Úloha 1: Uvažujme systém čísel s pohyblivou čárkou o základu $\beta = 10$, s rozsahem exponentů ± 20 a se strojovým epsilon $\varepsilon_{mach} = 10^{-5}$. Dále předpokládejme, že tento systém nepodléha normě IEEE a nepoužívá subnormální čísla. Z definice strojového epsilon $\varepsilon_{mach} = \beta^{1-t}$ plyne $t = 6$, kde t je počet čísel mantisy. V tomto systému provedeme následující výpočty:

a) $1 + 10^{-7}$

$\mathrm{fl}(1+10^{-7}) = 1,00000 \cdot 10^0 + 1,00000 \cdot 10^{-7} \approx 1,00000 \cdot 10^0 + 0,00000 \cdot 10^0 = 1,00000 \cdot 10^0$

b) $1 + 10^3$

$\mathrm{fl}(1+10^3) = 1,00000 \cdot 10^0 + 1,00000 \cdot 10^3 = 0,00100 \cdot 10^3 + 1,00000 \cdot 10^3 = 1,00100 \cdot 10^3$

c) $1 + 10^7$

$\mathrm{fl}(1+10^7) = 1,00000 \cdot 10^0 + 1,00000 \cdot 10^7 \approx 0,00000 \cdot 10^7 + 1,00000 \cdot 10^7 = 1,00000 \cdot 10^7$

d) $10^{10} + 10^3$

$\mathrm{fl}(10^{10} + 10^3) = 1,00000 \cdot 10^{10} + 1,00000 \cdot 10^3 \approx 1,00000 \cdot 10^{10} + 0,00000 \cdot 10^{10} = = 1,00000 \cdot 10^{10}$

e) $10^{10}/10^{-15}$

$\mathrm{fl}(10^{10}/10^{-15}) = (1,00000 \cdot 10^{10})/(1,00000 \cdot 10^{-15}) = 1,00000 \cdot 10^{25} = 100000 \cdot 10^{20}$

f) $10^{-10} \cdot 10^{-15}$

$\mathrm{fl}(10^{-10} \cdot 10^{-15}) = 1,00000 \cdot 10^{-10} \cdot 1,00000 \cdot 10^{-15} = 1,00000 \cdot 10^{-25} = 0,00001 \cdot 10^{-20}$

Nyní určíme, který z následujících výpočtů v systému čísel s pohyblivou řádovou čárkou povede k překročení číselného rozsahu směrem k nule, jestliže uvažujeme systém bez subnormálních čísel s UFL$= 10^{-38}$.

a) $a = \sqrt{b^2 + c^2}$, kde $b = 1, c = 10^{-25}$.

Dosazením hodnot za b a c máme

$$a = \sqrt{1^2 + (10^{-25})^2} = \sqrt{1 + 10^{-50}}$$

Vidíme, že zde dochází k překročení číselného rozsahu směrem k nule u čísla c^2. Nahradíme-li jej při výpočtu nulou, nedopustíme se „moc velké chyby“, neboť číslo 10^{-50} je o něco blíž k nule než číslo 10^{-38}. Položíme-li tedy c^2 rovnu nule, tak ve výsledku k podtečení nedojde.

b) $a = \sqrt{b^2 + c^2}$, kde $b = c = 10^{-25}$.

Dosazením hodnot za b a c máme

$$a = \sqrt{(10^{-25})^2 + (10^{-25})^2} = \sqrt{2 \cdot 10^{-50}}$$

Opět zde dochází k překročení číselného rozsahu směrem k nule a to u čísla 10^{-50}. Nahradíme-li nyní hodnotu 10^{-50} nulou dostname výsledek výpočtu $a = 0$. Což je naprostý nesmysl, neboť platí $a = \sqrt{2 \cdot 10^{-50}} = \sqrt{2} \cdot 10^{-25}$. Takže není zde rozumné hodnotu 10^{-50} nahrazovat nulou a výsledkem výpočtu bude překročení číselného rozsahu směrem k nule.

c) $u = (v \cdot w)/(y \cdot z)$, kde $v = 10^{-15}, w = 10^{-30}, y = 10^{-20}, z = 10^{-25}$.

Dosazením hodnot za b a c máme

$$a = (10^{-15} \cdot 10^{-30})/(10^{-20} \cdot 10^{-25}) = 10^{-45}/10^{-45}.$$

I zde tomu není jinak, než v předchozích případech. Dochází zde k překročení číselného rozsahu směrem k nule. Nahrazovat zde hodnoty 10^{-45} nulou nemá cenu, neboť bychom dostali neurčitý výraz $0/0$. Přesto vidíme, že výsledek výpočtu je roven číslu 1. Předpokládáme-li, že se výrazy vyhodnocují zleva do prava a násobení má nejvyšší prioritu operací, pak by měl být výsledek výpočtu přinejmenším varovné hlášení, že došlo k podtečení nebo, že výsledek je nedefinovaný (neurčitý) výraz. Ovšem toto závisí na daném tvůrci softwaru, jak tyto „nekalosti“ ošetří.