Testovanie prvočíselnosti I.

Kubo Kováč

19. apríla 2010

Οἱ πρῶτοι ἀριθμοὶ πλείους
εἰσὶ παντὸς τοῦ προτεθέντος
πλήθους πρώτων ἀριθμῶν.
Euklides

Prvočísla tu máme od čias antiky, keď Euklides z Alexandrie krásnym jednoduchým argumentom dokázal, že ich je nekonečne veľa: ak $p_1, p_2, \ldots, p_n$ sú prvočísla, tak číslo $N = p_1 \cdot p_2 \cdots p_n + 1$ má (ako každé iné číslo) nejakého prvočíselného deliteľa, ale žiadne z čísiel $p_1, \ldots, p_n$ číslo N nedelí (zvyšok je 1) – takže každý konečný zoznam prvočísiel je naprd (vieme k nemu vyrobiť číslo N, ktorého prvočíselný deliteľ na zozname chýba). Prvočísiel je nekonečne veľa.

V súčasnosti je najväčšie *známe* prvočíslo $2^{43\,112\,609} - 1$ (nájdené v auguste 2008) – má 12 978 189 cifier. A keďže ľudia majú radi výzvy a radi prekonávajú svoje možnosti, hľadajú prvočísla väčšie a väčšie. Pritom Euklides už dávno vedel, že táto zábavka im nejaký ten čas vydrží.

V tomto článku (a jeho pokračovaní) si povieme, ako sa dá zistiť, či nejaké číslo je, alebo nie je prvočíslo (a ako sa to dá urobiť čo možno najrýchlejšie). Na druhej strane, nebudeme hovoriť o tom, ako

1. testovať prvočísla „masovo“ – napr. nájsť všetky prvočísla v rozsahu od 1 po N – na to dal celkom úspešný recept Eratostenes z Kyrény;

2. rozložiť číslo na súčin prvočísel, resp. nájsť deliteľa nejakého čísla. Ak sa práve v tejto chvíli háčite: „Ako zistím, že číslo je zložené bez toho, aby som našiel nejakého deliteľa?“, ubezpečujem vás, že ste na správnom mieste a naozaj, v súčasnosti to vyzerá tak, že vieme rýchlo rozhodnúť, či je číslo zložené, alebo prvočíslo, ale ak je číslo zložené, nevieme vždy rýchlo nájsť nejakého deliteľa (táto úloha je obzvlášť ťažká, ak je číslo iba súčinom dvoch veľkých prvočísiel; to sa potom využíva v kryptológií – na tejto našej neschopnosti faktorizovať (rozkladať čísla na súčin prvočísiel) stojí napríklad šifrovací systém RSA).

Ako teda zistiť, či dané číslo n je, alebo nie je prvočíslo?

1 Z definície

Hovoríme, že číslo $p \in \mathbb{N}$, $p \geq 2$ je *prvočíslo*, ak nemá žiadneho netriviálneho deliteľa. Každé číslo $n \in \mathbb{N}$ je deliteľné jednotkou a samým sebou – týchto dvoch deliteľov označujeme za *triviálnych*; ak okrem nich n žiadneho iného deliteľa nemá, je to prvočíslo.

Potom je to ľahké – stačí vyskúšať všetky čísla od 2 po $n-1$ a ak žiadne z nich nedelí n, tak n je prvočíslo. Alebo zlepšovák: stačí vyskúšať čísla po $n-2$ (lebo $n-1$ nedelí n pre $n > 2$). Ešte lepšie: stačí skúšať dvojku a potom už iba nepárne čísla, pretože nepárne číslo nebude deliteľné párnym. V zásade vlastne stačí skúšať iba deliteľnosť prvočíslami (ak ovšem vieme, ktoré sú to): každé zložené číslo sa totiž dá rozložiť na súčin prvočísiel (a teda najmenší (netriviálny) deliteľ bude práve prvočíslo).

