définitions de l' ε -machine et de l'unité d'arrondi u . Donner les valeurs de ces deux quantités pour le format `binary32` de l'IEEE-754 et préciser les cas où ces deux quantités sont différentes.
2. Donner un exemple de *catastrophic cancellation*. Donner un exemple de correction de ce phénomène.
3. Je résous un problème de conditionnement d'ordre 10^5 avec un algorithme inverse stable. Quelle est la précision relative que je peux attendre du résultat calculé en `binary32`? et du résultat calculé en `binary64`?
4. Expliquer le rôle du *sticky-bit* dans les différentes étapes de l'addition IEEE-754.

Exercice 2.

On rappelle l'algorithme de sommation compensée de W. Kahan (1965) appliqué à n nombres flottants x_i ($i = 1, \dots, n$) et qui calcule s .

```
s = x1;  
c = 0;  
for i = 2 : n do  
    y = xi - c;  
    t = s + y;  
    c = (t - s) - y;  
    s = t;  
end
```

1. Quel est son intérêt?
2. Quel est son principe?
3. Quelles sont ses faiblesses?

Exercice 3.

On rappelle que l'exception `inexact` signale que $\circ(x) \neq x$ (pour $x \neq \text{NaN}$).

Donner les principes d'une mise en œuvre du contrôle de cette condition — afin de lever ou non l'exception `inexact`.

Exercice 4.

Les opérations arithmétiques suivantes sont dites *triviales* :

- addition ou soustraction avec 0 ($x \pm 0$), addition trivialement nulle ($x - x$),
- multiplication ou division par 1 ($x \times 1, x/1$),
- racine carrée de 0.

On considère *uniquement* le format `binary32` de l'IEEE-754 avec arrondi au plus près et on note u l'unité d'arrondi correspondante. Les opérations arithmétiques de l'IEEE-754 sont ici notées avec les symboles mathématiques classiques : $+$, $-$, \times , $/$, $\sqrt{}$. On note $s = x + y$ et $p = x \times y$. On note e l'erreur absolue définie comme la différence entre le résultat exact et le résultat calculé.

Donner, lorsqu'ils existent, deux exemples **non triviaux** d'opérandes `binary32`, x_i et y_i ($i = 1, 2$), qui vérifient chacune des situations suivantes.

Noter \emptyset si aucun opérande existe, et inversement `b32` si tout `binary32` convient. On évitera de répéter le même exemple valable pour des situations différentes. Les réponses seront données en complétant le tableau suivant (à insérer dans la copie avec un signe distinctif mais anonyme répété sur la copie et sur la feuille).

| Situation : | x_1 | y_1 | x_2 | y_2 |
|--------------------------------|-------|-------|-------|-------|
| $s = 2, e = 0$ | | | | |
| $s = 2, e = 2u$ | | | | |
| $s = 2, e = u$ | | | | |
| $s = 2, e = u/2$ | | | | |
| $s = 2, 0 < e < u/2$ | | | | |
| $s = 2^{-126}, e = 0$ | | | | |
| $s = 2^{-127}, e = 0$ | | | | |
| $s = 2^{-149}, e = 0$ | | | | |
| $s = 2^{-150}, e = 0$ | | | | |
| $s = 2^{-149}, e = 2^{-149}$ | | | | |
| $s = 2^{-149}, e = 2^{-150}$ | | | | |
| $1 + x = 1$ | | | | |
| $1 + (x + y) \neq (1 + x) + y$ | | | | |
| $x - y = 0, e \neq 0$ | | | | |
| $p = 1, e = 0$ | | | | |
| $p = 1, e = u$ | | | | |
| $p = 1 - u, e = 0$ | | | | |
| $x/y, e = 0$ | | | | |
| $1/x, e = 0$ | | X | | X |
| $1/x = 1/y$ ($x \neq y$) | | | | |