KLASICKÉ DÍLO V ČEŠTINĚ

AKTUALIZOVÁNO A PŘIZPŮSOBENO SOUČASNÝM PROGRAMÁTORSKÝM TRENDŮM

Umění programování

1. díl

Základní algoritmy

DONALD E. KNUTH

Donald E. Knuth

Umění programování
1. díl, Základní algoritmy

Computer Press, a.s.
Brno
2008

Umění programování
1. díl, Základní algoritmy

Donald E. Knuth

Computer Press, a.s., 2008. Vydání první.

Překlad: David Krásenský
Odborná korektura: Pavel Töpfer, Jan Kučera
Jazyková korektura: Alena Láníčková
Vnitřní úprava: Petr Sojka
Sazba: Petr Sojka
Rejstřík: David Krásenský, Petr Sojka
Obálka: Ivana Mitáčková

Komentář na zadní straně obálky: Radek Hylmar
Technická spolupráce: Jiří Matoušek, Petr Sojka, David Krásenský, Jiří Rybička, Petr Klíma
Odpovědný redaktor: Radek Hylmar
Technický redaktor: Jiří Matoušek
Produkce: Daniela Nečasová

Computer Press, a. s.,
Holandská 8, 639 00 Brno

Objednávky knih:
http://knihy.cpress.cz
distribuce@cpress.cz
tel.: 800 555 513

ISBN 978-80-251-2025-5
Prodejní kód: K1463
Vydalo nakladatelství Computer Press, a. s., jako svou 2992. publikaci.

Tuto sérii knih s láskou a vzpomínkami věnuji
počítači Typu 650, který byl kdysi instalován
na Case Institute of Technology
a se kterým jsem prožil mnoho nezapomenutelných večerů.

PŘEDMLUVA

Tady je knížka, o jejíž vydání jste nás žádali v tisících dopisů. Její sestavení nám trvalo mnoho let, kdy jsme prověřovali bezpočet různých receptů a vybírali z nich jen ty nejlepší, ty nejzajímavější a ty nejdokonalejší. Nyní se můžeme bez jediného stínu pochybnosti zaručit za jeden každý z nich: budete-li se doslova držet předepsaných postupů, dosáhnete přesně stejného výsledku jako my, přestože jste třeba nikdy dosud nevařili.

— *McCallova kuchařka* (1963)

Proces přípravy programů pro číslicové počítače je zvláště zajímavý, a to nejen pro své ekonomické a vědecké přínosy, ale také pro své estetické stránky, díky nimž se podobá skládání veršů nebo hudby. Tato kniha je prvním z několika svazků, ve kterých čtenář pozná různé dovednosti zdatného programátora.

Následující kapitoly *nejsou* ovšem pojaty jako úvod do programování počítačů; předpokládáme, že již čtenář v tomto smyslu jisté zkušenosti má. Požadavky ke studiu knihy jsou fakticky velmi jednoduché, ale začátečník musí věnovat jistou práci a čas základnímu seznámení s číslicovým počítačem. Čtenář této knihy by měl:

a) Mít nějakou představu o práci číslicového počítače s uloženým programem; není nutné znát jeho elektroniku, ale spíše vědět, jak se instrukce ukládají do paměti počítače a následně vykonávají.

b) Být schopen převést řešení problémů do natolik explicitních pojmů, že jim dokáže „porozumět“ i počítač. (Počítače totiž nemají žádný „zdravý rozum“; udělají přesně to, co jim řekneme, nic méně a nic více. Tento princip je při prvním setkání s počítačem nejobtížnější.)

c) Mít alespoň nějaké znalosti základních počítačových technik, jako je tvorba cyklů (opakované provádění jisté množiny instrukcí), volání podprogramů a práce s indexovanými proměnnými.

d) Znát alespoň trochu počítačovou hantýrku – vědět, co je „paměť“, „registry“, „bity“, „pohyblivá řádová čárka“, „přetečení“, „software“. K většině výrazů, které nejsou definovány přímo v textu, je uvedena stručná definice v rejstříku na konci svazku.

Tyto čtyři předpoklady bychom možná mohli shrnout do jediného požadavku, a sice že čtenář již napsal a otestoval dejme tomu nejméně čtyři programy pro nejméně jeden počítač.

Celou tuto sérii knih se snažím psát tak, aby sloužila několika účelům. Knihy jsou především referenčními publikacemi, které shrnují poznatky z několika důležitých oblastí. Knihy je také možné využít jako učebnice k samostatnému studiu nebo pro vysokoškolské kursy počítačů a informatiky. Pro splnění obou těchto cílů jsem do textu začlenil velké množství cvičení a většinu z nich jsem doplnil i o odpovědi. Mojí snahou bylo také zaplnit stránky fakty, a nikoli vágním, obecným výkladem.

Série knih je určena pro čtenáře, kteří se budou o počítače zajímat hlouběji, nemusí to ale nutně být počítačoví specialisté. Jedním z mých hlavních cílů bylo skutečně zpřístupnit tyto techniky programování i lidem, kteří pracují v jiných oblastech a mohou počítače vhodně využít, ale kteří na druhé straně nemají čas hledat všechny potřebné informace v různých odborných časopisech.

Téma těchto knih bychom mohli označit jako „nenumerickou analýzu“. Počítače bývaly tradičně spojovány s řešením různých numerických problémů, jako je výpočet kořenů rovnice, numerické interpolace a integrace atd., ale podobným tématům se zde nijak do hloubky nevěnujeme. Numerické programování počítačů je velice zajímavou a rychle se rozvíjející oblastí poznání a je mu věnována celá řada knih. Od začátku 60. let 20. století se ale počítače stále častěji používají i při řešení problémů, v nichž se s čísly pracuje téměř jen náhodou; zde již nevyužíváme schopnosti počítačů provádět aritmetické výpočty, ale schopnosti jejich rozhodování. Sčítání a odčítání i v těchto nenumerických oblastech přece jen využijeme, násobení a dělení již méně. Samozřejmě že i člověk, který se zajímá především o numerické programování počítačů, využije znalostí nenumerických technik, protože ty jsou zde uvedeny na pozadí numerických programů.

Výsledky výzkumů v oblasti nenumerické analýzy najdete v různých odborných časopisech. Mojí snahou bylo vyhledat v této rozsáhlé literatuře ty nejzákladnější techniky, které se při programování dají uplatnit v mnoha různých situacích. Pokusil jsem se tyto myšlenky sjednotit do jakési „teorie“, ale zároveň jsem se pokusil ukázat i její aplikaci na různé praktické problémy.

„Nenumerická analýza“ je pro tento obor samozřejmě neobyčejně nelichotivým názvem; mnohem vhodnější je charakterizovat jej pomocí výstižného pozitivního označení. Pojem „zpracování informací“ je pro uvažovanou publikaci příliš široký a „techniky programování“ jsou jako pojem naopak příliš úzké. Rád bych proto navrhl, aby se témata probíraná v těchto knihách označovala jako *analýza algoritmů*. Tento název by měl mít význam: „teorie vlastností určitých počítačových algoritmů“.

Celá série knih označená titulem *Umění programování* je vytvořena podle následující obecné osnovy:

Svazek 1. Základní algoritmy

Kapitola 1. Základní principy
Kapitola 2. Informační struktury

Svazek 2. Seminumerické algoritmy

Kapitola 3. Náhodná čísla

Svazek 4 hovoří o tak rozsáhlém tématu, že jej ve skutečnosti tvoří tři samostatné knihy (svazky 4A, AB a 4C). Připravuji také dva další svazky na speciální témata: Svazek 6, *Teorie jazyků* (Kapitola 11) a Svazek 7, *Kompilátory* (Kapitola 12).

Knihu s uvedenými kapitolami jsem začal psát v roce 1962, a to v jediném svazku, ale brzy jsem zjistil, že bude lépe se věnovat jednotlivým tématům do hloubky, než je jen přelétnout po povrchu. Výsledný rozsah textu znamená, že jedna každá z kapitol obsahuje víc než dost materiálu pro celý semestrální kurs na vysoké škole; má proto smysl vydat celou sérii v několika svazcích. Vím, že je neobvyklé mít v celé knize jen dvě kapitoly, ale rozhodl jsem se pro zachování původního číslování kapitol, a tím pádem pro snazší křížové odkazy. Plánuji také zkrácenou verzi svazků 1 až 5, která by měla sloužit jako obecnější reference a/nebo jako studijní materiál pro počítačové kursy řádného studia; jejím obsahem bude podmnožina materiálů z celé série a speciální informace budou vynechány. I ve zkrácené verzi bude zachováno stejné číslování kapitol jako v kompletním díle.

Tento svazek můžete považovat za jakýsi „průnik“ celé série, protože obsahuje základní materiál, který se používá v dalších svazcích. Zbývající svazky 2 až 5 se oproti tomu dají číst nezávisle na sobě. Svazek 1 je nejen referenční příručkou pro studium ostatních svazků, ale zároveň může sloužit pro studijní kursy nebo pro samostatné studium k tématu *datových struktur* (tím zdůrazňujeme látku kapitoly 2) nebo jako text z *diskrétní matematiky* (kde zdůrazňujeme látku kapitol 1.1, 1.2, 1.3.3 a 2.3.4) nebo jako výklad tématu *programování ve strojovém jazyce* (tentokrát klademe důraz na kapitoly 1.3 a 1.4).

Při psaní těchto kapitol jsem zvolil jiný přístup, než jaký přebírá většina současných knih o programování počítačů: nepokouším se naučit čtenáře pracovat se softwarem někoho jiného. Namísto toho chci, aby se naučil sám psát lepší software.

Mým původním cílem bylo přivést čtenáře k hranicím poznatků ve všech rozebíraných tématech. Držet krok s oborem, který je ekonomicky tak ziskový, je ale neobyčejně těžké, a vzhledem k rychlému rozmachu informatiky je tento sen doslova nemožný. Tento obor se stal pestrou mozaikou, do jejíchž desetitisíců

kamínků přispěly svými drobnými výsledky desetitisíce lidí po celém světě. Vzal jsem si proto nový cíl: soustředit se na „klasické“ techniky, které budou důležité i za několik desítek let, a popsat je tak, jak nejlépe umím. Zejména jsem se pokusil sledovat historii každého z témat a tím vytvořit pevné základy pro budoucí pokrok. Pokusil jsem se volit takovou terminologii, která je stručná a je v souladu se současným stavem. A pokusil jsem se do textu zahrnout všechny známé myšlenky o programování sekvenčních počítačů, které jsou krásné a dají se snadno vyslovit.

Nyní je na místě několik slov k matematickému obsahu celé série knih. Látka je uspořádána takovým způsobem, že ji dokáže strávit každý čtenář se znalostmi středoškolské algebry (snad jen přeskočí některé náročnější části); matematicky zdatný čtenář se zde dozví spoustu zajímavých matematických postupů souvisejících s diskrétní matematikou. Této dvojí úrovně výkladu jsem dosáhl jednak ohodnocením každého ze cvičení, mezi nimiž jsou zvlášť vyznačena cvičení matematického charakteru, a také díky uspořádání většiny částí, kdy jsou nejdůležitější matematické výsledky uvedeny *před* příslušným důkazem. Vlastní důkazy jsou buďto ponechány jako cvičení (odpovědi jsou shrnuty do samostatné části textu), nebo jsou uvedeny na konci výkladu.

Čtenář, kterého zajímá spíše programování než matematika, může po dosažení obtížnějších částí matematického výkladu zbytek příslušné části přeskočit. Na druhé straně matematicky orientovaný čtenář zde najde množství zajímavého materiálu. Významná část publikovaných matematických teorií k programování počítačů byla přitom chybná, a proto je jedním z úkolů této knížky nasměrovat čtenáře matematicky správným směrem. A protože se sám považuji za matematika, je mojí povinností udržovat v co největší možné míře také matematickou správnost textu.

Pro většinu matematického výkladu v těchto svazcích stačí základní znalosti diferenciálního a integrálního počtu, protože většina ostatních potřebných teorií je odvozena přímo v textu. Ve výkladu ale občas používám složitější konstrukce teorie funkcí komplexní proměnné, teorie pravděpodobnosti, teorie čísel a další; v takovém případě odkazuji čtenáře na příslušnou učebnici.

Nejtěžší rozhodnutí, jaké jsem musel při přípravě těchto svazků učinit, se týkalo způsobu výkladu různých postupů. Výhody vývojových diagramů a neformálního popisu algoritmů krok za krokem jsou dostatečně známé; podrobnější rozbor je uveden například v článku “Computer-Drawn Flowcharts” z časopisu ACM *Communications*, ročník 6 (září 1963), stránky 555–563. Při popisu jakéhokoli počítačového algoritmu je nicméně potřeba také formální, přesný jazyk, a proto jsem se musel rozhodnout, jestli pro tyto účely použít vyšší, algebraický jazyk jako ALGOL nebo FORTRAN, nebo strojově orientovaný jazyk. Mnozí z dnešních počítačových expertů nebudou zřejmě s tímto názorem souhlasit, ale z následujících důvodů jsem nakonec dospěl k definitivnímu přesvědčení, že je toto rozhodnutí správné:

a) Každý programátor je silně ovlivněn jazykem, ve kterém píše programy; lidé mají obrovskou tendenci preferovat takové konstrukce, které jsou v daném

jazyce nejjednodušší, nikoli ty, které jsou nejlepší pro daný počítač. Porozumí-li programátor strojově orientovanému jazyku, bude používat mnohem efektivnější metody a výrazně se přiblíží realitě.

b) Všechny programy, které společně napíšeme, jsou až na několik málo výjimek krátké, takže by je měl být schopen zpracovat prakticky jakýkoli vhodný počítač.

c) Jazyky vyšší úrovně jsou nevhodné pro diskusi různých důležitých otázek nízké úrovně, jako je linkování koprogramů, generování náhodných čísel, aritmetika s vícenásobnou přesností a řada problémů souvisejících s efektivním využíváním paměti.

d) Člověk s vážnějším zájmem o počítače by měl mít dobré znalosti strojového jazyka, protože ten je jednou z nejdůležitějších součástí počítače.

e) Výstupem mnoha softwarových programů, popsaných v řadě příkladů, bude tak jako tak nějaký kód ve strojovém jazyce.

f) Zhruba každých pět roků přicházejí a vycházejí z módy nějaké nové algebraické jazyky, zatímco já zde hovořím o základních, nadčasových principech.

Na druhé straně přiznávám, že v programovacích jazycích vyšší úrovně se programy píší o něco snáze a že se také výrazně snáze ladí. Zhruba od roku 1970 píšu své vlastní programy ve strojových jazycích nízké úrovně jen výjimečně, protože dnešní počítače jsou dostatečně velké a rychlé. My se ale v této knížce zaměřujeme na celou řadu problémů, u kterých má význam programování jako umění. Některé kombinatorické výpočty se například musí opakovat biliónkrát a každá mikrosekunda, o kterou zkrátíme vnitřní smyčku cyklu, znamená úsporu zhruba 11,6 dne výpočetního času. Podobně má smysl věnovat větší úsilí i při psaní jakéhokoli softwaru, který se bude používat mnohokrát denně v instalacích na mnoha počítačích, protože napsat jej musíme vždy pouze jednou.

Máme-li tedy za sebou rozhodnutí pro strojově orientovaný jazyk, který z jazyků to má být? Mohl jsme zvolit jazyk nějakého konkrétního počítače X, ale potom by si lidé, kteří k počítači X nemají přístup, mohli myslet, že je tato kniha jen pro majitele systému X. Navíc počítač X bude mít nejspíše spoustu různých zvláštností, které jsou vzhledem k látce uváděné v této knize naprosto irelevantní a které bychom si takto navíc museli vysvětlit; a nakonec výrobce počítače X přijde za dva roky s počítačem $X + 1$ nebo dokonce $10X$ a počítač X již nebude nikoho zajímat.

Tomuto dilematu jsem se vyhnul tím, že jsem se pokusil vytvořit „ideální“ počítač, který pracuje podle velmi jednoduchých pravidel (ta se můžete naučit dejme tomu za pouhou hodinu), ale který zároveň velice připomíná skutečné počítače. Není důvod, proč by se studenti měli bát učit se vlastnosti více než jednoho počítače; jakmile zvládnou jeden strojový jazyk, přizpůsobí se ostatním již snadno. Také profesionální programátor se ve své kariéře setkává s řadou různých strojových jazyků. Jedinou zbývající nevýhodou tohoto „mytického“ počítače tedy je, že programy pro něj napsané si jen velmi těžko vyzkoušíme.

Naštěstí to ale velký problém není, protože řada dobrovolníků napsala simulátory tohoto hypotetického počítače. Takovéto simulátory jsou ideální pro výuku, protože se s nimi pracuje ještě snáze než se skutečnými počítači.

Ke každému tématu jsem se pokusil citovat ty nejlepší první dokumenty a vybíral jsem také některé novější práce. V odkazech na literaturu používám standardní zkratky názvů periodik; výjimkou jsou nejčastěji citované časopisy, které zkracuji následujícím způsobem:

CACM = Communications of the Association for Computing Machinery
JACM = Journal of the Association for Computing Machinery
Comp. J. = The Computer Journal (British Computer Society)
Math. Comp. = Mathematics of Computation
AMM = American Mathematical Monthly
SICOMP = SIAM Journal on Computing
FOCS = IEEE Symposium on Foundations of Computer Science
SODA = ACM-SIAM Symposium on Discrete Algorithms
STOC = ACM Symposium on Theory of Computing
Crelle = Journal für die reine und angewandte Mathematik

Citaci uvedenou o dvě stránky zpět zkrátím tak například na zápis "*CACM* **6** (1963), 555–563". Výrazem "*CMath*" označuji dále knihu *Concrete Mathematics*, kterou cituji v úvodu do části 1.2.

Velká část odborného obsahu těchto knih je soustředěna do cvičení. Pokud jsou za nějakým netriviálním cvičením skryté cizí myšlenky, snažím se uvést jejich autora. Příslušné odkazy na literaturu jsou obvykle uvedeny v doprovodném textu dané části knihy nebo v odpovědi na cvičení, ale v řadě případů jsou cvičení založena na různém nepublikovaném materiálu, ke kterému ani nelze uvést další odkazy.

Při přípravě těchto knih mi samozřejmě po celou tu dobu pomáhala spousta jiných lidí, kterým jsem nesmírně zavázán. Děkuji především své ženě Jill za její nekonečnou trpělivost, za přípravu několika ilustrací a za nevýslovnou další pomoc. Děkuji také Robertu W. Floydovi, který v šedesátých letech výrazně pomohl zlepšit tento materiál. Svojí pomocí přispěly i tisíce jiných – sepsat jen jejich jména by zabralo celou další knížku! Mnozí z nich laskavě svolili k využití jejich dosud nepublikované práce. Moje výzkumné práce na Caltech a Stanfordu podporovala po dlouhá léta nadace National Science Foundation a Office of Naval Research. Vynikající pomoc a spolupráci mi poskytuje také vydavatelství Addison-Wesley, a to hned od zahájení projektu v roce 1962. Nejlépe ale všem poděkuji, když jim dnes mohu ukázat, že jsem s jejich podklady mohl napsat právě takovou knihu, jakou ode mne očekávali.

Předmluva ke třetímu vydání

Po deseti rocích vývoje systémů počítačové sazby TEX a METAFONT si dnes mohu splnit sen, se kterým jsem tyto práce začínal: mohu tyto systémy aplikovat

na celé dílo *Umění programování*. Přinejmenším celý text této knihy se nachází uvnitř mého osobního počítače v takové elektronické podobě, kterou bude možné snadno přizpůsobit i budoucím změnám v technologiích zobrazování a tisku. Díky novému uspořádání jsem do textu mohl začlenit doslova tisíce zlepšení, na která jsem už dlouho čekal.

