Edita Pelantová; Miloslav Znojil

Můžeme věřit své vlastní kalkulačce?

Rozhledy matematicko-fyzikální, Vol. 85 (2010), No. 4, 11–18

Persistent URL: http://dml.cz/dmlcz/146379

Literatura

[1] Avizienis, A.: Signed-digit number representations for fast parallel arithmetic. *IRE Transactions on Electronic Computers* **10** (1961), str. 389–400.
[2] Colson, J.: A short account of negativo-affirmative arithmetic. *Philosophical Transaction of the Royal Society* **34** (1726), str. 161–173.
[3] Cauchy, A. L.: Sur les moyens d'éviter les erreurs dans les calculs. *Oeuvres complètes* **1** (November 1940), str. 431–442.
[4] Hayes, B.: The third base. *American Scientist* (November/December 2001), str. 490–494.
[5] Müller, J.-M.: Ordinateurs en quête d'arithmétique. *La Recherche* **26** (1995), str. 90–96.

Můžeme věřit své vlastní kalkulačce?

Edita Pelantová, FJFI, ČVUT Praha
Miloslav Znojil, Ústav jaderné fyziky, AV ČR, Praha

Abstract. The article examines the influence of rounding-off necessarily made by calculators and computers on the correctness of the result. An example of a particular task is shown to illustrate that the use of computational technology without considering its accuracy may result in errors of any size.

I v běžných matematických výpočtech se často setkáváme s iracionálními čísly. Výpočet obsahu kruhu vyžaduje použití Ludolfova čísla π, stanovení výšky úroků vyžaduje logaritmování, kde základem je Eulerovo číslo e. Je dávno známo, že obě zmíněné konstanty jsou iracionální čísla, to znamená, že jejich zápis v desítkové soustavě není konečný – jak je tomu např. u čísla $\frac{1}{5} = 0{,}2$ – ani od jistého místa periodický – jak je tomu u čísla $\frac{1}{6} = 0{,}166\,6\ldots$ Zápisy čísel π a e mají tvar:

$$\pi = 3{,}141\,592\,653\,589\,793\,238\,462\,643\,383\,279\,502\,884\,197\,169\,399\,375\ldots$$

$$\mathrm{e} = 2{,}718\,281\,828\,459\,045\,235\,360\,287\,471\,352\,662\,497\,757\,247\,093\,699\ldots$$

Výpočty, které provádíme pomocí kalkulačky či počítače, jsou omezeny pouze na racionální čísla. Počet platných číslic, se kterými kalkulačka a počítač pracují, je limitován a závisí na typu přístroje. Velikosti čísel, se kterými pracuje člověk, nejsou v principu omezeny, ale ruku na srdce, kolikamístná čísla byste byli ochotni násobit s tužkou na papíře vy?

Nemožnost pracovat s přesnou hodnotou Ludolfova čísla a Eulerova čísla nás většinou netrápí. Při školních výpočtech si vystačíme s přibližnou hodnotou čísla $\pi \doteq 3{,}14$ nebo $\pi \doteq \frac{22}{3}$. Chyba, která vznikne při výpočtu obsahu kruhu, je tak menší než jedno promile.

Je-li však úloha složitější a vyžaduje větší počet operací, může chyba přerůst všechny meze. Konstrukce mnohých technických zařízení vyžaduje komplikovaný výpočet. Nebudeme dopodrobna popisovat úlohu ze skutečné praxe – zabíhali bychom do podrobností nedůležitých pro vysvětlení vzniku chyb. Následující vymyšlený příběh dobře ilustruje podstatu problému. Pro úplnost dodejme, že matematické jádro příběhu bylo inspirováno přednáškou Francouze Jeana-Michela Mullera.

Otce právě narozeného syna, pana Nešetřila, upoutá ve výloze banky reklama „Iracionálně ke štěstí“. Banka nabízí rodičům, aby založili pro narozené dítě účet, na který vloží e korun, tedy iracionální částku. Banka slibuje, že po každém roce odečte z účtu jednu korunu jako poplatek za vedení účtu a vynásobí zbytek počtem let od založení účtu. V den 25. narozenin banka dítěti vyplatí jmění, které rodiče našetřili.