No a nakoniec brutálny zlepšovák: stačí skúšať iba po $\lfloor\sqrt{n}\rfloor$. Prečo? Nuž všetky delitele čísla n sa vyskytujú po dvojiciach – ak a je deliteľ, potom aj $b = n/a$ je deliteľ; $n = a \cdot b$. Napríklad $12 = 1 \cdot 12 = 2 \cdot 6 = 3 \cdot 4$; alebo $45 = 1 \cdot 45 = 3 \cdot 15 = 5 \cdot 9$. Všimnite si, že z každej dvojice je jedno číslo menšie (alebo rovné) odmocnine a druhé väčšie (alebo rovné) odmocnine. Naozaj, keby $n = a \cdot b$ a a aj b by boli obe väčšie ako $\sqrt{n}$, tak $a \cdot b > \sqrt{n} \cdot \sqrt{n} = n$, teda $a \cdot b > n$, čo je blbosť, keďže $a \cdot b = n$. (Rovnaký argument ukazuje, že a a b nemôžu byť obe menšie ako $\sqrt{n}$.) Z toho teda vyplýva, že ak je číslo n zložené, má deliteľa menšieho alebo rovného $\lfloor\sqrt{n}\rfloor$. Stačí teda hľadať deliteľov od 2 po $\lfloor\sqrt{n}\rfloor$. Napríklad ak chceme o nejakom čísle do 100 zistiť, či je prvočíslo, stačí otestovať, či je deliteľné 2, 3, 5, alebo 7. Ak ho žiadne z týchto čísiel nedelí, je to na 100% prvočíslo. Ak chceme o čísle do milión zistiť, či je prvočíslo, stačí testovať deliteľnosť prvočíslami do 1000 (tých je 168). Ak chceme otestovať 19 ciferné číslo, „stačí“ postupne predeliť prvočíslami do 10^{10} a tých „iba“ $455\,052\,511\ldots$ To už na dnešných počítačoch niekoľko sekúnd potrvá a to sme len pri 19-ciferných číslach. Takýmto spôsobom sa nedostaneme ani len k 30-ciferným číslam (pre porovnanie, v RSA sa používajú vyše 600-ciferné (2048-bitové) prvočísla).

Uvedený postup je pomalý, pretože musí kontrolovať *veľa* čísiel – dokonca exponenciálne veľa od dĺžky čísla n. To si nemôžeme dovoliť.

Čo teda môžeme robiť lepšie? Pokúsime sa nájsť nejakú charakterizáciu prvočísiel, ktorá by sa testovala jednoduchšie.

2 Malá Fermatova veta

Et cette proposition est généralement vraie
en toutes progressions et en tous nombres premiers;
de quoi je vous envoierois la démonstration,
si je n'appréhendois d'être trop long.
Pierre de Fermat

Francúzsky matematik (povolaním právnik) Pierre de Fermat je známy skôr svojou poznámkou na okraji Diofantovej knihy Arithmetica: „našiel som nádherný dôkaz, ale tento okraj je naň priúzky“ (reč je o takzvanej Veľkej Ferma-

tovej vete: $a^n + b^n = c^n$ nemá pre $n > 2$ riešenie v prirodzených číslach). My však budeme potrebovať jeho tzv. Malú Fermatovu vetu. Tá tvrdí, že ak p je prvočíslo, tak p delí číslo $a^p - a$ pre ľubovoľné $a \in \mathbb{Z}$. Inými slovami zvyšok po delení $a^p - a$ je nula, čo je ekvivalentné tvrdeniu, že čísla a^p a a dávajú po delení p rovnaký zvyšok. Napísané v modernej matematickej notácii: Ak p je prvočíslo, tak

$$p \setminus a^p - a, \qquad \text{resp. } a^p - a \equiv 0 \pmod{p}, \qquad \text{resp. } a^p \equiv a \pmod{p}$$

(uvedomte si, že všetky tri formulácie hovoria presne to isté).

Navyše, ak sú a a p nesúdeliteľné (teda a nie je násobok p-čka), platí

$$a^{p-1} \equiv 1 \pmod{p}. \tag{$*$}$$

Napríklad vezmime také prvočíslo $p = 47$ a $a = 2$. Platí:

$$2^{46} = 70368744177664 = 47 \times 1497207322929 + 1;$$

inými slovami, 2^{46} dáva zvyšok 1 po delení 47, teda naozaj $2^{47-1} \equiv 1 \pmod{47}$ a kongruencia $(*)$ platí pre $a = 2$ a $p = 47$. Podobne pre $a = 3$:

$$3^{46} = 8862938119652501095929 = 47 \times 188573151481968108424 + 1,$$

teda $3^{47-1} \equiv 1 \pmod{47}$. Funguje to[1]. No a takto to funguje s úplne každým prvočíslom p a každým číslom a (okrem násobkov p).

Poďme si Fermatovu vetu dokázať. A nech nežerem, ukážeme si hneď dva dôkazy. Najskôr si však uvedomme, že kongruenciu $a^{p-1} \equiv 1 \pmod{p}$ stačí dokázať pre a od 1 po $p-1$ (keďže pracujeme iba so zvyškami po delení p; ďalej by sa už zvyšky opakovali).

