

# 1 IEEE 754 zapis realnih brojeva u pokretnom zrezu sa binarnom osnovom u jednostrukoj tačnosti (binary32)

Broj koji se zapisuje je oblika

$$\pm (1.d_{-1}d_{-2} \dots d_{-23})_2 \cdot 2^{exp}$$

Karakteristike *binary32* zapisa:

- 1 bit za zapis znaka broja: 0 ako je broj pozitivan ili 1 ako je broj negativan
- 8 bitova za zapis eksponenta: eksponent se uvek zapisuje sa uvećanjem 127 (maksimalna dekadna vrednost je 127, a minimalna -126)
- 23 bita za zapis frakcije: zapisuju se samo cifre iza decimalne tačke, a implicitna jedinica se podrazumeva

## 1.1 Zadaci

1. Zapisati broj  $-111.625$  po IEEE 754 standardu sa binarnom osnovom u jednostrukoj tačnosti.

$$(111)_{10} = (1101111)_2 \text{ jer je}$$

111	55	27	13	6	3	1	0
1	1	1	1	0	1	1	

smer čitanja ←

$$\text{ili: } (111)_{10} = (64 + 32 + 8 + 4 + 2 + 1)_{10} = (1101111)_2$$

$$(0.625)_{10} = (0.101)_2 \text{ jer je}$$

0.625	0.250	0.5	0
0	1	0	1

smer čitanja →

$$\text{ili: } (0.625)_{10} = (0.5 + 0.125)_{10} = (0.101)_2$$

$$\Rightarrow -(111.625)_{10} = -(1101111.101)_2 = -(1.101111101) \cdot 2^6$$

bit za znak: 1 (broj je negativan)

eksponent:  $6 + 127 = 133 = 128 + 5 = (10000101)_2$

frakcija:  $101111101 \underbrace{00 \dots 00}_{14}$

konačno:  $1 \ 10000101 \ 101111101 \underbrace{00 \dots 00}_{14}$

2. Zapisati broj  $-1124$  po IEEE 754 standardu sa binarnom osnovom u jednostrukoj tačnosti.

$$(1124)_{10} = (1024 + 64 + 32 + 4)_{10} = (10001100100)_2$$

$$\Rightarrow -(1124)_{10} = -(10001100100)_2 = -(1.00011001) \cdot 2^{10}$$

bit za znak: 1

eksponent:  $10 + 127 = 137 = 128 + 9 = (10001001)_2$

frakcija:  $00011001 \underbrace{00 \dots 00}_{15}$

konačno:  $1 \ 10001001 \ 00011001 \underbrace{00 \dots 00}_{15}$

3. Zapisati broj  $0.4375$  po IEEE754 standardu u formatu binary32.

$$(0.4375)_{10} = (0.25 + 0.125 + 0.0625)_{10} = (0.0111)_2$$

$$\Rightarrow (0.4375)_{10} = (0.0111)_2 = (1.11)_2 \cdot 2^{-2}$$

bit za znak: 0  
eksponent:  $-2 + 127 = 125 = (01111111)_2 - (00000010)_2 = (01111101)_2$   
frakcija:  $11 \underbrace{00 \dots 00}_{21}$

konačno:  $0 \ 01111101 \ 1 \underbrace{00 \dots 00}_{21}$

4. Zapisati broj 168.125 po IEEE 754 standardu u formatu binary32.

$$(168.125)_{10} = (10101000.001)_2 = (1.0101000001) \cdot 2^7$$

bit za znak: 0  
eksponent:  $7 + 127 = 128 + 6 = (10000110)_2$   
frakcija:  $0101000001 \underbrace{00 \dots 00}_{13}$

konačno:  $0 \ 10000110 \ 0101000001 \underbrace{00 \dots 00}_{13}$

5. Zapisati broj  $-79.5 \cdot 2^{-21}$  po IEEE 754 standardu u formatu binary32.

$$\begin{aligned} -(79.5)_{10} &= -(1001111.1)_2 = -(1.0011111) \cdot 2^6 \\ -79.5 \cdot 2^{-21} &= -(1.0011111) \cdot 2^{-15} \end{aligned}$$

bit za znak: 1  
eksponent:  $-15 + 127 = (01111111)_2 - (00001111)_2 = (01110000)_2$   
frakcija:  $0011111 \underbrace{00 \dots 00}_{16}$

