2. Seštevanje celih števil, Zapis v plavajoči vejici

Seštevanje celih števil

- Prenos in pravilnost rezultata pri seštevanju **nepredznačenih števil**

  Pravilnost rezultata pri seštevanju nepredznačenih števil določa prenos (*ang. carry*) na najpomembnejšem bitu C_{MSB}. Če je prenos C_{MSB} enak 0, je rezulat pravilen. Če je prenos 1, je rezultat nepravilen. V tem primeru se rezultata ne da predstaviti z danim številom bitov. Torej:

  $$C_{MSB} = \begin{cases} 0, & Rezultat\ seštevanja\ je\ pravilen \\ 1, & Rezultat\ seštevanja\ ni\ pravilen \end{cases},$$

  - **Primer**: Opazujte prenos pri seštevanju 190 in 70 :

    Najprej pretvorimo števili v dvojiški sistem:

    $$190_{(10)} = 128 + 32 + 16 + 8 + 4 + 2 \rightarrow 1011\,1110_{(2)}$$
    $$70_{(10)} = 64 + 4 + 2 \rightarrow 0100\,0110_{(2)}$$

    Seštejemo binarni števili po modulu 2 (1+1 = 0 -> prenos v naslednji bit enak 1)

    $$\begin{array}{rr} & 1011\,1110 \\ + & 0100\,0110 \\ \hline \\ C_{MSB} = 1 & 0000\,0010 \end{array}$$

    C_{MSB} je enak 1, torej rezultat ni pravilen.

- Prenos in pravilnost rezultata pri seštevanju **predznačenih števil**

  Veljavnost rezultat pri seštevanju predznačenih števil določa bit preliva (*ang. overflow*) V. Če sta oba znaka seštevancev enaka in je znak rezultata različen potem je V enak 1 ter je rezultat nepravilen. Drugače pa je $V = 0$ in je rezultat pravilen. To lahko predstavimo z naslednjo tabelo :

op1	op2	rez	V
+	-	+ / -	0
-	+	+ / -	0
+	+	+	0
+	+	-	1
-	-	-	0
-	-	+	0

$$V = \begin{cases} 0, & Rezultat\ seštevanja\ pravilen \\ 1, & Rezultat\ seštevanja\ ni\ pravilen \end{cases},$$

  - **Primer**: Opazujte prenos pri seštevanju 123 in (-123) :

    Pretvorba v dvojiški sistem (Dvojiški komplement):

$$124_{(10)} \rightarrow 0111\ 1100_{(2)}$$
$$-123_{(10)} \rightarrow 1000\ 0101_{(2)}$$

    Seštejemo po modulu 2

$$\begin{array}{rr} & 0111\ 1100 \\ + & 1000\ 0101 \\ \hline C_{MSB} = 1 & 0000\ 0001 \end{array}$$

    Seštevanca in rezultat so pozitivni -> $V = 0$ -> Rezultat je pravilen čeprav je $C_{MSB} = 1$. Pri seštevanju predznačenih števil upoštevamo samo bit preliva V.

  - **Primer**: Opazujte prenos pri seštevanju -80 in (-60) :

    Pretvorba v dvojiški sistem (Drugi komplement):

$$-80_{(10)} \rightarrow 1011\ 0000_{(2)}$$
$$-60_{(10)} \rightarrow 1100\ 0100_{(2)}$$

Seštejemo po modulu 2

$$\begin{array}{rr} & 1011\,0000 \\ + & 1100\,0100 \\ \hline C_{MSB} = 1 & 0111\,0100 \end{array}$$

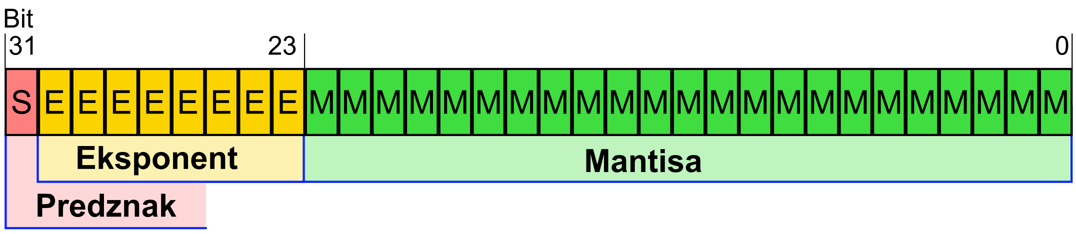
Seštevanci sta pozitivna, rezultat je negativen -> $V = 1$ -> Rezultat je nepravilen.