Pan Nešetřil se zamyslí, zda by neměl už teď pamatovat na štěstí svého syna. Rozhodne se o nabídce uvažovat a začne počítat: po prvním roce je na účtě $p_1 = \mathrm{e} - 1$ korun, po druhém roce $p_2 = 2(p_1 - 1) = 2(\mathrm{e} - 2)$ korun, po třetím roce $p_3 = 3(p_2 - 1)$ atd. Protože má po ruce mobil, začne počítat na kalkulačce. Hodnotu Eulerova čísla si nepamatuje ani přibližně, ale ve výloze banky je jako dekorace uvedeno číslo e s více než sto místy. Kalkulačka však dovolí naťukat pouze 9 platných míst. Proto správně zaokrouhlí a počítá částky p_n. Když mu mobil ukáže $p_{25} = 0{,}239 \cdot 10^{17}$, je celý bez sebe a spěchá oznámit svůj plán manželce. Ta je rozvážnější (a taky ví, že jméno Nešetřil nedostala manželova rodina náhodou) a udělá kontrolní výpočet na kalkulačce počítače, která pracuje s přesností 16 míst. Paní Nešetřilová provádí stejný výpočet a dostane $p_{25} = -0{,}365 \cdot 10^{10}$. Vyleká se a okamžitě telefonuje zpátky manželovi, že bankéři jsou vydřiduši a že jejich syn by po 25. narozeninách byl tak nanejvýš velkým dlužníkem, rozhodně ne boháčem. Pobouřená paní Nešetřilová manželovi vynadá a navrhuje banku žalovat pro klamavou reklamu. Pan Nešetřil se neodvažuje manželce odporovat, ale dřív, než podá žalobu, vezme tužku a papír a začne počítat v ruce. Vidí, že kalkulačkám věřit nelze. Když zjistí, že faktická částka, kterou by syn k narozeninám dostal, by byla kladná – něco kolem jedné koruny – koupí manželce iracionálně za posledních e korun kytku a spěchá domů, aby stihl vykoupat syna Bohuslava.

Pro čtenáře, který si bude provádět kontrolní výpočet, upřesněme, že uvedené částky p_{25} jsme získali na kalkulačce mobilu značky Motorola a na kalkulačce zabudované v příslušenství k Windows.

Rekonstruujme úvahy, které zkušený matematik pan Nešetřil s tužkou v ruce provedl. Číslo e je definováno pomocí posloupnosti $a_n = \left(1 + \frac{1}{n}\right)^n$. Tato posloupnost je rostoucí, tj. $a_n < a_{n+1}$, a právě reálné číslo, ke kterému se s rostoucím n hodnoty a_n přibližují, bylo na počest Eulera označeno e. Číslo e je limitou posloupnosti a_n, takže zapisujeme

$$\mathrm{e} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n .$$

Už samotný Euler věděl, že e lze vyjádřit i jako limitu jiné rostoucí posloupnosti, totiž posloupnosti

$$b_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{(n-1)!} + \frac{1}{n!}.$$

Tato posloupnost se pro naše účely bude hodit více. Pro zajímavost srovnejme prvních pár členů obou posloupností:

n	a_n	b_n
1	2	2
2	2,25	2,5
3	2,370 370 ...	2,666 666 ...
4	2,441 406 ...	2,708 333 ...
5	2,488 32	2,716 666 ...
$\vdots$		
9	2,581 174 ...	2,718 281 ...
$\vdots$		

Srovnáme-li hodnoty a_n a b_n v tabulce s uvedenou hodnotou e, vidíme, že posloupnost b_n se ke své limitě přibližuje rychleji*).

Obecně lze ukázat nerovnosti $a_n < b_n < \mathrm{e}$.

*) Např. b_9 se shoduje s číslem e na prvních 6 místech za desetinnou čárkou, zatímco a_9 na žádném.

