Mathématiques appliquées à l'informatique
1 – Arithmétique : numération

Bertrand LIAUDET

SOMMAIRE

Dernière édition : octobre 2017

CHAPITRE 1 - NUMERATION

0 - Préambule

Principes

Partir d'une application pour remonter à l'objet mathématique théorique qu'on étudie abstraitement ensuite.

Ensuite on fait des exercices théoriques (des gammes), des exercices appliqués et des traductions algorithmiques.

Le but est de rentrer dans une logique de mathématiques appliquées, c'est-à-dire : les maths comme un outil et pas une abstraction creuse.

Objectif : ne pas avoir peur des maths ! Les voir comme un outils pratique qui permet de régler des problèmes efficacement.

Références

Mathématique pour l'informatique – BTS SIO – Dunod – 2015 : Chapitre 1, pp. 3-33.

Méthodes mathématiques pour l'informatique – IUT-Licence-Ecole d'ingénieurs-CNAM – Dunond 2013.

Le problème : calcul d'un développement limité en C

En mathématiques, un développement limité d'une fonction en un point est une approximation polynomiale de cette fonction, c'est-à-dire l'écriture de cette fonction sous la forme d'une somme allant à l'infini.

Par exemple : $\sin(x) = x - x^3 / 3! + x^5/5 ! - x^7 / 7! + x^9/9 ! \ldots$

Si on veut calculer sin(x) avec un ordinateur, le problème est de savoir quand on arrête la somme.

Autrement dit, à partir de quelle valeur de N, N ! est trop grand pour être représenté, ou 1/N ! est trop petit pour être représenté.

En langage C, il existe un fichier float.h qui définit la constante FLT_EPSILON

FLT_EPSILON plus petit nombre e tel que $1.0 + e \neq 1.0$

On veut écrire un algorithme qui permettent de calculer sin(x).

Pour bien comprendre cet exercice, mieux vaut comprendre le standard IEEE 754

Exercices d'approche

Codage de l'information dans les ordinateurs

1. Soit les 16 bits suivants : 1110 1001 0111 0000

   Combien valent-ils en ASCII étendu Latin 1 (cf. rappels de cours plus bas) ?

   Combien valent-il en tant qu'entier codé sur 2 octets en complément à 2. Ecrire les détails de votre calcul.

   Résumé de cours pour cet exercice :

   Le complément à 2 consiste à utiliser le bit de poids fort comme bit de signe (0 pour positif). Le reste correspond à l'entier s'il est positif. Pour les entiers négatifs, on commence par inverser tous les bits (complément à 1) puis on ajoute 1 (complément à 2). De ce fait, -1 est codé avec des 1 pour tous les bits.

   Le code ASCII des minuscules se situe entre $(97 \text{ et } 122)_{10}$ soit $(61\text{-}7A)_{16}$ et entre $(65 \text{ et } 90)_{10}$ soit $(41\text{-}5A)_{16}$) pour les majuscules. On précise qu'on a les lettres suivantes à partir de 230 (code ascii étendu, Latin 1, ISO 8859-1) : 'æ', 'ç', 'è', 'é', 'ê'

2. Donnez la traduction à laquelle correspond le mot de 4 octets codé en hexadécimal suivant : 49 55 50 31selon qu'on le lit comme :

- un entier signé - un entier représenté en complément à 2,
- un nombre représenté en virgule flottante simple précision suivant la norme IEEE 754,
- une suite de caractères ASCII (représentés chacun sur 8 bits, le bit de plus fort poids étant inutilisé et codé à 0)

3. Convertir le nombre décimal 8,625 en virgule flottante suivant la norme IEEE 754 simple précision.

4. Convertissez le nombre 011001010 11100010 10101011 11000101 représenté sous forme d'un nombre entier binaire signé de 4 octets en un nombre réel représenté selon la norme IEEE 754 simple précision.

Convertissez le nombre réel obtenu en entier binaire signé codé sur 32 bits et comparez le résultat obtenu avec le nombre entier de départ.

Quelles conclusions en déduisez-vous ?

5. Calcul de FLT_EPSILON

Le langage C définit une constante : FLT_EPSILON. C'est la différence entre 1 et la plus petite valeur représentable supérieure à 1. Autrement dit, c'est la différence la plus petite qu'il puisse y avoir entre deux nombres X1 et X2. Donc, avec X1 ≥ X2 si X1-X2 < FLT_EPSILON, on peut considérer que X1 = X2.

En général, FLT_EPSILON = 1,19209 e^{-07}

La précision est de 6 chiffres : elle est fixée par une autre constante du C : FLT_DIG

On va calculer « logiquement » FLT_EPSILON

- ***Représentation de 1 en float***

  Convertissez 1 en float : $1_{(float)}$

- ***Représentation « logique » de 1 + FLT_EPSILON en float***

  A partir de la définition de FLT_EPSILON et $1_{(float)}$, déduisez $(1+FLT\_EPSILON)_{(float)}$

- ***Valeur de FLT_EPSILON en binaire***

  A partir de $(1+FLT\_EPSILON)_{(float)}$, déduisez $FLT\_EPSILON_{(2)}$

- ***Convertissez $FLT\_EPSILON_{(2)}$ en $FLT\_EPSILON_{(float)}$***

- ***Valeur de FLT_EPSILON en décimal***

  A partir de $FLT\_EPSILON_{(2)}$, calculez $FLT\_EPSILON_{(10)}$: vous pouvez utiliser une calculatrice !

- ***Valeur de FLT_EPSILON en binaire à partir de FLT_EPSILON en décimal***

  A partir de $FLT\_EPSILON_{(10)}$, calculez $FLT\_EPSILON_{(2)}$: vous pouvez utiliser une calculatrice ! Que constatez-vous ?

1 - Systèmes de numération des entiers naturels

Système de numération

Un système de numération décrit la façon avec laquelle les nombres sont représentés.

Système « naturel » = système décimal : la base 10

Présentation

- Les entiers naturels en base 10 sont 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, etc.
- Base 10 : c'est la base « naturelle ». 10 correspond au nombre de chiffres et au premier nombre à s'écrire avec 2 chiffres.
- 10 chiffres : { 0 ; 1 ; 2 ; 3 ; 4 ; 5 ; 6 ; 7 ; 8 ; 9 }
- Les chiffres sont aux entiers ce que les lettres sont aux mots.
- Le système de numération avec 10 chiffres, c'est le système décimal, ou système naturel.

Addition des puissances de la base

Un entier peut s'écrire sous la forme d'une addition de puissances de la base :

$$4134 = 4x10^3 + 1x10^2 + 3x10^1 + 4x10^0$$

De façon générale, on a :

$$(a_n, .., a_1, a_0) = a_n*10^n + .. + a_1*10^1 + a_0*10^0 = \Sigma_{i=0}^{n} a_i * 10^i$$

Système positionnel

Le système décimal est un **système positionnel** : chaque position du chiffre dans le nombre possède un poids (unité, dizaine, centaine, etc.) qui correspond à un numéro de position et à la puissance de 10 associée :

$$4134 = 4x10^3 + 1x10^2 + 3x10^1 + 4x10^0$$

position :	3	2	1	0
chiffre :	4	1	3	4

Le chiffre en position 0 est le **chiffre de poids faible**.

Le chiffre en position la plus grande est le c**hiffre de poids fort.**

Avantages - inconvénients

La base 10 permet de faire facilement les opérations en connaissant les tables d'opération élémentaires.

Base 2 : système binaire

Présentation

- Les nombres binaires (en base 2) sont 0, 1, 10, 11, 100, 101, 110, 111, 1000, 1001, 1010, etc.
- Base 2 : c'est le binaire. 2 correspond au nombre de chiffres. 2 vaut 10 en binaire et est le premier nombre à s'écrire avec 2 chiffres.
- 2 chiffres : {0 ; 1 }

Conversion en base 10 : Addition des puissances de la base

Un nombre binaire peut s'écrire en base 10 sous la forme d'une addition de puissances de la base :

$$\begin{aligned} 10101_{(2)} &= 1x2^4 + 0x2^3 + 1x2^2 + 0x2^1 + 1x2^0 \\ &= 16 + 0 + 4 + 0 + 1 \\ &= 21_{(10)} \end{aligned}$$

De façon générale, on a :

$$(a_n, .., a_1, a_0) = a_n*2^n + .. + a_1*2^1 + a_0*2^0 = \Sigma_{i=0}^{n}\ a_i * 2^i$$

Simplifications

- ***Principes***

  A la place de $1x2^n$ on peut écrire 2^n

  A la place de $0x2^n$ on n'écrit rien

  A noter que $2^0 = 1$

- ***Exemples***

  $10101 = 2^4 + 2^2 + 2^0 = 16 + 4 + 1 = 21$

Système positionnel

Le système binaire est un **système positionnel** : chaque position du chiffre dans le nombre possède un poids qui correspond à un numéro de position et à la puissance de 2 associée :

$10101_{(2)} =$	$1x2^4$	$+ 0x2^3$	$+ 1x2^2$	$+ 0x2^1$	$+ 1x2^0$
position :	4	3	2	1	0
chiffre :	1	0	1	0	1

Le chiffre en position 0 est le **chiffre de poids faible**.

