PRESEK
List za mlade matematike, fizike, astronome in računalnikarje

ISSN 0351-6652
Letnik **25** (1997/1998)
Številka 3
Strani 162–166

Roman Drnovšek:

PRAŠTEVILSKI DVOJČKI IN PROCESOR PENTIUM

Ključne besede: računalništvo, matematika, Pentium, teorija števil, praštevilski dvojčki, Brunova vrsta.

Elektronska verzija: http://www.presek.si/25/1335-Drnovsek.pdf

PRAŠTEVILSKI DVOJČKI IN PROCESOR PENTIUM

Oktobra leta 1994 je Thomas Nicely, profesor matematike na Lynchburg Collegu v ameriški zvezni državi Virginia, na internetu objavil vest, da ima procesor Pentium težave pri deljenju nekaterih števil. Njegov računalnik je namreč izračunal obratni vrednosti števil 824 633 702 441 in 824 633 702 443 le na 9 decimalnih mest natančno, čeprav je procesorjev proizvajalec Intel jamčil za natančnost 19 decimalnih mest. Za Intelove strokovnjake je novica pomenila katastrofo, saj njihov procesor ni obvladal računske operacije, ki se je vsi naučimo že v osnovni šoli. Nedvomno je zanimivo poznati približno ozadje te zgodbe.

Procesor Pentium, ki je bil predstavljen leta 1993, je občutno povečal hitrost in zmogljivost osebnih računalnikov. Do jeseni 1994 je Intel prodal skoraj milijon primerkov, ki so jih IBM, Packard Bell in drugi vgradili v svoje osebne računalnike. Intel je zato upravičeno pričakoval velik prodajni uspeh. Potem pa ga je pretresla novica, da Pentium ni zanesljiv pri deljenju števil.

Pravzaprav noben računalnik realnih števil ne deli natančno. (Drugače je pri celoštevilskem deljenju, pri katerem sta celoštevilski kvocient in ostanek izračunana natančno.) Pri deljenju realnih števil – ta so v računalniku predstavljena v zapisu s premično piko – pogosto pride do majhne napake. To se sicer dogaja tudi pri seštevanju, odštevanju in množenju. Razlog za to napako je v tem, da računalnik ne računa z natančnimi števili, temveč z njihovimi približki, ki imajo natančnih le določeno število mest. Število $1/9 = 0.111111\ldots$ na primer ni shranjeno kot neskončni, temveč le kot končni niz enic. (V resnici je zapis števil v računalniku še nekoliko bolj zapleten, saj so števila zapisana v dvojiškem in ne v desetiškem sistemu.) Prav tako je rezultat vsake računske operacije zaokrožen na vnaprej določeno število mest. Vendar so te zaokrožitvene napake predvidljive. Znani so postopki, kako jih pri posameznih obsežnejših računih obvladamo. Kljub temu je rezultat računanja nezanesljiv, če je računalnik pri tem izvedel ogromno število računskih operacij. Če na primer seštejemo 10^{10} števil na računalniku, ki uporablja desetiško aritmetiko, potem je natančnost te vsote vprašljiva, saj se zaokrožitvene napake lahko seštevajo. Bralcu, ki bi se rad seznanil z osnovnimi problemi numeričnega računanja, priporočamo knjigo Z. Bohte: *Uvod v numerično računanje*, Knjižnica Sigma, DMFA Slovenije 1995.

Vendar so bile težave pri procesorju Pentium bistveno hujše. Nekatera števila (ki imajo v dvojiški predstavitvi, ki jo uporablja procesor, določeno obliko) je delil napačno. Oglejmo si tedaj najbolj znani primer

napačnega deljenja. Na 6 decimalnih mest natančna vrednost kvocienta 4 195 835/3 145 727 je 1.33382, medtem ko je Pentium izračunal 1.33374, kar pomeni 0.006% relativno napako. Ta primer lahko predstavimo še bolj impresivno. Če definiramo

$$x := 4\,195\,835\,, \quad y := 3\,145\,727\ (= 3 \cdot 2^{20} - 1) \quad \text{in} \quad z := x - (x/y) \cdot y\,,$$

je Pentium namesto $z = 0$ (oziroma števila zelo blizu 0) dobil $z = 256$ $(= 2^8)$. Vendar je bilo zelo malo verjetno, da bi kdo želel izračunati kvocient, ki bi ga Pentium izračunal napačno. Zato ni čudno, da Intelovi strokovnjaki pri testiranjih izdelka, preden so ga poslali na trg, te napake niso odkrili. Zanimivo pa je, da so poleti 1994 že vedeli za napako, vendar o tem niso obvestili lastnikov računalnikov. Ocenili so, da povprečnemu uporabniku preglednic (npr. Excel ali Quattro Pro) procesor izračuna napačen rezultat enkrat v 27 000 letih. Poleg tega je tudi malo verjetno, da bi kdo podvomil v rezultat. Kako pa je potem profesor Nicely odkril napako na procesorju?