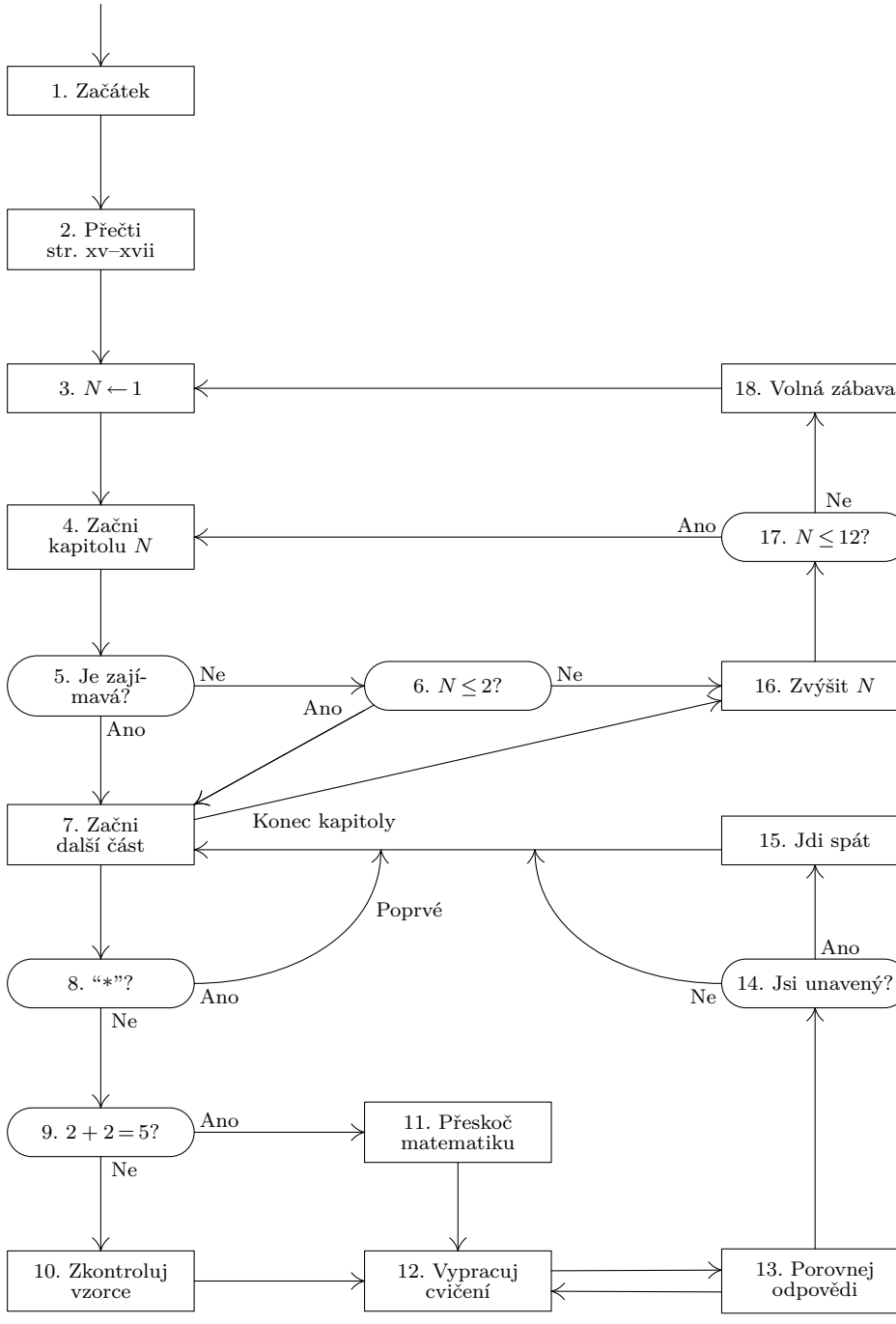
V tomto novém vydání jsem znovu prošel celý text slovo za slovem, pokusil jsem se zachovat mladickou nevázanost původních vět a zároveň do nich vnést kousek úsudku zralého muže. Přidal jsem desítky nových cvičení a k desítkám jiných jsem doplnil nové a rozšířené odpovědi.

Kniha *Umění programování* je však stále ve vývoji. U některých částí uvidíte proto symbol „ve výstavbě“, kterým se omlouvám, že daný materiál není zcela aktuální. Šuplíky mám plné důležitého materiálu, který chci začlenit do posledního, nejslavnějšího, čtvrtého vydání Svazku 1, ale nejprve musím dokončit svazky 4 a 5 a jejich vydání nechci odkládat více, než kolik je absolutně nezbytné.

Většinu náročných prací na přípravě nového vydání zajistili Phyllis Winkler a Silvio Levy, kteří zodpovědně přepsali a upravili text druhého vydání, a dále Jeffrey Oldham, jenž převedl téměř všechny původní ilustrace do formátu METAPOST. Opravil jsem veškeré chyby, na které mne čtenáři druhého vydání upozornili (a také několik chyb, kterých si nikdo nevšiml), a pokoušel jsem se nezanést nové chyby s novým materiálem. Přesto vím, že v knize nějaké závady být mohou, a každou z nich chci opravit co nejdříve. Za každou odbornou, typografickou nebo historickou chybu proto s radostí vyplatím 2,56 dolaru tomu, kdo ji nalezne jako první. Internetová stránka `http://www-cs-faculty.stanford.edu/~knuth/taocp.html` obsahuje informace o této knize (včetně aktuálního seznamu všech dosud oznámených chyb) a o související literatuře.

Stanford, California D. E. K.
duben 1997

Za poslední dvě desetiletí se věci změnily.
— BILL GATES (1995)



Vývojový diagram pro čtení této série knih.

Procedura pro čtení této série knih

1. Začněte číst tuto proceduru, pokud jste ji ještě nezačali číst. *Pečlivě pokračujte podle následujících kroků.* (Obecný formát procedury a doprovodného vývojového diagramu budeme používat i na jiných místech knížky.)

2. Přečtěte si Poznámky ke cvičením na stranách xv–xvii.

3. Přiřaďte N rovno 1.

4. Začněte číst kapitolu N. *Nečtěte* citáty uvedené na začátku kapitoly.

5. Je pro vás téma kapitoly zajímavé? Pokud ano, přejděte na krok 7, pokud ne, přejděte na krok 6.

6. Je $N \leq 2$? Pokud ne, přejděte na krok 16; pokud ano, projděte přesto text kapitoly. (Kapitoly 1 a 2 obsahují důležitý úvodní materiál a také přehled základních technik programování. Přinejmenším prolétněte část kapitoly o používaných notacích a o počítači `MIX`.)

7. Začněte se čtením další části kapitoly; pokud jste už ale došli na konec kapitoly, přejděte na krok 16.

8. Je číslo části kapitoly označeno znakem „*“? Pokud ano, můžete tuto část při prvním čtení vynechat (věnuje se jistému speciálnímu tématu, které je zajímavé, ale není úplně podstatné); vraťte se na krok 7.

9. Jste matematicky zdatný? Pokud je pro vás matematika „španělská vesnice“, přejděte na krok 11; jinak pokračujte krokem 10.

10. Zkontrolujte matematická odvození v této části kapitoly (a případné chyby oznamte autorovi). Přejděte na krok 12.

11. Pokud je tato část kapitoly plná matematických výpočtů, raději popsaná odvozování vůbec nečtěte. Seznamte se ale s nejdůležitějšími výsledky, které jsou obvykle uvedeny na začátku, nebo *šikmým písmem* naopak na konci obtížnější části.

12. Vypracujte doporučená cvičení k této části kapitoly, a to v souladu s pokyny uvedenými v Poznámkách ke cvičení (ty jste si přečetli v kroku 2).

13. Jakmile vypracujete cvičení ke své spokojenosti, porovnejte svou odpověď s odpovědí uvedenou v příslušném oddíle zadní části knihy (je-li k danému problému nějaká odpověď popsána). Přečtěte si také odpovědi ke cvičením,

na jejichž vypracování jste neměli čas. *Poznámka:* Ve většině případů je nejrozumnější si nejprve přečíst odpověď ke cvičení n a teprve poté začít pracovat na cvičení $n+1$, takže kroky 12–13 se obvykle provádějí souběžně.

14. Jste unaveni? Pokud ne, vraťte se na krok 7.

15. Jděte spát. Poté vstaňte a vraťte se na krok 7.

16. Zvyšte číslo N o jedničku. Pokud je $N = 3$, 5, 7, 9, 11 nebo 12, začněte číst další svazek z celé série.

17. Pokud je N menší nebo rovno 12, vraťte se na krok 4.

18. Blahopřejeme. Nyní se pokuste přesvědčit také své přátele, aby si koupili Svazek 1 a začali jej číst. Dále se vraťte na krok 3.

Běda tomu, který čte jen jednu knihu.
— GEORGE HERBERT, *Jacula Prudentum*, 1144 (1640)

Le défaut unique de tous les ouvrages
c'est d'être trop longs.
(Společným nedostatkem všech děl
je jejich přílišná délka.)
— VAUVENARGUES, *Réflexions*, 628 (1746)

Knihy jsou malicherné. Jen život je nádherný.
— THOMAS CARLYLE, *Journal* (1839)

POZNÁMKY KE CVIČENÍM

Cvičení v této sérii knih jsou vhodná jak k samostatnému studiu, tak i pro výuku. Naučit se nějaké téma jen s tím, že si něco přečteme, ale nepokusíme se uplatnit poznatky na konkrétní problémy, je velmi obtížné, ne-li nemožné, protože teprve při cvičení plně pochopíme, co všechno jsme se naučili. Navíc nejlépe se naučíme to, na co přijdeme sami. Proto tvoří cvičení velkou část tohoto díla; pokoušel jsem se v nich ale zachovat co nejvíce informativní hodnoty a současně volit takové problémy, které jsou zajímavé a poučné.

V mnoha knihách jsou lehká cvičení nahodile zamíchána mezi ta nejobtížnější. To bývá někdy nešťastné, protože čtenáři rádi dopředu vědí, kolik času jim řešení problému asi zabere – jinak mohou všechny problémy nakonec vynechat. Klasickým příkladem této situace je kniha *Dynamic Programming*, kterou napsal Richard Bellman; je to důležité, průkopnické dílo, v němž jsou ovšem problémy na konci některých kapitol sepsány pod jednotný nadpis "Exercises and Research Problems", tedy „Cvičení a výzkumné problémy"; zde jsou až krajně triviální otázky zařazeny přímo doprostřed složitých a dosud nevyřešených problémů. Říká se, že se prý jednou někdo Dr. Bellmana zeptal, jak pozná běžné cvičení od výzkumného problému; on tehdy odpověděl: „Pokud dokážete otázku vyřešit, znamená to, že je to cvičení; jinak je to výzkumný problém."

Dávat do knihy podobného druhu jak výzkumné problémy, tak i velmi snadná cvičení, je ovšem z řady důvodů rozumné; abych čtenáři ušetřil nepříjemné dilema, které otázky jsou které, označil jsem je *číselným hodnocením*, jež definuje úroveň obtížnosti. Číselné stupně mají následující obecný význam:

Hodnocení *Význam*

00 Velice snadné cvičení, na které může čtenář, jenž dobře porozuměl probírané látce, odpovědět okamžitě; takovéto cvičení stačí prakticky vždy udělat jen „z hlavy".

10 Jednoduchý problém, u něhož se musíte zamyslet nad přečtenou látkou, ale který ještě zdaleka není obtížný. Takovéto otázky byste měli zvládnout zhruba do jedné minuty a při řešení vám může pomoci tužka a papír.

20 Průměrně složitý problém, na kterém si vyzkoušíte základní porozumění probírané látce; k úplné odpovědi můžete ale potřebovat patnáct až dvacet minut.

30 Středně obtížný a/nebo složitý problém; k plně uspokojivé odpovědi se dopracujete i více než po dvou hodinách, případně po delší době, pokud u toho máte zapnutou televizi.

40 Dosti obtížný nebo rozsáhlý problém, který může být ve výuce vhodným námětem pro semestrální projekt. Student by měl být schopen problém vyřešit v rozumném čase, ale řešení není triviální.

50 Výzkumný problém, který dosud nebyl uspokojivě vyřešen (pokud bylo autorovi známo v době vzniku textu), i když se o jeho řešení mnozí pokoušeli. Pokud se vám podaří najít odpověď k takovémuto typu problému, rozhodně ji publikujte; také autora těchto stránek velice potěší, jestliže mu o řešení dáte co nejdříve vědět (za podmínky, že je řešení správné).

Interpolací této „logaritmické“ stupnice můžeme snadno vysvětlit i jiné číselné hodnocení. Výraz *17* bude například označovat cvičení, které je o něco více než průměrně jednoduché. Problémy s hodnocením *50*, které později nějaký čtenář vyřeší, mohou být v dalších vydáních knihy i v erratech na Internetu (viz strana iv) označeny číslem *45*.

Zbytek po dělení číselného hodnocení pěti vyjadřuje objem potřebné rutinní práce. Vyřešit cvičení označené číslem *24* může tedy trvat déle než cvičení s hodnocením *25*, ale ke cvičení s číslem *25* je potřeba více tvořivosti.

Autor se pokusil přiřadit číselné hodnocení jednotlivých cvičení co nejpřesněji, ale pro člověka, který nějaký problém vymyslí, je samozřejmě obtížné odhadnout, jak těžké bude hledat řešení pro někoho jiného; každý má navíc přirozené vlohy pro určité typy problémů a už ne pro jiné. Doufám, že je toto hodnocení dobrým odhadem obtížnosti; čísla berte ale jen jako rámcové vodítko, nikoli jako nějaký absolutní indikátor.

Knihu jsem psal pro čtenáře s různým stupněm matematického vzdělání a nadání; některá ze cvičení jsou proto určena jen pro matematicky zdatné studenty. Cvičení, k jehož řešení je potřeba více matematického myšlení či motivace, než kolik je běžné u lidí, kteří se zajímají jen o programování samotných algoritmů, má k hodnocení připojeno písmeno *M*. Pokud jsou ke cvičení nezbytné znalosti diferenciálního počtu nebo jiné vyšší matematiky, která není v této knize odvozena, je označeno písmeny „*VM*“. Samotné označení „*VM*“ ovšem *neznamená*, že by cvičení nutně muselo být obtížné.

Před některými cvičeními je uvedena šipka, „▸“; ta uvádí problémy, které jsou obzvláště názorné a jejichž řešení je možné čtenáři doporučit. Neočekávám samozřejmě, že každý čtenář či student vypracuje úplně *všechna* cvičení, a proto jsem takto označil ta nejcennější. (Tím vás ale nechci odradit od těch ostatních!) Každý čtenář by se ovšem měl přinejmenším pokusit o vyřešení všech problémů s hodnocením *10* a menším; podle šipek se může rozhodnout, kterým ze cvičení o vyšším hodnocení dá přednost.

Řešení většiny cvičení je uvedeno v sekci odpovědí. Ty prosím používejte s rozumem a na odpověď se nedívejte, pokud se nepokusíte problém poctivě vyřešit sami nebo pokud na vyřešení tohoto konkrétního problému skutečně nemáte čas. Až *poté*, co přijdete na svoje vlastní řešení, nebo se o to alespoň pokusíte, může pro vás být vzorová odpověď poučením a návodem. Řešení uvedená v knize jsou často velmi stručná a detailní postupy jsou postaveny na předpokladu, že jste

se nejprve poctivě pokusili problém vyřešit vlastními silami. Někdy se z řešení dozvíte méně informací, než kolik bylo v otázce požadováno, jindy naopak více. Je také docela možné, že se vám podaří najít lepší odpověď, než jaká je v knize uvedena; v tomto případě bude autor potěšen, pokud se s ním o řešení podělíte. V dalších vydáních knihy se pak podle možností objeví zdokonalené řešení spolu se jménem řešitele.

Při práci na cvičeních můžete obvykle využívat odpovědi na předchozí cvičení, pokud to není výslovně zakázáno. Také číselná hodnocení jsou definována s ohledem na toto pravidlo; je proto možné, že cvičení $n + 1$ bude mít nižší hodnocení než cvičení n, i když ve svém speciálním případě využívá výsledků cvičení n.

Shrnutí kódů:			
		00	Okamžitá odpověď
		10	Jednoduchá (do jedné minuty)
		20	Průměrná (čtvrt hodiny)
▸	Doporučené	*30*	Středně obtížná
M	Matematicky orientované	*40*	Semestrální projekt
VM	Vyžaduje „vyšší matematiku“	*50*	Výzkumný problém

CVIČENÍ

▸ **1.** [*00*] Co znamená hodnocení „*M20*“?

2. [*10*] Jakou hodnotu mohou mít cvičení v učebnici pro jejího čtenáře?

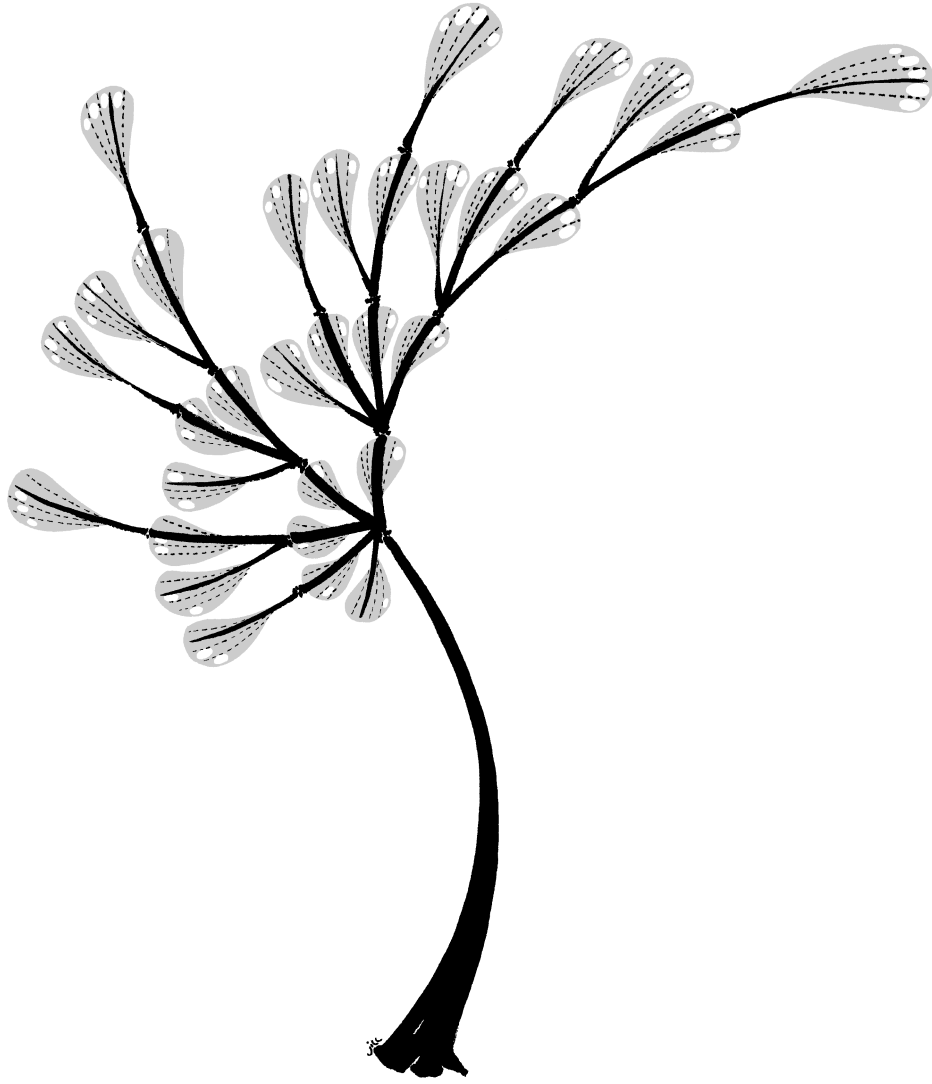
3. [*14*] Dokažte, že $13^3 = 2197$. Svoji odpověď zobecněte. [To je příklad hrozivého typu problémů, kterým se autor pokouší vyhýbat.]

4. [*VM45*] Dokažte, že pro žádné celé číslo n, kde $n > 2$, nemá rovnice $x^n + y^n = z^n$ žádné řešení v množině kladných celých čísel x, y, z.

Můžeme našemu problému čelit.
Můžeme takovéto skutečnosti zařídit
s náležitým řádem a postupem.

— HERCULE POIROT, v *Vražda v Orient Expressu* (1934)

OBSAH



KAPITOLA PRVNÍ

ZÁKLADNÍ PRINCIPY

Mnozí lidé, kteří nejsou zběhlí v matematických disciplínách,
se domnívají, že pokud je jeho úkolem [Babbageova Analytického Počítače]
dávat výsledky v číselné notaci, musí mít i jeho procesy aritmetickou
a číselnou povahu, nikoli algebraickou a analytickou. To je ovšem chyba.
Stroj dokáže uspořádat a zkombinovat svoje číselné veličiny
přesně stejně, jako by to byla písmena nebo jiné obecné symboly;
a ve skutečnosti může po přijetí odpovídajících opatření podávat výsledky
i v algebraické notaci.

— AUGUSTA ADA, hraběnka z Lovelace (1844)

Získej cvik, probůh, v malých věcech; potom pokračuj v těch velkých.

— EPIKTÉTOS (*Epiktétovy rozpravy* IV.i)

1.1. ALGORITMY

Slovo *algoritmus* je při programování počítačů natolik základním pojmem, že musíme začít jeho pečlivým rozborem.