Dôkaz #1. Všimnime si, že keď zoberieme násobky čísla a, konkrétne

$$1 \times a, \qquad 2 \times a, \qquad 3 \times a, \qquad \ldots, \qquad (p-1) \times a$$

a pozrieme sa na ich zvyšky po delení p, dostaneme práve čísla $1, 2, 3, \ldots, (p-1)$ (v nejakom poradí). Napríklad vezmime $p = 7$ a $a = 5$; násobky päťky sú:

$$5, \qquad 10, \qquad 15, \qquad 20, \qquad 25, \qquad 30,$$

ak sa pozrieme na zvyšky po delení 7, dostaneme

$$5, \qquad 3, \qquad 1, \qquad 6, \qquad 4, \qquad 2,$$

teda naozaj sú to čísla od 1 po $p - 1$ (iba v inom poradí).

Prečo je to tak? Z jednoduchého dôvodu: Keby boli i-te a j-te zvyšky rovnaké, teda $i \cdot a \equiv j \cdot a \pmod{p}$, potom by p delilo rozdiel $i \cdot a - j \cdot a = a \cdot (i - j)$. Ale keďže p je prvočíslo a nedelí a, musí deliť $i - j$. Rozdiel $i - j$ však nie je

[1] ešte že máme tie kalkulačky

väčší ako $p-1$ a preto jediná možnosť, ako môže p deliť $i-j$ je, že $i-j=0$, inými slovami $i=j$. Pre $i \neq j$ sa teda $i \cdot a \not\equiv j \cdot a \pmod{p}$.

Všetky zvyšky sú teda rôzne, no a keďže žiadne z $i \cdot a$ nedáva zvyšok 0 (i ani a nie je deliteľné p), musia to byť práve čísla $1, 2, 3, \ldots, (p-1)$ (môžno v inom poradí). Ak teraz všetky zvyšky vynásobíme, dostaneme

$$\begin{aligned}(1 \cdot a) \times (2 \cdot a) \times (3 \cdot a) \times \cdots \times (p-1) \cdot a &\equiv 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (p-1) \pmod{p} \\ a^{p-1}(1 \cdot 2 \cdot \cdots \cdot (p-1)) &\equiv (1 \cdot 2 \cdot \cdots \cdot (p-1)) \pmod{p} \\ a^{p-1}(p-1)! &\equiv (p-1)! \pmod{p}\end{aligned}$$

Vydelením[2] $(p-1)!$ dostávame kýžený výsledok

$$a^{p-1} \equiv 1 \pmod{p}. \qquad \square$$

Dôkaz #2. Všetci poznáme binomickú vetu:

$$(a+b)^n = a^n + \binom{n}{1}a^{n-1}b^1 + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n-1}ab^{n-1} + b^n.$$

Keď však umocňujeme na p-tu (p prvočíslo) a pozeráme sa iba na zvyšok po delení p, výsledok je oveľa jednoduchší:

$$(a+b)^p \equiv a^p + b^p \pmod{p}. \qquad (\dagger)$$

Všetky prostredné členy $\binom{p}{1}, \binom{p}{2}, \ldots, \binom{p}{p-1}$ sú deliteľné p (teda dávajú zvyšok 0 po delení p) – tým pádom zmiznú a ostanú iba dva najkrajnejšie.

Dokážme, že p naozaj delí $\binom{p}{k}$ pre $0 < k < p$. Platí:

$$\binom{p}{k} = \frac{p!}{k!(p-k)!}.$$

V čitateli máme p a v menovateli sú všetky čísla menšie ako p, takže p sa nevykráti a $\binom{p}{k}$ je násobok p.

S pomocou (†) je už dôkaz ľahký; (†) sa dá totiž zjavne zovšeobecniť na

$$(a_1 + a_2 + \cdots + a_n)^p \equiv a_1^p + a_2^p + \cdots + a_n^p \pmod{p}.$$

Odtiaľ:

$$a^p = (\underbrace{1+1+\cdots+1}_{a})^p \equiv \underbrace{1^p + 1^p + \cdots + 1^p}_{a} = \underbrace{1+1+\cdots+1}_{a} = a \pmod{p} \qquad \square$$

[2]Treba si dať pozor, že pri kongruenciách vo všeobecnosti nemôžeme len tak deliť. V tomto prípade môžeme, keďže p a $(p-1)!$ sú nesúdeliteľné. Podrobnejší argument by vyzeral asi takto: $a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}$ je to isté ako povedať, že p delí $(a^{p-1}-1)(p-1)!$; ale keďže p nedelí $(p-1)!$, musí deliť $a^{p-1}-1$.

