B4B33PSY: Seminar 6

Bc. Štěpán Pressl

1. Fixed point aritmetika

- Uvažujeme číslo s pevným počtem desetinných míst
- Pokud máme m bitů pro počet míst za desetinnou čárkou, pak nejmenší realizovatelné číslo je 2^{-m}.
- Výhody: jednodušše realizovatelné, stačí hodnoty držet např v `uint32_t`
- Nevýhody: vždy stejná chyba, pro realizaci malých přírůstků by bylo potřeba hodně bitů

Příklad: Předpokládejte fixed point reprezentaci čísla $[0.0, 1.0]$ pomocí 8 bitů.

1. Jak by vypadala reprezentace čísla 0.521?
2. Sečtěte 0.2 a 0.1.
3. Vynásobte 0.2 a 0.1.

2. Ad IEEE754

- číslo s plovoucí desetinnou čárkou (přesnost se může měnit)
- denormalizovaná čísla: reprezentace nejmenších čísel, hodnota exponentu 0:

$$\text{DENORMALIZED} = (-1)^{\text{sign}} \cdot 2^{-126} \cdot (m_1 2^{-1} + m_2 2^{-2} + ...),$$

- Nula v IEEE754: exponent 0, mantissa 0. Znaménko jakékoli $\Rightarrow$ IEEE754 má dvě nuly (potřeba řešit),
- $\pm\infty$: exponent 255, mantissa 0,
- NaN: exponent 255, mantissa cokoliv od 0.

2. Převod čísel IEEE754

Na 32 bitů. https://www.h-schmidt.net/FloatConverter/IEEE754.html

1. 96.75
2. −163.1205
3. 6828.76873
4. 0.1

3. Vlastnosti IEEE754 (bonus bod)

1. Jaká je nejmenší hodnota realizovatelná pro normalizovaná čísla v absolutní hodnotě?
2. Pro denormalizovaná čísla (vyjma 0)?

4. Vlastnosti IEEE754 (bonus bod)

Navrhněte jednoduchý algoritmus, který má na vstupu dvě IEEE754 čísla. Algoritmus vytiskne to větší z nich.

5. Aritmetika v IEEE754

Sčítání (pro normalizovaná)

1. Obě čísla převést na stejný exponent.

   K zamyšlení: na jaký?
2. Sečíst mantissy.
3. Opětovně normalizovat

Příklad:

1. `0x41bc126f` + `0x3e374bc7` (cca $23.509 + 0.179$)
2. `0x4428d5d7` + `0x3700000f` (cca $675.34125 + 0.000007629408$)

5. Aritmetika v IEEE754

- U sčítání je potřeba brát na vědomí nejmenší přírůstek a uvažovat chybu při sčítání
- Obecně nelze v IEEE754 realizovat všechna čísla

Násobení

1. Znaménko: $S_1 \oplus S_2$,
2. Nový exponent: součet dvou exponentů
3. Vynásobit mantissy
4. Normalizovat

Příklad: 0.5 (`0x3f000000`) $\cdot$ -0.4375 (`0xbee00000`)