Už samotné slovo „algoritmus“ je docela zajímavé: na první pohled to vypadá, jako by někdo chtěl napsat „logaritmus“, ale první čtyři písmena trochu zpřeházel. V anglickém slovníku *Webster's New World Dictionary* se toto slovo objevilo až v roce 1957; do té doby jsme zde nalezli jen jeho starší podobu „algorism“ s ještě starším významem, kterým bylo provádění aritmetických výpočtů s arabskými číslicemi. Za středověku totiž abakisté počítali na abaku a algoristé pomocí algorismů. V období renesance byl ale původ tohoto slova nejasný a první lingvisté se pokoušeli vysvětlit je různými kombinacemi, například *algiros* [bolestný]+*arithmos* [číslo]; jiní tvrdili, že slovo pochází od kastilského krále Algora. Nakonec ale historikové matematiky nalezli skutečný původ slova algorism: pochází ze jména slavného perského matematika a autora učebnic, který se jmenoval Abū ʻAbd Allāh Muḥammad ibn Mūsā al-Chwārizmī (c. 825), doslova „otec Abdulláha, Mohammeda, syn Mojžíšův, narozený v Chorezmu“. Aralskému moři ve Střední Asii se totiž kdysi říkalo jezero Chorezm (Chvárizm, v anglické transkripci Khwārizm) a oblast Chorezm se nachází v povodí řeky Amu-Darji jižně od tohoto moře. Al-Chwārizmī napsal slavný arabský text *Kitāb al-jabr wa'l-muqābala* („Pravidla pro odvozování a srovnávání“); ze samotného názvu knihy, která se systematicky věnovala řešení lineárních a kvadratických rovnic, pak pochází další důležité slovo, „algebra“. [Bližší pojednání o životě a díle al-Chwārizmīho najdete v článku H. Zemanek, *Lecture Notes in Computer Science* **122** (1981), 1–81.]

Tvar i význam původního slova *algorism* se ale postupně posunul a – jak vysvětluje slovník *Oxford English Dictionary* – slovo „prošlo množstvím pseudoetymologických zkomolení, mimo jiné nedávným slovem *algoritmus*, které se vědecky zaměňuje“ s řeckým původem slova *aritmetika*. Jestliže si uvědomíme, že lidé dávno zapomněli originální původ slova, není těžké tento posun od slova „algorism“ k „algoritmus“ pochopit. Jeden z prvních německých matematických slovníků, *Vollständiges mathematisches Lexicon* (Leipzig: 1747), definuje slovo *Algoritmus* následovně: „Pod tímto pojmem vyskytují se významy čtyř typů aritmetických výpočtů, jmenovitě sčítání, odčítání, násobení a dělení.“ Latinským výrazem *algorithmus infinitesimalis* se v té době označovaly „způsoby výpočtů s nekonečně malými veličinami, které vynalezl Leibniz“.

Do roku 1950 bylo slovo algoritmus spojováno nejčastěji s Euklidovým algoritmem, který slouží ke zjištění největšího společného dělitele dvou čísel a který je uveden v Euklidově díle *Základy* (Kniha 7, věta 1 a 2). Není od věci si nyní Euklidův algoritmus uvést:

Algoritmus E (*Euklidův algoritmus*). Jsou-li dána dvě kladná celá čísla m a n, najděte jejich *největšího společného dělitele*, tedy největší kladné celé číslo, kterým jsou beze zbytku dělitelná čísla m i n.

E1. [Nalezení zbytku.] Vydělte m číslem n a nechť r je zbytek. (Znamená to, že $0 \leq r < n$.)

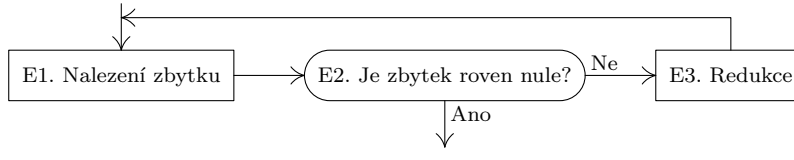
E2. [Je zbytek roven nule?] Pokud $r = 0$, algoritmus končí a n je odpověď.

E3. [Redukce.] Přiřaďte $m \leftarrow n$, $n \leftarrow r$, a vraťte se na krok E1. ∎

Euklides samozřejmě neuváděl svůj algoritmus v přesně tomto znění. Na tomto zápise si pouze ilustrujeme, jakým způsobem budeme uvádět i ostatní algoritmy v této knize.

Každý rozebíraný algoritmus zde dostane písmenné označení (v předcházejícím příkladu to bylo E) a jednotlivé kroky mají označení stejného písmene následovaného číslicí (E1, E2, E3). Kapitoly knihy jsou rozděleny do částí a v každé části stačí k identifikaci algoritmu jen jeho písmeno; pokud se ale odvoláváme na algoritmus z jiné části knihy, doplníme i číslo části. Nyní se například nacházíme v části 1.1, a proto ve stejné části knihy budeme Euklidův algoritmus nazývat jen jako Algoritmus E, zatímco v dalších částech to bude Algoritmus 1.1E.

Každý krok algoritmu, jako například krok E1, začíná krátkou frází v hranatých závorkách, která co nejstručněji vystihuje podstatný význam kroku. Stejnou frázi najdeme obvykle také v doprovodném *vývojovém diagramu*, jehož příklad je na obrázku 1, takže čtenář si činnost algoritmu může lépe představit.



Obr. 1. Vývojový diagram Algoritmu E

Za tímto stručným shrnutím následuje popis, který pomocí slov a symbolů vyjadřuje prováděnou *akci* či operaci nebo nějaké rozhodování. Mohou zde být uvedeny také *komentáře* v závorkách, jako je v kroku E1 druhá věta. Ty obsahují různé vysvětlující informace a často popisují jisté invarianty neboli neměnné vlastnosti proměnných, případně cíle, které je třeba v tomto kroku splnit. Nepopisují tak akce vlastního algoritmu, ale co nejvíce pomáhají čtenáři při jeho pochopení.

Šipka „$\leftarrow$“ v kroku E3 je ze všeho nejdůležitější operace, a sice *nahrazení*, někdy nazývaná také *přiřazení* nebo *substituce*: „$m \leftarrow n$“ znamená, že hodnotu proměnné m nahradíme aktuální hodnotou proměnné n. Při zahájení algoritmu E obsahují hodnoty proměnných m a n původně zadaná čísla, ale při jeho skončení již v nich budou obvykle jiné hodnoty. Šipka tak odlišuje operaci nahrazení od relace rovnosti: neříkáme „Nechť $m = n$“, ale mohli bychom se zeptat, „Je $m = n$?“ Znaménko „$=$“ označuje tedy podmínku, na kterou se můžeme ptát (testovat), zatímco „$\leftarrow$“ vyjadřuje akci (operaci), kterou je možné provést. Operaci *zvýšení n o jedničku* zapíšeme tudíž výrazem „$n \leftarrow n + 1$“ (čtete „n bude nahrazeno za $n + 1$“ nebo „do n zapíšeme $n + 1$“). Výraz „proměnná $\leftarrow$ $\leftarrow$ vzorec“ pak obvykle znamená, že vypočteme vzorec na základě současných hodnot zapsaných v proměnných a poté výsledkem nahradíme dosavadní hodnotu proměnné uvedené vlevo od šipky. Lidé bez znalostí práce s počítačem mívají sklony vyjadřovat operaci zvýšení n o jedničku větou „n bude rovno $n + 1$“ a zapisovat ji jako „$n \rightarrow n+1$“; tato symbolika vede ovšem jen ke zmatení, protože je v rozporu se standardními zvyklostmi, a proto ji nepoužívejte.

Všimněte si, že v kroku E3 je významné i pořadí jednotlivých akcí: „nechť $m \leftarrow n$, $n \leftarrow r$“ má naprosto jiný význam než „nechť $n \leftarrow r$, $m \leftarrow n$“, protože u druhého uvedeného zápisu se předchozí hodnota n ztratí, a nemůžeme ji tedy zapsat do m. Druhá sekvence operací je tedy ekvivalentní zápisu „nechť $n \leftarrow r$, $m \leftarrow r$“. Přiřazení stejné veličiny do několika proměnných můžeme vyjádřit několika šipkami po sobě a například místo „$n \leftarrow r$, $m \leftarrow r$“ zapsat „$n \leftarrow m \leftarrow$ $\leftarrow r$“. Výměnu hodnot dvou proměnných můžeme stručně zapsat „Vyměnit $m \leftrightarrow$ $\leftrightarrow n$“; stejou operaci můžeme definovat také pomocí nové proměnné t, se kterou napíšeme: „Nechť $t \leftarrow m$, $m \leftarrow n$, $n \leftarrow t$.“

Každý algoritmus začíná krokem o nejnižším čísle, obvykle krokem 1, a další kroky provádí v běžném sekvenčním pořadí, není-li řečeno jinak. V kroku E3 je například uveden příkaz „vraťte se na krok E1“, který zřejmým způsobem definuje pořadí kroků ve výpočtu. V kroku E2 je před operací uvedena podmínka „Pokud $r = 0$“; je-li tedy $r \neq 0$, zbytek příkazu již neplatí a žádná operace se neprovádí. Mohli bychom doplnit další větu, která je již ale redundantní: „Pokud $r \neq 0$, přejděte na krok E3.“

Silná vertikální čára „∎“ na konci kroku E3 označuje konec algoritmu, za kterým opět pokračuje běžný text.

Nyní jsme tedy rozebrali prakticky všechny konvence používané v této knize pro zápis algoritmů, jen s výjimkou notace indexovaných položek, které definují prvky uspořádaného pole. Dejme tomu, že máme n veličin, $v_1, v_2, \ldots, v_n$; namísto zápisu v_j budeme často j-tý element označovat pomocí notace $v[j]$. Podobně

zápis $a[i, j]$ nahrazuje někdy notaci s dvojitým indexem, a_{ij}. Proměnné bývají také označovány identifikátorem z několika písmen, například `TEMP` může být název proměnné pro dočasné uložení vypočtené hodnoty, `PRIME[K]` může být `K`-té prvočíslo apod.

Tolik tedy k *formě* algoritmů a nyní zkusíme jeden *provést*. Na tomto místě je ale třeba čtenáře upozornit, že se algoritmy *nedají* číst jako texty v krásné literatuře; při takovéto lehké četbě bychom těžko pochopili, o co v něm jde. Algoritmu musíme věřit a nejlépe jeho význam pochopíme tak, že si jej vyzkoušíme. Čtenář by si měl ke každému algoritmu vzít tužku a papír a zkusit si jej ihned projít. Obvykle bude načrtnut hotový funkční příklad, jindy si jej čtenář snadno sestaví sám. Je to jednoduchý a bezbolestný způsob, jak daný algoritmus pochopit; jakýkoli jiný postup bývá obvykle neúspěšný.

Projděme si tedy nyní příklad Algoritmu E. Nechť je dáno $m = 119$ a $n = 544$; můžeme tedy začít krokem 1. (Čtenář může laskavě sledovat činnost algoritmu, podle známého „škola hrou“.) Dělení m číslem n je v tomto případě jednoduché, snad až příliš jednoduché, protože neúplný podíl je roven nule a zbytek je 119. Máme tedy $r \leftarrow 119$. Pokračujeme krokem E2, a protože $r \neq 0$, neprovádíme žádnou operaci. V kroku E3 přiřadíme $m \leftarrow 544$, $n \leftarrow 119$. Je zřejmé, že pokud bylo původně $m < n$, bude první podíl v kroku E1 vždy nula, a algoritmus tak v dalším kroku fakticky vymění proměnné m a n, i když poněkud těžkopádným způsobem. Mohli bychom doplnit nový krok:

E0. [Zajistit $m \geq n$.] Pokud je $m < n$, vyměňte $m \leftrightarrow n$.

Pro samotný algoritmus by to neznamenalo žádnou podstatnou změnu, jen bychom ho tím mírně prodloužili a zhruba v polovině případů bychom zároveň zkrátili dobu jeho provádění.

Vrátíme-li se zpět do kroku E1, zjistíme, že $544/119 = 4 + 68/119$, takže $r \leftarrow 68$. Podmínka v E2 tak opět neplatí a v kroku E3 přiřadíme $m \leftarrow 119$, $n \leftarrow 68$. Při dalším průchodu dostáváme $r \leftarrow 51$ a následně $m \leftarrow 68$, $n \leftarrow 51$. Ještě jeden průchod dá výsledek $r \leftarrow 17$ a $m \leftarrow 51$, $n \leftarrow 17$. Nakonec dělíme 51 číslem 17 a konečně máme $r \leftarrow 0$, takže algoritmus v kroku E2 skončí. Největší společný dělitel čísel 119 a 544 je 17.

To je tedy algoritmus. Moderní význam slova algoritmus je hodně podobný výrazům jako *recept, proces, metoda, technika, procedura, postup, bláboly,* jen slovo „algoritmus“ znamená přece jen trošku něco jiného. Algoritmus je nejen konečnou množinou pravidel, která popisují posloupnost operací pro řešení jistého typu problémů, ale zároveň musí splňovat pět důležitých vlastností:

1) *Konečnost.* Algoritmus musí vždy po konečném počtu kroků skončit. Popsaný algoritmus E tuto podmínku splňuje, protože po provedení kroku E1 je hodnota proměnné r *menší* než n; pokud tedy $r \neq 0$, hodnota proměnné n se v příštím průchodu krokem E1 *sníží*. Každá klesající posloupnost kladných celých čísel musí jednou skončit, a proto se i krok E1 při každé dané hodnotě n provádí jen po konečný počet opakování. Poznamenejme ale, že počet kroků (průchodů) může být libovolně velký; při vhodně zvolených velkých hodnotách m a n se krok E1 může provést třeba více než milionkrát.

(Proceduru, která vykazuje všechny charakteristické vlastnosti algoritmu, ale která nemusí splňovat podmínku konečnosti, můžeme nazývat *výpočetní metoda*.Euklides původně vytvořil nejen algoritmus výpočtu největšího společného dělitele dvou čísel, ale také velmi podobný geometrický postup pro konstrukci „největší společné míry“ délek dvou úseček; pokud jsou ale délky obou zadaných úseček nesouměřitelné, tato výpočetní metoda neskončí. Jiným příkladem neukončené výpočetní metody je *reaktivní proces*, který nepřetržitě reaguje na podněty ze svého prostředí.)

2) *Určitost.* Každý krok algoritmu musí být přesně definován a pro každý případ v něm musí být s určitostí a jednoznačností popsány prováděné operace. Algoritmy v této knížce budou (doufejme) tuto podmínku splňovat, jsou ale napsány v přirozeném jazyce, takže se může stát, že čtenář neporozumí přesně záměrům autora. Pro překonání těchto obtíží s jednoznačnou interpretací se algoritmy zapisují ve formálně definovaných *programovacích jazycích* neboli *počítačových jazycích*, kde má každý příkaz přesně určený význam. Celá řada algoritmů bude v této knize popsána jak v běžném, tak i v počítačovém jazyce. Vyjádření jisté výpočetní metody v počítačovém jazyce je označováno jako *program*.

Kritérium určitosti u kroku E1 v algoritmu E vyjadřuje, že čtenář přesně rozumí, co znamená vydělit číslo m číslem n a stanovit zbytek. Pokud ale m a n nejsou kladná celá čísla, není přesný výklad této operace jednotný: jaký je zbytek po dělení -8 děleno $-\pi$? A jaký je zbytek $59/13$ děleno nulou? Kritérium určitosti zde tedy mimo jiné znamená, že musíme před provedením kroku E1 zkontrolovat, zda jsou hodnoty m a n vždy kladná celá čísla. Na začátku je tato podmínka splněna díky předpokladu; po provedení kroku E1 je již r určitě nezáporné celé číslo, a pokud se dostaneme na krok E3, musí být zároveň nenulové. Takže čísla m a n jsou skutečně při dělení vždy kladná a celá, jak potřebujeme.

3) *Vstup.* Každý algoritmus má nula nebo více *vstupů*: to jsou veličiny, které do algoritmu zadáme před jeho zahájením nebo které načteme dynamicky za běhu. Tyto vstupy se přebírají z určené množiny objektů. V algoritmu E máme například dva vstupy, jmenovitě m a n, a volíme je z množiny *kladných celých čísel*.

4) *Výstup.* Algoritmus má také jeden nebo více *výstupů*: to jsou veličiny, které mají zadaný vztah ke vstupům. Algoritmus E má jeden výstup, a to proměnnou n definovanou v kroku E2, která nabude hodnoty největšího společného dělitele obou vstupních veličin.

(Že je toto číslo skutečně největším společným dělitelem, to můžeme snadno *dokázat*; podívejme se na důkaz. Po provedení kroku E1 dostáváme

$$m = qn + r,$$

pro nějaké celé číslo q. Pokud $r = 0$, je m násobkem čísla n a v takovém případě je samozřejmě zároveň největším společným dělitelem čísel m a n. Jestliže $r \neq 0$, pak jakékoli číslo, které je dělitelem m a n, musí beze zbytku dělit také $m - qn = r$ a jakékoli číslo, které je dělitelem n a r, musí beze zbytku dělit $qn + r = m$; množina společných dělitelů čísel $\{m, n\}$ je tudíž stejná jako množina společných

dělitelů čísel $\{n, r\}$. Zejména pak *největší* společný dělitel čísel $\{m, n\}$ je stejný jako největší společný dělitel čísel $\{n, r\}$. Krok E3 proto nezmění odpověď na náš původní problém.)

5) *Efektivita.* Algoritmus by měl být zároveň *efektivním*, což znamená, že všechny jeho operace musí být v rozumné míře jednoduché, takže by je v principu měl být schopen přesně a za konečnou dobu provést kdokoli s tužkou a papírem. Algoritmus E používá jen operace dělení jednoho kladného celého čísla druhým, testování rovnosti celého čísla s nulou a přiřazení hodnoty jedné proměnné do druhé proměnné. Tyto operace jsou efektivní, protože celá čísla se dají na papíře vyjádřit konečným způsobem a protože existuje nejméně jedna metoda („algoritmus dělení“) pro dělení jednoho čísla druhým. Stejné operace ale *nebudou* efektivní, pokud za obě hodnoty zvolíme libovolná reálná čísla s nekonečným desetinným rozvojem nebo pokud místo nich zadáme délku úsečky (kterou také nelze přesně stanovit). Jiným příkladem neefektivního kroku může být: „Pokud je 4 největší celé číslo n, pro které existuje řešení rovnice $w^n + x^n + y^n = z^n$ v kladných celých číslech w, x, y a z, přejděte na krok E4.“ Takováto věta nemůže být efektivní operací, dokud někdo úspěšně nevytvoří algoritmus, kterým by zjistil, jestli 4 je, nebo není největším číslem s uvedenou vlastností.

Porovnejme nyní pojem algoritmu s receptem z kuchařské knihy. Každý recept je s nejvyšší pravděpodobností konečný (i když se říká, že hrnec, na který se jen díváme, nezačne vařit), má nějaké vstupy (vejce, mouku atd.) a také výstupy (večeře podle televize apod.), zároveň mu ale zoufale chybí určitost. V kuchařských knihách najdeme až příliš často různé neurčité instrukce: „Přidejte špetku soli.“ Taková „špetka“ je definována jako „méně než $1/8$ kávové lžičky“ a definice soli by měla být jasná každému; ale kam máme sůl přidat – nahoru? nebo na stranu? Podobné instrukce jako „zlehka promíchávejte, dokud není směs hladká“, nebo „ohřejte koňak v malé pánvi“ jsou dostatečným vysvětlením snad jen pro zkušeného šéfkuchaře; algoritmus musí být ale specifikován do takového stupně přesnosti, že podle něj může pracovat i počítač. I počítačový programátor se nicméně může studiem kuchařky hodně naučit. (Autor jen těžko odolával pokušení nazvat tento svazek „Kuchařkou programátora“; možná se jednoho dne pokusí napsat knihu s názvem „Algoritmy do kuchyně“.)

Nutno poznamenat, že samotná podmínka konečnosti není pro praktické použití dostatečně silná. Aby byl algoritmus užitečný, musí skončit nejen po konečném, ale po *velmi* konečném, tedy rozumném počtu kroků. Existuje například algoritmus, který zjistí, jestli bílý může šachovou partii vždy vyhrát, pokud neudělá chybu (viz cvičení 2.2.3–28). Tento algoritmus dokáže tedy vyřešit problém, o který se živě zajímají tisíce lidí na celém světě, ale přesto se můžeme v klidu vsadit, že za svého života odpověď nepoznáme; provedení algoritmu totiž trvá přímo neskutečné množství času, i když je sám konečný. Výklad na téma tak velkých čísel, která jsou již za hranicemi normálního chápání, najdete v kapitole 8.