3 Frematov test

Ako sa dá Fermatova veta použiť na testovanie prvočíselnosti? Fermatova veta tvrdí, že každé prvočíslo má vlastnosť $(*)$; ak číslo n vlastnosť $(*)$ nemá, tak to *nemôže byť prvočíslo*! Napríklad

$$2^{48} = 281474976710656 = 49 \times 5744387279809 + 15,$$

teda $2^{49-1} \equiv 15 \pmod{49}$ a nie 1; číslo 49 nemá vlastnosť $(*)$, teda nemôže to byť prvočíslo.

Môžeme teda zostrojiť nasledovný test (tzv. Frematov test):

> Chceme zistiť, či n je prvočíslo. Zvolíme nejaké $1 < a < n$ a zistíme, či platí $(*)$ t.j., či
>
> $$a^{n-1} \equiv 1 \pmod{n}.$$
>
> Ak nie, vieme, že číslo n je zložené.

Pred chvíľou sme Fermatov test vykonali pre $n = 49$, $a = 2$ a zistili sme, že $2^{48} \not\equiv 1 \pmod{49}$ a preto 49 nie je prvočíslo. Všimnite si, že sme dokázali, že číslo 49 je zložené a pritom sme nenašli žiadneho jeho deliteľa! Prinajmenšom pozoruhodné, nie?

Iný príklad: Je 341 prvočíslo? Skúsme Fermatov test pre $a = 2$. Zistíme, že $2^{340} \equiv 1 \pmod{341}$. Číslo 341 teda prešlo Fermatovým testom pre $a = 2$; je to prvočíslo? Možno. Skúsme $a = 3$. Zistíme, že $3^{340} \equiv 56 \pmod{341}$; 341 teda neprešlo testom – nie je to prvočíslo! Naozaj $341 = 11 \times 31$ je súčinom dvoch prvočísiel. Číslo 3 v tomto prípade nazveme *svedok* – trojka totiž našu 341 usvedčila z toho, že je zložená. Zato dvojka nám v tomto prípade bohapusto „klamala“ – „presviedčala“ nás, že 340 je prvočíslo (veď prešlo testom), a pritom je to zložené číslo. Dvojku preto v tomto prípade nazveme *klamárom*.

Vo všeobecnosti, ak nejaké zložené číslo n *neprejde* Fermatovým testom, tak číslo a voláme *svedok* (a dosvedčí, že n je zložené). Na druhej strane, ak je n zložené a *napriek tomu prejde* Fermatovým testom, a voláme *klamár* a číslo n nazývame *pseudoprvočíslo* (pri báze a). Pseudoprvočíslom teda nazývame také zložené číslo, ktoré sa (pri niektorom a-čku) „tvári“ ako prvočíslo.

Ak v tomto momente zhruba chápeme Frematov test, naskytá sa viacero otázok:

- Ako otestujem, či $a^{n-1} \equiv 1 \pmod{n}$; nie je to príliš prácne? Vypočítať 2^{48} sa ešte s trochou námahy dá, ale neboli našou ambíciou aspoň 600-ciferné čísla spomínané v súvislosti s RSA? Ako umocníme 2 na 600-ciferné číslo (hoci s pomocou počítača)?
- Čo ak vyskúšam niekoľko a-čok, a zistím, že rovnosť (resp. kongruencia) platí? Znamená to, že n je prvočíslo? (Alebo sú všetci klamári?) Respektíve koľko a-čok mám asi vyskúšať, kým nájdem nejakého svedka a ktoré to sú? Stačí ak vyskúšam $a = 2$? Alebo $a = 2$ a $a = 3$? Alebo treba všetky a-čka medzi 1 a n?

Tak najprv prvá otázka: Ako vyhodnotiť $a^t \bmod n$ rýchlo? Naivné by bolo zobrať jednotku, t-krát ju vynásobiť a-čkom, výsledok vydeliť n-kom a zistiť zvyšok. To by sme sa ďaleko nedostali: Po prvé môže byť t také veľké (stačí 30-ciferné), že sa to nespočíta ani za miliardy rokov, po druhé a^t bude také obrovské, že sa nezmestí ani na disk. Ukážeme si preto dve finty, ako to spočítať rýchlo.

Zlepšovák #1. Ak chceme vypočítať iba zvyšok po delení číslom n, stačí stále (aj pri medzivýsledkoch) počítať iba so zvyškami modulo n, teda číslami menšími ako n. Ak niekedy počas výpočtu dostaneme číslo väčšie ako n, zoberieme zvyšok po delení n a pokračujeme s ním. Takto nám počas celého výpočtu nevzniknú obrovské čísla.