Čtenáři, který by se chtěl ve větších podrobnostech věnovat číslu e, doporučujeme učebnici [2]. Pro ilustraci toho, jak lze od posloupnosti a_n dospět k posloupnosti b_n, odvodíme první ze dvou nerovností. Použijeme známou binomickou větu

$$(A+B)^n = A^n + \binom{n}{1}A^{n-1}B^1 + \binom{n}{2}A^{n-2}B^2 + \binom{n}{3}A^{n-3}B^3 + \\ + \dots + \binom{n}{n-1}A^1B^{n-1} + B^n,$$

kde koeficienty $\binom{n}{k}$ jsou kombinační čísla definovaná předpisem

$$\binom{n}{k} = \frac{n!}{(n-k)!k!},$$

nebo chcete-li, jsou to prvky z n-tého řádku Pascalova trojúhelníku. Binomickou větu použijeme na výpočet hodnoty a_n, kde za A dosadíme 1 a za B dosadíme $\frac{1}{n}$. Dostaneme

$$a_n = 1 + \binom{n}{1}\frac{1}{n} + \binom{n}{2}\frac{1}{n^2} + \binom{n}{3}\frac{1}{n^3} + \dots + \binom{n}{n-1}\frac{1}{n^{n-1}} + \frac{1}{n^n}.$$

Obecný tvar sčítance v předchozím součtu můžeme upravit:

$$\binom{n}{k}\frac{1}{n^k} = \frac{n!}{(n-k)!k!} \cdot \frac{1}{n^k} = \frac{1}{k!} \cdot \frac{n(n-1)(n-2)\dots(n-k+1)}{n^k}$$

Poněvadž každý člen součinu

$$\frac{n(n-1)(n-2)\dots(n-k+1)}{n^k} = \\ = \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \ldots \cdot \frac{n-k+2}{n} \cdot \frac{n-k+1}{n}$$

je nanejvýš jedna, lze odhadnout

$$\binom{n}{k}\frac{1}{n^k} \leq \frac{1}{k!}$$

a celkově

$$a_n < 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{(n-1)!} + \frac{1}{n!} = b_n.$$

Vzhledem ke zmíněné nerovnosti $a_n < b_n < \mathrm{e}$ je limita posloupnosti b_n rovna číslu e. Číslo e si tedy můžeme představit jako výsledek nekonečného součtu

$$\mathrm{e} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \cdots$$

Začněme nyní vyjadřovat částky p_n naspořené v bance:

$$p_1 = 1 \cdot (e - 1) = \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \cdots$$

$$p_2 = 2 \cdot (p_1 - 1) = 2\left(\frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \cdots\right)$$

$$p_3 = 3 \cdot (p_2 - 1) = 3 \cdot 2\left(\frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \cdots\right)$$

$$p_4 = 4 \cdot (p_3 - 1) = 4 \cdot 3 \cdot 2\left(\frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \frac{1}{7!} + \cdots\right)$$

$$p_5 = 5 \cdot (p_4 - 1) = 5!\left(\frac{1}{5!} + \frac{1}{6!} + \frac{1}{7!} + \frac{1}{8!} + \cdots\right)$$

$$\vdots$$

$$p_{24} = 24 \cdot (p_{23} - 1) = 24!\left(\frac{1}{24!} + \frac{1}{25!} + \frac{1}{26!} + \frac{1}{27!} + \cdots\right)$$

$$p_{25} = 25 \cdot (p_{24} - 1) = 25!\left(\frac{1}{25!} + \frac{1}{26!} + \frac{1}{27!} + \frac{1}{28!} + \cdots\right)$$

Je jasné, že částka naspořená po 25. roce musí být kladná a větší než jedna. Odhadněme, jak je opravdu veliká. Po zkrácení faktoriálů ve vyjádření p_{25} dostaneme

$$p_{25} = 1 + \frac{1}{26} + \frac{1}{26 \cdot 27} + \frac{1}{26 \cdot 27 \cdot 28} + \frac{1}{26 \cdot 27 \cdot 28 \cdot 29} + \cdots <$$
$$< 1 + \frac{1}{26} + \frac{1}{26 \cdot 26} + \frac{1}{26 \cdot 26 \cdot 26} + \frac{1}{26 \cdot 26 \cdot 26 \cdot 26} + \cdots.$$

