

<p style="text-align: center;">math202 : mathématiques pour le numérique TD 1 : binaire, codage, nombres</p>
--

Pierre Hyvernats
Laboratoire de mathématiques de l'université de Savoie
bâtiment Chablais, bureau 17, poste : 94 22
email : Pierre.Hyvernats@univ-savoie.fr
www : <http://www.lama.univ-savoie.fr/~hyvernats/>

Partie 1 : Rappels : base 2 et complément à 2

Exercice 1 : Nombres négatifs

Question 1. Donnez la représentation en complément à 2 sur 8 bits des nombres suivants :

12, -25, 100, -100, 1000.

Question 2. Donnez la valeur (représentation en base 10) des nombres suivants, donnés en complément à 2 sur 8 bits

00101010₂, 11010101₂, 11111111₂, 01111111₂.

Question 3.

- Donnez la représentation binaire du plus grand entier représentable avec n bits, ainsi que sa valeur.
- Donnez la représentation binaire du plus petit entier représentable avec n bits, ainsi que sa valeur.
- Lorsqu'un entier est encodé en complément à deux, comment déterminer s'il s'agit d'un nombre pair ou impair ?
- Lorsqu'un entier est encodé en complément à deux, comment déterminer s'il s'agit d'un nombre positif ou négatif ?

Question 4. Quand on veut trouver la représentation en base 2 d'un nombre écrit en base 10, est-il plus simple de commencer par chercher les bits de poids fort, ou ceux de poids faible.

Décrivez (et programmez) un algorithme pour effectuer cette transformation.

Partie 2 : exponentielle et logarithme binaires

Question 1. Sur Youtube, en 2014, la vidéo Gangnam Style a été visionnée un nombre de fois assez grand pour dépasser la capacité d'un entier signé 32 bits. Si Youtube avait remplacé son compteur par un entier non signé de même taille, est-ce qu'il pourrait contenir le nombre actuel de vues, soit 3060995409 ?

Question 2. La solution adoptée par Youtube fut de remplacer l'entier signé 32 bits par un entier signé 64 bits. Sachant qu'il y a environ 7 milliards d'êtres humains sur terre, combien de fois chaque personne devrait regarder ce vidéo pour dépasser la capacité de ce nouveau compteur ?

Question 3. Avec le protocole IPv4, chaque hôte se voit attribué une adresse IP. Ces adresse sont habituellement représentées par 4 nombres séparés par des points, par exemple : 23.235.33.67. Sachant que chacun de ces nombres peut prendre une valeur allant de 0 à 255, combien de bits sont nécessaire pour coder une adresse IP en IPv4 ?

Question 7. Que se passe-t'il si on demande d'afficher 1/5 avec 60 chiffres ?

```
>>> print("1/5 = {0:.60f}".format(1.0/5))
```

Question 8. Expliquez le résultat de l'expérience suivante :

```
>>> 0.1
0.1
>>> 0.2
0.2
>>> 0.1+0.2
0.30000000000000004
>>> (0.1 + 0.2) + 0.3
0.6000000000000001
>>> 0.1 + (0.2 + 0.3)
0.6
```

Question 9. Pour les flottants IEEE754 (32 bits), les valeurs des paramètres sont :

- 8 bit pour l'exposant (avec un biais de 127),
- 23 bit pour la mantisse.

On rappelle que lorsque l'exposant est 0000000_2 et que la mantisse contient au moins un 1, on parle de nombre *sous-normal*.

Question 10. Quelle est la représentation du plus grand flottant IEEE754 ? Estimez sa valeur.

Question 11. Quelle est la représentation du plus petit flottant IEEE754 strictement positif ? Estimez sa valeur.

Que se passe-t'il si vous ajoutez ce nombre à celui de la question précédente ?

Partie 4 : caractères : UTF-8

On rappelle que les caractères UTF-8 (UCS Transformation Format 8 bit, UCS = Universal Character Set) sont codés sur un, deux, trois ou quatre octets.

Les codes UTF-8 sont toujours de la forme :

- 0..... : caractère ASCII sur un seul octet,
- 110..... 10..... : caractère sur deux octets,
- 1110.... 10..... 10..... : caractère trois sur octets,
- 11110... 10..... 10..... 10..... : caractère sur quatre octets.

Question 1.

- Combien de caractères peuvent être codés en ASCII ?
- Combien de caractères peuvent être codés sur les extensions 8 bits de l'ASCII ?
- Combien de caractères peuvent être codés en UTF-8 ?

Question 2. Le texte du roman "Les Misérables" forme une chaîne de 3 096 302 caractères. En UTF-8, 3011682 de ces caractères sont encodés sur un octet alors que les 84620 restants sont encodés sur 2 octets. Combien de bits sont utilisés pour encoder l'intégralité du roman en UTF-8 ?

Question 3. L'ensemble des caractères encodés par l'UTF-8 contient, entre autres, les 128172 caractères Unicode. Si on utilisait un code à taille fixe permettant d'encoder tous les caractères Unicode, combien de bits seraient nécessaires pour encoder l'intégralité du roman "Les Misérables" ?