V praxi ovšem požadujeme nejen nějaké algoritmy, ale takové algoritmy, které jsou v jakémsi volně definovaném estetickém smyslu *dobré*. Jedním z kri-

térií dobrého algoritmu je doba nezbytná k jeho provedení; můžeme ji vyjádřit také počtem opakování každého kroku. Dalším kritériem může být adaptabilita algoritmu pro různé typy počítačů, jeho jednoduchost a elegance atd.

Často je řešením jednoho stejného problému hned několik algoritmů, z nichž musíme vybrat ten nejlepší. Tím se dostáváme k jedné velice zajímavé a současně důležité oblasti *analýzy algoritmů*: je-li dán nějaký algoritmus, stanovte jeho výkonovou charakteristiku.

Uvažujme například znovu Euklidův algoritmus a předpokládejme tuto otázku: „Pokud je hodnota čísla n pevně daná, ale m může nabývat libovolného kladného celého čísla, jakým *průměrným* počtem opakování T_n projde v algoritmu E krok E1?“ Nejprve se musíme přesvědčit, jestli má tato otázka vůbec smysluplnou odpověď, protože se pokoušíme stanovit průměr z nekonečně mnoha hodnot m. Je ale zřejmé, že již po prvním provedení kroku E1 je podstatný pouze zbytek po dělení m číslem n. Pro zjištění čísla T_n nám tedy stačí vyzkoušet algoritmus pro $m = 1$, $m = 2$, …, $m = n$, spočítat celkový počet provedení kroku E1 a výsledek podělit číslem n.

Nyní přichází důležitá otázka stanovení *povahy* čísla T_n; bude přibližně rovno $\frac{1}{3}n$, nebo třeba $\sqrt{n}$? Stanovit odpověď na takovouto otázku je neobyčejně obtížný a fascinující matematický problém, který dosud není úplně vyřešen; podrobněji se mu věnujeme v části 4.5.3. Pro velké hodnoty čísla n je možné dokázat, že je T_n přibližně rovno $\big(12(\ln 2)/\pi^2\big)\ln n$, tedy že je přímo úměrné *přirozenému logaritmu* čísla n, přičemž násobící konstantu by sotva kdo dokázal přímo odhadnout! Podrobnější informace k Euklidově algoritmu a k dalším způsobům výpočtu největšího společného dělitele najdete v části 4.5.2.

Výrazem *analýza algoritmů* popisuje autor různá šetření podobného charakteru. Základní myšlenkou je vzít určitý algoritmus a stanovit jeho kvantitativní chování; příležitostně také zkoumáme, jestli daný algoritmus je nebo není v jistém slova smyslu optimální. *Teorie algoritmů* je zcela jiné téma, které se zabývá především existencí nebo neexistencí efektivního algoritmu pro výpočet určitých veličin.

Zatím byl náš výklad algoritmů poměrně málo přesný a matematicky orientovaný čtenář se může po právu domnívat, že předcházející text je jen velmi chatrným základem pro výstavbu jakékoli teorie algoritmů. Tuto část textu uzavřeme proto stručným popisem jisté metody, pomocí níž můžeme pojem algoritmu pevně zakotvit do sféry matematické teorie čísel. Definujeme proto formálně *výpočetní metodu* jako čtveřici (Q, I, Ω, f), kde Q je množina obsahující podmnožiny z I a Ω, a f je zobrazení z množiny Q do sebe sama. Navíc f musí být na podmnožině Ω identické, tedy $f(q)$ musí být rovno q pro všechna q z Ω. Uvedené čtyři veličiny Q, I, Ω, f po řadě reprezentují stavy výpočtu, vstupy, výstupy a výpočetní pravidla. Každý vstup x v množině I definuje následujícím způsobem *výpočetní posloupnost* $x_0, x_1, x_2, \ldots$:

$$x_0 = x \qquad \text{a} \qquad x_{k+1} = f(x_k) \quad \text{pro} \quad k \geq 0. \tag{1}$$

Říkáme, že výpočetní posloupnost *končí v k krocích*, pokud je k nejmenší celé číslo, pro které x_k náleží do Ω; zároveň říkáme, že z x je odvozen výstup x_k.

(Poznamenejme, že pokud x_k náleží do Ω, náleží sem i x_{k+1}, protože v takovém případě je $x_{k+1} = x_k$.) Některé výpočetní posloupnost nemusí nikdy skončit; *algoritmus* je pak taková výpočetní posloupnost, která pro všechna x z I skončí v konečně mnoha krocích.

Algoritmus E můžeme například formalizovat následujícím způsobem: nechť Q je množina všech jednoprvkových množin (n), všech uspořádaných dvojic (m,n) a všech uspořádaných čtveřic $(m,n,r,1)$, $(m,n,r,2)$ a $(m,n,p,3)$, kde m, n a p jsou kladná celá čísla a r je nezáporné celé číslo. Nechť I je podmnožina všech dvojic (m,n) a nechť Ω je podmnožina všech jednoprvkových množin (n). Dále nechť funkce f je definována následovně:

$$\begin{aligned} f\big((m,n)\big) &= (m,n,0,1); \qquad f\big((n)\big) = (n);\\ f\big((m,n,r,1)\big) &= (m,\ n,\ \text{zbytek z } m \text{ děleno } n,\ 2);\\ f\big((m,n,r,2)\big) &= (n) \quad \text{pokud} \quad r=0, \qquad (m,n,r,3) \quad \text{jinak};\\ f\big((m,n,p,3)\big) &= (n,p,p,1). \end{aligned} \tag{2}$$

Vzájemná korespondence mezi touto notací a algoritmem E je evidentní.

Součástí takto formulovaného algoritmu ještě ale není výše uvedená podmínka efektivity. Z množiny Q mohou například vyplývat nekonečné posloupnosti, které se nedají za pomoci tužky a papíru spočítat, nebo součástí funkce f mohou být operace, které normální smrtelník nezvládne. Pokud bychom chtěli omezit pojem algoritmu jen na elementární operace, mohli bychom pro Q, I, Ω a f vyslovit například následující omezující podmínky: Nechť A je konečná množina písmen a A^* je množina všech řetězců složených z prvků množiny A (tedy množina všech uspořádaných posloupností $x_1x_2\ldots x_n$, kde $n \geq 0$ a x_j náleží do A pro každé $1 \leq j \leq n$). Myšlenkou je zakódovat stavy výpočtu do reprezentací pomocí řetězců z A^*. Nyní uvažujme, že N je nezáporné celé číslo a že Q je množina všech (σ,j), kde σ náleží do A^* a j je celé číslo, $0 \leq j \leq N$; nechť I je podmnožina množiny Q pro $j=0$ a Ω je její podmnožina pro $j=N$. Pokud θ a σ jsou řetězce z A^*, říkáme, že se θ vyskytuje v σ, má-li σ formu $\alpha\theta\omega$ pro nějaké řetězce α a ω. Pro dokončení naší definice potřebujeme ještě funkci f, která bude následujícího typu, přičemž ji budou definovat řetězce θ_j, ϕ_j a celá čásla a_j, b_j pro jakékoli $0 \leq j < N$:

$$\begin{aligned} f\big((\sigma,j)\big) &= (\sigma, a_j) && \text{pokud se } \theta_j \text{ nevyskytuje v } \sigma;\\ f\big((\sigma,j)\big) &= (\alpha\phi_j\omega, b_j) && \text{pokud } \alpha \text{ je nejkratší možný řetězec, pro který } \sigma = \alpha\theta_j\omega;\\ f\big((\sigma,N)\big) &= (\sigma,N). \end{aligned} \tag{3}$$

Takováto výpočetní metoda je již zřetelně efektivní a zkušenosti ukazují, že je také dostatečně silná a umí provést cokoli, co zvládneme ručně. Pojem efektivní výpočetní metody je možné formulovat mnoha v podstatě ekvivalentními způsoby (například pomocí Turingova stroje). Výše uvedená formulace je prakticky shodná s definicí, kterou popsal A. A. Markov ve své knize *Teorie algoritmů* [*Trudy Mat. Inst. Akad. Nauk* **42** (1954), 1–376], kterou později upravil a rozšířil

N. M. Nagornyj (Moskva: Nauka, 1984; anglické vydání, Dordrecht: Kluwer, 1988).

CVIČENÍ

1. [*10*] V textu jsme si ukázali, jak pomocí notace s nahrazováním vyměnit hodnoty proměnných m a n, a sice v operacích $t \leftarrow m$, $m \leftarrow n$, $n \leftarrow t$. Ukažte, jak v posloupnosti operací nahrazení vyměnit hodnoty *čtyř* proměnných (a, b, c, d) na (b, c, d, a). Jinými slovy novou hodnotou proměnné a bude původní hodnota b atd. Pokuste se použít co nejmenší počet nahrazení.

2. [*15*] Dokažte, že na začátku kroku E1 je m vždy větší než n, s možnou výjimkou prvního provedení kroku.

3. [*20*] Upravte Algoritmus E a pro zlepšení jeho efektivity z něj odstraňte všechny triviální operace nahrazení jako „$m \leftarrow n$“. Tento nový algoritmus zapište v podobném formátu jako Algoritmus E a nazvěte jej Algoritmus F.

4. [*16*] Jaký je největší společný dělitel čísel 2166 a 6099?

▶ **5.** [*12*] Dokažte, že „Procedura pro čtení této série knih“, uvedená v předmluvě tohoto svazku, nesplňuje celé tři z pěti podmínek pro správný algoritmus! Uveďte také některé rozdíly mezi jejím formátem a zápisem Algoritmu E.

6. [*20*] Čemu se rovná T_5, tedy průměrný počet provedení kroku E1, při hodnotě $n = 5$?

▶ **7.** [*M21*] Dejme tomu, že m je dané a že n může nabývat libovolného kladného celého čísla; nechť dále U_m je průměrný počet provedení kroku E1 v Algoritmu E. Dokažte, že je veličina U_m dobře definována. Má U_m nějaký vztah k T_m?

8. [*M25*] Popište „efektivní“ formální algoritmus pro výpočet největšího společného dělitele dvou kladných celých čísel m a n, ve kterém určíte θ_j, ϕ_j, a_j, b_j jako ve vztazích (3). Nechť je vstup vyjádřen řetězcem $a^m b^n$, tedy m písmen a a za nimi n písmen b. Pokuste se o co nejjednodušší řešení. [*Nápověda:* Vyjděte z Algoritmu E, ale namísto dělení proveďte v kroku E1 přiřazení $r \leftarrow |m - n|$, $n \leftarrow \min(m, n)$.]

▶ **9.** [*M30*] Uvažujte dvě výpočetní metody $C_1 = (Q_1, I_1, \Omega_1, f_1)$, $C_2 = (Q_2, I_2, \Omega_2, f_2)$. Metoda C_1 může například označovat Algoritmus E ze vztahů (2), v němž je ale omezena velikost proměnných m a n, a metoda C_2 tvoří implementaci Algoritmu E v počítačovém programu. (To znamená, že Q_2 může být množina všech stavů počítače, tedy všech konfigurací paměti a registrů; f_2 bude pak definice jednotlivých strojových operací a I_2 bude počáteční stav počítače, jehož součástí je program pro stanovení největšího společného dělitele i hodnoty m a n.)

V jazyku teorie množin zformulujte definici tvrzení „C_2 je reprezentací C_1“ neboli „C_2 simuluje C_1“. Intuitivně můžeme tento výrok vysvětlit tak, že jakoukoli výpočetní posloupnost C_1 je možné napodobit i v C_2, pouze C_2 může při výpočtu provést větší počet kroků a může si ve svých stavech zapamatovat více informací. (Tak dostaneme přísnou interpretaci tvrzení „Program X je implementací Algoritmu Y“.)

1.2. MATEMATICKÉ ZÁKLADY

V TÉTO ČÁSTI TEXTU se budeme zabývat různými matematickými notacemi, které používáme v celé knize *Umění programování*, a odvodíme si několik základních vzorců, jež budeme opakovaně používat. I čtenář, který se o složitá matematická odvození hlouběji nezajímá, by se měl přinejmenším seznámit s *významem* různých vzorců, aby tak mohl využívat výsledky odvození.

Matematická notace se v této knize používá ke dvěma hlavním účelům: jednak pro popis částí algoritmu, jednak pro analýzu výkonových charakteristik algoritmu. Notace pro popis algoritmů je poměrně jednoduchá a vysvětlili jsme si ji již v předcházející části textu. Při analýze výkonnosti algoritmů budeme potřebovat další speciální notace.

Většina zde rozebíraných algoritmů je doplněna o různé matematické výpočty, které předurčují očekávanou rychlost jejich zpracování. Tyto výpočty čerpají snad ze všech možných odvětví matematiky a pro odvození všech matematických pojmů, které na různých místech výkladu používáme, bychom museli napsat samostatnou knížku. Drtivou většinu výpočtů je ale možné provést se znalostmi běžné vysokoškolské algebry a čtenář se základními znalostmi diferenciálního a integrálního počtu by měl porozumět prakticky veškeré matematické teorii. Někdy budeme potřebovat hlubší výsledky teorie komplexní proměnné, teorie grup, teorie čísel, teorie pravděpodobnosti atd.; v takovém případě bude příslušné téma vysvětleno pokud možno jednoduchým způsobem nebo budou uvedeny odkazy na jiné zdroje informací.

Matematické techniky, které jsou součástí analýzy algoritmů, jsou obvykle poněkud zvláštní. Často budeme například pracovat s konečnými součty racionálních čísel nebo s řešením rekurentních relací. Takováto témata dostávají v matematických kurzech zpravidla jen málo prostoru, a proto je smyslem následujících částí textu nejen důkladný dril používání definovaných notací, ale také podrobná ilustrace různých typů výpočtů a technik, které využijeme nejčastěji.

Důležitá poznámka: I když následující části textu podávají poměrně rozsáhlý výklad v matematických dovednostech, které jsou pro studium počítačových algoritmů nezbytné, většina čtenářů na první pohled nepochopí, jak silné jsou vazby mezi touto látkou a programováním počítačů (s výjimkou části 1.2.1). Čtenář může následující text pročíst skutečně pečlivě a důvěřovat tak autorovi, že je probíraná látka skutečně velmi důležitá, z motivačního hlediska je ale vhodnější *tuto část nejprve jen zběžně prolétnout a až později* (až uvidíte, kolik aplikací popisovaných technik se skrývá v dalších kapitolách) *se k ní vrátit a prostudovat ji důkladněji.* Pokud někdo do tohoto materiálu zabředne příliš hned napoprvé, nemusí se k tématům z programování počítačů už vůbec dostat! Čtenář by nicméně měl být přinejmenším obeznámen s celkovým obsahem těchto částí výkladu a měl by i při prvním čtení zkusit vypracovat některá ze cvičení. Zvláštní pozornost si zasluhuje část 1.2.10, ze které vychází většina později odvozeného teoretického materiálu. Část 1.3, která následuje po části 1.2, již rychle opouští království „čisté matematiky“ a vstupuje do říše „čistého programování počítačů“.

Rozšířený, volnější výklad velké části této látky najdete ve druhém vydání knihy *Concrete Mathematics*, kterou napsali Graham, Knuth a Patashnik, druhé vydání (Reading, Mass.: Addison-Wesley, 1994). Tuto knihu budeme v následných odkazech označovat jednoduše jako *CMath*.

1.2.1. Matematická indukce

Nechť $P(n)$ je nějaké tvrzení o celém čísle n; výrok $P(n)$ může být například „n krát $(n+3)$ je sudé číslo“ nebo „jestliže $n \geq 10$, pak $2^n > n^3$“. Dejme tomu, že nyní chceme *dokázat pravdivost tvrzení pro všechna kladná celá čísla* n. To můžeme udělat pomocí jednoho velmi důležitého postupu:

a) Dokažte, že tvrzení $P(1)$ je pravdivé.

b) Dokažte, že je pravdivé také tvrzení „jsou-li pravdivá všechna tvrzení $P(1)$, $P(2), \ldots, P(n)$, pak je pravdivé i tvrzení $P(n+1)$“, přičemž důkaz bude platný pro libovolné kladné celé číslo n.

Jako příklad uvažujme následující sérii rovností, na kterou od dávných dob přišla nezávisle na sobě spousta lidí:

$$\begin{aligned} 1 &= 1^2, \\ 1+3 &= 2^2, \\ 1+3+5 &= 3^2, \\ 1+3+5+7 &= 4^2, \\ 1+3+5+7+9 &= 5^2. \end{aligned} \tag{1}$$

Obecný vztah můžeme nyní formulovat takto:

$$1+3+\cdots+(2n-1) = n^2. \tag{2}$$

Nazvěme tento vztah tvrzením $P(n)$; chceme dokázat, že je $P(n)$ pravdivé pro všechna kladná celá n. Podle výše uvedeného postupu dokážeme:

a) „$P(1)$ je pravdivé, protože $1 = 1^2$.“

b) „Jsou-li pravdivá všechna tvrzení $P(1), \ldots, P(n)$, pak je pravdivé zejména i $P(n)$, takže platí vztah (2); jestliže k oběma stranám přičteme $2n+1$, dostaneme

$$1+3+\cdots+(2n-1)+(2n+1) = n^2+2n+1 = (n+1)^2,$$

čímž jsme dokázali, že je pravdivé i tvrzení $P(n+1)$.“

Tuto metodu můžeme považovat za *algoritmickou proceduru důkazu.* Pomocí následujícího algoritmu můžeme skutečně vytvořit důkaz tvrzení $P(n)$ pro libovolné kladné celé číslo n, pokud jsme ovšem již vypracovali kroky (a) a (b):

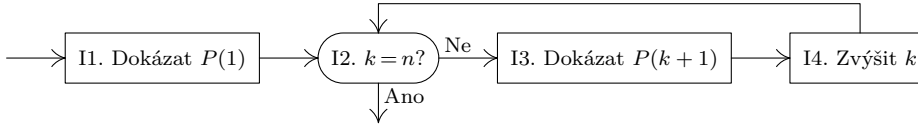
Algoritmus I (*Konstrukce důkazu*). Pro dané kladné celé číslo n vytvoří tento algoritmus důkaz, že je tvrzení $P(n)$ pravdivé.

I1. [Dokázat $P(1)$.] Přiřaďte $k \leftarrow 1$ a podle bodu (a) vypište důkaz tvrzení $P(1)$.

I2. [$k = n$?] Pokud $k = n$, algoritmus končí a výstupem je požadovaný důkaz.

I3. [Dokázat $P(k+1)$.] V souladu s bodem (b) vypište důkaz tvrzení: „Jsou-li pravdivá všechna tvrzení $P(1), \ldots, P(k)$, pak je pravdivé i tvrzení $P(k+1)$.“ Dále vypište větu: „Dokázali jsme již $P(1), \ldots, P(k)$; tudíž i $P(k+1)$ je pravdivé.“

I4. [Zvýšit k.] Zvětšete k o 1 a vraťte se na krok I2. ∎



Obr. 2. Algoritmus I: Matematická indukce

Protože tento algoritmus jasně předkládá důkaz tvrzení $P(n)$ pro libovolné dané n, je i technika důkazu složená z kroků (a) a (b) logicky platná. Nazýváme ji *důkaz metodou matematické indukce*.

Princip matematické indukce je ale třeba odlišit od obecné indukce jako metody vědeckého zkoumání. Vědec vychází z jistých pozorování a pomocí „indukce“ vytvoří obecnou teorii neboli hypotézu, která je vyjádřením těchto faktů; můžeme například pozorovat pět relací ve vztahu (1) a na jejich základě zformulovat hypotézu (2). V tomto slova smyslu není indukce ničím jiným než jakýmsi zdravým úsudkem z dané situace; matematik by ji ale označil za empirický výsledek nebo dohady.

Všechno si ozřejmíme na dalším příkladu. Nechť $p(n)$ označuje počet *rozkladů* čísla n, tedy počet různých způsobů, jakými je možné zapsat n ve formě součtu kladných celých čísel, nezávisle na jejich pořadí. Protože číslo 5 je možné takto rozložit právě sedmi způsoby, a sice

$$1+1+1+1+1 = 2+1+1+1 = 2+2+1 = 3+1+1 = 3+2 = 4+1 = 5,$$

dostáváme $p(5) = 7$. Několik prvních hodnot nahlédneme velice snadno:

$$p(1) = 1, \quad p(2) = 2, \quad p(3) = 3, \quad p(4) = 5, \quad p(5) = 7.$$

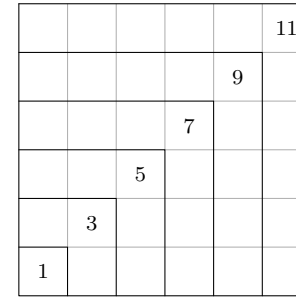
Na tomto místě bychom mohli podlehnout pokušení zformulovat na základě indukce hypotézu, že posloupnost $p(2), p(3), \ldots$ generuje *prvočísla*. Pro její ověření vypočteme $p(6)$, a ejhle, $p(6) = 11$, naše domněnka se potvrzuje!