Zapis realnih števil s plavajočo vejico

Zapis s plavajočo vejico v formatu IEEE 754 uporabljamo za predstavitev realnih števil v računalništvu. Obstajaja več formatov zapisov števil v plavajoči vejici. Najpogostojše se uporabljata dva zapisa:

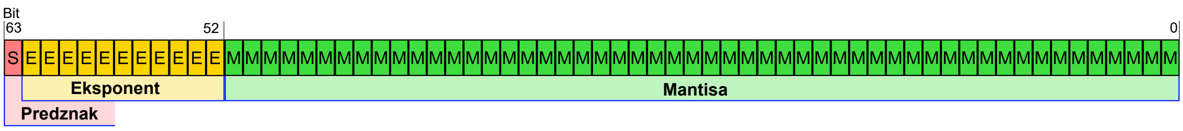
- IEEE 754 z enojno natančnostjo (podatkovni tip *float* - 32 bitov):

    1. Znak - 1 bit
    2. Eksponent - 8 bitov
    3. Mantisa - 23 bitov



- IEEE 754 z dvojno natančnostjo (podatkovni tip *double* - 64 bitov):

    1. Znak - 1 bit
    2. Eksponent - 11 bitov
    3. Mantisa - 52 bitov



- Primer: Število –210,5937510 najprej zapišimo v binarni obliki s plavajočo vejico, nato pa še šestnajstiško v prestavitvi IEEE 754 z enojno natančnostjo.

$$-210,5937510_{(10)} \rightarrow ?_{(IEEE754\ z\ enojno\ natančnostjo)}$$

    Postopek:

    1. Pretvorba v **dvojiški sistem**:

$$-210,5937510_{(10)} \rightarrow -1101\,0010,1001\,1_{(2)}$$

To ni IEEE 754 zapis !!!!

2. Normalizacija -> Pretvorimo v zapis $(1,m \cdot 2^e)$

$$-1101\,0010,1001\,1 = -1,1010\,0101\,0011 \cdot 2^7$$

   pri čemer:

   - $m = 1010\,0101\,0011$,
   - $e = 7$

3. Zapis v IEEE 754 z enojno natačnostjo

   - Predznak: Število je negativno $\rightarrow s = 1_{(2)}$
   - Eksponent: $E = e + 127 = 134_{(10)} \rightarrow 1000\,0110_{(2)}$
   - Mantisa: $m = 1010\,0101\,0011\,0000\,0000\,000_{(2)}$

4. Rešitev:

   $1\,1000\,01101010\,0101\,0011\,0000\,0000\,000_{IEEE754} = 0\mathrm{x}C352\,9800_{IEEE754}$

- Primer: Število $0\mathrm{x}BF580000$ je zapisano v IEEE 754 z enojno natančnostjo. Zapišimo desetiško vrednost.

$$0\mathrm{x}BF580000_{(IEEE754\ z\ enojno\ natančnosti)} \rightarrow ?_{(10)}$$

  Zapišimo podano število v dvojiškem sistemu:

$$0\mathrm{x}BF580000 = \underbrace{1}_{Predznak}\ \overbrace{011\,1111\,0}^{Eksponent}\underbrace{101\,1000\,0000\cdots0}_{Mantisa}$$

  - Predznank
    - $s = 1 \rightarrow$ Število je negativno
  - Eksponent
    - $E = 126 \rightarrow e = E - 127 = -1$
  - Mantisa
    - $m = 1011$
  - Končni rezultat
    - $$(-1)^s \cdot 1,m \cdot 2^e = -1.1011_2 \cdot 2^{-1} = -0,84375$$