konačno:  $1 \ 01110000 \ 0011111 \underbrace{00 \dots 00}_{16}$

6. Koji dekadni broj je predstavljen IEEE 754 zapisom sa binarnom osnovom:

a) 11000011010010000000000000000000

Zapis razdvajamo na znak, eksponent i frakciju:  
1 10000110 100100000000000000000000

znak: -  
eksponent:  $(10000110)_2 - (01111111)_2 = (00000111)_2 = 7$   
ili  
 $(10000110)_2 - 127 = 134 - 127 = 7$   
frakcija:  $1.100100000000000000000000$

dekadna vrednost:  
 $-(1.1001)_2 \cdot 2^7 = -(11001000)_2 = -200$

b) 11000011000010011010100000000000

Zapis razdvajamo na znak, eksponent i frakciju:  
1 10000110 000100110101000000000000

znak: -  
eksponent:  $(10000110)_2 - 127 = 134 - 127 = 7$   
ili  
 $(10000110)_2 - (01111111)_2 = (00000111)_2 = 7$   
frakcija:  $1.000100110101000000000000$

dekadna vrednost:  
 $-(1.000100110101)_2 \cdot 2^7 = -(10001001.10101)_2 = -137.65625$

c) 00011011110010100000000000000000

Zapis razdvajamo na znak, eksponent i frakciju:  
0 00110111 100101000000000000000000

znak: +

$$\text{eksponent: } (00110111)_2 - (01111111)_2 = (10111000)_2 = -72$$

ili

$$(00110111)_2 - 127 = 55 - 127 = -72$$

frakcija: 1.100101000000000000000000

dekadna vrednost:

$$+(1.100101)_2 \cdot 2^{-72} = (1100101)_2 \cdot 2^{-78} = 101 \cdot 2^{-78}$$

d) 01000011101001000000000000000000

Zapis razdvajamo na znak, eksponent i frakciju:  
0 10000111 010010000000000000000000

znak: +

$$(10000011)_2 - (01111111)_2 = (00000100)_2 = 4$$

ili

$$\text{eksponent: } (10000011)_2 - 127 = 131 - 127 = 4$$

frakcija: 1.010010000000000000000000

dekadna vrednost:

$$+(1.01001)_2 \cdot 2^4 = (10100.1)_2 = 20.5$$

e) 11001000101101000000000000000000

Zapis razdvajamo na znak, eksponent i frakciju:  
1 10010001 011010000000000000000000

znak: -

$$\text{eksponent: } (10010001)_2 - (01111111)_2 = (00010010)_2 = 18$$

ili

$$(10010001)_2 - 127 = 145 - 127 = 18$$

frakcija: 1.011010000000000000000000

dekadna vrednost:

$$-(1.01101)_2 \cdot 2^{18} = -(101101)_2 \cdot 2^{13} = -45 \cdot 2^{13}$$

## 2 Zapis specijalnih vrednosti

- Pozitivna nula: `0 00000000 000000000000000000000000`
- Negativna nula: `1 00000000 000000000000000000000000`
- $+\infty$ : `0 11111111 000000000000000000000000`
- $-\infty$ : `1 11111111 000000000000000000000000`
- sNaN vrednost: bit znaka može biti 0 ili 1, eksponent je 11111111, prvi bit frakcije je 0, a bar jedan od preostalih bitova frakcije mora biti 1

Na primer:

```
1 11111111 011000000000000000000000
0 11111111 00000000000110010000010
0 11111111 011111111111111111111111
```

- qNaN vrednost: bit znaka može biti 0 ili 1, eksponent je 11111111, prvi bit frakcije je 1, a preostali bitovi mogu uzeti proizvoljne vrednosti

Na primer:

```
1 11111111 100110000000000000000000
0 11111111 11000011000000110001110
0 11111111 100000000000000000000000
```

- subnormalne (denormalizovane) vrednosti: bit znak može biti 0 ili 1, eksponent je 00000000, a frakcija mora biti različna od nule.

Na primer:

```
0 00000000 101011000000000000000000
0 00000000 000000000000000000000011
1 00000000 010000100000000000000000
```

Implicitni bit frakcije u ovom slučaju je 0, a dekadna vrednost eksponenta uvek -126.