Napríklad ak počítame $2^{48} \bmod 49$, môžeme počítať asi takto:

2^1	2^2	2^3	2^4	2^5	2^6	2^7	2^8
2	4	8	16	32	$64 \equiv 15$	30	$60 \equiv 11$
2^9	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}	2^{15}	2^{16}
22	44	$88 \equiv 39$	$78 \equiv 29$	$58 \equiv 9$	18	36	$72 \equiv 23$
2^{17}	2^{18}	2^{19}	2^{20}	2^{21}	2^{22}	2^{23}	2^{24}
46	$92 \equiv 43$	$86 \equiv 37$	$74 \equiv 25$	$50 \equiv 1$	2	4	8
2^{25}	2^{26}	2^{27}	2^{28}	2^{29}	2^{30}	2^{31}	2^{32}
16	32	$64 \equiv 15$	30	$60 \equiv 11$	22	44	$88 \equiv 39$
2^{33}	2^{34}	2^{35}	2^{36}	2^{37}	2^{38}	2^{39}	2^{40}
$78 \equiv 29$	$58 \equiv 9$	18	36	$72 \equiv 23$	46	$92 \equiv 43$	$86 \equiv 37$
2^{41}	2^{42}	2^{43}	2^{44}	2^{45}	2^{46}	2^{47}	2^{48}
$74 \equiv 25$	$50 \equiv 1$	2	4	8	16	32	$64 \equiv 15$

A tak sme opäť dostali správny výsledok $2^{48} \equiv 15 \pmod{49}$, pričom sme nemuseli počítať $2^{48} = 281474976710656$ – počítali sme iba so zvyškami modulo 49. (Mimochodom, akonáhle sme vypočítali $2^{21} \bmod 49 = 1$, mohli sme skončiť – zvyšky sa totiž ďalej opakujú s periódou 21. Vďaka prvým 21 číslam v tabuľke vieme povedať zvyšok 2^k po delení 49 pre ľubovoľne veľké k – stačí vypočítať zvyšok k po delení 21 (perióda), a pozrieť sa na príslušné políčko do tabuľky; $48 = 2 \times 21 + 6$, takže $2^{48} \equiv 2^6 \equiv 15 \pmod{49}$.)

V predchádzajúcom výpočte pri umocnení na 48 sme použili 47 násobení. Nedalo by sa to lepšie? Odpoveď: Dalo!

Zlepšovák #2. Budeme postupne umocňovať na druhú (samozrejme, modulo 49). Ukážeme si to na príklade výpočtu $2^{48} \bmod 49$:

2^2	$(2^2)^2 = 2^4$	$(2^4)^2 = 2^8$	$(2^8)^2 = 2^{16}$	$(2^{16})^2 = 2^{32}$
4	$4^2 = 16$	$16^2 = 256 \equiv 11$	$11^2 = 121 \equiv 23$	$23^2 = 529 \equiv 39$

Keďže 48 sa dá napísať ako $32 + 16$, platí: $2^{48} = 2^{32+16} = 2^{32} \cdot 2^{16}$, čo je, podľa toho, čo sme už vypočítali, kongruentné $39 \cdot 23 = 897 \equiv 15 \pmod{49}$. Super, nie? Stačilo nám 6 násobení!

Vo všeobecnosti budeme jednoducho počítať (modulo n)

$$a, a^2, a^4, a^8, a^{16}, a^{32}, a^{64}, a^{128}, \ldots$$

(nasledujúce číslo je vždy predošlé umocnené na druhú). Každé číslo sa dá napísať ako súčet rôznych mocnín dvojky (teda ako súčet niektorých čísiel z postupnosti $1, 2, 4, 8, 16, 32, 64, 128, \ldots$); z takto vypočítaných zvyškov vieme teda poskladať a^t (stačí vynásobiť príslušné zvyšky). Napríklad $87 = 64+16+4+2+1$, teda $a^{87} = a^{64+16+4+2+1} = a^{64} \cdot a^{16} \cdot a^4 \cdot a^2 \cdot a^1$. Toto zjavne súvisí so zápisom čísla v dvojkovej sústave: $87 = (1010111)_2 = 1 \cdot 64 + 0 \cdot 32 + 1 \cdot 16 + 0 \cdot 8 + 1 \cdot 4 + 1 \cdot 2 + 1 \cdot 1$.

Iný pohľad na to isté je rekurzívny: ako vypočítame $a^t \pmod{n}$? Nuž ak je t párne, stačí vypočítať $a^{t/2}$ a toto číslo umocniť na druhú. Naopak, ak je t nepárne, tak $t - 1$ je párne; stačí teda vypočítať $a^{(t-1)/2}$, výsledok umocniť na druhú a vynásobiť a-čkom. Inými slovami: ak $t = 2k$ (párne t), tak $a^t = (a^k)^2$; ak $t = 2k + 1$ (nepárne t), tak $a^t = a \cdot (a^k)^2$.