[Další číslo $p(7)$ je ale bohužel rovno 15, takže celá hypotéza se nám hroutí a musíme začít znovu. O číslech $p(n)$ je známo, že jsou dosti komplikovaná, i když S. Ramanujanovi se podařilo o nich odhadnout a dokázat celou řadu pozoruhodných tvrzení. Další informace najdete v knize G. H. Hardy, *Ramanujan* (London: Cambridge University Press, 1940), kapitoly 6 a 8. Viz též část 7.2.1.4.]

Matematická indukce se od výše uvedené obecné indukce podstatně liší. Nepracuje s žádnými dohady, ale s přesvědčivým důkazem tvrzení; skutečně pomocí ní dokážeme hned nekonečně mnoho tvrzení, pro každé n jedno. „Indukcí“ je nazývána jen proto, že se nejprve musíme nějak rozhodnout, *co* budeme

dokazovat, a teprve *poté* ji můžeme aplikovat. V této knize budeme důkazy matematickou indukcí označovat právě pomocí samotného slova indukce.

Vztah (2) můžeme dokázat také geometricky. Na obrázku 3 vidíme pro $n = 6$ celkem n^2 buněk, rozdělených do skupin po $1 + 3 + \cdots + (2n - 1)$ buňkách. V závěrečné analýze můžeme ale tento obrázek pokládat za „důkaz“ jen tehdy, pokud ukážeme, že je možné tuto konstrukci provést pro všechna n, a tento postup je v podstatě stejný jako důkaz indukcí.



Obr. 3. Součet lichých čísel je vždy čtvercem

Při našem důkazu vztahu (2) jsme použili jen speciální případ (b); ukázali jsme pouze, že z pravdivosti tvrzení $P(n)$ vyplývá i pravdivost $P(n+1)$. To je důležitý jednoduchý případ, se kterým se budeme setkávat často, ale hned v dalším příkladu si ukážeme sílu celé metody ještě lépe. Nadefinujeme si *Fibonacciho posloupnost* F_0, F_1, F_2, ... podle pravidla, že $F_0 = 0$, $F_1 = 1$ a každé další číslo je součtem dvou předcházejících. Celá posloupnost začíná tudíž čísly 0, 1, 1, 2, 3, 5, 8, 13, ...; podrobněji ji budeme vyšetřovat v části 1.2.8. Nyní dokážeme, že pokud je ϕ rovno $(1 + \sqrt{5})/2$, platí

$$F_n \le \phi^{n-1} \tag{3}$$

pro každé kladné celé číslo n. Tento vztah označíme $P(n)$.

Jestliže $n = 1$, pak $F_1 = 1 = \phi^0 = \phi^{n-1}$, takže krok (a) je splněn. V kroku (b) si nejprve uvědomíme, že platí také $P(2)$, protože $F_2 = 1 < 1.6 < \phi^1 = \phi^{2-1}$. Nyní, pokud jsou pravdivá všechna tvrzení $P(1)$, $P(2)$, ..., $P(n)$ a pokud je $n > 1$, víme zejména, že je pravdivé $P(n-1)$ a $P(n)$, takže $F_{n-1} \le \phi^{n-2}$ a $F_n \le \phi^{n-1}$. Tyto dvě nerovnosti dále sečteme a dostaneme:

$$F_{n+1} = F_{n-1} + F_n \le \phi^{n-2} + \phi^{n-1} = \phi^{n-2}(1 + \phi). \tag{4}$$

Číslo ϕ má jednu důležitou vlastnost, pro kterou jsme si je ostatně v tomto příkladu vybrali:

$$1 + \phi = \phi^2. \tag{5}$$

Dosadíme-li nyní (5) do (4), dostáváme $F_{n+1} \le \phi^n$, což je tvrzení $P(n+1)$. Krok (b) jsme tedy zvládli a pomocí matematické indukce jsme dokázali i požadované tvrzení (3). Všimněte si, že krok (b) jsme zde prováděli dvěma různými způsoby: pro $n = 1$ jsme platnost tvrzení $P(n+1)$ dokázali *přímo*, zatímco pro $n > 1$ jsme již využili induktivní metodu. Tento postup byl nutný proto, že při $n = 1$ bychom se odvolávali na tvrzení $P(n-1) = P(0)$, které není definováno.

Pomocí matematické indukce můžeme také dokazovat různé vlastnosti *algoritmů*. Uvažujme následující zobecnění Euklidova algoritmu:

Algoritmus E (*Rozšířený Euklidův algoritmus*). Jsou-li dána dvě kladná celá čísla m a n, vypočteme jejich největšího společného dělitele a současně dvě celá, nikoli však nutně kladná čísla a a b, pro která platí $am + bn = d$.

E1. [Inicializace.] Přiřaďte $a' \leftarrow b \leftarrow 1$, $a \leftarrow b' \leftarrow 0$, $c \leftarrow m$, $d \leftarrow n$.

E2. [Dělení.] Nechť q a r budou neúplný podíl a zbytek (v tomto pořadí) při dělení c číslem d. (To znamená, že $c = qd + r$ a $0 \le r < d$.)

E3. [Je zbytek nulový?] Pokud $r = 0$, algoritmus končí; v tomto případě je $am + bn = d$, jak jsme potřebovali.

E4. [Nový cyklus.] Přiřaďte $c \leftarrow d$, $d \leftarrow r$, $t \leftarrow a'$, $a' \leftarrow a$, $a \leftarrow t - qa$, $t \leftarrow b'$, $b' \leftarrow b$, $b \leftarrow t - qb$ a jděte zpět na krok E2. ∎

Jestliže z tohoto algoritmu vypustíme proměnné a, b, a' a b' a jestliže namísto pomocných proměnných c a d použijeme m a n, dostáváme náš starý známý algoritmus 1.1E. Nová verze umí ovšem o něco více, protože stanovuje koeficienty a a b. Dejme tomu, že $m = 1769$ a $n = 551$; z algoritmu postupně dostáváme (vždy po kroku E2):

a'	a	b'	b	c	d	q	r
1	0	0	1	1769	551	3	116
0	1	1	−3	551	116	4	87
1	−4	−3	13	116	87	1	29
−4	5	13	−16	87	29	3	0

Odpověď je správná: $5 \times 1769 - 16 \times 551 = 8845 - 8816 = 29$, což je největší společný dělitel čísel 1769 a 551.

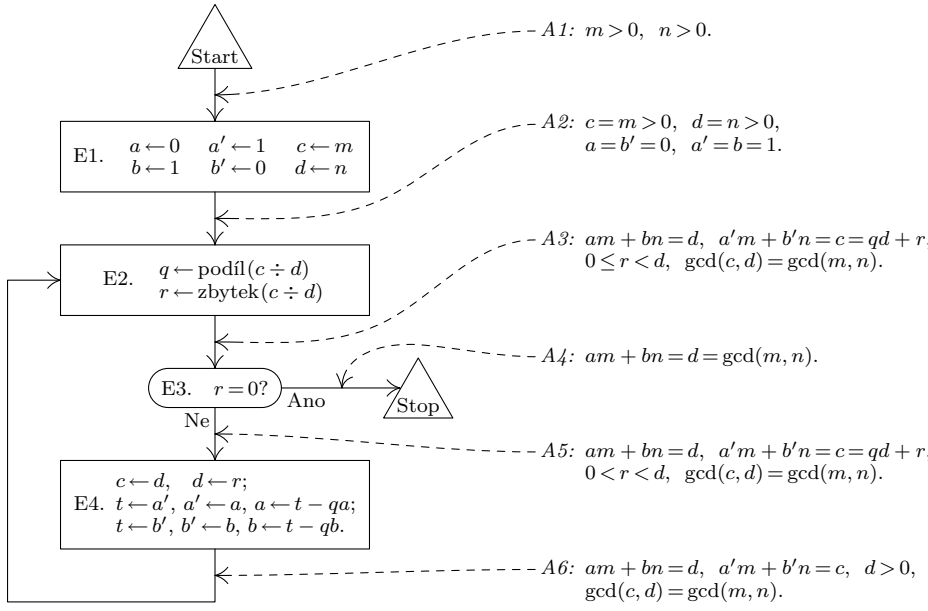
Problém je ovšem *dokázat*, že tento algoritmus funguje správně pro všechna čísla m a n. Mohli bychom se pokusit o uplatnění metody matematické indukce, přičemž za $P(n)$ bychom považovali výrok: „Algoritmus E funguje správně pro všechna n a pro všechna celá m.“ Tento postup ale není úplně jednoduchý a musíme při něm dokázat ještě několik tvrzení navíc. Po chvilce zkoumání přijdeme na to, že musíme dokázat jisté tvrzení o proměnných a, b, a' a b', a sice že po provedení kroku E2 vždy platí rovnosti:

$$a'm + b'n = c, \qquad am + bn = d. \tag{6}$$

Platnost těchto vztahů můžeme dokázat tak, že si ověříme jejich platnost při prvním zahájení kroku E2, a poté prokážeme, že se v kroku E4 jejich platnost nezmění. (Důkaz viz cvičení 6.)

Nyní již můžeme indukcí nad proměnnou n dokázat platnost algoritmu E: Pokud je m násobkem n, funguje algoritmus zřejmě správně, protože skončíme hned po prvním průchodu kroku E3. Tento případ nastává také při $n = 1$. Jediný zbývající případ je, kdy $n > 1$ a m není násobkem n. V takovém případě algoritmus pokračuje a po prvním průchodu přiřadí $c \leftarrow n$, $d \leftarrow r$; protože $r < n$, *můžeme podle indukce předpokládat*, že konečná hodnota d je největším společným dělitelem čísel n a r. Podle zdůvodnění v části 1.1 mají dvojice $\{m, n\}$ a $\{n, r\}$ stejné společné dělitele, a zejména pak největšího společného dělitele. Tudíž d je největším společným dělitelem m a n, a $am + bn = d$ podle (6).

Ve výše uvedeném důkazu jsme kurzívou zvýraznili jednu frázi, která ukazuje, proč se v důkazech indukcí tak často používá přirozený jazyk. Při indukčním kroku (b) neříkáme totiž: „Nyní budeme předpokládat $P(1), P(2), \ldots, P(n)$ a na



Obr. 4. Vývojový diagram Algoritmu E doplněný o tvrzení, která dokazují jeho správnost

základě těchto předpokladů dokážeme $P(n+1)$“, ale často jen jednoduše: „Nyní dokážeme $P(n)$; *podle indukce můžeme předpokládat*, že platí $P(k)$ pro každé $1\le k<n$“.

Jestliže se na tento postup podíváme z trochu jiného pohledu, můžeme vytvořit obecnou metodu, pomocí níž dokážeme platnost *jakéhokoli* algoritmu. Myšlenka spočívá v tom, že vezmeme vývojový diagram určitého algoritmu a každou ze šipek označíme jistým tvrzením o aktuálním stavu, které platí při průchodu výpočtu po šipce. Na obrázku 4 byla tato tvrzení označena jako *A1*, *A2*, ..., *A6*. (Všechna tato tvrzení vycházejí kromě toho také z předpokladu, že proměnné mají celočíselné hodnoty; ten jsme pro úsporu místa vynechali.) Tvrzení *A1* definuje prvotní předpoklady při vstupu do algoritmu, zatímco *A4* popisuje tvrzení, které chceme dokázat o výstupních hodnotách a, b a d.

Obecná metoda znamená dokázat pro každý rámeček ve vývojovém diagramu, že:

> *pokud je tvrzení u šipky, která vede do rámečku, platné před provedením operace v rámečku, pak jsou po provedení operace platná také všechna tvrzení na odpovídajících šipkách vedoucích ven z rámečku.* (7)

To znamená, že musíme například dokázat, že pokud před provedením E2 platí *A2* nebo *A6*, pak po E2 platí *A3*. (V tomto případě je *A2* dokonce silnější tvrzení než *A6*, čili z *A2* přímo vyplývá *A6*. Stačí nám tedy dokázat, že pokud platí *A6* před E2, vyplývá z toho *A3* po E2. Všimněte si přitom, že podmínka

$d > 0$ je v tvrzení *A6* nutná jen proto, aby operace E2 měla smysl.) Dále musíme ukázat, že z *A3* a $r = 0$ vyplývá *A4*, že z *A3* a $r \neq 0$ vyplývá *A5* atd. Všechny tyto potřebné důkazy jsou velice jednoduché.

Jakmile dokážeme výrok (7) *pro každý rámeček, znamená to, že jsou všechna uvedená tvrzení platná při každém průchodu algoritmem.* Nyní můžeme provést indukci podle počtu kroků výpočtu, tedy podle počtu šipek procházených ve vývojovém diagramu. Při průchodu první šipky, která vede z bodu „Start“, je tvrzení *A1* pravdivé, protože předpokládáme, že vstupní hodnoty vždy odpovídají specifikacím; je tedy i tvrzení u první procházení šipky správné. Je-li pravdivé tvrzení u n-té šipky, pak podle (7) je pravdivé i tvrzení uvedené u $(n + 1)$-ní šipky.

Vyjdeme-li z tohoto obecného postupu, je zřejmé, že problém důkazu správnosti určitého algoritmu se v podstatě redukuje na nalezení vhodných tvrzení, kterými označíme šipky ve vývojovém diagramu. Po provedení tohoto úvodního kroku již není nijak těžké dokázat, že z každého tvrzení u šipky vedoucí do rámečku logicky vyplývá také platnost tvrzení u šipky vedoucí z rámečku ven. Ve skutečnosti je i vytvoření těchto tvrzení poměrně snadným, rutinním úkolem, musíme pouze zvládnout několik obtížnějších; jestliže tak v našem příkladu máme dána tvrzení *A1*, *A4* a *A6*, je již jednoduché napsat i *A2*, *A3* a *A5*. Nejvíce práce nám dá vytvořit tvrzení *A6*; zbytek již můžeme doplnit mechanicky. O podrobné formální důkazy algoritmů, tedy o důkazy s takovou mírou podrobnosti jako na obrázku 4, se proto ve zbytku knížky nesnažíme. Stačí vždy vyslovit jen klíčová úvodní tvrzení, která uvádíme buďto ve výkladu za algoritmem, nebo v závorkách jako poznámky v textu samotného algoritmu.

Tento postup důkazu správnosti algoritmů má ale ještě jednu důležitější stránku: *zrcadlí se v něm způsob, jakým celý algoritmus chápeme.* Vzpomeňte si, že už v části 1.1 jsme čtenáře upozornili, aby algoritmus nečetli jako texty z krásné literatury; doporučujeme jednou nebo dvakrát pročíst vhodný algoritmus s vzorkem dat. Při takovém průchodu algoritmem již člověk může lépe zformulovat potřebná tvrzení. Záměrem autora je ukázat, že správnosti algoritmu porozumíme až v okamžiku, kdy si v hlavě dáme dohromady všechna tvrzení, stejně jako na obrázku 4. Tento úhel pohledu má významné psychologické důsledky pro správné sdělení algoritmu od jedné osoby ke druhé: vyplývá odtud, že klíčová tvrzení, která nelze snadno odvodit žádným automatem, musíme při vysvětlování algoritmu druhému člověku vždy jasně, explicitně vyslovit. Vrátíme-li se nyní k Algoritmu E, je třeba připomenout také tvrzení *A6*.

Vnímavý čtenář si ale jistě všiml jedné trhliny v našem posledním důkazu algoritmu E. Neukázali jsme totiž, že algoritmus skončí; pouze jsme dokázali, že *pokud* opravdu skončí, dá správný výsledek!

(Poznamenejme například, že Algoritmus E má smysl i v případě, že v něm budou proměnné m, n, c, d a r moci nabývat hodnot ve formě $u + v\sqrt{2}$, kde u a v jsou celá čísla. Proměnné q, a, b, a', b' nabývají i nadále celočíselných hodnot. Pokud metodu zahájíme třeba s hodnotami $m = 12 - 6\sqrt{2}$ a $n = 20 - 10\sqrt{2}$, vypočteme pomocí ní „největšího společného dělitele“ $d = 4 - 2\sqrt{2}$, kde $a = +2$, $b = -1$. Důkazy tvrzení *A1* až *A6* zůstávají platné i při takto rozšířených

předpokladech; pravdivá jsou tedy i všechna tvrzení při jakémkoli provádění procedury. Jestliže ale začneme s hodnotami $m = 1$ a $n = \sqrt{2}$, výpočet nikdy neskončí (viz cvičení 12). Z důkazu tvrzení *A1* až *A6* tudíž logicky *nevyplývá*, že by byl algoritmus konečný.)

Důkaz ukončení algoritmu se obvykle provádí samostatně. Ve cvičení 13 si ale ukážeme, že výše uvedenou metodu *můžeme* v řadě důležitých případů rozšířit takovým způsobem, že důkaz konečnosti dostaneme jako její vedlejší produkt.

Nyní jsme tedy dvakrát dokázali platnost Algoritmu E. Z přísně logického pohledu bychom se měli pokusit také dokázat platnost prvního algoritmu této části textu, tedy Algoritmu I; pomocí Algoritmu I jsme ve skutečnosti ukázali správnost jakéhokoli důkazu indukcí. Pokusíme-li se ale *dokázat*, že Algoritmus I pracuje správně, dostáváme se k rozporu – to můžeme dokázat opět jedině indukcí! Ocitáme se tak v kruhu.

V poslední analýze vidíme, že *každou* vlastnost celých čísel musíme dokázat indukcí, protože samotná celá čísla jsou v podstatě *definována* indukcí. Větu, že libovolné kladné celé číslo n se buďto rovná 1, nebo se k němu dostaneme od 1 opakovaným přičítáním 1, můžeme tudíž považovat za axiom; to už stačí k dokázání správnosti Algoritmu I. [Podrobný výklad základních principů celých čísel najdete v článku "On Mathematical Induction", který napsal Leon Henkin, *AMM* **67** (1960), 323–338.]

Myšlenka matematické indukce je tudíž úzce spjata s pojmem samotného čísla. V Evropě aplikoval metodu matematické indukce na přesné důkazy italský vědec Francesco Maurolico, a to v roce 1575. S dalšími zdokonaleními přišel začátkem 17. století Pierre de Fermat, který ji nazval „metodou nekonečného poklesu". Pojem také ve svých pozdějších pracích zřetelně používá Blaise Pascal (1653). Samotný výraz „matematická indukce" vytvořil zřejmě A. De Morgan začátkem 19. století. [Viz *The Penny Cyclopædia* **11** (1838), 465–466; *AMM* **24** (1917), 199–207; **25** (1918), 197–201; *Arch. Hist. Exact Sci.* **9** (1972), 1–21.] Další výklad matematické indukce najdete v knize G. Pólya: *Induction and Analogy in Mathematics* (Princeton, N.J.: Princeton University Press, 1954), kapitola 7.

Výše popsaná formulace důkazu algoritmů pomocí tvrzení a indukce je v podstatě dílem R. W. Floyda. Ten poukázal, že sémantickou definici každé operace v programovacím jazyku je možné formulovat jako logické pravidlo, které přesně stanovuje, z jakých tvrzení platných před jeho provedením dokážeme jaká tvrzení po jeho dokončení [viz "Assigning Meanings to Programs", *Proc. Symp. Appl. Math.*, Amer. Math. Soc., **19** (1967), 19–32]. Podobné myšlenky vyslovil nezávisle Peter Naur, *BIT* **6** (1966), 310–316, který tato tvrzení nazýval "general snapshots", „obecné snímky". Důležitým zpřesněním je pojem „invariantů", který zavedl C. A. R. Hoare; viz například *CACM* **14** (1971), 39–45. Pozdější autoři se rozhodli Floydův směr myšlenek obrátit a navrhli vycházet z tvrzení, které musí platit *po* dokončení operace, a odvodit z něj „nejslabší podmínku", která musí platit naopak *před* operací. Díky tomuto postupu je možné objevit i nové algoritmy, které jsou zaručeně správné; stačí vyjít ze specifikací požadovaných výsledků a postupovat zpět. [Viz E. W. Dijkstra, *CACM* **18** (1975), 453–457; *A Discipline of Programming* (Prentice-Hall, 1976).]