### 2.1 Zadaci

1. Koji broj je predstavljen IEEE 754 zapisom sa binarnom osnovom:

a) `11111111110010011001000000000000`

Zapis razdvajamo na znak, eksponent i frakciju:  
`1 11111111 100100110010000000000000`

U eksponentu su sve jedinice, pa je u pitanju specijalna vredost: kako je frakcija različita od nule i prvi bit 1 u pitanju je qNaN vrednost.

b) `10000000000000010001000000000000`

Zapis razdvajamo na znak, eksponent i frakciju:  
`1 00000000 0000010001000000000000`

U eksponentu su samo nule, pa je u pitanju specijalna vredost: kako je frakcija različita od nule u pitanju je zapis subnormalnog broja:

znak: -  
eksponent: -126  
frakcija: `0.000000100010000000000000`

dekadna vrednost:  
 $-(0.00000010001)_2 \cdot 2^{-126} = -(10001)_2 \cdot 2^{-137} = -17 \cdot 2^{-137}$

c) 00000000000000000000000000000000

Zapis razdvajamo na znak, eksponent i frakciju:  
0 00000000 000000000000000000000000

U eksponentu su sve nule i u frakciji su sve nule pa je u pitanju vrednost +0.

d) 11111111110010011001000000000000

Zapis razdvajamo na znak, eksponent i frakciju:  
1 11111111 100100110010000000000000

U eksponentu su sve jedinice, pa je u pitanju specijalna vredost: kako je frakcija različita od nule i prvi bit 1 u pitanju je qNaN vrednost.

e) 00000000011010000000000000000000

Zapis razdvajamo na znak, eksponent i frakciju:  
0 00000000 110100000000000000000000

U eksponentu su samo nule, pa je u pitanju specijalna vredost: kako je frakcija različita od nule u pitanju je zapis subnormalnog broja:

znak: +  
eksponent: -126  
frakcija: 0.110100000000000000000000

dekadna vrednost:  
 $+(0.1101)_2 \cdot 2^{-126} = (1101)_2 \cdot 2^{-130} = 13 \cdot 2^{-130}$

f) 11111111101100000000000000000011

Zapis razdvajamo na znak, eksponent i frakciju:  
1 11111111 011000000000000000000011

U eksponentu su sve jedinice, pa je u pitanju specijalna vredost: kako je frakcija različita od nule i prvi bit 0 u pitanju je sNaN vrednost.

g) 11111111100000000000000000000000

Zapis razdvajamo na znak, eksponent i frakciju:  
1 11111111 000000000000000000000000

U eksponentu su sve jedinice, pa je u pitanju specijalna vredost: frakcija je jednaka nuli pa je u pitanju vrednost  $-\infty$ .

### 3 IEEE 754 zapis realnih brojeva u pokretnom zarezu sa binarnom osnovom u dvostrukoj tačnosti (binary64)

Broj koji se zapisuje je oblika

$$\pm (1.d_{-1}d_{-2}\dots d_{-52})_2 \cdot 2^{exp}$$

Karakteristike *binary64* zapisa:

- 1 bit za zapis znaka broja: 0 ako je broj pozitivan ili 1 ako je broj negativan
- 11 bitova za zapis eksponenta: eksponent se uvek zapisuje sa uvećanjem 1023 (maksimalna dekadna vrednost je 1023, a minimalna -1022)
- 52 bita za zapis frakcije: zapisuju se samo cifre iza decimalne tačke, a implicitna jedinica se podrazumeva

#### 3.1 Zadaci

1. Zapisati broj 48.125 po IEEE 754 standardu sa binarnom osnovom u dvostrukoj tačnosti.

$$48.125 = (110000.001)_2 = (1.10000001)_2 \cdot 2^5$$

bit za znak: 0

$$\text{eksponent: } 5 + 1023 = 1028 = 1024 + 4 = (1000000100)_2$$

$$\text{frakcija: } 10000001\underbrace{00\dots 00}_{44}$$

$$\text{konačno: } 0\ 1000000100\ 10000001\underbrace{00\dots 00}_{44}$$

2. Zapisati broj  $-1780.53125$  po IEEE 754 standardu sa binarnom osnovom u dvostrukoj tačnosti.

$$\begin{aligned} -1780.53125 &= -(11011110100.10001)_2 = -(1.101111010010001)_2 \cdot 2^{10} \\ 0.53125 &= 0.5 + 0.03125 = 2^{-1} + 2^{-5} = (0.10001)_2 \end{aligned}$$

bit za znak: 1

$$\text{eksponent: } 10 + 1023 = 1033 = 1024 + 9 = (10000001001)_2$$

$$\text{frakcija: } 101111010010001\underbrace{00\dots 00}_{37}$$

$$\text{konačno: } 1\ 10000001001\ 101111010010001\underbrace{00\dots 00}_{37}$$

3. Zapisati broj  $-0.4375$  po IEEE 754 standardu sa binarnom osnovom u dvostrukoj tačnosti.

$$-0.4375 = -(0.25 + 0.125 + 0.0625)_{10} = (0.0111)_2 = -(1.11)_2 \cdot 2^{-2}$$