Napríklad:

$$\begin{aligned} a^{37} &= a \cdot a^{36} = a \cdot \left(a^{18}\right)^2 = a \cdot \left(\left(a^9\right)^2\right)^2 = a \cdot \left(\left(a \cdot a^8\right)^2\right)^2 \\ &= a \cdot \left(\left(a \cdot \left(a^4\right)^2\right)^2\right)^2 = a \cdot \left(\left(a \cdot \left(\left(a^2\right)^2\right)^2\right)^2\right)^2 \end{aligned}$$

takže vidíme, že na výpočet a^{37} stačí 7 násobení.

Vo všeobecnosti na umocnenie a na t stačí rádovo $\mathcal{O}(\log t)$ násobení (a modulovaní). (Vďaka tomu vieme umocňovať aj na 600-ciferné t-éčka; hoci 2048-bitové číslo je obrovské, na umocnenie stačí 4096 násobení.)

Algoritmus 1 Implementácia umocňovania (mod n) v Pythone: `modexp (a, t, n)` vypočíta hodnotu $a^t \bmod n$. Program vľavo počíta a^t rekurzívne: najskôr vypočíta $a^{\lfloor t/2 \rfloor}$, výsledok umocní na druhú (prípadne vynásobí a). Program vpravo sa pozerá na jednotlivé cifry t v dvojkovej sústave a násobí príslušné mocniny a-čka. (Znak `%` znamená modulo.)

```
def modexp (a, t, n):
   if t == 0: return 1
   x = modexp (a, t/2, n)
   if t % 2 == 0:
      return (x * x) % n
   else:
      return (a * x * x) % n
```

```
def modexp (a, t, n):
   x = 1
   while t > 0:
      if t % 2 == 1:
         x = (x * a) % n
      t = t / 2
      a = (a * a) % n;
   return x
```

Ak v tomto momente vieme, ako efektívne vypočítať $a^{n-1} \bmod n$ a vykonať

Fermatov test, môžeme sa pokúsiť zodpovedať druhú otázku. Ktoré a-čka máme vo Fermatovom teste voliť? Má vôbec *každé* zložené číslo nejakého svedka?

☺ *Dobrá správa #1.* Áno má. Ak je číslo n zložené, potom má nejakého deliteľa d (iného ako 1 a n). My tvrdíme, že d je svedok. Prečo? Vykonajme Fermatov test pre d; môže sa stať, že n delí $d^{n-1}-1$? To sotva – odtiaľ by totiž vyplývalo, že aj d delí $d^{n-1}-1$ a to je blbosť: d predsa delí d^{n-1} a nemôže zároveň deliť aj číslo o 1 menšie.

Platí dokonca silnejšie tvrdenie: Ak a a n majú spoločného deliteľa $d>1$, tak a je svedkom zloženosti n:

$$d \setminus n, \qquad n \setminus a^{n-1}-1 \quad \Longrightarrow \quad d \setminus a^{n-1}-1$$

ale keďže $d \setminus a$, tak $d \setminus a^{n-1}$ a preto nemôže deliť $a^{n-1}-1$.

To je trochu upokojujúce: ak by sme vykonali Fermatov test pre všetky $1<a<n$, respektíve stačí pre $a<\lfloor\sqrt{n}\rfloor$, tak s istotou môžeme tvrdiť, či je číslo prvočíslo alebo nie: Ak prejde všetkými testami (pre všetky a), je to prvočíslo, ak čo i len jedným neprejde, je to zložené číslo.

Na druhej strane, toto nie je žiadny úspech. Ak chceme skúšať *všetky* a-čka od 2 po $\lfloor\sqrt{n}\rfloor$, tak to už môžeme rovno skúšať obyčajnú deliteľnosť tak, ako sme to robili v prvej kapitole. Nemusíme sa obťažovať nejakým umocňovaním na $n-1$. Ak chceme rýchly test prvočíselnosti, nemôžeme si dovoliť testovať $\sqrt{n}$ čísiel.

Nedalo by sa teda jednoducho zobrať malú „hŕstku“ a-čok, otestovať nimi n a ak prejde cez všetky testy, vyhlásiť n za prvočíslo?

Nestačilo by zobrať $a=2$? Odpoveď: Nestačilo. Už sme videli číslo 341, ktoré je pseudoprvočíslom pri báze 2 (teda prejde Fermatovým testom pre $a=2$; mimochodom 341 je najmenšie zložené číslo, pri ktorom dvojka zaklame). Dobre. Nestačilo by potom zobrať $a=3$? Odpoveď: Nestačilo. Číslo $91=7\times13$ je pseudoprvočíslo pri báze 3. Situácia je však oveľa horšia:

☹ *Zlá správa #1.* Jeden test nikdy nestačí. Pre každé $a\geq2$ existuje nekonečne veľa zložených čísiel, ktoré pre dané a prejdú testom. Inými slovami, existuje nekonečne veľa pseudoprvočísiel pri každej báze. Každé a-čko pri nekonečne veľa n-kách zaklame.