Princip úvodních tvrzení se v jisté zárodečné podobě objevil již roku 1946, kdy současně H. H. Goldstine a J. von Neumann představili vývojové diagramy. Ty ve své původní formě obsahovaly „rámečky tvrzení“, “assertion boxes”, které velmi připomínají tvrzení na obrázku 4. [Viz John von Neumann, *Collected Works* **5** (New York: Macmillan, 1963), 91–99. Viz též první poznámky A. M. Turinga k verifikaci, uvedené v článku *Report of a Conference on High Speed Automatic Calculating Machines* (Cambridge Univ., 1949), 67–68 a obrázky; přetištěno s komentářem od F. L. Morrise a C. B. Jonese v *Annals of the History of Computing* **6** (1984), 139–143.]

Při vysvětlení teorie programů výrazně pomůže,
pokud během konstrukce vyslovíme jedno nebo dvě tvrzení
o stavu stroje ve vhodně vybraných místech. . . .
V extrémní podobě teoretické metody
jsou tvrzení podložená přesným matematickým důkazem.
V extrémní podobě experimentální metody
je program spuštěn na stroji s různými počátečními podmínkami
a je prohlášen za správný, pokud tvrzení platí v každém z případů.
Obě metody mají své slabé stránky.

— A. M. TURING, Ferranti Mark I Programming Manual (1950)

CVIČENÍ

1. [*05*] Vysvětlete, jak upravit postup důkazu matematickou indukcí v případě, že budeme chtít dokázat nějaké tvrzení $P(n)$ pro všechna *nezáporná* celá čísla – tedy nikoli pro $n = 1, 2, 3, \ldots$, nýbrž pro $n = 0, 1, 2, \ldots$

▸ **2.** [*15*] V následujícím důkazu musí být chyba. Najdete ji? „**Věta.** Nechť a je libovolné kladné číslo. Pro všechna kladná čísla n je $a^{n-1} = 1$. *Důkaz.* Je-li $n = 1$, pak $a^{n-1} = a^{1-1} = a^0 = 1$. Pokud dále podle indukce předpokládáme, že věta platí pro 1, 2, . . . , n, dostáváme

$$a^{(n+1)-1} = a^n = \frac{a^{n-1} \times a^{n-1}}{a^{(n-1)-1}} = \frac{1 \times 1}{1} = 1;$$

takže věta platí i pro $n + 1$.“

3. [*18*] Následující důkaz indukcí je zdánlivě správný, ale pro $n = 6$ dává z nějakého důvodu na levé straně $\frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \frac{1}{20} + \frac{1}{30} = \frac{5}{6}$ a na pravé straně $\frac{3}{2} - \frac{1}{6} = \frac{4}{3}$. Najdete, kde je v něm chyba? „**Věta.**

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

Důkaz. Provedeme indukcí podle n. Pro $n = 1$ je zřejmě $3/2 - 1/n = 1/(1 \times 2)$; a za předpokladu, že pro n věta platí, je

$$\frac{1}{1 \times 2} + \cdots + \frac{1}{(n-1) \times n} + \frac{1}{n \times (n+1)}$$
$$= \frac{3}{2} - \frac{1}{n} + \frac{1}{n(n+1)} = \frac{3}{2} - \frac{1}{n} + \left(\frac{1}{n} - \frac{1}{n+1}\right) = \frac{3}{2} - \frac{1}{n+1}.\text{“}$$

4. [*20*] Dokažte, že Fibonacciho čísla splňují kromě vztahu (3) také podmínku $F_n \geq \geq \phi^{n-2}$.

5. [*21*] *Prvočíslo* je celé číslo > 1, které je beze zbytku dělitelné jen jedničkou a sebou samým. Na základě této definice a metody matematické indukce dokažte, že každé celé číslo > 1 je možné zapsat jako součin jednoho nebo více prvočísel. (Prvočíslo považujeme za „součin“ jediného prvočísla, a sice sebe samého.)

6. [*20*] Dokažte, že pokud platí vztah (6) bezprostředně před krokem E4, platí také po jeho provedení.

7. [*23*] Zformulujte a indukcí dokažte pravidlo pro součty 1^2, $2^2 - 1^2$, $3^2 - 2^2 + 1^2$, $4^2 - 3^2 + 2^2 - 1^2$, $5^2 - 4^2 + 3^2 - 2^2 + 1^2$ atd.

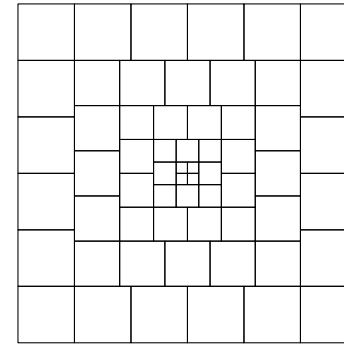
▸ **8.** [*25*] (a) Indukcí dokažte následující větu, kterou okolo roku 100 vyslovil Nicomachus: $1^3 = 1$, $2^3 = 3 + 5$, $3^3 = 7 + 9 + 11$, $4^3 = 13 + 15 + 17 + 19$ atd. (b) Pomocí tohoto výsledku dokažte pozoruhodný vztah $1^3 + 2^3 + \cdots + n^3 = (1 + 2 + \cdots + n)^2$.

[*Poznámka:* Zajímavou geometrickou interpretaci tohoto vzorce přinesl Warren Lushbaugh; je znázorněna na obrázku 5 a byla zveřejněna v *Math. Gazette* **49** (1965), 200. Myšlenka souvisí s Nicomachovou větou a obrázkem 3. Další důkazy „na první pohled“ najdeme v literatuře: Martin Gardner, *Knotted Doughnuts* (New York: Freeman, 1986), kapitola 16; J. H. Conway a R. K. Guy, *The Book of Numbers* (New York: Copernicus, 1996), kapitola 2.]

$$\text{Strana} = 5+5+5+5+5+5 = 5\cdot(5+1)$$

$$\text{Strana} = 5+4+3+2+1+1+2+3+4+5 = 2\cdot(1+2+\cdots+5)$$

$$\text{Plocha} = 4\cdot1^2+4\cdot2\cdot2^2+4\cdot3\cdot3^2+4\cdot4\cdot4^2+4\cdot5\cdot5^2 = 4\cdot(1^3+2^3+\cdots+5^3)$$



Obr. 5. Geometrická verze cvičení 8(b)

9. [*20*] Dokažte indukcí, že pokud je $0 < a < 1$, pak $(1 - a)^n \geq 1 - na$.

10. [*M22*] Dokažte indukcí, že pokud je $n \geq 10$, pak $2^n > n^3$.

11. [*M30*] Nalezněte a dokažte jednoduchý vzorec pro součet

$$\frac{1^3}{1^4+4} - \frac{3^3}{3^4+4} + \frac{5^3}{5^4+4} - \cdots + \frac{(-1)^n(2n+1)^3}{(2n+1)^4+4}.$$

12. [*M25*] Ukažte, jak je možné zobecnit Algoritmus E podle popisu v textu tak, že do něj budou vstupovat hodnoty tvaru $u + v\sqrt{2}$, kde u a v jsou celá čísla, přičemž výpočty bude možné opět provádět elementárním způsobem (tedy bez nekonečného desetinného rozvoje $\sqrt{2}$). Dokažte ale, že pro $m = 1$ a $n = \sqrt{2}$ výpočet neskončí.

▸ **13.** [*M23*] Rozšiřte Algoritmus E: přidejte do něj novou proměnnou T a na začátek každého kroku doplňte operaci „$T \leftarrow T+1$“. (Proměnná T tvoří tudíž něco jako hodiny a počítá počet provedených kroků.) Na začátku předpokládejte, že je T rovno nule, takže tvrzení *A1* na obrázku 4 se změní na „$m > 0$, $n > 0$, $T = 0$“. Podobně k tvrzení *A2* přidáme podmínku „$T = 1$“. Ukažte, jak do jednotlivých tvrzení doplnit další podmínky, aby z libovolného tvrzení *A1*, *A2*, ..., *A6* vyplývalo $T \leq 3n$ a aby nadále bylo možné provést důkaz indukcí. (Jinými slovy: výpočet musí skončit za nejvýše $3n$ kroků.)

14. [*50*] (R. W. Floyd.) Vytvořte počítačový program, který na vstupu převezme program v nějakém programovacím jazyce a nepovinně také jistá tvrzení, a dále se pokusí doplnit zbývající tvrzení potřebná k důkazu správnosti programu. (Tento program by se například měl pokusit dokázat platnost Algoritmu E, přičemž vyjde jen z tvrzení *A1*, *A4*, a *A6*. Další výklad viz příspěvky R. W. Floyd a J. C. King, IFIP Congress proceedings, 1971.)

▸ **15.** [*VM28*] (*Zobecněná indukce.*) V textu jsme si ukázali, jak dokázat výrok $P(n)$ závislý na jediném celém čísle n, ale neukázali jsme si, jak dokázat výrok $P(m,n)$ závislý na dvou celočíselných proměnných. Tyto důkazy se často provádějí nějakou formou „dvojité indukce“, která bývá na pohled dosti nesrozumitelná. Ve skutečnosti existuje ale jeden důležitý princip, který je obecnější než jednoduchá indukce a který je možné použít nejen na tuto situaci, ale i na důkaz výroků nad nespočetnými množinami, například dokázat $P(x)$ pro všechna reálná x. Tomuto obecnému principu se říká *dobré uspořádání*.

Nechť „$\prec$“ je relace v množině S, která splňuje následující vlastnosti:

i) Pokud x, y a z náleží do S a pokud $x \prec y$ a $y \prec z$, pak i $x \prec z$.
ii) Pokud x a y náleží do S, platí právě jedno z následujících tří tvrzení: $x \prec y$, $x = y$ nebo $y \prec x$.
iii) Je-li A libovolná neprázdná podmnožina S, existuje prvek x z A takový, že $x \preceq y$ (tedy $x \prec y$ nebo $x = y$) pro všechna y z A.

Tuto relaci nazýváme dobrým uspořádáním množiny S. Je například zřejmé, že množina kladných celých čísel je dobře uspořádaná v obyčejné relaci „menší než“, $<$.

a) Ukažte, že množina *všech* celých čísel již není v relaci $<$ dobře uspořádaná.
b) Definujte relaci dobrého uspořádání na množině všech celých čísel.
c) Je množina všech nezáporných reálných čísel dobře uspořádaná relací $<$?
d) (*Lexikografické řazení.*) Nechť S je dobře uspořádaná s relací $\prec$ a nechť pro každé $n > 0$ je T_n množina všech n-rozměrných vektorů $(x_1, x_2, \ldots, x_n)$ prvků x_j z S. Definujeme $(x_1, x_2, \ldots, x_n) \prec (y_1, y_2, \ldots, y_n)$, pokud existuje takové k, $1 \le k \le n$, že $x_j = y_j$ pro $1 \le j < k$, ale $x_k \prec y_k$ v S. Je $\prec$ dobré uspořádání množiny T_n?
e) Pokračujeme v části (d) a řekněme, že $T = \bigcup_{n \ge 1} T_n$; definujeme $(x_1, x_2, \ldots, x_m) \prec \prec (y_1, y_2, \ldots, y_n)$, pokud je $x_j = y_j$ pro $1 \le j < k$ a $x_k \prec y_k$ pro některé $k \le \min(m, n)$, nebo pokud je $m < n$ a $x_j = y_j$ pro $1 \le j \le m$. Je relace $\prec$ dobrým uspořádáním T?
f) Ukažte, že $\prec$ je dobrým uspořádáním množiny S tehdy a jen tehdy, pokud splňuje podmínky (i) a (ii) a pokud neexistuje žádná nekonečná posloupnost x_1, x_2, x_3, $\ldots$, kde $x_{j+1} \prec x_j$ pro všechna $j \ge 1$.
g) Nechť S je dobře uspořádaná s relací $\prec$ a nechť $P(x)$ je výrok o prvku x z S. Ukažte, že pokud je možné dokázat $P(x)$ za předpokladu, že platí $P(y)$ pro všechna $y \prec x$, pak $P(x)$ platí pro *všechna* x z S.

[*Poznámky:* Část (g) je slíbené zobecnění jednoduché indukce; pokud je S = množina kladných celých čísel, jedná se o stejnou matematickou indukci, jakou jsme v textu probírali. V takovém případě máme dokázat, že platí $P(1)$, jestliže platí $P(y)$ pro všechna kladná celá čísla $y < 1$; to je totéž, jako když máme dokázat $P(1)$, protože $P(y)$ platí pro všechna taková y triviálně. Jinými slovy: v mnoha situacích není nutné výrok $P(1)$ dokazovat zvláštním postupem.

Část (d) nám spolu s částí (g) dává silnou metodu indukce nad n-rozměrným vektorem, pomocí níž můžeme dokazovat výroky $P(m_1, \ldots, m_n)$ o n kladných celých číslech $m_1, \ldots, m_n$.

Část (f) má u počítačových algoritmů další uplatnění: Podaří-li se nám zobrazit každý stav výpočtu x do prvku $f(x)$ z jisté dobře uspořádané množiny S takovým způsobem, že při každém kroku výpočtu se stav x změní do stavu y s tím, že $f(y) \prec \prec f(x)$, pak algoritmus musí skončit. Tento princip je zobecněním tvrzení o klesající posloupnosti hodnot n, pomocí něhož jsme dokázali konečnost Algoritmu 1.1E.]

1.2.2. Čísla, mocniny a logaritmy

Začněme nyní výklad numerické matematiky podrobným pohledem na čísla, s nimiž pracujeme. *Celá čísla* jsou všechna celá čísla

$$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$$

(záporná, nula a kladná). *Racionální číslo* je podílem dvou celých čísel p/q, kde q je kladné. *Reálné číslo* je pak veličina x s *desetinným rozvojem*

$$x = n + 0.d_1d_2d_3\ldots, \tag{1}$$

kde n je celé číslo, každé z d_i je číslice od 0 do 9 a posloupnost číslic nekončí nekonečně mnoha devítkami. Vyjádření (1) znamená, že

$$n + \frac{d_1}{10} + \frac{d_2}{100} + \cdots + \frac{d_k}{10^k} \le x < n + \frac{d_1}{10} + \frac{d_2}{100} + \cdots + \frac{d_k}{10^k} + \frac{1}{10^k}, \tag{2}$$

pro všechna kladná celá čísla k. Příklady reálných čísel, která nejsou racionální, jsou

$\pi = 3,14159265358979\ldots,$ podíl obvodu kruhu a jeho průměru;

$\phi = 1,61803398874989\ldots,$ takzvaný *zlatý řez* $(1+\sqrt{5}\,)/2$ (viz část 1.2.8).

V příloze A je uvedena tabulka důležitých konstant s přesností čtyřiceti desetinných míst. O obecně známých vlastnostech sčítání, odčítání, násobení, dělení a porovnávání reálných čísel nemusíme podrobněji hovořit.

Obtížné problémy z celých čísel se často dají vyřešit přechodem do množiny reálných čísel a obtížné problémy z reálných čísel se často dají vyřešit v ještě obecnější třídě, které se říká komplexní čísla. *Komplexní číslo* je veličina z ve tvaru $z = x + iy$, kde x a y jsou reálná čísla a i je speciální jednotka, která splňuje rovnost $i^2 = -1$. Čísla x a y označujeme za *reálnou část* a *imaginární část* čísla z; jeho absolutní hodnotu definujeme jako

$$|z| = \sqrt{x^2 + y^2}. \tag{3}$$

Číslo komplexně sdružené k z je $\overline{z} = x - iy$, přičemž $z\overline{z} = x^2 + y^2 = |z|^2$. Teorie komplexních čísel je v řadě ohledů jednodušší a krásnější než teorie reálných čísel, ale obvykle je považována za látku pro pokročilé. V této knize se proto soustředíme na reálná čísla a výjimkou budou jen situace, kdy by reálná čísla byla zbytečně komplikovaná.

Jsou-li u a v reálná čísla, přičemž $u \le v$, pak je *uzavřený interval* $[u\,..\,v]$ množina takových reálných čísel x, pro něž $u \le x \le v$. *Otevřený interval* $(u\,..\,v)$ je podobně množina reálných čísel x, pro něž $u < x < v$. Analogicky můžeme definovat i *polootevřené intervaly* $[u\,..\,v)$ a $(u\,..\,v]$. Na otevřené straně intervalu

může být také u rovno $-\infty$ nebo v rovno ∞, takže interval nemá dolní, resp. horní hranici; zápis $(-\infty\,..\,\infty)$ vyjadřuje tudíž množinu všech reálných čísel a interval $[0\,..\,\infty)$ představuje všechna nezáporná reálná čísla.

V této části textu budeme písmenem b označovat kladné reálné číslo. Pokud je n celé číslo, definujeme mocninu b^n podle známých pravidel:

$$b^0 = 1, \qquad b^n = b^{n-1}b \quad \text{pokud} \quad n > 0, \qquad b^n = b^{n+1}/b \quad \text{pokud} \quad n < 0. \tag{4}$$

Indukcí snadno dokážeme platnost *pravidel umocňování*:

$$b^{x+y} = b^x b^y, \qquad (b^x)^y = b^{xy}, \tag{5}$$

pro každá celá čísla x a y.

Pokud je u kladné reálné číslo a m je kladné celé číslo, pak vždy existuje právě jedno kladné reálné číslo v takové, že $v^m = u$; nazýváme je *m-tá odmocnina* čísla u a označujeme $v = \sqrt[m]{u}$.

Nyní následujícím způsobem definujeme výraz b^r pro racionální číslo $r = p/q$:

$$b^{p/q} = \sqrt[q]{b^p}. \tag{6}$$

Tato definice, kterou vyslovil Mikuláš z Oresme (okolo 1360), je správná, protože $b^{ap/aq} = b^{p/q}$ a protože pravidla umocňování platí i pro libovolně zvolená racionální čísla x a y (viz cvičení 9).

Nakonec definujeme b^x pro všechna reálná čísla x. Nejprve uvažujme, že $b > 1$; pokud je x dáno vztahem (1), chceme, aby platilo

$$b^{n+d_1/10+\cdots+d_k/10^k} \le b^x < b^{n+d_1/10+\cdots+d_k/10^k+1/10^k}. \tag{7}$$

Tím jednoznačně *definujeme* b^x jako kladné reálné číslo, protože rozdíl mezi pravou a levou stranou nerovnosti ve vztahu (7) je $b^{n+d_1/10+\cdots+d_k/10^k}(b^{1/10^k}-1)$ a podle níže uvedeného cvičení 13 je tento rozdíl menší než $b^{n+1}(b-1)/10^k$, a při dostatečně velkém k vypočteme b^x s libovolnou požadovanou přesností.

Takto můžeme například zjistit, že

$$10^{0,30102999} = 1,9999999739\ldots, \quad 10^{0,30103000} = 2,0000000199\ldots; \tag{8}$$

a tudíž pokud je $b = 10$ a $x = 0,30102999\ldots$, známe hodnotu čísla 10^x s lepší přesností než jedna ku 10 miliónům (i když zatím nevíme, jestli je desetinný rozvoj čísla 10^x roven $1,999\ldots$ nebo $2,000\ldots$).

Je-li $b < 1$, definujeme $b^x = (1/b)^{-x}$; pro $b = 1$ je $b^x = 1$. Na základě těchto definicí již můžeme dokázat, že pravidla umocňování (5) platí i pro libovolná *reálná* čísla x a y. Tyto principy definování b^x formuloval poprvé John Wallis (1655) a Isaac Newton (1669).

Nyní přejdeme k jedné důležité otázce. Předpokládejme, že je dáno kladné reálné číslo y; umíme nalézt takové reálné x, pro něž je $y = b^x$? Odpověď zní „ano“ (za předpokladu že $b \ne 1$); je-li dáno $b^x = y$, stačí použít vztah (7) *obráceně* a již zjistíme n a $d_1, d_2, \ldots$ Výsledné číslo x se nazývá *logaritmus čísla y při základu b*, přičemž zapisujeme $x = \log_b y$. Podle této definice máme

$$x = b^{\log_b x} = \log_b(b^x). \tag{9}$$

Jako příklad můžeme z rovnic (8) odvodit, že

$$\log_{10} 2 = 0,30102999\ldots \tag{10}$$

Z pravidel umocňování vyplývá, že

$$\log_b(xy) = \log_b x + \log_b y, \quad \text{pokud} \quad x > 0,\ \ y > 0 \tag{11}$$

a dále

$$\log_b(c^y) = y\ \log_b c, \quad \text{pokud} \quad c > 0. \tag{12}$$

Vztah (10) ilustruje takzvané *dekadické logaritmy*, které dostáváme při základu rovném 10. Možná si řeknete, že v počítačích budou užitečné spíše *dvojkové logaritmy* (o základu 2), protože většina počítačů pracuje s dvojkovou neboli binární aritmetikou. Dvojkové logaritmy jsou skutečně velmi užitečné, ale především z jiného důvodu: počítačové algoritmy provádějí často rozhodování mezi dvěma větvemi výpočtu. S dvojkovými logaritmy se setkáváme tak často, že pro ně zavedeme zkrácenou notaci a budeme psát

$$\lg x = \log_2 x, \tag{13}$$

ve tvaru, který zavedl Edward M. Reingold.