Odgovor na to vprašanje je povezan s praštevilskimi dvojčki. Par števil $(p, p+2)$ imenujemo **praštevilski dvojček**, če sta p in $p+2$ praštevili. Prvih 5 praštevilskih dvojčkov je tako

$$(3,5)\,,\ (5,7)\,,\ (11,13)\,,\ (17,19)\,,\ (29,31)\,.$$

Dogovorimo se še za oznaki. Število vseh praštevil, ki ne presegajo realnega števila $x \geq 2$, označimo s $\pi(x)$, število vseh praštevilskih dvojčkov, od katerih manjše praštevilo ne presega realnega števila $x \geq 2$, pa zaznamujemo s $\pi_2(x)$. Tako je $\pi(10) = 4$, saj so 2, 3, 5 in 7 edina praštevila, ki ne presegajo 10. Iz zgornjega seznama praštevilskih dvojčkov pa razberemo, da je $\pi_2(6) = \pi_2(10) = 2$. Naslednja tabela prikazuje vrednosti funkcij π in π_2 za nekatere x, ki so potence števila 10:

x	$\pi(x)$	$\pi_2(x)$
10^2	25	8
10^3	168	35
10^4	1 229	205
10^5	9 592	1 224
10^6	78 498	8 169
10^7	664 579	58 980
10^8	5 761 455	440 312
10^9	50 847 534	3 424 506
10^{10}	455 052 511	27 412 679
10^{11}	4 118 054 813	224 376 048

Na tem mestu omenimo znameniti **praštevilski izrek**, ki pravi, da je $\pi(x)$ približno enako kvocientu $x/\ln x$ (glej npr. članek G. Pavlič: *Porazdelitev praštevil v množici* $\mathbb{N}$, Presek **24** (1996/97), štev. 3, str. 140–143). Praštevilski izrek je posledica naslednjih zanimivih neenakosti, odkritih leta 1962. Za vse $x \geq 2$ velja zgornja ocena

$$\pi(x) < \frac{x}{\ln x}\left(1 + \frac{3}{2\ln x}\right),$$

za $x \geq 59$ pa velja spodnja ocena

$$\pi(x) > \frac{x}{\ln x}\left(1 + \frac{1}{2\ln x}\right).$$

Ker že od Evklida naprej vemo, da je praštevil neskončno mnogo, se pojavi vprašanje, ali je tudi praštevilskih dvojčkov neskončno. Odgovor na to vprašanje še vedno ni znan, čeprav obstaja domneva, da je pritrdilen. Leta 1919 je norveški matematik Viggo Brun dokazal, da je

$$\pi_2(x) \leq \frac{100\,x}{(\ln x)^2}$$

za velika števila x. (Ta zgornja meja je bila pred leti znižana na $6x/(\ln x)^2$.) Toda nihče še ni našel (podobne) spodnje ocene za $\pi_2(x)$, iz katere bi sledilo, da je praštevilskih dvojčkov neskončno mnogo. Leta 1922 sta Hardy in Littlewood postavila domnevo, da je $\pi_2(x)$ približno proporcionalno kvocientu $x/(\ln x)^2$. Večina matematikov je prepričana, da je resnična. Zato kar nekaj rezultatov iz teorije števil predpostavlja veljavnost Hardy-Littlewoodove domneve.

Zgornji izrek pa ni edini rezultat o praštevilskih dvojčkih, ki ga je dokazal Viggo Brun. Izračunajmo obratne vrednosti vseh praštevilskih dvojčkov in jih seštejmo, torej sestavimo vrsto

$$\left(\frac{1}{3}+\frac{1}{5}\right)+\left(\frac{1}{5}+\frac{1}{7}\right)+\left(\frac{1}{11}+\frac{1}{13}\right)+\left(\frac{1}{17}+\frac{1}{19}\right)+\left(\frac{1}{29}+\frac{1}{31}\right)+\cdots.$$

To vrsto imenujemo **Brunova vrsta**. Če je praštevilskih dvojčkov končno mnogo, potem je Brunova vrsta običajna vsota s končno mnogo členi. Denimo sedaj, da je praštevilskih dvojčkov neskončno mnogo. Ali tedaj obstaja vsota Brunove vrste oziroma ali vrsta konvergira? O konvergentnih (neskončnih) vrstah je Presek že pisal (glej npr. članek A. Cedilnik: *Geometrijska in harmonična vrsta*, Presek **23** (1995/96), štev. 1, str. 40–45).

Viggo Brun je pokazal, da je njegova vrsta konvergentna. Če s (p_n, p_n+2) označimo n-ti praštevilski dvojček in uvedemo delne vsote

$$B_n := \left(\frac{1}{3}+\frac{1}{5}\right)+\left(\frac{1}{5}+\frac{1}{7}\right)+\left(\frac{1}{11}+\frac{1}{13}\right)+\ldots+\left(\frac{1}{p_n}+\frac{1}{p_n+2}\right),$$

potem to pomeni, da obstaja realno število B, od katerega se števila B_n z dovolj velikimi indeksi n razlikujejo tako malo, kot le hočemo. Zanimivo je, da to ne velja, če v zgornji vrsti vzamemo vsa praštevila in se torej ne omejimo le na praštevilske dvojčke. Vrsta

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{11}+\frac{1}{13}+\frac{1}{17}+\cdots$$

namreč ne konvergira, saj njene delne vsote sčasoma postanejo poljubno velike. (Radovedni bralec lahko najde dokaz tega dejstva npr. v knjigi H. Rademacher: *Lectures on Elementary Number Theory*, Blaisdell 1964 ali v knjigi G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, Oxford University Press, 1979 (izrek 19 stran 17).