Le chiffre en position la plus grande est le c**hiffre de poids fort.**

Tableau des puissances

Position	10	9	8	7	6	5	4	3	2	1	0
Puissance	2^{10}	2^9	2^8	2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
Valeur en base 10	1024	512	256	128	64	32	16	8	4	2	1

Tableau de conversion

Base 10	0	1	2	3	4	5	6	7	8	9	10	11	Etc.
Base 2	0	1	10	11	100	101	110	111	1000	1001	1010	1011	Etc.

Utilisations

La base 2 est utilisée dans les ordinateurs.

On retrouve des puissances de 2 un peu partout dans l'informatique :

- Mémoire des téléphones portable : 32 GO, 128 GO, etc.
- Barrettes de RAM : 8GO, 16GO, etc.
- Clé USB : 16 GO, 32 GO, etc.

Ce sont toujours des multiples de 2.

Base 16 : système hexadécimal

Présentation

- Les nombres hexadécimaux (en base 2) sont 0, 1, 2, …, 8, 9, A, B, C, D, E, F, 10, 11, 12, …, 19, 1A, 1B, 1C, 1D, 1E, 1F, 20, 21, 22, …, 29, 3A, 3B, etc.
- Base 16 : c'est l'hexadécimal ou hexa. 16 correspond au nombre de chiffres. 16 vaut 10 en hexa et est le premier nombre à s'écrire avec 2 chiffres.
- 16 chiffres : {0 ; 1 ; 2 ; 3 ; 4 ; 5 ; 6 ; 7 ; 8 ; 9 ; A ; B ; C ; D ; E ; F }
- A=10, B=11, C=12, D=13, E=14, F=15

Conversion en base 10 : Addition des puissances de la base

Un nombre binaire peut s'écrire en base 10 sous la forme d'une addition de puissances de la base :

$$\begin{aligned} 1A2E_{(16)} &= 1x16^3 + Ax16^2 + 2x16^1 + Ex16^0 \\ &= 1x4096 + 10x256 + 2x16 + 14x1 \\ &= 6072_{(10)} \end{aligned}$$

De façon générale, on a :

$$(a_n, .., a_1, a_0) = a_n*16^n + .. + a_1*16^1 + a_0*16^0 = \Sigma_{i=0}^{n}\ a_i * 16^i$$

Simplification

A noter que $16^0 = 1$

Système positionnel

Le système binaire est un **système positionnel** : chaque position du chiffre dans le nombre possède un poids qui correspond à un numéro de position et à la puissance de 2 associée :

$1A2E_{(16)} = 1x16^3 + Ax16^2 + 2x16^1 + Ex16^0$

position :	3	2	1	0
chiffre :	1	A	2	E

Le chiffre en position 0 est le **chiffre de poids faible**.

Le chiffre en position la plus grande est le c**hiffre de poids fort.**

Tableau des puissances

Position	6	5	4	3	2	1	0
Puissance	16^6	16^5	16^4	16^3	16^2	16^1	16^0
Valeur en base 10	2^{24}	2^{20}	2^{16}	4096	256	16	1

Tableau de conversion

Base 10	0	1	..	9	10	..	15	16	17	..	25	26	Etc.
Base 16	0	1	..	9	A	..	F	10	11	..	19	1A	Etc.

Notation $(n)_p$ ou $n_{(p)}$

- On utilise la notation $(n)_p$ pour indiquer la base du nombre : $(57)_{10}$, $(1A)_{16}$, $(101101)_2$
  $(23)_{10} = (1011\ 1)_2 = (17)_{16}$
- On peut aussi utiliser la notation $n_{(p)}$ pour indiquer la base du nombre : $57_{(10)}$, $1A_{(16)}$, $101101_{(2)}$
  $23_{(10)} = 1011\ 1_{(2)} = 17_{(2)}$
- On écrit parfois $(17)_h$ à la place de $(17)_{16}$

Puissances de 2 remarquables

On a intérêt à connaître la série des puissances de 2, et particulièrement 2^8 et 2^{10}

2^1	2^2	2^3	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}	2^{20}	2^{30}
2	4	8	16	32	64	128	256	512	1024	$\approx 10^6$	$\approx 10^9$

Remarques sur les bases 2 et 16

2 et 16 sont des puissances de 2 : $2 = 2^1$, $16 = 2^4$

Pour traduire de l'hexadécimal en décimal, on peut travailler avec les puissances de 2.

Sachant que $(n^p)^q = n^{p*q}$, on a, par exemple : $16^3 = (2^4)^3 = 2^{4*3} = 2^{12} = 4096$

Base 2 vers la base 16 et inversement

Etant donné que 2^4=16, les quatre premiers chiffres d'un nombre en base 2 correspondent au premier chiffre du même nombre en base 16, et ainsi de suite.

Idem entre la base 2 et la base 8 : 2^3=8. Les 3 premiers chiffres d'un nombre en base 2 correspondent au premier chiffre du même nombre en base 8, et ainsi de suite.

On va donc utiliser un tableau de conversion entre les chiffres en base 16 et le nombre correspondant en base 2 :

Base 16	0	1	2	3	4	5	6	7
Base 2	0000	0001	0010	0011	0100	0101	0110	0111

Base 16	8	9	A	B	C	D	E	F
Base 2	1000	1001	1010	1011	1100	1101	1110	1111

Exemples de conversion

- ***Méthode***

  On regroupe les chiffres du nombre binaire par paquets de 4 chiffres qui correspondent à 1 chiffre en hexadécimal.

  On peut séparer les paquets de 4 par des tirets pour faciliter la lecture

- ***Exemple 1***

  $(1001101)_2 = (0100\text{-}1101)_{2\,=} (4D)_{16}$

Le premier paquet de 4 : (0100) vaut 4 en hexa

Le deuxième paquet de 4 : (1101) vaut D en hexa

Vérification

$(1001101)_2 = 2^6 + 2^3 + 2^2 + 2^0 = 64 + 8 + 4 + 1 = 77$

$(4D)_{16} = 4x16^1 + Dx16^0 = 64 + 13 = 77$

- ***Exemple 2***

$(7C2)_{16} = (0111\text{-}1100\text{-}0010)_2 = (11111000010)_2$

7 en hexa vaut (0111) en binaire

C en hexa vaut 12 en base 10 et (1100) en binaire

2 en hexa vaut (0010) en binaire

Vérification

A faire en exercice !

Exercices – série 1 – Bases 10, 2, 16

Base 10

Ecrire les nombres suivants sous la forme d'une addition des puissances de 10 :

7324 100234 10101010 1234,56 1000000,001

Base 2

- ***Pour les nombres suivants :***

  $101010_{(2)}$ $(11011101)_2$ $101110011_{(2)}$ $101\ 1011\ 0111_{(2)}$

  - Les écrire sous la forme d'une addition des puissances de la base.
  - Donner leur valeur en base 10.

- ***Combien valent précisément et/ou approximativement :***

  $2^8, 2^{10}, 2^{20}, 2^{30}$

Base 16

- ***Pour les nombres suivants :***

  $1F2_{(16)}$ $(B4)_{(16)}$ $101A_{(16)}$ $(E0B7)_{(16)}$ $2A8F_{(16)}$

  Les écrire sous la forme d'une addition des puissances de la base.

  Donner leur valeur en base 10.

Base 2 vers base 16 et inversement

- Faire la vérification de l'exemple 2 : vérifier que $(7C2)_{16} = (11111000010)_2$
- Donner la valeur de $30A_{(16)}$ et de $(C8DE)_{16}$ en base 2 et vérifiez le.
- Donner la valeur de $10010111_{(2)}$ et de $(110\ 0111\ 0100\ 0101)_2$ en base 16 et vérifiez le.