Přichází tedy otázka, jestli je nějaký vztah také mezi $\lg x$ a $\log_{10} x$; naštěstí takový vztah existuje,

$$\log_{10} x = \log_{10}(2^{\lg x}) = (\lg x)(\log_{10} 2),$$

podle (9) a (12). Je tedy $\lg x = \log_{10} x / \log_{10} 2$, a obecně platí

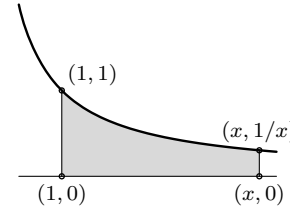
$$\log_c x = \frac{\log_b x}{\log_b c}. \tag{14}$$

Vztahy (11), (12) a (14) tvoří základní pravidla pro manipulaci s logaritmy.

Ukazuje se ovšem, že se ve většině případů nepočítá nejlépe ani se základem 10, ani se základem 2. Jednodušší vlastnosti mají logaritmy při základu jistého reálného čísla označovaného jako $e = 2,718281828459045\ldots$ Logaritmy při základu e se nazývají *přirozené logaritmy* a zapisujeme je jako

$$\ln x = \log_e x. \tag{15}$$

Tato dosti zvláštní definice (fakticky jsme ani nedefinovali e) nebude zřejmě čtenáři připadat jako příliš „přirozený“ logaritmus; přesto je ale funkce $\ln x$ skutečně mnohem přirozenější, a budeme s ní proto pracovat. Přirozené logaritmy (s mírnými úpravami a bez spojení s mocninami) objevil John Napier před rokem 1590, dávno před objevením jiných druhů logaritmů. Následující dva příklady, které čtenář najde v každé učebnici diferenciálního počtu, nám poněkud osvětlí, proč si Napierovy logaritmy skutečně zasluhují označení „přirozené“: (a) Na obrázku 6 tvoří stínovaná plocha $\ln x$. (b) Jestliže banka vyplácí z vkladu složené úroky s mírou r, přičemž je připisuje pololetně, je



Obr. 6. Přirozený logaritmus

roční výnosnost každého dolaru vkladu rovna $(1+r/2)^2$ dolarům; při čtvrtletním připisování dostáváme $(1+r/4)^4$ dolarů a při denním připisování $(1+r/365)^{365}$ dolarů. Pokud by bylo možné připisovat úroky *spojitě*, dostali bychom za každý dolar vkladu přesně e^r (pomineme-li zaokrouhlovací chyby). Dnes, ve věku počítačů, již mnozí bankéři skutečně dosáhli omezující hranice.

Zajímavou historii principů logaritmů a umocňování napsal v sérii článků F. Cajori, *AMM* **20** (1913), 5–14, 35–47, 75–84, 107–117, 148–151, 173–182, 205–210.

Na závěr této části textu si ukážeme, jak logaritmy *počítat*. Jedna metoda přímo vyplývá ze vztahu (7): Vyjdeme-li z rovnosti $b^x = y$ a umocníme obě její strany na 10^k, zjistíme, že je

$$b^m \le y^{10^k} < b^{m+1}, \tag{16}$$

pro nějaké celé číslo m. Pro zjištění logaritmu y nám nyní stačí umocnit y na toto velké číslo a zjistit, mezi kterými mocninami $(m,\ m+1)$ čísla b leží výsledek; výraz $m/10^k$ pak tvoří odpověď s přesností na k desetinných míst.

Mírnými úpravami této zjevně nepraktické metody dostaneme jednodušší a rozumnější postup. Ukážeme si, jak vypočítat $\log_{10} x$ a jak odpověď vyjádřit ve *dvojkové* soustavě, tedy

$$\log_{10} x = n + b_1/2 + b_2/4 + b_3/8 + \cdots \tag{17}$$

Nejprve posuneme desetinnou čárku v čísle x doleva nebo doprava tak, abychom měli $1 \le x/10^n < 10$; tím zjistíme celou část výsledku, n. Pro zjištění b_1, b_2, ... položíme nyní $x_0 = x/10^n$ a pro $k \ge 1$,

$$\begin{aligned} &b_k = 0, \quad x_k = x_{k-1}^2, && \text{pokud} \quad x_{k-1}^2 < 10; \\ &b_k = 1, \quad x_k = x_{k-1}^2/10, && \text{pokud} \quad x_{k-1}^2 \ge 10. \end{aligned} \tag{18}$$

Správnost této procedury vyplývá ze vztahu

$$1 \le x_k = x^{2^k} \big/ 10^{2^k(n+b_1/2+\cdots+b_k/2^k)} < 10, \tag{19}$$

který platí pro $k = 0, 1, 2, \ldots$, jak snadno dokážeme indukcí.

V praxi se ovšem musíme spokojit jen s konečnou přesností, a nemůžeme tedy stanovit přesně $x_k = x_{k-1}^2$. Namísto toho položíme $x_k = x_{k-1}^2$ *zaokrouhlené* nebo *oříznuté* na jistý počet desetinných míst. Takto například vypočteme $\log_{10} 2$ na čtyři významné číslice:

$$\begin{array}{llllllll} x_0 = 2{,}000; & & & & & & & \\ x_1 = 4{,}000, & b_1 = 0; & x_6 = 1{,}845, & b_6 = 1; & & & & \\ x_2 = 1{,}600, & b_2 = 1; & x_7 = 3{,}404, & b_7 = 0; & & & & \\ x_3 = 2{,}560, & b_3 = 0; & x_8 = 1{,}159, & b_8 = 1; & & & & \\ x_4 = 6{,}554, & b_4 = 0; & x_9 = 1{,}343, & b_9 = 0; & & & & \\ x_5 = 4{,}295, & b_5 = 1; & x_{10} = 1{,}804, & b_{10} = 0; & \text{atd.} & & & \end{array}$$

Díky výpočetním chybám se ve výpočtu kumuluje odchylka; skutečná zaokrouhlená hodnota x_{10} je 1,798. Tak nakonec vypočteme b_{19} nesprávně a místo správné

hodnoty z (10) dostaneme dvojkové číslo $(0,01001101000100000011\ldots)_2$, které odpovídá desítkovému vyjádření $0,301031\ldots$

U jakékoli potřebné metody je nutné zjistit, k jak velké výpočetní chybě vedou daná omezení. Ve cvičení 27 zjistíme horní hranici chyby; při výpočtu na čtyři platné číslice je tak chyba v hodnotě logaritmu zaručeně menší než 0,00044. My jsme ovšem získali přesnější odpověď, a to zejména proto, že čísla x_0, x_1, x_2 a x_3 byla stanovena *přesně*.

Tato metoda je jednoduchá a docela zajímavá, nepředstavuje ale nejlepší způsob výpočtu logaritmů na počítači. Jinou metodu si ukážeme ve cvičení 25.

CVIČENÍ

1. [*00*] Jaké je nejmenší kladné racionální číslo?

2. [*00*] Je $1 + 0,239999999\ldots$ platný desetinný rozvoj?

3. [*02*] Kolik je $(-3)^{-3}$?

▶ **4.** [*05*] Kolik je $(0,125)^{-2/3}$?

5. [*05*] Definovali jsme reálná čísla pomocí desetinného rozvoje. Uveďte, jak by bylo možné je definovat pomocí dvojkového rozvoje, a napište definici, kterou byste nahradili vztah (2).

6. [*10*] Nechť $x = m + 0,d_1d_2\ldots$ a $y = n + 0,e_1e_2\ldots$ jsou reálná čísla. Uveďte pravidlo, podle něhož z desetinného rozvoje zjistíte, jestli je $x = y$, $x < y$ nebo $x > y$.

7. [*M23*] Jsou-li dána dvě celá čísla x a y, dokažte pravidla umocňování; vyjděte z definice ve vztahu (4).

8. [*25*] Nechť m je kladné celé číslo. *Dokažte*, že každé kladné reálné číslo u má právě jednu m-tou odmocninu, a uveďte metodu pro postupný výpočet hodnot n, d_1, d_2, ..., v desítkovém rozvoji odmocniny.

9. [*M23*] Jsou-li dána dvě racionální čísla x a y, dokažte pravidla umocňování z předpokladu, že tato pravidla platí pro celá čísla x a y.

10. [*18*] Dokažte, že $\log_{10} 2$ není racionální číslo.

▶ **11.** [*10*] Pokud je $b = 10$ a $x \approx \log_{10} 2$, na kolik desetinných míst přesnosti musíme znát hodnotu x, abychom mohli určit první tři desetinná místa desítkového rozvoje b^x? [*Poznámka:* Můžete využít výsledků cvičení 10.]

12. [*02*] Vysvětlete, proč vztah (10) vyplývá ze vztahu (8).

▶ **13.** [*M23*] (a) Je-li dáno kladné reálné číslo x a kladné celé číslo n, dokažte nerovnost $\sqrt[n]{1+x} - 1 \le x/n$. (b) Na základě tohoto tvrzení vysvětlete poznámky za vztahem (7).

14. [*15*] Dokažte vztah (12).

15. [*10*] Dokažte nebo vyvraťte:

$$\log_b x/y = \log_b x - \log_b y, \quad \text{je-li} \quad x, y > 0.$$

16. [*00*] Jak můžeme vyjádřit $\log_{10} x$ s využitím $\ln x$ a $\ln 10$?

▶ **17.** [*05*] Kolik je $\lg 32$? $\log_\pi \pi$? $\ln e$? $\log_b 1$? $\log_b(-1)$?

18. [*10*] Dokažte, nebo vyvraťte: $\log_8 x = \frac{1}{2} \lg x$.

▶ **19.** [*20*] Pokud je n celé číslo s desítkovou reprezentací o 14 číslicích, vejde se jeho hodnota do počítačového slova s kapacitou 47 bitů a jedním bitem znaménka?

20. [*10*] Je nějaký jednoduchý vztah mezi čísly $\log_{10} 2$ a $\log_2 10$?

21. [*15*] (*Logaritmy logaritmů.*) Vyjádřete $\log_b \log_b x$ pomocí $\ln\ln x$, $\ln\ln b$ a $\ln b$.

▸ **22.** [*20*] (R. W. Hamming.) Dokažte, že

$$\lg x \approx \ln x + \log_{10} x,$$

s menší chybou než 1 %! (Pomocí tabulky přirozených logaritmů a dekadických logaritmů můžeme tudíž zjistit i přibližné hodnoty dvojkových logaritmů.)

23. [*M25*] Uveďte *geometrický* důkaz rovnosti $\ln xy = \ln x + \ln y$; vyjděte z obrázku 6.

24. [*15*] Vysvětlete, jak metodu výpočtu logaritmů při základu 10 na konci této části textu upravíte, aby vypočítala logaritmus o základu 2.

25. [*22*] Uvažujme binární (dvojkový) počítač a číslo x, kde $1 \le x < 2$. Ukažte, že pomocí následujícího algoritmu, který pracuje jen s operacemi posuvu, sčítání a odčítání, úměrně počtu míst požadované přesnosti, můžeme vypočítat přibližnou hodnotu $y = \log_b x$:

L1. [Inicializace.] Přiřaďte $y \leftarrow 0$, $z \leftarrow x$ posun vpravo o 1, $k \leftarrow 1$.

L2. [Test ukončení.] Je-li $x = 1$, skončete.

L3. [Porovnání.] Je-li $x - z < 1$, přiřaďte $z \leftarrow z$ posun vpravo o 1, $k \leftarrow k + 1$, a opakujte tento krok.

L4. [Zmenšení hodnot.] Přiřaďte $x \leftarrow x - z$, $z \leftarrow x$ posun vpravo o k, $y \leftarrow y + \log_b(2^k/(2^k - 1))$ a jděte na L2. ∎

[*Poznámky:* Tato metoda je velmi podobná postupu, kterým se provádí dělení v hardwaru počítačů. Autorem podstaty této myšlenky je Henry Briggs, který takto počítal tabulky logaritmů (v desítkové, nikoli dvojkové podobě) a vydal je roku 1624. Potřebujeme u ní pomocnou tabulku konstant $\log_b 2$, $\log_b(4/3)$, $\log_b(8/7)$ atd., a to na tolik míst, jaká je přesnost počítače. Součástí algoritmu jsou úmyslné výpočetní chyby, protože čísla se posouvají doprava, takže x se nakonec sníží na jedničku a algoritmus skončí. Úkolem tohoto cvičení je vysvětlit, proč skončí a proč vypočte přibližnou hodnotu $\log_b x$.]

26. [*M27*] Najděte ostrou horní hranici chyby algoritmu z předcházejícího cvičení, odvozenou od přesnosti použité v aritmetických operacích.

▸ **27.** [*M25*] Uvažujte v textu probíranou metodu výpočtu $\log_{10} x$. Nechť x'_k označuje vypočtenou aproximaci x_k, stanovenou takto: $x(1-\delta) \le 10^n x'_0 \le x(1+\epsilon)$; při určení x'_k podle vztahů (18) se veličina y_k použije namísto $(x'_{k-1})^2$, kde $(x'_{k-1})^2(1-\delta) \le y_k \le (x'_{k-1})^2(1+\epsilon)$ a $1 \le y_k < 100$. Zde jsou δ a ϵ malé konstanty, do nichž se promítají horní a spodní hranice chyb vzniklých při zaokrouhlování nebo oříznutí. Jestliže $\log' x$ označuje výsledek výpočtu, ukažte, že po k krocích dostaneme

$$\log_{10} x + 2\log_{10}(1-\delta) - 1/2^k < \log' x \le \log_{10} x + 2\log_{10}(1+\epsilon).$$

28. [*M30*] (R. Feynman.) Vytvořte metodu výpočtu b^x pro $0 \le x < 1$, přičemž použijete jen operace posunu, sčítání a odčítání (podobně jako v algoritmu ze cvičení 25) a analyzujte jeho přesnost.

29. [*VM20*] Nechť x je reálné číslo větší než 1. (a) Pro jaké reálné číslo $b > 1$ je $b\log_b x$ minimální? (b) Pro jaké *celé číslo* $b > 1$ je minimální? (c) Pro jaké celé $b > 1$ je $(b+1)\log_b x$ minimální?

30. [*12*] Zjednodušte výraz $(\ln n)^{\ln n/\ln\ln n}$ za předpokladu že $n > 1$.

1.2.3. Součty a součiny

Nechť $a_1, a_2, \ldots$ je libovolná posloupnost čísel. Často nás budou zajímat různé součty neboli sumy, například $a_1 + a_2 + \cdots + a_n$; tento výraz zapíšeme úsporněji pomocí následujících dvou ekvivalentních notací:

$$\sum_{j=1}^{n} a_j \qquad \text{nebo} \qquad \sum_{1\le j\le n} a_j. \tag{1}$$

Pokud je n rovno nule, je i hodnota součtu definována jako nulová. Obecně pokud je $R(j)$ libovolná relace s proměnnou j, vyjadřuje symbol

$$\sum_{R(j)} a_j \tag{2}$$

součet všech takových a_j, pro něž celé číslo j splňuje podmínku $R(j)$. Jestliže takové celé číslo neexistuje, vyjadřuje zápis (2) nulu. Písmeno j ve vztahu (1) a (2) označujeme jako *pomocný index* neboli *indexovou proměnnou*, která je zavedena jen pro účely této notace. Jako indexové proměnné používáme zpravidla písmena i, j, k, m, n, r, s, t (někdy také s dolním indexem, čárkou nebo jiným znaménkem). Velké sumační symboly jako ve vztahu (1) a (2) můžeme také zapsat úsporněji jako $\sum_{j=1}^{n} a_j$ nebo $\sum_{R(j)} a_j$. Sumaci s definovanými hranicemi, vyjádřenou symbolem $\sum$ a indexovými proměnnými, zavedl J. Fourier v roce 1820.

Přesněji řečeno notace $\sum_{1\le j\le n} a_j$ není jednoznačná, protože z ní není jasné, jestli sumace probíhá podle proměnné j nebo n. Tento konkrétní případ by bylo samozřejmě hloupé interpretovat jako součet všech hodnot $n \ge j$; snadno ale můžeme zkonstruovat smysluplné příklady, v nichž indexová proměnná není jasně určena, například $\sum_{j\le k} \binom{j+k}{2j-k}$. V takovém případě musíme z kontextu upřesnit, která z proměnných je pomocná a která je významná i mimo sumační vzorec. Sumu jako $\sum_{j\le k} \binom{j+k}{2j-k}$ bychom měli zapisovat jen tehdy, pokud má proměnná j, nebo k (nikoli obojí) i nějaký vnější význam.

Většinou budeme notaci (2) používat jen u *konečné* sumy – tedy v případě, že relaci $R(j)$ a podmínku $a_j \ne 0$ splňuje jen konečný počet hodnot j. Potřebujeme-li zapsat nekonečnou sumu, například

$$\sum_{j=1}^{\infty} a_j \;=\; \sum_{j\ge 1} a_j \;=\; a_1 + a_2 + a_3 + \cdots$$

s nekonečně mnoha nenulovými členy, musíme již použít techniky diferenciálního počtu; přesný význam vztahu (2) je pak

$$\sum_{R(j)} a_j = \left(\lim_{n\to\infty} \sum_{\substack{R(j)\\ 0\le j<n}} a_j\right) + \left(\lim_{n\to\infty} \sum_{\substack{R(j)\\ -n\le j<0}} a_j\right), \tag{3}$$

tedy za předpokladu, že obě limity existují. Pokud některá z limit neexistuje, neexistuje ani nekonečný součet a suma je *divergentní*; jinak hovoříme o *konvergentní* sumě.

Jsou-li pod sumačním znakem $\sum$ uvedeny dvě nebo více podmínek, jako například v (3), musí být splněny *všechny zároveň*.

Nad sumami uvedeme nyní čtyři jednoduché a velmi důležité algebraické operace, které usnadňují řešení mnoha problémů. Podívejme se tedy na ně.

a) *Distributivní zákon*, součin sum:

$$\left(\sum_{R(i)} a_i\right)\left(\sum_{S(j)} b_j\right) = \sum_{R(i)}\left(\sum_{S(j)} a_i b_j\right). \tag{4}$$

Tuto rovnost si vysvětlíme na speciálním případu

$$\begin{aligned}\left(\sum_{i=1}^{2} a_i\right)\left(\sum_{j=1}^{3} b_j\right) &= (a_1+a_2)(b_1+b_2+b_3)\\ &= (a_1b_1+a_1b_2+a_1b_3)+(a_2b_1+a_2b_2+a_2b_3)\\ &= \sum_{i=1}^{2}\left(\sum_{j=1}^{3} a_i b_j\right).\end{aligned}$$

Závorky na pravé straně vztahu (4) se obvykle vynechávají a dvojitou sumu $\sum_{R(i)}\bigl(\sum_{S(j)} a_{ij}\bigr)$ zapisujeme jednoduše jako $\sum_{R(i)}\sum_{S(j)} a_{ij}$.

b) *Záměna proměnné:*

$$\sum_{R(i)} a_i = \sum_{R(j)} a_j = \sum_{R(p(j))} a_{p(j)}. \tag{5}$$

Tato rovnost vyjadřuje dva typy transformací. V prvním případě jednoduše změníme název indexové proměnné z i na j. Druhý případ je zajímavější: zde je $p(j)$ funkcí proměnné j, která vyjadřuje *permutaci* příslušných hodnot; přesněji řečeno pro každé celé i splňující relaci $R(i)$ musí existovat právě jedno celé číslo j, které splňuje relaci $p(j) = i$. Tato podmínka je vždy splněna mimo jiné v důležitých případech $p(j) = c + j$ a $p(j) = c - j$, kde c je celé číslo nezávislé na j; tyto konstrukce se často používají v různých aplikacích. Například

$$\sum_{1\le j\le n} a_j = \sum_{1\le j-1\le n} a_{j-1} = \sum_{2\le j\le n+1} a_{j-1}. \tag{6}$$

Uvedený příklad by si čtenář měl důkladně prostudovat.