Oceniti vsoto Brunove vrste je vse prej kot enostavno. Tudi če poznamo vse praštevilske dvojčke do milijarde in iz njih izračunamo delno vsoto Brunove vrste (to je število $B_{\pi_2(10^9)}$), ni lahko povedati, kako blizu števila B s tem pridemo. Videti je, da je brez predpostavke o resničnosti Hardy-Littlewoodove domneve težko odgovoriti na to vprašanje. Zato privzemimo, da je ta domneva resnična. Pri tem privzetku so matematiki izpeljali naslednjo trditev, ki jo le grobo opišimo. Ker pri n-ti delni vsoti B_n upoštevamo samo prvih n praštevilskih dvojčkov (zadnjega kot prej označimo s (p_n, p_n+2)), ji prištejemo člen, proporcionalen številu $(\ln p_n)^{-1}$. Tako dobimo približek števila B, katerega napaka je proporcionalna številu $(\sqrt{p_n}\,\ln p_n)^{-1}$. S pomočjo te trditve sta leta 1974 Daniel Shanks in John Wrench, potem ko sta določila vse praštevilske dvojčke izmed prvih dveh milijonov praštevil, za vsoto Brunove vrste dobila približek 1.90218. Za napako približka sta v svojem članku zapisala, da je kvečjemu $2\cdot 10^{-5}$. Seveda pri pogoju, da velja Hardy-Littlewoodova domneva. Podobno je leta 1976 Richard Brent določil vse praštevilske dvojčke do 10^{11} in za vsoto Brunove vrste dobil približek 1.9021605.

Leta 1993 se je problema ponovno lotil v začetku prispevka omenjeni Thomas Nicely. Odločil se je, da samo z uporabo osebnih računalnikov (torej brez superračunalnikov, ki so dostopni le nekaterim) določi še nadaljnjih nekaj decimalk števila B. Da bi se izognil vsem možnim

pastem, je Brunovo vsoto računal na dva načina in hkrati uporabljal več osebnih računalnikov. Na začetku je uporabljal 5 osebnih računalnikov, ki so imeli Intelov procesor 486. Marca 1994 je v računanje vključil še računalnik s procesorjem Pentium, s katerim je računanje potekalo precej hitreje. Kljub temu delo ni bilo opravljeno čez noč. Tudi ko je že vedel, da je z rezultati, ki jih je izračunal Pentium, nekaj narobe, je potreboval še nekaj časa, da je izoliral napako. Zaradi neznanega vzroka je njegov Pentium napačno izračunal obratni vrednosti praštevilskih dvojčkov 824 633 702 441 in 824 633 702 443. Takoj je o svojem odkritju obvestil Intelove strokovnjake. Ker ni dobil odgovora, je novico objavil na internetu, kjer je hitro pritegnila zanimanje drugih uporabnikov Intelovih procesorjev ter končno tudi medijev.

Ko je bila širša javnost obveščena o napaki, jo je priznal tudi Intel, vendar je vztrajal, da se skoraj nihče od imetnikov procesorja Pentium ne bo nikdar "srečal" s to napako. Zato je Intel sprva ponudil zamenjavo procesorjev le tistim, ki so lahko dokazali, da potrebujejo veliko natančnost pri zapletenih računanjih (npr. numerični matematiki). Temu je sledilo hudo ogorčenje javnosti, zato so bili pri Intelu na koncu prisiljeni zamenjati procesor vsakemu, ki je za to prosil. Kljub temu da je bil Intel zaradi tega spodrsljaja udeležen v mnogih šalah, je bila nadaljnja prodaja procesorjev Pentium zelo uspešna. Sredi leta 1995 je prodaja presegla 10 milijonov, s čimer je Intel brez težav pokril stroške, ki so nastali zaradi napake.

To ni bila niti prva niti zadnja napaka pri (osebnih) računalnikih. Vsekakor pa je bila ena od najbolj osupljivih. Ker je pri računalnikih (še posebej pa pri programski opremi) kar nekaj stvari, ki lahko gredo narobe, uporabniki namreč pričakujemo zanesljivost vsaj pri strojni opremi, ki opravlja osnovne računske operacije. Na srečo ta napaka v procesorju Pentium ni imela hujših posledic. Upamo lahko, da je še dodatno opozorila izdelovalce različnih naprav, ki so delno ali v celoti odvisne od računalnikov, na možnost napak s hujšimi posledicami.

Navkljub težavam je Thomas Nicely določil vse praštevilske dvojčke do 10^{14}, s pomočjo katerih je dobil približek

$$B \approx 1.9021605778\,.$$

Ta približek je seveda lahko napačen, če Hardy-Littlewoodova domneva ni resnična.

Roman Drnovšek