# Lecture Notes of Computer Architecture

## Zhou Fan ACM Class, SJTU

## Contents

1.1	=	 			 													
	1.1.1				 													
	1.1.2				 													
	1.1.3				 													
1.2	2	 			 													
	1.2.1				 													
	1.2.2				 													
	1.2.3				 													
	1.2.4				 													
1.3	3	 			 													
2.1		 			 													

1

1.1

### 1.1.1

1.1.2

**1** ( ) <sup>1</sup>

$$F(2, n, L, U) = \{\pm 0.a_1 a_2 \dots a_n \times 10^m\} \cup \{0\}$$

 $a_1 = 1 \ a_i \in \{0, 1\}. \quad m \ L \le m \le U. \quad n$ 

2 (IEEE)

1. : t = 24, L = -126, U = 127

2. t = 53, L = -1022, U = 1023

3. Underflow Limit:  $UFL = 0.1 \times 2^L$ . 0 < x < UFL fl(x) = 0.

4. Overflow Limit:  $OFL = 0.11 \dots 1 * 2^U$ . x > OFL  $fl(x) = \infty$ .

5. :  $UFL \le x \le OFL$  fl(x)  $x = 0.a_1a_2...a_n \cdots \times 2^m$ .  $a_{n+1} = 1$   $d_t + 1$ 

3 ( )

$$x - fl(x) = 2^m \times 0.0 \dots 0a_{n+2} \dots$$
$$= 2^m \times [2^{-(t+2)} + 2^{-(t+3)} + \dots]$$
$$= 2^m \times 2^{-(t+1)}$$

$$\frac{x - fl(x)}{x} < \frac{x - fl(x)}{0.5 \times 2^m} = 2^{-t}$$

 $\varepsilon$ 

4

 $fl(x) = x(1+\delta), \quad |\delta| \le \varepsilon$ 

<sup>1</sup>floating Number System

## 1.1.3

 $+ \Rightarrow$ 

- $5 \quad (x+y) \quad x, y \quad x+y$ 
  - 1. *m*
  - 2.
  - 3.
  - 4. .....
  - 5.

$$fl(x) + fl(y) = x(1 + \delta_x) + y(1 + \delta_y)$$

## 1.2

## 1.2.1

- 1.
- 2.
- $y = f(x) f(x^*)$  Taylor 3.
- rounding-off 4.

## 1.2.2

**6** ( ) 
$$x = x^*$$

$$e(x^*) = x^* - x.$$

$$|e(x^*)| \le \varepsilon^*$$

**7** ( ) 
$$x x^*$$

$$e_r(x^*) = \frac{x^* - x}{x}$$

 $\varepsilon_r^*$ 

$$|e_r(x^*)| \le \varepsilon_r^*$$

 $\boldsymbol{x}$ 

$$\bar{e}_r(x^*) = \frac{x^* - x}{x^*}$$

$$|\bar{e}_r(x^*)| \le \frac{\varepsilon^*}{|x^*|}.$$

 $\varepsilon^* \ll 1$ 

$$|e_r - \bar{e}_r| = O((\varepsilon_r^*)^2)$$

8 ( )  $x \in R x^*$   $x^* x n$  n

$$x^*$$
  $x$   $n$ 

$$n$$
  $n$ 

$$|x^* - x| \le \frac{1}{2} \times 10^{m-n}$$

 $x^*$  n

9 ( ) 
$$x = 0.a_1 a_2 \dots a_n \times 10^m n$$

$$\left| \frac{x^* - x}{x} \right| \le \frac{1}{2a_1} \times 10^{1-n}.$$

$$\left| \frac{x^* - x}{x} \right| \le \frac{1}{2(1 + a_1)} \times 10^{1-n},$$

 $x^*$  n

$$x \ge 0.a_1 \qquad a_1 \ne 0$$

#### 1.2.3

1. 
$$/: \varepsilon(x^* \pm y^*) \le \varepsilon_x^* + \varepsilon_y^*$$

2. 
$$\varepsilon(x^*y^*) \le |x^*|\varepsilon_y^* + |y^*|\varepsilon_x^*$$

3. : 
$$\varepsilon(\frac{x^*}{y^*}) = \frac{|x^*|\varepsilon_y^* + |y^*|\varepsilon_x^*}{|y^*|^2}$$

$$\begin{array}{cccc} x \ y & x^* \ y^* & x^* \pm y^* \ x \pm y \\ & |x^* \pm y^* - (x \pm y)| \leq |x^* - x| + |y^* - y| \leq \varepsilon_x^* + \varepsilon_y^* \\ \Rightarrow & \varepsilon(x^* \pm y^*) \leq \varepsilon_x^* + \varepsilon_y^* \end{array}$$

$$+1 - 1$$

11 ( ) 
$$A = f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) \mathbf{x}^* \mathbf{x}$$
 Peano Taylor  $A$ 

$$e(A^*) = f(\mathbf{x}^*) - f(\mathbf{x})$$

$$= \sum_{p=1}^{q} d^k f(\mathbf{x}^*) + o(||x^* - x||^q)$$

$$= 1$$

$$= \sum_{k=1}^{n} \partial_k f(\mathbf{x}^*)(x^* - x) + o(||x^* - x||^q)$$

$$\varepsilon(A^*) \approx \sum_{k=1}^n \partial_k f(\mathbf{x}^*) \varepsilon(x^*)$$
$$\varepsilon_r(A^*) = \frac{\varepsilon(A^*)}{|A^*|}$$

