

Representasi Floating Point

Pada

CPU

ARITMATICH

Representasi Floating Point

Untuk menuliskan bilangan floating point (bilangan pecahan) dilakukan dengan menuliskan dalam bentuk eksponensial.

Sehingga bilangan tersebut memiliki bilangan dasar, bilangan pemangkat dan basis bilangan tersebut

Penulisan Notasi Ilmiah

Contoh ; pada bil. Desimal

976.000.000.000.000 ditulis $9,76 \times 10^{14}$

0,000000000000976 ditulis $9,76 \times 10^{-12}$

- Representasi :

$$\pm S * B \pm E$$

Tanda : + atau –

Signifikan (S) disebut juga mantissa

Eksponen (E)

Base (B)

Representasi Floating Point

Ko-prosesor Aritmatika mendukung tiga repr floating point :

- Short (32 bit)/ presisi tunggal dengan bias 7Fh
- Long (64 bit)/presisi ganda dengan bias 3FFh

Exponensial

Penulisan bilangan floating point dengan cara exponensial dapat menyebabkan adanya kemungkinan sebuah bilangan ditulis dengan cara yang bermacam-macam

$$300 = 0,3 * 10^3 = 3 * 10^2 = 30 * 10^1 = 30^0 * 100 = 3000 * 10^{-1} \dots$$

demikian juga pada bilangan biner

$$11,11 = 11,11 * 2^0 = 1,111 * 2^1 = 0,1111 * 2^2$$

Standarisasi untuk penulisan bilangan

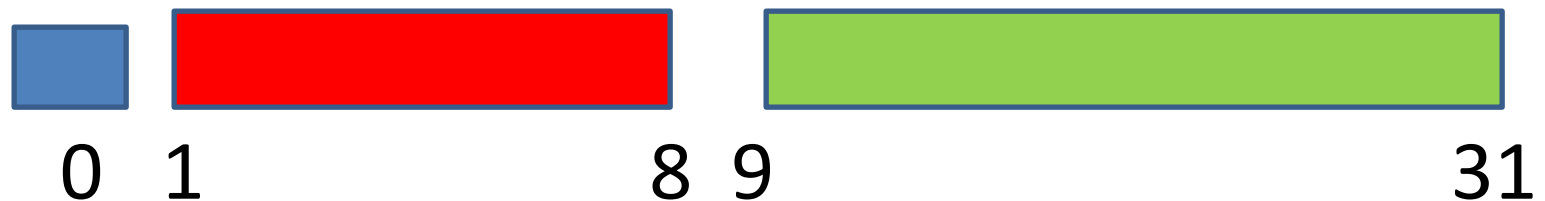
- Normalisasi bentuknya adalah #

$$\pm 1,bbb\dots b \times 2^{\pm E}$$

- Bit pertama significand selalu 1 sehingga tidak perlu disimpan (implisit) dalam field significand.
Dan b adalah bilangan biner (1 atau 0).
- Untuk keperluan yang luas, maka diadakan standar bagi representasi bilangan floating point ini, yaitu standard IEEE 754. Standar ini juga mendefinisikan operasi aritmetiknya

- Format dari standar IEEE 754 tahun 1985
Sign bit, Eksponen+bias, signifikan

➤ format floating point 32 bit :



➤ Pada 64 bit:

-1 - | | --- 11 bit ----- | | ----- 52 bit ----- |

Format penulisan menurut standar IEEE 754

| | | |
|---|-----------------|------------------------|
| 1 bit | 8 bit | 23 bit |
| tanda | biased exponent | Fraction / significand |
| (a) Format tunggal (single format) 32 bit | | |
| 1 bit | 11 bit | 52 bit |
| | biased exponent | Fraction / significand |
| (b) Format ganda (double format) 64 bit | | |

- Pada format tunggal, bit paling kiri digunakan untuk representasi tanda 0 jika positif dan 1 jika negatif, sedangkan 8 bit berikutnya adalah pangkat (exponen) yang direpresentasikan dalam bentuk bias.
- Bagian 23 bit terakhir digunakan untuk menunjukkan bit dari bilangan fraction nya

Contoh konversi bilangan ke format floating point IEEE

Bilangan= -113.3125

$$113_{10} = 1110001_2$$

sedangkan 0.3125 didapat:

$$0.3125 \times 2 = 0.625 \quad 0$$

$$0.625 \times 2 = 1.25 \quad 1$$

$$0.25 \times 2 = 0.5 \quad 0$$

$$0.5 \times 2 = 1.0 \quad 1$$

$$\text{jadi } 113.3125_{10} = 1110001.0101_2.$$

$$\text{di normalisasi menjadi: } 1110001.0101 \times 2^0 = 1.1100010101 \times 2^6$$

Pada penulisan SIGNIFIKAN yang ditulis setelah tanda koma,

1,xxxxx

ditulis xxxxx nya saja

Karena bit 1 adalah *implied one-bit*.

Contoh konversi bilangan ke format floating point IEEE

karena mantisa terdiri dari 23 bit maka akan menjadi 110001010100000000000000.

Bilangan pemangkat adalah 6 maka dirubah ke bias dengan cara $6 + 127 = 133 = 10000101_2$.

Karena negatif maka bilangan tanda = 1

| | | |
|---|----------|--------------------------|
| 1 | 10000101 | 110001010100000000000000 |
|---|----------|--------------------------|

Sama dengan

$C2E2A000_{16}$.

Referensi :

1. Riyanto Sigit, ST., MKom , Nur Rosyid Muhtada'i, SKom , Setiawardhana, ST, Hero Yudo Martono, ST , ITS
2. Dr. Wahyu Kusuma, staffsite.gunadarma.ac.id