Exercices – série 2 - Numération romaine

Présentation

http://chiffreromain.com

- Les nombres romains sont : I, II, III, IV, V, VI, VII, VIII, IX, X, XI, XII, etc.
- Le système constitué de 7 symboles

Romain	I	V	X	L	C	D	M
Système décimal	1	5	10	50	100	500	1000

- On peut mettre jusqu'à 3 symboles I, X, C, M à la suite. Pas plus. Par exemple : MMMCCCXXXIII pour 3333.
- Pour obtenir l'équivalent de 4 symboles I, X, C, on les fait suivre respectivement par V, L, D : IV, IX, XL, XC, CD, CM. Par exemple CDXLIV pour 444. CMXCIX pour 999.
- Le plus grand nombre exprimable est 3999 : MMMCMXCIX.

Avantages - inconvénients

Défaut majeur : ce système n'est pas pratique pour faire des opérations.

Exercices

- ***Ecrivez en numération romaine :***

  12, 37, 49, 75, 98, 437, 495, 2995

- ***Concevez un système***

  Imaginez une version adaptée du système romain qui permettent d'écrire les entiers jusqu'à l'infini. Demandez vous comment écrire 10 mille, 100 mille, 1 million, 100 millions, 1 milliard.

  Vous devez définir les règles du système.

  Ecrivez alors 2000, 3212, 397 742, 253 521 430.

Exercices – série 3 - Base 8 : système octal

Présentation

Base 8

8 chiffres : {0 ; 1 ; 2 ; 3 ; 4 ; 5 ; 6 ; 7 }

$(a_n, a_{n-1},.., a_1, a_0) = a_n*8^n + a_n*8^{n-1} +.. + a_1*8^1 + a_0*8^0 = \Sigma_{i=0}^{n} a_i*8^i$

Exemples

$25736_{(8)} = 2x8^4 + 5x8^3 + 7x8^2 + 3x8^1 + 6x8^0$

$= 2x4096 + 5x512 + 7x64 + 24 + 6 =$

$= 2x8192 + 2560 + 448 + 24 + 6 =$

$= 11230$

$9876_{(8)} = 9x8^3 + 8x8^2 + 7x8^1 + 6x8^0$

$= 9x512 + 8x64 + 56 + 6$

Tableau des puissances

Position	6	5	4	3	2	1	0
Puissance	8^6	8^5	8^4	8^3	8^2	8^1	8^0
Valeur en base 10	261144	32768	4096	512	64	8	1

Tableau de conversion

Base 10	0	1	..	7	8	9	12	15	16	17	..	23	Etc.
Base 8	0	1	..	7	10	11	..	17	20	21	..	27	Etc.

Base 2 vers la base 8 et inversement

Etant donné que $2^3=8$, les trois premiers chiffres d'un nombre en base 2 correspondent au premier chiffre du même nombre en base 8, et ainsi de suite.

On peut utiliser cette technique pour passer le la base 2 à la base 8 et inversement comme on l'a fait pour passer de la base 2 à la base 16.

Exercices

- ***Pour les nombres suivants :***

  $(54)_{(8)}$ $(107)_{(8)}$

  Les écrire sous la forme d'une addition des puissances de la base.

  Donner leur valeur en base 10.

- ***Base 2 vers base 8 et inversement***
  - Donner la valeur de $307_{(8)}$ en base 2 et vérifiez-le.
  - Donner la valeur de $10010111_{(2)}$ en base 8 et vérifiez-le.

2 - Conversion d'une base à une autre

Présentation

Il y a 3 cas de conversions :

- Des autres bases, et particulièrement 2, 8 et 16, vers la base 10 : on a traité ça au chapitre précédent.
- Des autres bases que 10 entre elles (et particulièrement de 2 vers 16) : on a traité ça au chapitre précédent.
- De la base 10 vers la les autres bases, et particulièrement 2, 8 et 16 : on va traiter ça dans ce chapitre

Convertisseur en ligne : ici ou là

Base 10 vers les bases 2, 8 ou 16

2 méthodes

Il existe 2 méthodes qui s'appuient sur la division euclidienne (division entière) :

- La méthode par la puissance
- La méthode par la division euclidienne

Rappel : la division euclidienne (ou division entière)

➢ ***Exemple***

La division euclidienne c'est une division dont le résultat, appelé quotient, est un entier (sans chiffres après la virgule). Il y a donc possiblement un reste.

Exemple : 17 divisé par 5 = 3 et il reste 2

➢ ***Opérateur DIV et MOD, // et %***

L'opérateur de division entière s'écrit DIV ou parfois //

On écrit 17 DIV 5 = 3 ou 17 // 5 = 3

Pour connaître le reste, l'opérateur est MOD ou %

On écrit 17 MOD 5 = 2 ou 17 % 5 = 2

➢ ***Propriété***

17 = 5 * 3 + 2

17 est le diviseur 5 est le dividende 3 est le résultat 2 est le reste

➢ ***Formulation théorique***

Soit A, B, Q et R : 4 entiers naturels. Effectuer la division euclidienne de A par B non nul c'est déterminer les uniques entiers Q (appelé quotient) et R (appelé reste) tels que :

$$A = B \times Q + R \text{ avec } R < B.$$

Principe des méthodes

On part d'une base source **BS** pour arriver à une base cible **BC**.

On part d'un $(\mathbf{nBS})_{bs}$ pour arriver à un $(\mathbf{nBC})_{bc}$ avec $(\mathbf{nBS})_{bs} = (\mathbf{nBC})_{bc}$

Notre **BS** sera toujours la base 10 : BS = 10.

Méthode par la puissance

Présentation

La méthode décrit un algorithme :

On cherche la plus grande puissance de de la base cible qui est inférieure au nombre de la base source.

Pour cela, on cherche à encadrer nBS par deux puissances successives.

Ensuite, on effectue la division euclidienne du **nBS** par la puissance inférieure (qui est donc la plus grande puissance inférieure possible).

On recommence en remplaçant le **nBS** par le reste, jusqu'à ce que la puissance vaille 0 ou que le reste vaille 0.

On écrit alors le nombre par une addition des puissances inférieures trouvées : c'est une écriture du nombre sous forme d'une addition des puissances de la base.

Les facteurs des puissances sont les chiffres du **nBC**

Exemple 1 : $(77)_{10}$ en base 2

nBS	Encadrement	Division euclidienne
77	$2^6 = 64 <= 77 < 2^7 = 128$	$77 = 2^6 + 13$
13	$2^3 = 8 <= 13 < 2^4 = 16$	$13 = 2^3 + 5$
5	$2^2 = 4 <= 5 < 2^3 = 8$	$5 = 2^2 + 1$
1	$2^0 = 1 <= 1 < 2^1 = 2$	$1 = 2^0 + 0$
0		

$(77)_{10} = 2^6 + 2^3 + 2^2 + 2^0 = (1001101)_2$

Les puissances de 2 donnent les positions des 1 dans le nombre en base 2 : 6, 3, 2, 0. Les autres positions valent 0.

Exemple 2 : $(1986)_{10}$ en base 16

nBS	Encadrement	Division euclidienne
1986	$16^2 = 2^{4*2} = 256 <= 1986 < 16^3 = 2^{4*3} = 4096$	$1986 = 7 \times 16^2 + 194$
194	$16^1 = 16 <= 194 < 16^2 = 2^{4*2} = 256$	$194 = 12 \times 16^1 + 2$
2	$16^0 = 1 <= 2 < 16^1 = 16$	$2 = 2 \times 16^0 + 0$
0		

$(1986)_{10} = 7*16^2 + 12*16^1 * 2*16^0 = (7C2)_{16}$

Méthode par la division euclidienne

- ***Présentation***

  La méthode décrit un algorithme :

  On effectue la division euclidienne du nombre de la base source, **nBS**, par la valeur de la base cible, **BC**.

  On recommence avec le quotient jusqu'à obtenir 0 comme quotient.

  Les restes obtenus, en partant du dernier (celui du quotient à 0) au premier, celui du dividende de départ (le nombre de la base source), forment le nombre dans la base cible, de gauche à droite.

