Rappel : (vous aurez besoin de deux notions vues au dernier cours)

1.Retrouver le nombre sans exposant à partir d’un nombre normalisé

Ex : Écrire le nombre sans exposant

En base 10 : 6.5837 \* 102 = 658.37

4.5 \* 10-2 = .045

En base 2 : 1.1101 \* 21 = 11.101

* 1. \* 2-1 = .11

Remarque : en base 2, la mantisse d’un nombre normalisé commence toujours par 1. (Pourquoi ?)

Exercice : Écrire les nombres suivants sans exposant :

1. 5.678 \* 103
2. 2.4 \* 10-3
3. 1.11 \* 2-2
4. 1.0011 \* 22

Solution

1. 5678
2. .0024
3. .0111
4. 100.11

2. La transformation de la partie fractionnaire de binaire à décimal

On fait la somme de chaque chiffre binaire multiplié par sa valeur positionnelle (rang). Les rangs sont négatifs pour les chiffres après le point.

Ex :

0.0112 = ( )10

0 + (1 \* 2-2) + (1 \* 2-3)

= 1/22 + 1/23

= ¼ + 1/8

= 0.25 + 0.125 = 0.375

Réponse = 0.375

Exercice : Transformez les valeurs suivantes en décimal

1. 0.12
2. 0.10110

Solution

a) (1 \* 2-1) = 1/21 = 0.510

b) (1 \* 2-1) + (1 \* 2-3) = ½ +1/8 = 0.5 + 0.125 = 0.62510

Pour savoir quel nombre à virgule est représenté en mémoire, il faut faire l’inverse de la méthode précédente :

1. mettre en binaire le nombre hexadécimal
2. bits 23 à 30 inclusivement : exposant
3. bits 0 à 22 : mantisse
4. bit 31 : signe

Exemple :

Quel est le nombre à virgule représenté par C1DB9000 ?

1. C 1 D B 9 0 0 016 = 1100 0001 1101 1011 1001 0000 0000 00002

= 1 10000011 10110111001000000000000

1. Bits 23 à 30 = exposant = 100000112 = 13110

131 – 127 = 4

1. a) Bits 0 à 22 = mantisse = 10110111001

En utilisant l’exposant, on a : 1. 10110111001 \* 24

Il faut se débarrasser de l’exposant pour trouver la valeur du nombre : 11011.0111001

b) Il reste à mettre 11011.0111001 en décimal :

Partie entière : 110112 = 2710

Partie après le point : .0111001 = ¼ + 1/8 + 1/16 + 1/128 = .4453

Nombre à virgule en décimal : 27.4453

1. Comme il y a un 1 dans le bit de signe (le bit 31), la réponse est : -27.4453

En C, les types réel simple précision et réel double précision sont respectivement : FLOAT et DOUBLE.