bit za znak: 1

$$\text{eksponent: } -2 + 1023 = 1021 = 1024 - 3 = (10000000000)_2 - (00000000011)_2 = (0111111101)_2$$

$$\text{frakcija: } 11\underbrace{00\dots 00}_{50}$$

$$\text{konačno: } 1\ 0111111101\ 11\underbrace{00\dots 00}_{50}$$

4. Koji dekadni broj je predstavljen IEEE 754 zapisom sa binarnom osnovom u dvostrukoj tačnosti:

$$\text{a) } 1100000110011011\underbrace{00\dots 00}_{48}$$

Zapis razdvajamo na znak, eksponent i frakciju:

$$1\ 10000011001\ 1011\underbrace{00\dots 00}_{48}$$

znak: -

$$\text{eksponent: } (10000011001)_2 - 1023 = 1024 + 16 + 9 - 1023 = 26$$

ili  
 $(10000011001)_2 - (01111111111)_2 = (00000011010)_2 = 26$   
 frakcija:  $1.1011\underbrace{00\dots 00}_{48}$

dekadna vrednost:  
 $-(1.1011)_2 \cdot 2^{26} = -(11011)_2 \cdot 2^{22} = -27 \cdot 2^{22}$

b)  $100001100010100011\underbrace{00\dots 00}_{46}$

Zapis razdvajamo na znak, eksponent i frakciju:  
 $1\ 00001100010\ 100011\underbrace{00\dots 00}_{46}$

znak: -  
 eksponent:  $(00001100010)_2 - 1023 = 98 - 1023 = -925$   
 ili  
 $(00001100010)_2 - (01111111111)_2 = (10001100011)_2 = -925$   
 frakcija:  $1.100011\underbrace{00\dots 00}_{46}$

dekadna vrednost:  
 $-(1.100011)_2 \cdot 2^{-925} = -(1100011)_2 \cdot 2^{-931} = -99 \cdot 2^{-931}$

c)  $0111000010000000101\underbrace{00\dots 00}_{45}$

Zapis razdvajamo na znak, eksponent i frakciju:  
 $0\ 11100001000\ 0000101\underbrace{00\dots 00}_{45}$

znak: +  
 eksponent:  $(11100001000)_2 - 1023 = 1024 + 512 + 256 + 8 - 1023 = 777$   
 ili  
 $(11100001000)_2 - (01111111111)_2 = (01100001001)_2 = 777$   
 frakcija:  $1.0000101\underbrace{00\dots 00}_{45}$

dekadna vrednost:  
 $+(1.0000101)_2 \cdot 2^{777} = (10000101)_2 \cdot 2^{770} = 133 \cdot 2^{770}$

## 4 Zapis specijalnih vrednosti

- Pozitivna nula:  $0\ 000000000000\underbrace{00\dots 000}_{52}$
- Negativna nula:  $1\ 000000000000\underbrace{00\dots 000}_{52}$
- $+\infty$ :  $0\ 11111111111\underbrace{00\dots 000}_{52}$
- $-\infty$ :  $1\ 11111111111\underbrace{00\dots 000}_{52}$
- sNaN vrednost: bit za znaka može biti 0 ili 1, eksponent je 1111111111, prvi bit frakcije je 0, a bar jedan od preostalih bitova frakcije mora biti 1

Na primer:

$1\ 11111111111\ \underbrace{00\dots 00}_{20}\ \underbrace{11\ 00\dots 00}_{30}$   
 $0\ 11111111111\ 0111111111\ \underbrace{00\dots 00}_{44}$   
 $0\ 11111111111\ 010101\ \underbrace{11\dots 11}_{46}$

- qNaN vrednost: bit za znaka može biti 0 ili 1, eksponent je 1111111111, prvi bit frakcije je 1, a preostali bitovi mogu uzeti proizvoljne vrednosti

Na primer:

1 1111111111 1001111 00...00  
 0 1111111111 100...00  
 0 1111111111 11011011 00...00

- subnormalne (denormalizovane) vrednosti: bit za znak može bit 0 ili 1, eksponent je 0000000000, a frakcija mora biti različina od nule.

Na primer:

0 0000000000 101 00...00  
 0 0000000000 00...00 11  
 1 0000000000 1111111111 00...00

Imlicitni bit frakcije u ovom slučaju je 0, a dekadna vrednost eksponenta uvek -1022.