- ***Exemples***

  $(77)_{10}$ en base 2

Dividende : **nBS**	Diviseur (**BC**)	Quotient	Reste : Chiffres du **nBC**	Position
77	2	38	1	0
38	2	19	0	1
19	2	9	1	2
9	2	4	1	3
4	2	2	0	4
2	2	1	0	5
1	2	0	1	6

$(75)_{10} = (1001101)_2$

$(1986)_{10}$ en base 16

Dividende : **nBS**	Diviseur (**BC**)	Quotient	Reste : chiffres du **nBC**	Position
1986	16	124	2	0
124	16	7	12 = C	1
7	16	0	7	2

$(1986)_{10} = (7C2)_{16}$

Exercices – série 4 – Conversion d'une base à une autre

Conversions

Donner la valeur de $(387)_{10}$ et de $(238)_{10}$ en base 2. Utilisez les 2 méthodes.

Donner la valeur de $(1002)_{10}$ en base 8. Utilisez les 2 méthodes.

Donner la valeur de $(2387)_{10}$ et de $(8251)_{10}$ en base 16. Utilisez les 2 méthodes.

Programmation excel

- ***Première partie***

1. Codez sous excel la méthode par la division euclidienne pour des entiers en base 10. Le principe de résolution consiste à avoir un tableau avec dividende, diviseur, quotient et reste (le diviseur correspond à la base cible).

Exemple de résultat : on part de $43_{(10)}$

On arrive à $101011_{(2)}$

Dividende : nBS base 10	Diviseur (BC)	Quotient	Reste : Chiffres du nBC
43	2	21	1
21	2	10	1
10	2	5	0
5	2	2	1
2	2	1	0
1	2	0	1

Sous Excel, pour faire la division entière, on utilise la fonction QUOTIENT.

Pour le modulo, on utilise la fonction MOD.

- ***Deuxième partie***

Dans un second temps, on peut se doter d'une colonne avec le reste sous forme de caractère (pour gérer les chiffres de la base 16) et d'une colonne permettant de construire au fur et à mesure le résultat.

Exemple en base 2 :

Dividende : nBS base 10	Diviseur (BC)	Quotient	Reste : Chiffres du nBC	Position	Reste : en texte	Résultat en texte
129	2	64	1	0	1	1
64	2	32	0	1	0	01
32	2	16	0	2	0	001
16	2	8	0	3	0	0001
8	2	4	0	4	0	00001
4	2	2	0	5	0	000001
2	2	1	0	6	0	0000001
1	2	0	1	7	1	10000001

Exemple en base 16 :

Dividende : nBS base 10	Diviseur (BC)	Quotient	Reste : Chiffres du nBC	Position	Reste : en texte	Résultat en texte
6 750	16	421	14	0	E	E
421	16	26	5	1	5	5E
26	16	1	10	2	A	A5E
1	16	0	1	3	1	1A5E

Pour le reste en texte, on utilise la fonction CTXT qui transforme un entier en texte.
Pour obtenir A, B, C, etc. il faut faire des tests avec un SI.
Pour obtenir le résultat en texte on utilise la fonction CONCATENER.

3 - Opérations élémentaires en base 2, 8 et 16

Présentation

Addition, soustraction, multiplication, division des entiers et des réels : **les opérations se font en les posant de la même manière quelle que soit la base.**

Mais il faut connaître les tables d'addition et de multiplication dans la base !

Ou alors convertir en base 10, faire l'opération en base 10, puis reconvertir dans la base d'origine.

En base 2, c'est facile !

En base 8 ou 16, c'est plus compliqué !

Extraits de tables d'addition et de multiplication en base 8 et 16 :

$(7)_8 + (7)_8 = (16)_8$ $\quad$ $(7)_8 * (7)_8 = (61)_8$

$(F)_{16} + (F)_{16} = (1E)_{16}$ $\quad$ $(F)_{16} * (F)_{16} = (E1)_{16}$

Exemples d'opérations en base 2

1+0=1	1+1=10	10+1 = 11	11+1=100	100+1=101	101+1=110	110+1=111
1-0=1	1-1=0	10-1 = 1	11-1=10	100-1=11	101-1=100	110-1=101

$^{1}\,^{1}1\ 1$	1 plus 1 égale 10, je pose 0 et je retiens 1
$+\quad 0\ 1$	0 plus 1 égale 1, plus1 égale 10, je pose 0 et je retiens 1
$=\ 1\ 0\ 0$	1 plus 1 égal 1

$1\ 1^{1}0$	1 ôté de 0 : impossible
$-\quad 1$	1 ôté de 10 égale 1et je retiens 1
$=\ 1\ 0\ 1$	0+1=1 ôté de 1 égale 0
	0 ôté de 1 égale 1

$1\ 1$	$1\ 0\ 1\ 0 \mid 11$	En 101 combien de fois 11 : 1 fois
$\times\ 1\ 1$	$1\ 0\ 0 \mid 11$	1 fois 11, 11, oté de 101 il reste 10
$1\ 1$	$1 \mid$	En 100 combien de fois 11 : 1 fois
$1\ 1\ .$		1 fois 11, 11, oté de 100 il reste 1
$=1\ 0\ 0\ 1$		

Exercices – série 5 – Opérations élémentaires en bases 2, 8, 16 et numération babylonienne

Opérations en base 2, 8 et 16

1. Soit 2 entiers en base 2 : a=10110 et b=1101. Calculer a+b, a-b, a*b
2. Soit 2 réels en base 2 : a=110,01 + b=100,1. Calculer a+b, a-b, a*b
3. Soit 2 entiers en base 16 : a=D1 et b=19. Calculer a+b, a-b, a*b

Numération babylonienne (environ 2000 avant JC) : base 60

Les clous représentent les unités : 1

Les chevrons représentent les dizaines : <

On peut juxtaposer jusqu'à 9 clous.

On peut juxtaposer jusqu'à 5 chevrons : le système est à base 60.

Quand on arrive à 6 chevrons, on les remplace par un clou dans la colonne voisine à gauche : c'est un système en partie positionnel.

43 : <<<<111

73 : 1 <111 (la « dizaine » à 60 est décalé des « unités » : 60 + 10 + 3)

3673 : 1 1 <111 (60*60 = 3600 + 60 + 10 + 3)

7273 : 11 1 <111 (2*60*60 = 7200 + 60 + 10 + 3)

Ce système permet de faire des opérations :

			1	1	
		3	6	9	8
+		7	2	8	3
=	1	0	9	8	1

		1	<
	1	1	<<< 11111111
+	11	1	<< 111
=	111	111	1

Défaut majeur : il n'y a pas de chiffres, d'où la complexité des calculs.

La base 60 est trop grande pour avoir des tables faciles à apprendre.

Exercices

Ecrire 4215 et 382 en système babylonien.

Additionner 4215 + 382 en système babylonien

4 - Codage de l'information dans les ordinateurs

Notion de BIT

Les informations traitées par les ordinateurs sont de différentes natures : nombres, texte, images, sons, vidéos, programmes, etc.

Dans un ordinateur, elles sont toujours représentées sous **forme binaire : une suite de 0 et de 1**.

On appelle chaque information élémentaire (un 0 ou un 1) **BIT** (Binary digIT = chiffre binaire).

Représentation externe et représentation binaire

Le codage permet de passer d'une **représentation dite externe** (la représentation compréhensible directement par un humain : nombres, texte, images, sons, vidéos) à **représentation interne** sous forme de BIT. On parle alors de représentation binaire.

Représentation binaire et représentation physique

La représentation binaire est une représentation logique : elle tient encore du discours.

Cette représentation logique se traduit physiquement par des **états de la matière**.

Il existe plusieurs solutions de **représentation physique** de la représentation binaire :

Pour le stockage

La charge électrique (RAM : Condensateur-transistor, sans électricité on perd la mémoire !) : chargé (bit 1) ou non chargé (bit 0)

La magnétisation (Disque dur, disquette) : polarisation Nord (bit 1) ou Sud (bit 0)

Les alvéoles (CDROM) : réflexion (bit 1) ou pas de réflexion (bit 0)

Pour la circulation

Les câbles électriques : niveau haut ou bas de tension, présence ou absence de lumière.

Le débit d'information est donc le nombre de 0 et de 1 que l'on arrive à caser par seconde.

La numérisation

C'est la **conversion** des informations d'un support (texte, image, audio, vidéo) ou d'un signal électrique **en données numériques** que des dispositifs informatiques ou d'électronique numérique pourront traiter.