#### 1.2.4

**12** *n* 

$$S_n = \sum_{i=0}^n a_i$$

$$S_2^* = fl(a_1 + a_2) = (a_1 + a_2)(1 + \varepsilon_2), \quad |\varepsilon_2| \le \varepsilon = 2^{-t}$$

$$\dots$$

$$S_n^* = fl(S_{n-1}^* + a_n)(1 + \varepsilon_n), \quad |\varepsilon_n| \le \varepsilon$$

$$S_n^* \qquad \varepsilon_1 = 0$$

$$S_n^* = \sum_{k=1}^n a_k \prod_{p=k}^n (1 + \varepsilon_p)$$

$$\prod_{i=k}^{n} (1 + \varepsilon_k) \approx 1 + \sum_{i=k}^{n} \varepsilon_k$$

$$S_n^* \approx \sum_{k=1}^n a_k (1 + \sum_{p=k}^n \varepsilon_p)$$
$$= S_n + \sum_{k=1}^n a_k \sum_{p=k}^n \varepsilon_p$$

 $|\varepsilon_i| \le \varepsilon$ 

$$|S_n^* - S_n| \le \sum_{k=1}^n |a_k| \sum_{p=k}^n |\varepsilon_p| \le \varepsilon \sum_{k=1}^n |a_k| (n-k+1)$$

 $n - k + 1 k \qquad [18]$ 

1.3

1.

2.

3.

14 ( 
$$e^A$$
)  $e^A A \in \mathbb{R}^{n \times n}$ 

$$e^A = e^{(A/2^n)2^n} = B^{2^n}$$

B  $B^{2^n}$  B  $x \to 0$   $e^x$  Taylor

$$e^x = 1 + x + \dots + \frac{x^n}{n!} + \dots$$

 $n \qquad A/2^n \approx 0$ 

$$B \approx I + C + \frac{1}{2}C^2$$
,  $C = A/2^n$ 

 $B^2$ 

$$B^2 \approx I + 2(C + \frac{1}{2}C^2) + (C + \frac{1}{2}C^2)^2$$

**15** ( ) 
$$(1)$$
  $p(z), z \in \mathbf{R}$ 

$$p(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n$$
 (1)

 $b_n$ 

$$b_0 = a_0, \quad b_k = a_k + b_{k-1}z$$

 $b_n$ 

$$p'(z) = \sum_{k=0}^{n-1} b_k z^{n-1-k}$$

$$x-z$$
  $p(x)$   $b_n(x)$ 

$$p(x) = (x - z)q(x) + b_n(x),$$

$$x = z 0 p(z) = b_n(z)$$

**16** ( ) 
$$x_0 x_1 x$$
  $x_1 x_0 x$ 

$$\bar{x} = x_1 + \omega(x_1 - x_0)$$

x

17 
$$(\pi)$$
  $\pi \pi_n$   $2n$ 

$$\widetilde{\pi}_n = \frac{1}{3}(4\pi_{2n} - \pi_n)$$

$$\pi_n \widetilde{\pi}_n \pi$$

$$|\pi_n - \pi| = O(\frac{1}{n^2}), \quad |\widetilde{\pi}_n - \pi| = O(\frac{1}{n^4})$$

$$\pi_n$$
.

$$\pi_n = n \sin \frac{\pi}{n} = \pi - \frac{\pi^3}{3!} \frac{1}{n^2} + \frac{\pi^5}{5!} \frac{1}{n^4} - \dots \Rightarrow |\pi_n - \pi| = O(\frac{1}{n^2})$$

 $\widetilde{\pi}_n$ .

$$\widetilde{\pi}_n = \pi_{2n} + k(\pi_{2n} - \pi_n) = (1+k)\pi_{2n} - k\pi_n$$

$$= (1+k)(\pi - \frac{\pi^3}{3!} \frac{1}{4n^2} + \cdots) - k(\pi - \frac{\pi^3}{3!} \frac{1}{n^2} + \cdots)$$

$$= \pi - (\frac{k+1}{4} - k)\frac{\pi^3}{3!} \frac{1}{n^2} + O(\frac{1}{n^4})$$

$$k = \frac{1}{3}$$

$$|\widetilde{\pi}_n - \pi| = O(\frac{1}{n^4}) \quad \blacksquare$$

$$\pi_n \qquad \qquad n = 3 \quad 6 \qquad \qquad \pi_3 = 3\sqrt{3}/2$$

$$\pi_{2n} = \sqrt{2n(n - \sqrt{n^2 - \pi_n^2})},$$

 $\pi_2 n \ \pi_n$ 

2.1

18 ( ) 
$$\{a_n\} \{b_n\}$$

$$0 \le a_1 \le a_2 \le \dots \le a_n$$
  
$$0 \le b_1 \le b_2 \le \dots \le b_n$$

$$\sum_{i=1}^{n} a_i b_i \ge \sum_{i=1}^{n} a_i b_{k_i} \ge \sum_{i=1}^{n} a_i b_{n-i+1}$$

$$(a_1 a_2 \cdots a_n)^{1/n} \le \frac{a_1 + a_2 + \cdots + a_n}{n}$$

$$a_1 = a_2 = \dots = a_n$$

$$\prod a_i = 1$$

$$a_1 = \frac{\alpha_1}{\alpha_2}, \quad \dots, \quad a_{n-1} = \frac{\alpha_{n-1}}{\alpha_n}, \quad a_n = \frac{\alpha_n}{\alpha_1}$$
$$\frac{\alpha_1}{\alpha_2} + \dots + \frac{\alpha_n}{\alpha_1} \ge n$$

$$\alpha_1 \le \alpha_2 \le \dots \le \alpha_n$$

L.H.S 
$$\geq \alpha_1 \frac{1}{\alpha_1} + \dots + \alpha_n \frac{1}{\alpha_n} = n$$