Dobre. A čo ak vezmeme dve bázy? Stačilo by skontrolovať, či n prejde Fermatovým testom pre $a=2$ a $a=3$? Nie. Situácia je dokonca oveľa horšia! Ukázali sme, že každé zložené číslo n má svedkov – všetky a-čka, ktoré sú s n súdeliteľné (majú spoločného deliteľa >1) sú svedkovia.

☹ *Zlá správa #2.* Existujú však zložené čísla, ktoré žiadnych iných svedkov nemajú! Tzn. existujú zložené čísla, ktoré prejdú Fermatovým testom pre všetky $a\geq2$ nesúdeliteľné s n. Napríklad $561=3\times11\times17$ je zložené číslo. Všetky násobky trojky, jedenástky, alebo sedemnástky to dosvedčia. Avšak nikto iný! Alebo také $n=999\,623\,179\,387\,201=25\,541\times102\,161\times383\,101$. Jediní svedkovia zloženosti sú násobky jeho prvočíselných deliteľov a tých je iba 51 531 547 200. Možno sa to nezdá málo, ale znamená to, že

1. Keby sme si vybrali náhodné a (medzi 1 a n), na svedka by sme natrafili s pravdepodobnosťou iba takmer 1 ku 20-tisíc (0.05 promile)!

2. Keby sme postupne volili $a = 2, 3, 4, \ldots$, tak najmenší svedok je 25 541.

Takéto čísla (ktorých jediní svedkovia s n-kom zdieľajú spoločného deliteľa) sa volajú Carmichaelove čísla. Carmichaelove čísla sú nočná mora pre Fermatov test – pre ne nie je Fermatov test o nič lepší ako naivný test skúšaním všetkých deliteľov.

Klincom do rakvi Fermatovmu testu je nasledujúce tvrdenie (ktoré pomerne nedávno, v roku 1994, dokázali Alford, Granville a Pomerance):

☹ *Zlá správa #3.* Carmichaelových čísiel je nekonečne veľa. (Dokonca vieme, že počet Carmichaelových čísiel $\leq x$ je pre dostatočne veľké x aspoň $x^{2/7}$.)

Fermatov test sa teda *nedá* zachrániť tým, že by sme do programu jednoducho „zadrátovali“ Carmichaelove čísla ako výnimky – takýto zoznam výnimiek by bol nekonečne dlhý.

To boli zlé správy. Teraz sa však pokúsime ukázať, že Fermatov test je oproti naivnému skúšaniu všetkých deliteľov predsalen obrovský pokrok. Ukážeme, že hoci úlohu: „dané je n, zisti, či je to prvočíslo“, jednoducho nerieši spoľahlivo (čo ak nám niekto podhodí práve obrovské Carmichaelove prvočíslo?). Za to je „v praxi“ dosť úspešný. Najmä ak nám n-ko nepodhodí nejaký záškodník, ale dajme tomu je to nejaké náhodné číslo.

☺ *Dobrá správa #2.* Hoci je Carmichaelových čísiel nekonečne veľa, nevyskytujú sa až tak často. Prvých 10 Carmichaelových čísiel je

$$561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, 15841, 29341.$$

V prvej 10-tisícke je teda iba 7 Carmichaelových čísiel. V prvej 100-tisícke je ich 16, v prvom milióne sa nachádza iba 43. Carmichaelových čísiel menších ako 10^{20} je 8 220 777. To znamená, že keby sme si vymysleli náhodné 19-ciferné číslo, šanca, že je Carmichaelovo je asi

$$0.000000000008\%.$$

Dokonca ani zložených čísiel, ktoré prejdú Fermatovým testom pre $a = 2$ nie je až tak veľa. (Tá dvojka neklame zas až tak často.) Existujú 3 také čísla menšie ako 1000 a iba 245 menších ako milión. Z prvých 10^{15} čísiel je 1 801 533 Fermatových pseudoprvočísiel pri základe 2. To znamená, že ak zvolíme náhodné 14-ciferné číslo, šanca, že natrafíme na zložené číslo, ktoré prejde Fermatovým testom pre $a = 2$ je

$$0.00000018\%.$$

Ak počet týchto čísiel porovnáme s počtom prvočísiel, zistíme, že ak zvolíme náhodné 14-ciferné číslo a to prejde Fermatovým testom pre $a = 2$, tak je to prvočíslo s pravdepodobnosťou

$$99.999994\%.$$

Fermatov test sa používa na generovanie náhodných prvočísiel napríklad v programe PGP (pretty good privacy). Zvolí sa náhodné číslo n, spraví sa zopár rýchlych testov (ktoré väčšinu zložených čísiel vylúčia); ak potom n prejde 4-mi kolami Fermatovho testu (cez 4 a-čka), našli sme n, ktoré je s vysokou pravdepodobnosťou prvočíslo. Ak nie, zvolíme iné n a pokračujeme v hľadaní.