Notion d'octet

Présentation

Dans les ordinateurs, on travaille rarement directement au niveau du BIT. On travaille en général au niveau de l'octet. **Un octet c'est 8 BIT à la suite.**

Un octet permet donc de coder **256 possibilités** : 2^8

L'octet servira à coder les caractères alphanumériques et spéciaux (a, b, 1, 2, ;, -, +, etc.)

Notion de bit de poids fort ou faible

Dans un octet, le premier bit, dans une lecture de gauche à droite, qui représente 2^7, est appelé **bit de poids fort.** Le dernier bit, qui représente 2^0, est appelé **bit de poids faible.**

Cela vient de l'écriture des entiers en somme des puissances de la base :

$(a_n, .., a_1, a_0) = a_n*10^n + .. + a_1*10^1 + a_0*10^0 = \Sigma_{i=0}^{n} a_i*10^i$.

Dans ce cas, a_0 est le **chiffre de poids faible**. a_n est le c**hiffre de poids fort.**

ko, Mo, Go, To, etc.

un kilooctet = 1ko = 1000 octets

un megaoctet = 1Mo = 10^6 octets = 1 mega

un gigaoctet = 1Go = 10^9 octets = 1 giga

un teraoctet = 1To = 10^{12} octets = 1 tera = 1000 giga

un petaoctet = 1Po = 10^{15} octets

Attention : en 1998, l'IEC, organisme de standardisation, a fixé que un ko = 1000 octets et pas 1024 ou 2^{10} octets ! idem pour les suivants : un Mo = 10^6 octets et pas 2^{20}, un Go = 10^9 octets et pas 2^{30} octets.

Représentation d'un octet en hexadécimal

Un octet est souvent représenté sous la forme de 2 blocs de 4 bits, soit 2 blocs de 2^4=16 possibilités.

Le contenu d'un octet est donc représenté par deux entiers en base 16 : 4F, E6 par exemple.

4F, c'est 0100 et 1111 soit 1+2+4+8+64 = 79

E6, c'est 1110 et 0110 soit 2+4+32+64+128 = 230

Convertisseur en ligne : ici ou là

Exercice

1. Sachant qu'un livre de 200 pages contient environ 300 000 caractères espaces compris et qu'un caractère est codé sur un octet ; sachant que le traitement de texte triple en moyenne la taille du texte (1000 caractères deviennent 3000) ; calculer le nombre de livres que vous pourrez enregistrer sur une clé de 2 GO.

5 - Codage des caractères – ASCII – ISO 8859 – UNICODE – UTF8

Présentation

Il s'agit de coder les caractères alphanumériques : les chiffres (0,1,...,9), les lettres minuscules et majuscules (a,...z, A,...,Z), les signes de ponctuation (. , ; / = * - …), des symboles usuels (%, @, $, &, etc.), des caractères spéciaux (passage à la ligne, delete, bip sonore, etc.)

Cet ensemble de caractères contient : 10 chiffres, 52 lettres, une 20aine de signe de ponctuation ou de symboles usuels, ce qui fait déjà plus de 80 caractères.

Avec 6 bits, on peut coder $2^6 = 64$ caractères : c'est insuffisant. Avec 7 bits, on peut coder $2^7 = 128$ caractères. C'est suffisant.

Il existe finalement trois codes : le code ASCII, ASCII étendu ou ISO et l'UNICODE.

- Le code **ASCII** codifie 128 caractères sur 7 bits.
- Les codes **ASCII étendus ou codes ISO 8859** permettent de codifier 256 caractères sur 8 bits. Il en existe de nombreux, ce qui est source de problèmes avec les accents particulièrement.
- L'**UNICODE** permet de normaliser tous les caractères possibles, actuellement 128 172. Il est indépendant de la représentation physique (le nombre d'octets).

Le code ASCII

Présentation

Le code ASCII est un code standard qui définit la valeur du jeu courant de caractères sur 7 bits.

Sur un octet, il restera le bit de poids fort.

Les caractères sont donc codés de 0 à 127 :

- 0 à 31 : des caractères spéciaux
- 32 à 47 : des symboles et de la ponctuation : (espace, ! , « , #, etc.)
- 48 à 57 : les chiffres dans l'ordre : (0, … , 9)
- 58 à 64 : des symboles et de la ponctuation : (: , ; , < , etc.)
- 65 à 90 : les alphabets majuscules (A,...,Z)
- 97 à 122 : les alphabets minuscule (a,...z)
- 123 à 127 : des symboles et de la ponctuation : ([, \ , etc.)

Avantages

Il permet aux machines de communiquer entre elles, quels que soient les processeurs et quels que soit les système d'exploitation.

Défaut

American Standard Code for Information Interchange" signifie en français "Code américain normalisé pour l'échange d'information".

C'est un code basé sur l'alphabet américain : alphabet latin sans accents.

Table ASCII

Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char
0	00	Null	32	20	Space	64	40	@	96	60	`
1	01	Start of heading	33	21	!	65	41	A	97	61	a
2	02	Start of text	34	22	"	66	42	B	98	62	b
3	03	End of text	35	23	#	67	43	C	99	63	c
4	04	End of transmit	36	24	$	68	44	D	100	64	d
5	05	Enquiry	37	25	%	69	45	E	101	65	e
6	06	Acknowledge	38	26	&	70	46	F	102	66	f
7	07	Audible bell	39	27	'	71	47	G	103	67	g
8	08	Backspace	40	28	(	72	48	H	104	68	h
9	09	Horizontal tab	41	29	)	73	49	I	105	69	i
10	0A	Line feed	42	2A	*	74	4A	J	106	6A	j
11	0B	Vertical tab	43	2B	+	75	4B	K	107	6B	k
12	0C	Form feed	44	2C	,	76	4C	L	108	6C	l
13	0D	Carriage return	45	2D	-	77	4D	M	109	6D	m
14	0E	Shift out	46	2E	.	78	4E	N	110	6E	n
15	0F	Shift in	47	2F	/	79	4F	O	111	6F	o
16	10	Data link escape	48	30	0	80	50	P	112	70	p
17	11	Device control 1	49	31	1	81	51	Q	113	71	q
18	12	Device control 2	50	32	2	82	52	R	114	72	r
19	13	Device control 3	51	33	3	83	53	S	115	73	s
20	14	Device control 4	52	34	4	84	54	T	116	74	t
21	15	Neg. acknowledge	53	35	5	85	55	U	117	75	u
22	16	Synchronous idle	54	36	6	86	56	V	118	76	v
23	17	End trans. block	55	37	7	87	57	W	119	77	w
24	18	Cancel	56	38	8	88	58	X	120	78	x
25	19	End of medium	57	39	9	89	59	Y	121	79	y
26	1A	Substitution	58	3A	:	90	5A	Z	122	7A	z
27	1B	Escape	59	3B	;	91	5B	[	123	7B	{
28	1C	File separator	60	3C	<	92	5C	\	124	7C	|
29	1D	Group separator	61	3D	=	93	5D	]	125	7D	}
30	1E	Record separator	62	3E	>	94	5E	^	126	7E	~
31	1F	Unit separator	63	3F	?	95	5F	_	127	7F	□

Les codes ASCII étendus : codes ISO 8859

Présentation

L'objectif est d'élargir le code ASCII pour permettre par exemple l'introduction d'accents.

Le 8ème bit de l'octet du caractère va permettre cela. C'est ce qu'on appelle le code ASCII étendu. C'est le code ISO 8859 qui normalise ça.

ISO signifie International Organization for Standardization (le sigle ne correspond pas !)

Comme il y a beaucoup de système d'accentuation dans les langues latines, et qu'il y a aussi des langues non latines, le code ISO 8859 se décline en :

ISO 8859-1 (ou latin-1 ou européen occidental) : pour la plupart des langues européennes.

ISO 8859-2 (ou latin-2 ou européen central) : pour le le croate, le tchec, le hongrois, etc.

…

ISO 8859-6 : pour l'arabe

…

ISO 8859-15 : révision du 8859-1

…

https://fr.wikipedia.org/wiki/ISO/CEI_8859

Avantage

Intègre le standard ASCII.

Permet de standardiser les accents.

Défaut

Il y a plusieurs standards ! Et des révisions dans les standards.