Na pravé straně se objevil součet geometrické posloupnosti s kvocientem

$q = \frac{1}{26}$, viz [1]. Protože součet prvních N členů této posloupnosti je roven

$$\frac{1-q^N}{1-q}$$

a pro N blížící se k nekonečnu se naše q^N blíží k 0, dostaneme pro součet všech členů hodnotu

$$\frac{1}{1-q},$$

a tedy celkový odhad

$$p_{25} < \frac{1}{1-\frac{1}{26}} = \frac{26}{25} = 1{,}04.$$

Hodnotu p_{25} můžeme odhadnout zdola jednoduše součtem prvních dvou členů geometrické posloupnosti. Získáme tak dolní odhad

$$p_{25} > 1 + \frac{1}{26} = \frac{27}{26} = 1{,}038.$$

Vidíme, že skutečná výše uspořených peněz p_{25} se podstatně liší od výsledků, které dostaneme pomocí kalkulačky mobilu nebo kalkulačky počítače. Paradoxní je, že i když např. displeje různých mobilů umožňují pracovat se stejným počtem platných míst, můžeme dostat různé výsledky (zkuste si výpočet p_{25} i s mobily svých spolužáků). Jak je to možné? Můžeme vůbec důvěřovat výsledkům, které dostaneme pomocí výpočetní techniky?

Prvním krokem k pochopení zdánlivého paradoxu je zjištění, že zabudované programy pro násobení a sčítání jsou různé. Mohou se lišit v tom, jak a kdy zaokrouhlují – zda před převodem do dvojkové soustavy, ve které většina algoritmů pracuje, nebo až po tomto převodu. Vyskytují se i kalkulačky, které pracují pouze v desítkové soustavě. Při výpočtu mohou dále používat více platných cifer, než kolik se jich nakonec ukáže na displeji, atp. Zjistit, jakou zaokrouhlovací taktiku vlastně váš počítač používá, je téměř nemožné.

I našeho čtenáře jistě napadne, že v předchozím příkladu by k vyřešení celé záhady mohlo pomoci použití algoritmu, který by počítal prostě na více platných cifer. Při práci v pohyblivé desetinné čárce dnes samozřejmě existuje možnost měnit přesnost výpočtů velmi snadno. Pro naše potřeby jsme proto použili program *Maple*, který byl vyvinut na univerzitě v kanadském Waterloo (proto název *javorový* list – v angličtině *maple* leaf). V následující tabulce znamená symbol D počet platných míst, které jsme programu Maple předepsali pro provedení výpočtu:

D	p_{25}	D	p_{25}	D	p_{25}
17	$-0{,}54848 \cdot 10^9$	22	$-4457{,}9$	27	1,0189
18	$0{,}71967 \cdot 10^8$	23	195,37	28	1,0344
19	$-0{,}55884 \cdot 10^7$	24	40,262	29	1,0406
20	615990	25	$-6{,}2713$	30	1,0399
21	$-4457{,}9$	26	1,4842	31	1,0399

Z hodnot p_{25} napočítaných Maplem uvádíme v tabulce pouze prvních pět platných míst*).

Při pohledu na naši tabulku se musíme zeptat: Podle jakého obecného pravidla máme tedy *předem* zvolit počet platných míst pro výpočet, abychom se vyvarovali velkých chyb? Nebo naopak: Co se nám při dané přesnosti výpočtu vlastně na displeji počítače ukazuje, když ne správné hodnoty?

Označme tyto hodnoty z displeje $\tilde{p}_1, \tilde{p}_2, \tilde{p}_3, \ldots, \tilde{p}_{25}$ a zkusme odvodit, s jakou přesností by musel počítač pracovat, aby se zobrazená částka $\tilde{p}_{25}$ lišila od skutečné p_{25} o méně než korunu. Odchylku výsledku operace prováděné počítačem od výsledku operace, který bychom dostali při absolutně přesných výpočtech, budeme označovat písmenem ε a to bude mít v indexu pořadové číslo příslušné operace. Odchylka ε_n může být kladná nebo záporná podle toho, zda se zaokrouhlí nahoru nebo dolů. Tedy:

