

# Cours Algorithmique Numérique

Antoine Lambert

September 24, 2017

## 0.1 Introduction

Les exercices des TP's sont un bon moyen de se préparer à l'examen. (ils sont en annexe dans un autre PDF). L'année passée 50% des questions étaient des questions de programmation

# Chapter 1

## Working with real numbers

### 1.1 Internal representations of numbers

In Java : byte(8 bit), short(16 bit), int(32 bit), long(64 bit).

In c : char(at least 8 bits), short(at least 16 bits),...

#### 1.1.1 Unsigned type in c

$[0, 2^8-1]$

$$222 = 2^7 + 2^6 + 2^4 + 2^3 + 2^2 + 2^1$$

donc on obtiens : 11011110

#### 1.1.2 Two's Complement 8-bits type in java

$[-2^7, 2^7-1]$

$-x+1$  pour les nombres négatifs

$$34 = 2^5 + 2^1 = 00100010$$

$$-34 + 1 = 2^7 + 2^6 + 2^4 + 2^3 + 2^2 + 2^0 + 1 = 11011110$$

#### 1.1.3 Fixed-point Representation

$2^k.x$  (x:un nombre réel, k:indice de précision)

$$x + y = (2^k.x) + (2^k.y)$$

$$x.y = \frac{(2^k.x).(2^k.y)}{2^k}$$

La partie décimale est donc représentée par pas de  $\frac{1}{2^k}$

**+ and -**

+ : facile à implémenter

- : perte d'espace quand on traite avec des grands ou de petits nombres

#### **1.1.4 Floating-point representation**

Séparation du signifiant et de l'exposant :  $3 \cdot 10^9$  ou  $3 \cdot 10^{-9}$

#### **1.1.5 IEEE754 Single Précision(32bits)**

$b_{31}$  : sign  $b_{30}$  à  $b_{23}$  : Exposant  $b_{22}$  à  $b_0$  : Signifiant

le nombre est exprimé suivant la formule :

$$(-1)^{b_{31}} \cdot (1.b_{22} \dots b_0)_2 \cdot 2^{(b_{30} \dots b_{23})_2 - 127}$$

### **1.2 Finite precision**

Quand on fait appel à la précision finie, les résultats peuvent être approximatés et mener à des erreurs. c'est pour cette raison qu'il ne faut jamais comparer 2 floating point numbers ensemble directement. à la place, on utilise un Epsilon pour vérifier que l'un soit égal à l'autre à une approximation epsilon près ( $\text{Math.abs}(a-b) < \text{Epsilon}$ )

#### **1.2.1 Representing Fractions**

On peut représenter la fraction  $\frac{a}{b}$  par 2 entiers et calculer ainsi sans approximations, mais ces fractions doivent être réduites, sinon les nombres a et b deviennent trop grands.

## Chapter 2

# Working with matrices

## Chapter 3

# Solving linear systems

## Chapter 4

# Linear regression

## Chapter 5

# Matrix Norm and Condition



## Chapter 6

# Interpolation

## Chapter 7

# Cubic splines and b-Splines

## Chapter 8

# Numerical integration

## Chapter 9

# Numerical differential equations

## Chapter 10

# Solving systems of differential equations

## Chapter 11

# Probably : intro to 3D graphics

# Contents

0.1	Introduction . . . . .	1
<b>1</b>	<b>Working with real numbers</b>	<b>2</b>
1.1	Internal representations of numbers . . . . .	2
1.1.1	Unsigned type in c . . . . .	2
1.1.2	Two's Complement 8-bits type in java . . . . .	2
1.1.3	Fixed-point Representation . . . . .	2
1.1.4	Floating-point representation . . . . .	3
1.1.5	IEEE754 Single Précision(32bits) . . . . .	3
1.2	Finite precision . . . . .	3
1.2.1	Representing Fractions . . . . .	3
<b>2</b>	<b>Working with matrices</b>	<b>4</b>
<b>3</b>	<b>Solving linerar systems</b>	<b>5</b>
<b>4</b>	<b>Linear regression</b>	<b>6</b>
<b>5</b>	<b>Matrix Norm and Condition</b>	<b>7</b>
<b>6</b>	<b>Interpolation</b>	<b>8</b>
<b>7</b>	<b>Cubic splines and b-Splines</b>	<b>9</b>
<b>8</b>	<b>Numerical integration</b>	<b>10</b>
<b>9</b>	<b>Numerical differential equations</b>	<b>11</b>
<b>10</b>	<b>Solving systems of differential equations</b>	<b>12</b>
<b>11</b>	<b>Probably : intro to 3D graphics</b>	<b>13</b>