L'UNICODE et UTF8

UNICODE

http://www.tuteurs.ens.fr/faq/utf8.html#unicode

Le Standard Unicode est produit par une organisation à but non lucratif (le Consortium Unicode) ayant pour objectif d'attribuer un numéro à tout caractère utilisé dans une langue humaine.

Un caractère Unicode est un caractère défini dans le Standard Unicode. On y fait référence par son numéro écrit en hexadécimal précédé de «U+». Par exemple, la lettre latine « a » correspond à U+0041.

Unicode couvre actuellement 128 172 caractères.

http://www.unicode.org/charts/

UTF 8

L'UTF8 est le système d'encodage qui transforme une code Unicode en octets : de 1 à 4 octets.

L'UTF8 est compatible avec les caractères ASCII. Particulièrement, les caractères de 32 à 127 seront codés sur 1 octet.

Avantage

Un seul encodage pour toutes les langues : c'est LE système universel.

Inconvénients

La taille des caractères n'est pas fixe, ce qui complique certains calculs.

Usages

Généralisé ! Dans les sites web, sur les terminaux d'ordinateur, etc.

- ***Aparté technique :***

L'encodage des terminaux windows fonctionne avec des « page de code ». La page de code n°850 est celle d'Europe occidentale. On peut voir le n° avec la commande chcp.

https://fr.wikipedia.org/wiki/Page_de_code

Sous linux, la commande xxd permet de voir l'encodage d'un texte : xxd-entrée-texte-entrée-ctrl-d. Sont affichés les caractères du texte en hexadécimal avec un 0a à la fin pour le line feed. Si les caractères accentués sont sur 2 octets, c'est que l'encodage est en UTF8.

Exercices – série 6 – ASCII et ASCII étendu

1. Que vaut en binaire ASCII « m », « R », « 9 ». Quelle est la valeur correspondante en Hexa ?
2. Soit les 8 bits suivants : 0101 1100

Combien valent-ils en ASCII. Donnez aussi la valeur en Hexa.

3. Soit les 16 bits suivants : 1110 1001 0111 0000

Combien valent-ils en ASCII étendu Latin 1 (cf. rappels de cours plus bas) ?

Le code ASCII des minuscules se situe entre $(97 \text{ et } 122)_{10}$ soit $(61\text{-}7A)_{16}$ et entre $(65 \text{ et } 90)_{10}$ soit $(41\text{-}5A)_{16}$) pour les majuscules. On précise qu'on a les lettres suivantes à partir de 230 (code ascii étendu, Latin 1, ISO 8859-1) : 'æ', 'ç', 'è', 'é', 'ê'

6 - Numération des réels

Propriété : écriture des réels en somme des puissances de la base

Un réel de n'importe quelle base peut s'écrire, en base 10, sous la forme d'une addition de puissances de la base :

$(a_n .. a_0 , b_1 .. b_m)_p = a_n*p^n + .. + a_1*p^1 + a_0*p^0 + b_1/p^1 + .. + b_m/p^m = \Sigma_{i=0}^{n} a_i*p^i + \Sigma_{i=1}^{m} b_i/p^i$

Conversions vers la base 10 : facile

$(4,25)_{10} = 4*10^0 + 2/10^1 * 5/10^2 = 4 + 0,2 + 0,05$

$(1.011)_2 = 1*2^0 + 0/2^1 + 1/2^2 + 1/2^3 = 1 + 0 + 0,25 + 0,125 = 1,375$

$(1.A2)_{16} = 1*16^0 + 10/16^1 + 2*16^2 = 1+ 5/8 + 2/256$

Exercices

Donner la valeur de $(10,101)_2$ en base 10.

Donner la valeur de $(1A,B1F)_{16}$ en base 10.

Conversions de la base 10 vers la base 2

Objectif

Traduire $(23,375)_{10}$ en binaire.

Méthode

La partie entière se gère comme un entier.

Pour la partie décimale : on multiplie la partie décimale par 2 (la base cible). Et on répète l'opération avec la partie décimale du résultat de la multiplication.

Quand on arrive à 1, on s'arrête. Si on retombe sur une opération déjà faite, on s'arrête.

La suite des chiffres des unités des résultats donne la partie décimale en binaire.

avec une série de chiffres répétés à l'infini si on est retombé sur une opération déjà faite pour finir.

Exemple 1

$(0,375)_{10} * 2 = 0,750$

$0,750 * 2 = 1,5$

$0,5 * 2 = 1,0$

$(0,375)_{10} = (0,011)_2$

Vérification :

$(0,011)_2 = 1/2^2 + 1/2^3 = 1/4 + 1/8 = 0,25 + 0,125 = 0,375$

Exemple 2

$(0,3)_{10} * 2 = 0,6$

$0,6 * 2 = 1,2$

$0,2 * 2 = 0,4$

$0,4 * 2 = 0,8$

$0,8 * 2 = 1,6$

$0,6 * 2 = 1,2$

etc.

$(0,3)_{10} = (0,0[1001])_2$

- ***Vérification sur 6 chiffres dont 5 après la virgule***

$(0,3)_{10} = (0,01001)_2$

$= 1/2^2 + 1/2^5$

$= 1/4 + 1/32$

$= 0,25 + 0,03125$

$= 0,28125$

- ***Vérification sur 10 chiffres dont 9 après la virgule***

$(0,3)_{10} = (0,010011001)_2$

$= 1/2^2 + 1/2^5 + 1/2^6 + 1/2^9$

$= 1/4 + 1/32 + 1/64 + 1/512$

$= 0,25 + 0,03125 + 0,015625 + 0,001953$

$= 0,298828$

Conclusion

On voit ici que le résultat change selon le nombre de chiffres qu'on peut mettre dans le nombre.

Avec 5 chiffres après la virgule en base 2, on obtient 0, 28125.

Avec 9 chiffres, on obtient 0,298828.

En augmentant le nombre de chiffres en base 2, on se rapprocherait de 0,3 !

Dans les ordinateurs, ce nombre de chiffre sera limité : 23 bits pour un float, 52 bits pour un double quand même ! On obtiendra donc 0,3.

Toutefois, on constate que les calculs peuvent être approximatifs.

7 - Arrondi et précision

Définition

On peut **arrondir** les réels à une certaine **précision.**

La précision « p » appartient à l'ensemble {… 100 ; 10 ; 1 ; 0,1 ; 0,01 ; .. }, cela quelle que soit la base. « p » c'est 1, ou 10, ou 0,1, etc.

Chercher l'arrondi d'un réel X à une certaine précision « p », c'est :

Chercher **les 2 multiples de la précision les plus proches de X** tels que X soit compris entre ces deux multiples, autrement dit, trouver l'entier « n » tel que $n*p \leq X < (n+1)*p$

L'arrondi est alors le multiple le plus proche de X.

S'il existe 2 nombres possibles, on choisit **le plus grand par convention.**

Méthode pour trouver A, l'arrondi

BI (borne supérieure) = X avec tous les chiffres de poids plus petit que celui du 1 de p ramenés à 0 (BI = 10 avec X=12 et p = 10)

BS (borne supérieure) = BI +p et $BI \leq X < BS$

DP (demi-précision) = p/2

Si BI+DP > X alors A = BI, sinon A = BS (si 12 < 10 + 5 alors A = 10 sinon (X=17) A=20)

Exemples en base 10

- ***Arrondi de $(432,614)_{10}$ à une précision de 10.***

  BI = 430 BS=440 DP = 5

  BI+DP= 435 > 432,614 donc BI est l'arrondi : **430**

- ***Arrondi de $(432,614)_{10}$ à une précision de 1***

  BI=432 BS 433 DP = 0,5

  BI+DP= 432,5 <= 432,614 donc BS est l'arrondi : **433**

- ***Arrondi de $(432,614)_{10}$ à une précision de 0,1.***

  BI=432,6 BS 432,7 DP = 0,05

  BI+DP= 432,65 > 432,614 donc BI est l'arrondi : **432,6**

Exemples en base 2

Rappel des correspondances entre la base 2 et la base 10

Base 2	0	1	10	11	100	101	110	111	1000	1001	1010	1011	Etc.
Base 10	0	1	2	3	4	5	6	7	8	9	10	11	Etc.