Treba však pamätať na to, že týmto spôsobom *nedokážeme*, že číslo je skutočne prvočíslo. Pamätať na to treba najmä vtedy, keď číslo n nie je náhodné, ale niekto nám n podhodí.

☺ *Dobrá správa #3.* Ak zložené číslo n *nie je* Carmichaelove, tak aspoň polovica a-čiek ho usvedčí! To ale znamená, že ak nemáme tú strašnú smolu, že n je Carmichaelove (a teda pomocou Fermata ním aj tak veľmi nepohneme), môžeme vyskúšať niekoľko *náhodných* a-čiek ($1 < a < n$). Keďže svedkov je aspoň polovica, šanca, že natrafíme (náhodou) na klamára je menej ako 50%. Šanca, že natrafíme na klamára dvakrát je menej ako 25%; že trikrát, menej ako 12.5%. Je to podobné ako s hádzaním mincou: šanca, že mi 10-krát po sebe padne hlava je iba

$$\underbrace{\frac{1}{2} \times \frac{1}{2} \times \cdots \times \frac{1}{2}}_{10} = \frac{1}{2^{10}} = \frac{1}{1024},$$

menej ako promile. Rovnaká je šanca, že z 10 náhodných výberov 10-krát po sebe zrovna natrafím na klamára. Ak pokus zopakujem 50-krát (vždy s náhodným a-čkom), pričom n je zložené, ale nie Carmichaelove, potom šanca, že sme nenašli žiadneho svedka je menej ako $1/2^{50}$, čo je jedna ku 1 125 899 906 842 624 (slovom: mrňavá; oveľa oveľa menšia ako vyhrať jackpot v lotérii).

Ak teda n prejde 50 testami (s náhodne vyberanými a-čkami), môžeme ho pomaly vyhlásiť za prvočíslo. Pomýlime sa, iba ak máme strašnú smolu (a 50-krát natrafíme na klamára), alebo je to Carmichaelove číslo (teda máme strašnú smolu).

Algoritmus 2 Implementácia Fermatovho testu (50 pokusov s náhodným a-čkom). Ak číslo n nie je Carmichaelovo, funkcia vráti zlú hodnotu iba s pravdepodobnosťou menšou ako $1/2^{50}$.

```
from random import randint

def fermat (n):
   for i in xrange(0,50):
      a = randint(2, n-1)
      if modexp(a, n-1, n) != 1:
         return False              # zlozene
   return True                     # asi prvocislo
```

Na záver kapitoly si ešte našu Dobrú správu #3 dokážme: Ak n *nie je* Carmichaelovo, svedkov je aspoň toľko, čo klamárov (teda aspoň polovica).

Dôkaz. Ak n nie je Carmichaelove číslo, potom má svedka ω, ktorý je s n *nesúdeliteľný*. Nech

$$\alpha_1, \alpha_2, \ldots, \alpha_k$$

je zoznam všetkých klamárov (modulo n; pričom každý je na zozname *práve* raz). Tvrdíme, že potom

$$\omega \times \alpha_1,\ \omega \times \alpha_2,\ \ \ldots,\ \omega \times \alpha_k$$

sú všetko *rôzni* svedkovia. Odtiaľ už vyplýva, že svedkov je aspoň toľko, čo klamárov.

Ako vieme, že hore-uvedení svedkovia sú rôzni (samozrejme, modulo n)? Nuž ak $\omega\alpha_i \equiv \omega\alpha_j \pmod{n}$, tak n delí $\omega(\alpha_i - \alpha_j)$; tu využijeme, že n a ω sú nesúdeliteľné; odtiaľ potom vyplýva, že n delí $\alpha_i - \alpha_j$, teda $\alpha_i \equiv \alpha_j \pmod{n}$, ale to je možné iba vtedy, ak $i = j$. Preto ak $i \neq j$, tak $\omega\alpha_i \not\equiv \omega\alpha_j \pmod{n}$, teda uvedení svedkovia sú všetci rôzni. □