EISTI
CERGY - PAU

Correction du TP1 : Erreur due à la conversion en binaire

1. On a $x = 0.1_{10}$. Le nombre en binaire obtenu est :

$$y_2 = 0, 0\underline{0011}...{}_2$$

2. La valeur décimale de y est obtenue par la fonction `function [y]=BinaireDecimal(x)` à partir de l'initialisation de x réalisée par la fonction `function [x]=Initx()`. La variable x est le vecteur des décimales de 0.1 écrit en binaire. Le programme principal qui vous est fourni renvoie deux valeurs de z : l'une sans format d'écriture, l'autre avec. Comparez les résultats obtenus pour vous convaincre de l'utilité de l'écriture avec format.

3. Une simple boucle suffit pour calculer les valeurs de som et som1, sans oublier une initialisation de la variable som (resp. som1) à 0. En effet certains compilateurs n'initialisant pas à zéro les variables lors de leur déclaration (qui n'existe pas dans Scilab), il est toujours préférable d'initiliser à zéro.

Les résultats donnent :

```
valeur de som :  9.9999999999998670000e-001
valeur de som1 :  9.9999999999999890000e-001
différence entre som et som1 :  -1.2212453270876722000e-015
```

On observe une différence entre les deux valeurs qui correspond à :

$$\begin{aligned} som - som1 &= 10z - 10x \\ &= 10(z - x) \simeq -1.2.10^{-15} \end{aligned}$$

Si on appelle f l'application qui convertit de décimal en binaire, et par suite sa réciproque f^{-1} l'application qui convertit de binaire en décimal, on devrait avoir :

$$z = (f^{-1} o f)(x) = x$$

En fait $z \neq x$ car la conversion de décimal en binaire est assortie d'une troncature de la valeur exacte $0, 0\underline{0011}...{}_2$ en $0, \underbrace{000110011...001}_{\text{sur 52 bits}}$ donc en convertissant on a $m(y) < y$. Ensuite on applique f^{-1} qui est monotone croissante puisque f l'est aussi, et donc

$$z = f^{-1}\left[m(y)\right] < f^{-1}\left(y\right) = x$$

ce qui correspond bien au fait que l'on observe que $som - som1 = 10(z - x) < 0$. On peut même préciser qu'ici

$$z - x = -1.2.10^{-16}$$