- ***Arrondi de X = $(101,10011)_2$ à une précision de $(10)_2$***

  BI : 100 BS : 100+10=110 DP = 10/10 = 1 (on divise par 2, mais en base 2, donc par 10)

  BI + DP = 100 + 1 = 101 <= X, donc l'arrondi c'est BS : **110**

- ***Arrondi de X = $(101,10011)_2$ à une précision de $(1)_2$***

  BI : 101 BS : 101+1=110 DP = 1/10 = 0,1 (on divise par 2, mais en base 2, donc par 10)

  BI + DP = 101 + 0,1 = 101,1 <= X, donc l'arrondi c'est BS : **110**

- ***Arrondi de X = $(101,10011)_2$ à une précision de $(0,1)_2$.***

  BI : 101,1 BS : 101,1+0,1=110 DP = 0,1/10 = 0,01

  BI + DP = 101,1 + 0,01 = 101,11 > X, donc l'arrondi c'est BI : **101,1**

Exercices

1. Quel est l'arrondi de $1101011_{(2)}$ à $100_{(2)}$ près ?
2. Quel est l'arrondi de $10,011_{(2)}$ à $0,1_{(2)}$ près ?

8 - Codage des entiers dans les ordinateurs

Présentation

Les entiers se codent sur un ou plusieurs octets.

Si on code des entiers négatifs (entiers relatifs, Z), on utilise le bit de poids fort pour le signe.

Il existe 3 solutions de codages des entiers relatifs :

- le code binaire signé,
- le complément à 1,
- le complément à 2.

Code binaire signé : (bs)

Principes sur un octet

- Le bit de poids fort sert pour le signe (0 pour +, 1 pour -).
- Les 7 autres bits servent pour la valeur absolue du nombre.

Résultats

- Sur un octet, on peut coder de -127 $(1111\ 1111)_{bs}$ à +127 $(0111\ 1111)_{bs}$
  Autrement dit de -2^7-1 à $+2^7-1$
- Avec ce système, on a 2 codages pour 0 : $(0000\ 0000)_{bs}$ et $(1000\ 0000)_{bs}$

Conséquences

- Les opérations sont un peu compliquées :
  Addition et soustraction diffèrent selon le signe : additionner un nombre négatif à un nombre positif revient à le soustraire.

Généralisation à n octets

Même principe : on peut coder de $-2^{8*n-1}-1$ à $+2^{8*n-1}-1$

- ***Soit sur 2 octets***

de $-2^{8*2-1}-1$ à $+2^{8*2-1}-1$ soit de $-2^{15}-1$ à $+2^{15}-1$ de -32767 à +32767

En code binaire signé, de $(1111\ 1111\ 1111\ 1111)_{bs}$ à de $(0111\ 1111\ 1111\ 1111)_{bs}$

Complément à 1 : (ca1)

Exemple sur un octet

1. Pour les nombres positifs, on utilise le code binaire signé

   Donc $(0xxx\ xxxx)_{bs} = (0xxx\ xxxx)_{ca1}$ avec $x \in \{0, 1\}$

   - ***Par exemple***

     $(1)_{10} = (0000\ 0001)_{bs} = (0000\ 0001)_{ca1}$

2. Pour les nombres négatifs, de -1 à -127, on inverse les tous les bits du nombre positif correspondant.

   - ***Par exemple :***

     $(1)_{10} = (0000\ 0001)_{ca1}$. Donc $(-1)_{10} = (1111\ 1110)_{ca1}$

     $(127)_{10} = (0111\ 1111)_{ca1}$. Donc $(-127)_{10}\ (1000\ 0000)_{ca1}$

3. Avec ce système, on a 2 codages pour 0 : $(0000\ 0000)_{ca1}$ et $(1111\ 1111)_{ca1}$

Avantage : les opérations sont simplifiées

- ***Les additions sont simplifiées***

  On peut additionner tous les bits sans se soucier du signe.

- ***Par exemple :***

  $(1)_{10}+(-1)_{10} = (0000\ 0001)_{ca1}+(1111\ 1110)_{ca1} = (1111\ 1111)_{ca1} = (0)_{10}$

  $(69)_{10} = (0100\ 0101)_{ca1}$ Donc $(-69)_{10} = (1011\ 1010)_{ca1}$

  $(33)_{10} + (-69)_{10} = (0010\ 0001)_{ca1}+(1011\ 1010)_{ca1} = (1101\ 1011)_{ca1} = (0010\ 0100)_{bs} = (32+4)_{10} = (36)_{10}$

- ***La soustraction c'est l'addition de l'opposé.***

  L'opposé d'un nombre, c'est l'inversion de tous les bits.

  La soustraction c'est l'addition de l'opposé : $5-3 \Leftrightarrow 5+(-3)$

Complément à 2 : c'est la représentation dans les ordinateurs

Exemple sur un octet

- ***Pour les nombres positifs, on utilise le complément à 1 (donc le code binaire signé) :***

  Donc $(0xxx\ xxxx)_{bs} = (0xxx\ xxxx)_{ca1} = (0xxx\ xxxx)_{ca2}$ avec $x \in \{0, 1\}$

- ***Pour les nombres négatifs, on ajoute 1 au complément à 1 :***

  $(1)_{10} = (0000\ 0001)_{ca1}$ donc $(-1)_{10} = (1111\ 1110)_{ca1} = (1111\ 1111)_{ca2}$: 1111 1110 + 1

  $(127)_{10} = (0111\ 1111)_{ca1}$ donc $(-127)_{10} = (1000\ 0000)_{ca1} = (1000\ 0001)_{ca2}$: 1000 0000+ 1

- ***Avec ce système, on a 1 seul codage pour 0 : $(0000\ 0000)_{ca2}$***

  La deuxième valeur pour 0 en complément à 1 : (1111 1111) vaut -1 en complément à 2.

- ***On peut coder de -127 à +127 mais il nous reste une valeur de codage. Est-ce -128 ou +128 ?***

  On peut constater que :

  $(128)_{10} = (127)_{10} + (1)_{10} = (0111\ 1111)_{ca2} + (0000\ 0001)_{ca2} = (1000\ 0000)_{ca2}$

  $(-128)_{10} = (-127)_{10} + (-1)_{10} = (1000\ 0001)_{ca2} + (1111\ 1111)_{ca2} = (1000\ 0000)_{ca2}$

  Les deux opérations fonctionnent.

  Mais (1000 0000) est a son bit de poids fort qui vaut 0 : c'est donc un nombre négatif : -128.

  Donc on va donc coder en complément à 2 de -128 à +127.

Avantage : les opérations sont simplifiées et il n'y a qu'un seul 0

Addition et soustraction sont simplifiées : elles sont indépendantes du signe.

Par exemple : 127 – 127 c'est $(0111\ 1111)_2 + (1000\ 0001)_{ca2} = (0000\ 0000)_{ca2}$

Le complément à 2 pour les entiers relatifs, c'est le mode de représentation utilisé dans les ordinateurs et les langages informatiques.

Conclusion : représentation logique, représentation physique et optimisation

La représentation externe est celle en base 10.

La représentation logique est celle en base 2 : c'est la représentation binaire signée avec un bit de signe (le bit de poids fort).

La représentation en complément à 2 est la représentation finale utilisée par les ordinateurs. Cette représentation c'est celle qui est adaptée aux contraintes physique des ordinateurs pour optimiser les traitements.

C'est la plus éloignée de la représentation externe et la plus proche de la représentation physique matérielle finale.

On parle de représentation physique pour cette représentation en complément à 2.

On a donc trois niveaux de représentation :

- Externe : base 10, image, son, etc.
- Logique : binaire
- Physique : ca2

  Vrai physique : magnétique, électrique, etc.

9 - Codage des nombres réels dans les ordinateurs

Présentation

Les réels se codent sur au moins 2 octets.

Il existe 2 solutions de codages des réels :

- le code à virgule fixe
- le code à virgule flottante.

Code à virgule fixe

Principes

Le code à virgule fixe traduit les nombre écrit sous la forme : $\pm(a_n .. a_0 , b_1 .. b_m)_p$

- 1 bit de signe, n bits de partie entière, m bits de partie fractionnaire.
- Le bit de signe et les n bits de partie entières sont codés comme des entiers relatifs en complément à 2.
- Les m bits de partie fractionnaire sont codés comme des entiers positifs en simple code binaire.

Conséquences

La précision dépend de m et donc du nombre de bit dédié à la partie fractionnaire.