$$\tilde{p}_1 = \mathrm{e} - 1 + \varepsilon_1 = p_1 + \varepsilon_1$$

$$\tilde{p}_2 = 2(\tilde{p}_1 - 1) + \varepsilon_2 = 2(p_1 + \varepsilon_1 - 1) + \varepsilon_2 = p_2 + 2\varepsilon_1 + \varepsilon_2$$

$$\tilde{p}_3 = 3(\tilde{p}_2 - 1) + \varepsilon_3 = 3(p_2 + 2\varepsilon_1 + \varepsilon_2 - 1) + \varepsilon_3 = p_3 + 3 \cdot 2\varepsilon_1 + 3\varepsilon_2 + \varepsilon_3$$

$$\tilde{p}_4 = 4(\tilde{p}_3 - 1) + \varepsilon_4 = p_4 + 4 \cdot 3 \cdot 2\varepsilon_1 + 4 \cdot 3\varepsilon_2 + 4\varepsilon_3 + \varepsilon_4$$

$$\cdots\cdots\cdots$$

$$\tilde{p}_{25} = 25(\tilde{p}_{24} - 1) + \varepsilon_{25} =$$

*) Na první pohled překvapí, že pro volbu přesnosti výpočtu $D = 21$ a $D = 22$ dostaneme stejnou (a přesto nesprávnou) hodnotu p_{25}. Čtenář si snad vysvětlí tento jev, když se podívá na 21. číslici za desetinnou čárkou v zápisu čísla e.

$$= p_{25} + 25!\varepsilon_1 + \frac{25!}{2}\varepsilon_2 + \frac{25!}{2\cdot 3}\varepsilon_3 + \frac{25!}{2\cdot 3\cdot 4}\varepsilon_4 + \cdots + \varepsilon_{25}$$

Rozdíl mezi skutečnou hodnotou p_{25} a vypočítanou hodnotou $\tilde{p}_{25}$ je

$$\text{chyba} = 25!\left(\varepsilon_1 + \frac{1}{2!}\varepsilon_2 + \frac{1}{3!}\varepsilon_3 + \frac{1}{4!}\varepsilon_4 + \ldots + \frac{1}{25!}\varepsilon_{25}\right).$$

Kdybychom měli zaručeno, že velikosti ε nepřesáhnou hodnotu 10^{-N} (to lze docílit tím, že budeme zaokrouhlovat na N platných desetinných míst), bude chyba v absolutní hodnotě:

$$|\text{chyba}| \leq 25!\left(1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots + \frac{1}{25!}\right)10^{-N} < 25!\cdot 10^{-N}\mathrm{e}$$

Bude-li tedy N aspoň tak velké, že

$$25!\cdot 10^{-N}\mathrm{e} < 1,$$

budeme mít zaručeno, že chyba výsledku je menší než jedna. Nejmenší takové N je 26. Tedy pro naši úlohu je třeba počítat s 26 místy za desetinnou čárkou.

A co dodat nakonec? Naše úloha byla velice jednoduchá, v podstatě se v ní 25krát násobilo. Skutečné úlohy z praxe konstruktérů letadel nebo jaderných elektráren, tvůrců modelů pro předpovědi počasí nebo pro plánování letů do kosmu jsou daleko složitější. Číslo 10^{10} prováděných operací není nijak neobvyklý počet. Proto důsledná analýza numerických chyb je nezbytnou součástí každého špičkového softwarového produktu. Komerčně nabízené programy nám velice usnadňují práci, musíme si však být vědomi i jejich – ne vždy na první pohled zřejmých – omezení.

Traduje se, že věhlasný matematik Householder – zakladatel oboru numerická matematika – nikdy necestoval letadlem. Prý si byl až moc dobře vědom, kde všude mohlo při výpočtech spojených s konstrukcí letadel dojít k chybě.

Literatura

[1] Odvárko, O.: *Matematika pro gymnázia. Posloupnosti a řady.* Prometheus, Praha, 1995.

[2] Pošta, S., Pošta, P.: *Analýza v příkladech.* ČVUT, Praha, 2009, učební text.