Usage

A l'usage, **il faut connaître à l'avance le niveau de précision** dont on aura besoin dans les programmes pour positionner au mieux la virgule. Comme en général, on ne peut pas déterminer le niveau de précision nécessaire, on préfère utiliser le code à virgule flottante.

Cette connaissance a priori existe pour certaines applications, mais dans le cas général ce n'est pas le cas. Pour pallier à ce manque de flexibilité, le concept de virgule flottante a été introduit.

Code à virgule flottante

Le code à virgule flottante traduit les nombre écrit sous la forme : $\pm\ m\ .\ b^{e}_{(b)}$

- $\pm$ est le signe
- m est **la mantisse**. C'est un réel positif. $m \in [1 ; 10 [$. m est codé en virgule fixe. 10 est la valeur maximum, quelle que soit la base. La mantisse n'a qu'un chiffre avant la virgule.
- e est **l'exposant**. C'est un entier relatif codé en complément à 2.
- b est une **base** b : 2, 8, 10, 16,… Dans les ordinateurs, c'est 2.

Exemples

Base 10 : $5{,}425 \times 10^{2}$. Le nombre vaut : 542, 25 : en virgule fixe on a 2 chiffres après la virgule.

Base 2 : $1{,}01 \times 2^{-1}$. Le nombre vaut 0,101 : en virgule fixe on a 3 chiffres après la virgule.

Base 16 : $A{,}0B4 \times 16^{-2}$. Le nombre vaut 0,0A0B4 : en virgule fixe on a 5 chiffres après la virgule.

En virgule fixe : $0{,}25_{(10)} = 0{,}01_{(2)}$

En virgule flottante : $2{,}5 \times 10^{-1} = 1{,}0 \times 2^{-2}$

Représentation dans les ordinateurs

Répartition des bits

Signe	Exposant biaisé	Mantisse normalisée
1 bit	p bits	q bits

Notion de mantisse normalisé : M

La mantisse est un réel positif. En base 2, m ∈ [1 ; 2 [

Donc, la mantisse est de la forme : 1, …

Donc, la forme du code à virgule flottante devient : $\pm 1{,}M \cdot 2^E$

Le premier 1 devient une valeur par défaut et M est la mantisse normalisée.

M est donc un entier positif : son codage est donc simplifié. **C'est un simple codage binaire**. On n'a plus besoin de binaire signé ni ne complément à 1 ou à 2.

Notion d'exposant biaisé : E

- ***Présentation***

Dans la forme : $\pm\ m \cdot b^e_{(b)}$, « e » est l'exposant (petit e). C'est un entier relatif codé en binaire signé : il sera donc compris entre 2 bornes [-biais; +biais]

Dans la forme : $\pm 1{,}M \cdot 2^E$, « E » (grand E) est l'exposant biaisé.

On définit E ainsi : $E \geq 0$ et e + biais = E

- ***L'exposant « e »***

Si « e » est codé en binaire signé sur sur **p bits**

Alors $e \in [-2^{p-1}-1 ; 2^{p-1}-1]$.

Par exemple :

si « e » est codé en binaire signé sur sur 8 bits (cas des float)

Alors $e \in [-2^7-1 ; 2^7-1]$ soit [-127 ; 127]

- ***Le biais***

Biais = $2^{p-1}-1$

C'est un nombre fixé par le nombre de bits du codage de l'exposant.

- ***De l'exposant « e » à l'exposant biaisé « E »***

Pour passer de « e » à « E », on ajoute le biais « e ».

Biais = $2^{p-1}-1$

$E \in [0; +2^p-2]$.

Par exemple :

si « e » est codé en binaire signé sur sur 8 bits (cas des float)

Alors e ∈ [-2^7-1 ; 2^7-1] soit [-127 ; 127]

Le biais vaut 127

E ∈ [0 ; +254]

Cas particuliers

- ***infini et NaN***

  Si E ∈ [0 ; +254] sur 8 bits, il reste une valeur possible : 255 : que des 1 sur tous les octets de E de façon générale.

  Ce cas permet de coder deux cas particuliers :

  - E=255 et mantisse = 0 : pour coder l'infini
  - E=255 et mantisse != 0 : pour coder NaN (not a number, c'est-à-dire les erreurs)

- ***Coder 0***

  Avec une mantisse normalisée, on a un 1 par défaut. On ne peut donc pas coder le 0 avec la mantisse.

  On va donc utiliser une valeur spéciale de E pour coder le 0 : ce sera 0 (-127 pour e)

- ***Tous les cas***

  E ∈ [1 ; +254] : permet de coder les nombres (donc de -126 à +127)

  E = 255 : permet de coder l'infini (mantisse=0) et NaN (mantisse !=0)

  E = 0 : permet de coder le 0

Exemple

Que vaut en base 10 le codage suivant : 1 0 1 0 0 1 0 1 avec 3 bits d'exposant biaisé ?

- ***signe***

  1 : négatif

- ***exposant***

  nombre de bits = p = 3

  E = $(010)_{vf} = (2)_{10}$

  Biais = $2^{p-1}-1 = (3)_{10}$

  e + b = E donc e = $(-1)_{10}$

- ***mantisse***

  Mantisse normalisée = 0101 donc mantisse = $(1,0101)_2$

- ***nombre en base 2***

  Nombre en base 2 = $-1,0101 \times 2^{-1}$

  Sachant que $(a_n .. a_0 , b_1 .. b_m)_p = a_n*p^n + .. + a_1*p^1 + a_0*p^0 + b_1/p^1 + .. + b_m/p^m = \Sigma_{i=0}^{n} a_i*p^i + \Sigma_{i=1}^{m} b_i/p^i$

  $(-1,0101)_2 \times 2^{-1} = -0,10101_{(2)}$

  $= -(1/2^1 + 1/2^3 + 1/2^5)_{10} = -(1/2+1/8+1/32)_{10} = -0,65625_{(10)}$

Standard IEEE 754 (1985) – float et double

Simple précision : sur 32 bits (4 octets)

Signe	Exposant biaisé	Mantisse normalisée
1 bit	8 bits biais=2^7-1=127 e ∈ [-126 ; +127] E ∈ [1 ; +254]	23 bits M ∈ [0 ; 2^{23}] mantisse : 1 bit de plus m ∈ [1 ; 2^{24}]

2^{24} = 16 777 216 : on a donc 8 chiffres significatifs en base 10

2^{-126} = 1.1754944 e^{-38}

2^{+127} = 1.7014118 e^{+38}

Type C ou java : float

Double précision : sur 64 bits (8 octets)

Signe	Exposant biaisé	Mantisse normalisée
1 bit	11 bits (biais=2^{10}-1=1023)	52 bits
1 bit	11 bits biais=2^{10}-1=1023 e ∈ [-1022 ; +1023] E ∈ [1 ; +2046]	52 bits M ∈ [0 ; 2^{52}] mantisse : 1 bit de plus m ∈ [1 ; 2^{53}]

2^{53} = 9.0071993 e^{+15}: 0n a donc 16 chiffres significatifs en base 10

2^{-1022} = 2.225074 e^{-308}

2^{+1023} = 8.988466 e^{+307}

Type C ou java : double

Valeurs particulières

Exposant	**Mantisse**	**Valeur**
0	0	0
Que des 1	0	+ou- infini
Que des 1	Tout sauf 0	NaN : not a number : error !

Voir ici ou là et encore ici

Exemple

Ecrire $35,5_{(10)}$ en float.

$35_{(10)} = 32+2+1 = 2^5+2^1+2^0 = 10\ 0011_{(2)}$

$0,5_{(10)} = 0,1_{(2)}$ (Méthode : 0,5 * 2 = 1)

$35,5_{(10)} = 10\ 0011,1_{(2)} = 1,000\ 111*2^5{}_{(2)}$

Mantisse = 1,000 111. Mantisse normalisée = 000111 puis des 0 pour remplir les 23 bits.

Exposant = $5_{(10)}$. Biais = 127. Exposant biaisé : 5 + 127 = 132 = 128 + 4 = $(1000\ 0100)_2$

Signe	Exposant biaisé	Mantisse normalisée	Valeurs
1 bit	8 bits (biais=2^7-1=127)	23 bits	
0	5+127=132=128+4=1000 0100 $_{(2)}$	00011100000000000000000	1,000 111*$2^5{}_{(2)}$

$35,5_{(10)}$ s'écrit donc en float sur 4 octets :
0100 0010 – 0000 1110 – 0000 0000 – 0000 0000