Náhradu j za $p(j)$ nemůžeme ale provést pro všechny *nekonečné* sumy. Platí-li $p(j) = c \pm j$ jako výše, je tato operace vždy platná, ale v jiných případech musíme postupovat opatrně. [Viz například T. M. Apostol, *Mathematical Analysis* (Reading, Mass.: Addison-Wesley, 1957), kapitola 12. Postačující podmínkou obecné platnosti vztahu (5) pro libovolnou permutaci celých čísel $p(j)$ je, že existuje suma $\sum_{R(j)} |a_j|$.]

c) *Záměna pořadí sumace:*

$$\sum_{R(i)}\sum_{S(j)} a_{ij} = \sum_{S(j)}\sum_{R(i)} a_{ij}. \tag{7}$$

Uvažujme nyní velmi jednoduchý speciální případ této rovnosti:

$$\sum_{R(i)} \sum_{j=1}^{2} a_{ij} = \sum_{R(i)} (a_{i1} + a_{i2}),$$

$$\sum_{j=1}^{2} \sum_{R(i)} a_{ij} = \sum_{R(i)} a_{i1} + \sum_{R(i)} a_{i2}.$$

Podle (7) jsou tyto dvě sumy rovné; neříkáme tím nic jiného, než

$$\sum_{R(i)} (b_i + c_i) = \sum_{R(i)} b_i + \sum_{R(i)} c_i, \tag{8}$$

kde položíme $b_i = a_{i1}$ a $c_i = a_{i2}$.

Operace záměny pořadí sumace je velice užitečná, protože často dokážeme najít jednoduché vyjádření sumy $\sum_{R(i)} a_{ij}$, ale už ne $\sum_{S(j)} a_{ij}$. Pořadí sumace potřebujeme často zaměnit také v jiné obecnější situaci, kdy relace $S(j)$ závisí na i i na j. V takovém případě můžeme relaci zapsat jako „$S(i,j)$“. Záměnu sumací můžeme vždy provést, přinejmenším teoreticky, následujícím způsobem:

$$\sum_{R(i)} \sum_{S(i,j)} a_{ij} = \sum_{S'(j)} \sum_{R'(i,j)} a_{ij}, \tag{9}$$

kde $S'(j)$ je relace „existuje celé číslo i takové, že platí $R(i)$ a zároveň $S(i,j)$“; dále $R'(i,j)$ je relace „platí $R(i)$ a zároveň $S(i,j)$“. Máme-li například sumu $\sum_{i=1}^{n} \sum_{j=1}^{i} a_{ij}$, pak $S'(j)$ je relace „existuje celé číslo i takové, že platí $1 \le i \le$ $\le n$ a $1 \le j \le i$“, tedy $1 \le j \le n$, a $R'(i,j)$ je relace „$1 \le i \le n$ a $1 \le j \le i$“, čili $j \le i \le n$. Tudíž

$$\sum_{i=1}^{n} \sum_{j=1}^{i} a_{ij} = \sum_{j=1}^{n} \sum_{i=j}^{n} a_{ij}. \tag{10}$$

[*Poznámka:* Stejně jako v případě (b) ani operace záměny pořadí sumace *není pro nekonečné řady vždy platná.* Pokud je tato řada *absolutně konvergentní* – tedy pokud existuje $\sum_{R(i)} \sum_{S(j)} |a_{ij}|$ – můžeme dokázat, že platí i vztahy (7) a (9). Také pokud *libovolný jeden* z výrazů $R(i)$ nebo $S(j)$ určuje ve vztahu (7) *konečnou* sumu a pokud je každá nekonečná suma konvergentní, záměna pořadí opět platí. Pro konvergentní nekonečné sumy platí vždy zejména (8).]

d) *Úpravy definičního oboru.* Jsou-li $R(j)$ a $S(j)$ dvě libovolné relace, máme

$$\sum_{R(j)} a_j + \sum_{S(j)} a_j = \sum_{R(j) \text{ nebo } S(j)} a_j + \sum_{R(j) \text{ a } S(j)} a_j. \tag{11}$$

Například

$$\sum_{1 \le j \le m} a_j + \sum_{m \le j \le n} a_j = \left(\sum_{1 \le j \le n} a_j \right) + a_m, \tag{12}$$

za předpokladu, že $1 \le m \le n$. V tomto případě je podmínka „$R(j)$ a $S(j)$“ jednoduše „$j = m$“, takže jsme druhou sumu zredukovali na pouhé „a_m“. Ve

většině aplikací vztahu (11) jsou přitom relace $R(j)$ a $S(j)$ současně splněny jen pro jednu nebo dvě hodnoty j, jinak není možné, aby relace $R(j)$ a $S(j)$ platily zároveň pro stejné j. Ve druhém z uvedených případů tak druhá suma na pravé straně vztahu (11) zmizí.

Nyní jsme si tedy ukázali čtyři základní pravidla pro manipulaci se sumami a můžeme se podívat na další ukázky aplikace těchto technik.

Příklad 1.

$$\begin{aligned}
\sum_{0\le j\le n} a_j &= \sum_{\substack{0\le j\le n\\ j\text{ sudé}}} a_j + \sum_{\substack{0\le j\le n\\ j\text{ liché}}} a_j && \text{podle pravidla (d),}\\
&= \sum_{\substack{0\le 2j\le n\\ 2j\text{ sudé}}} a_{2j} + \sum_{\substack{0\le 2j+1\le n\\ 2j+1\text{ liché}}} a_{2j+1} && \text{podle pravidla (b),}\\
&= \sum_{0\le j\le n/2} a_{2j} + \sum_{0\le j<n/2} a_{2j+1}.
\end{aligned}$$

Poslední krok znamená pouhé zjednodušení relací pod sumačními znaky.

Příklad 2. Nechť

$$\begin{aligned}
S_1 = \sum_{i=0}^{n}\sum_{j=0}^{i} a_i a_j &= \sum_{j=0}^{n}\sum_{i=j}^{n} a_i a_j && \text{podle pravidla (c) [viz (10)],}\\
&= \sum_{i=0}^{n}\sum_{j=i}^{n} a_i a_j && \text{podle pravidla (b),}
\end{aligned}$$

takže jsme zaměnili názvy i a j s tím, že $a_j a_i = a_i a_j$. Označíme-li nyní druhou sumu jako S_2, máme

$$\begin{aligned}
2S_1 = S_1 + S_2 &= \sum_{i=0}^{n}\left(\sum_{j=0}^{i} a_i a_j + \sum_{j=i}^{n} a_i a_j\right) && \text{podle (8),}\\
&= \sum_{i=0}^{n}\left(\left(\sum_{j=0}^{n} a_i a_j\right) + a_i a_i\right) && \text{podle pravidla (d) [viz (12)],}\\
&= \sum_{i=0}^{n}\sum_{j=0}^{n} a_i a_j + \sum_{i=0}^{n} a_i a_i && \text{podle (8),}\\
&= \left(\sum_{i=0}^{n} a_i\right)\left(\sum_{j=0}^{n} a_j\right) + \left(\sum_{i=0}^{n} a_i^2\right) && \text{podle pravidla (a),}\\
&= \left(\sum_{i=0}^{n} a_i\right)^2 + \left(\sum_{i=0}^{n} a_i^2\right) && \text{podle pravidla (b).}
\end{aligned}$$

Odvodili jsme tudíž důležitou identitu

$$\sum_{i=0}^{n}\sum_{j=0}^{i} a_i a_j = \frac{1}{2}\left(\left(\sum_{i=0}^{n} a_i\right)^2 + \left(\sum_{i=0}^{n} a_i^2\right)\right). \tag{13}$$

Příklad 3 (*Součet geometrické posloupnosti*). Předpokládejme, že $x \neq 1$, $n \geq 0$. Potom je

$$\begin{aligned}
a + ax + \cdots + ax^n &= \sum_{0\leq j\leq n} ax^j && \text{podle definice (2),}\\
&= a + \sum_{1\leq j\leq n} ax^j && \text{podle pravidla (d),}\\
&= a + x\sum_{1\leq j\leq n} ax^{j-1} && \text{podle speciálního případu (a),}\\
&= a + x\sum_{0\leq j\leq n-1} ax^j && \text{podle pravidla (b) [viz (6)],}\\
&= a + x\sum_{0\leq j\leq n} ax^j - ax^{n+1} && \text{podle pravidla (d).}
\end{aligned}$$

Porovnáme-li první relaci s poslední, dostáváme

$$(1-x)\sum_{0\leq j\leq n} ax^j = a - ax^{n+1};$$

a nakonec tudíž obdržíme základní vzorec

$$\sum_{0\leq j\leq n} ax^j = a\Big(\frac{1-x^{n+1}}{1-x}\Big). \tag{14}$$

Příklad 4 (*Součet aritmetické posloupnosti*). Předpokládejme, že $n \geq 0$. Pak

$$\begin{aligned}
&a + (a+b) + \cdots + (a+nb)\\
&\quad= \sum_{0\leq j\leq n} (a+bj) && \text{podle definice (2),}\\
&\quad= \sum_{0\leq n-j\leq n} \big(a + b(n-j)\big) && \text{podle pravidla (b),}\\
&\quad= \sum_{0\leq j\leq n} (a+bn-bj) && \text{zjednodušením,}\\
&\quad= \sum_{0\leq j\leq n} (2a+bn) - \sum_{0\leq j\leq n} (a+bj) && \text{podle (8),}\\
&\quad= (n+1)(2a+bn) - \sum_{0\leq j\leq n} (a+bj),
\end{aligned}$$

protože v první sumě se jednoduše sčítá $(n+1)$ členů nezávislých na j. Položíme-li nyní rovnítko mezi první a poslední výraz a rovnost vydělíme 2, dostaneme

$$\sum_{0\le j\le n} (a+bj) = a(n+1) + \tfrac{1}{2}bn(n+1). \tag{15}$$

To je $n+1$ krát $\frac{1}{2}\big(a+(a+bn)\big)$, což můžeme přečíst jako počet členů krát průměr prvního a posledního členu.

Všimněte si, že všechny důležité vztahy (13), (14) a (15) jsme odvodili čistě pomocí jednoduchých úprav sum. Ve většině učebnic jsou nicméně tyto vztahy pouze *vysloveny* a poté se dokazují *indukcí*. Indukce je samozřejmě zcela správný postup, ale vůbec z ní nepochopíme, jak někdo mohl takovéto vztahy vymyslet, snad jen nějakou šťastnou náhodou. Při analýze algoritmů se potkáváme se stovkami sum, které nezapadají do žádného známého vzoru; pouhými úpravami či manipulacemi se často dostaneme k odpovědi i bez zázračného uhodnutí.

Celou řadu úprav sum a dalších vzorců si výrazně zjednodušíme, pokud zavedeme následující *závorkovou notaci*:

$$[\text{výrok}] \;=\; \begin{cases} 1, & \text{pokud je výrok pravdivý;} \\ 0, & \text{pokud je výrok nepravdivý.} \end{cases} \tag{16}$$

Potom můžeme psát například

$$\sum_{R(j)} a_j = \sum_j a_j \big[R(j)\big], \tag{17}$$

kde suma na pravé straně prochází přes *všechna* celá čísla j, protože při nesplnění podmínky $R(j)$ je příslušný člen nekonečné sumy nulový. (Zde ale předpokládáme, že a_j je definováno pro všechna j.)

S pomocí závorkové notace můžeme také zajímavým způsobem odvodit pravidlo (b) z pravidel (a) a (c):

$$\begin{aligned} \sum_{R(p(j))} a_{p(j)} &= \sum_j a_{p(j)} \big[R(p(j))\big] \\ &= \sum_j \sum_i a_i \big[R(i)\big] \big[i = p(j)\big] \\ &= \sum_i a_i \big[R(i)\big] \sum_j \big[i = p(j)\big]. \end{aligned} \tag{18}$$

Zbývající suma podle j je rovna 1, pokud je $R(i)$ pravdivé a pokud předpokládáme, že p je permutace příslušných hodnot, jak požaduje vztah (5); zůstává nám tudíž $\sum_i a_i[R(i)]$, což je $\sum_{R(i)} a_i$. Tím jsme dokázali (5). Pokud p *není* takovouto permutací, zjistíme z (18) skutečnou hodnotu $\sum_{R(p(j))} a_{p(j)}$.

Nejznámějším speciálním případem závorkové notace je takzvané *Kroneckerovo delta*

$$\delta_{ij} \;=\; [i=j] \;=\; \begin{cases} 1, & \text{pokud } i=j, \\ 0, & \text{pokud } i\ne j, \end{cases} \tag{19}$$

které zavedl Leopold Kronecker roku 1868. Obecnější notace jako například (16) představil K. E. Iverson v roce 1962; notace (16) se proto často nazývá *Iversonova konvence*. [Viz D. E. Knuth, *AMM* **99** (1992), 403–422.]

Analogicky k součtům existuje podobná notace pro součiny: symbol

$$\prod_{R(j)} a_j \tag{20}$$

vyjadřuje součin všech a_j, pro něž celé číslo j splňuje $R(j)$. Pokud takové celé j neexistuje, je hodnota součinu definována jako 1 (*nikoli* 0).

Operace (b), (c) a (d) platí pro součinovou notaci $\prod$ stejně jako pro součtovou $\sum$, s několika jednoduchými úpravami. Několik příkladů na součinovou notaci najdete ve cvičeních na konci této části textu.

Na závěr této části textu se zmíníme o jiné, často velice šikovné notaci pro násobné sumy: u jediného sumačního znaku $\sum$ je možné použít jednu nebo více relací nad *několika* indexovými proměnnými, přičemž suma prochází přes všechny možné kombinace proměnných, které splňují stanovené podmínky. Například

$$\sum_{0\le i\le n}\sum_{0\le j\le n} a_{ij} = \sum_{0\le i,j\le n} a_{ij}; \qquad \sum_{0\le i\le n}\sum_{0\le j\le i} a_{ij} = \sum_{0\le j\le i\le n} a_{ij}.$$

Podle této notace nemá žádný ze sumačních indexů přednost před druhým, takže můžeme novým způsobem odvodit vztah (10):

$$\sum_{i=1}^{n}\sum_{j=1}^{i} a_{ij} = \sum_{i,j} a_{ij}[1\le i\le n][1\le j\le i] = \sum_{i,j} a_{ij}[1\le j\le n][j\le i\le n] = \\ = \sum_{j=1}^{n}\sum_{i=j}^{n} a_{ij}\,,$$

neboť využijeme, že $[1\le i\le n][1\le j\le i] = [1\le j\le i\le n] = [1\le j\le n][j\le i\le n]$. Obecnější rovnost (9) vyplývá podobným způsobem z identity

$$\big[R(i)\big]\big[S(i,j)\big] \;=\; \big[R(i) \text{ and } S(i,j)\big] \;=\big[S'(j)\big]\big[R'(i,j)\big]\,. \tag{21}$$

Další příklad, na němž vidíme užitečnost sumace s několika indexy, je

$$\sum_{\substack{j_1+\cdots+j_n=n\\ j_1\ge\cdots\ge j_n\ge 0}} a_{j_1\ldots j_n}, \tag{22}$$

kde a je proměnná s n-rozměrným vektorem indexů; je-li například $n = 5$, vyjadřuje tato notace součet

$$a_{11111} + a_{21110} + a_{22100} + a_{31100} + a_{32000} + a_{41000} + a_{50000}.$$

(Viz poznámky k rozkladům čísel v části 1.2.1.)

CVIČENÍ – První skupina

1. [*01*] Co znamená zápis $\sum_{1\le j\le n} a_j$, je-li $n = 3,14$?

2. [*10*] Bez sumační notace napište ekvivalent výrazu

$$\sum_{0\le n\le 5} \frac{1}{2n+1},$$

a dále ekvivalent výrazu

$$\sum_{0\le n^2\le 5} \frac{1}{2n^2+1}.$$

▶ **3.** [*13*] Vysvětlete, proč jsou výsledky obou částí předchozího cvičení navzdory pravidlu (b) různé.

4. [*10*] Bez sumační notace napište ekvivalent obou stran rovnosti (10), a to jako sumu sum pro $n = 3$.

▶ **5.** [*VM20*] Dokažte, že pravidlo (a) platí pro libovolnou nekonečnou řadu, pokud tato řada konverguje.

6. [*VM20*] Dokažte, že pravidlo (d) platí pro libovolnou nekonečnou řadu, pokud existují libovolné tři ze čtyř sum.

7. [*VM23*] Je-li dáno celé číslo c, dokažte, že $\sum_{R(j)} a_j = \sum_{R(c-j)} a_{c-j}$, i když jsou obě řady nekonečné.

8. [*VM25*] Najděte příklad nekonečné řady, v níž neplatí rovnost (7).

▶ **9.** [*05*] Platí odvození vztahu (14) i pro $n = -1$?

10. [*05*] Platí odvození vztahu (14) i pro $n = -2$?

11. [*03*] Co bude na pravé straně vztahu (14), je-li $x = 1$?

12. [*10*] Kolik je $1 + \frac{1}{7} + \frac{1}{49} + \frac{1}{343} + \cdots + \left(\frac{1}{7}\right)^n$?

13. [*10*] Na základě rovnosti (15) a za předpokladu $m \le n$ vypočtěte $\sum_{j=m}^{n} j$.

14. [*11*] Na základě výsledku předchozího cvičení vypočtěte $\sum_{j=m}^{n} \sum_{k=r}^{s} jk$.

▶ **15.** [*M22*] Vypočtěte součet $1\times 2 + 2\times 2^2 + 3\times 2^3 + \cdots + n\times 2^n$ pro malé hodnoty n. Vidíte, jaký vzorek se mezi těmito čísly opakuje? Pokud ne, odvoďte jej pomocí podobných úprav, jaké vedly ke vztahu (14).

16. [*M22*] Dokažte, že

$$\sum_{j=0}^{n} jx^j = \frac{nx^{n+2} - (n+1)x^{n+1} + x}{(x-1)^2},$$

je-li $x \ne 1$, a to bez matematické indukce.

▶ **17.** [*M00*] Nechť S je množina celých čísel. Co znamená $\sum_{j\in S} 1$?

18. [*M20*] Ukažte, jak zaměnit pořadí sumace jako ve vztahu (9), je-li $R(i)$ relace „n je násobkem i“ a $S(i,j)$ je relace „$1 \le j < i$“.

19. [*20*] Čemu je rovno $\sum_{j=m}^{n} (a_j - a_{j-1})$?

▶ **20.** [*25*] Dr. I. J. Matrix pozoroval zajímavou řadu rovností:

$$9\times 1 + 2 = 11,\ 9\times 12 + 3 = 111,\ 9\times 123 + 4 = 1111,\ 9\times 1234 + 5 = 11111.$$

a) Napište tento vynikající objev pomocí sumační notace $\sum$.

b) V odpovědi na část (a) nepochybně používáte číslo 10 jako základ desítkové soustavy; vzorec zobecněte tak, aby platil pro soustavu o libovolném základu b.

c) Dokažte vzorec odvozený v (b) pomocí vzorců odvozených v textu nebo ve cvičení 16.

▸ **21.** [*M25*] Odvoďte pravidlo (d) ze vztahů (8) a (17).

▸ **22.** [*20*] Uveďte příslušné analogie k sumačním vztahům (5), (7), (8) a (11), které platí pro *součiny*.

23. [*10*] Vysvětlete, proč je v případě, kdy podmínku $R(j)$ nesplňuje žádné celé číslo, vhodné definovat $\sum_{R(j)} a_j$ jako nulu a $\prod_{R(j)} a_j$ jako jedničku.

24. [*20*] Předpokládejme, že $R(j)$ platí jen pro konečně mnoho j. Dokažte indukcí podle počtu celých čísel, která splňují $R(j)$, že $\log_b \prod_{R(j)} a_j = \sum_{R(j)} (\log_b a_j)$, pokud jsou všechna $a_j > 0$.

▸ **25.** [*15*] Uvažujte následující odvození; není v něm nějaká chyba?

$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{n} \frac{1}{a_j}\right) = \sum_{1\le i\le n} \sum_{1\le j\le n} \frac{a_i}{a_j} = \sum_{1\le i\le n} \sum_{1\le i\le n} \frac{a_i}{a_i} = \sum_{i=1}^{n} 1 = n.$$

26. [*25*] Ukažte, že součin $\prod_{i=0}^{n} \prod_{j=0}^{i} a_i a_j$ je možné vyjádřit pomocí součinů $\prod_{i=0}^{n} a_i$ a úprav součinové notace $\prod$, jak je uvedeno ve cvičení 22.

27. [*M20*] Zobecněte výsledek cvičení 1.2.1–9 a dokažte, že

$$\prod_{j=1}^{n} (1 - a_j) \ge 1 - \sum_{j=1}^{n} a_j$$

za předpokladu, že $0 < a_j < 1$.

28. [*M22*] Najděte jednoduchý vzorec pro součin $\prod_{j=2}^{n} (1 - 1/j^2)$.

▸ **29.** [*M30*] (a) Vyjádřete součet $\sum_{i=0}^{n} \sum_{j=0}^{i} \sum_{k=0}^{j} a_i a_j a_k$ pomocí notace násobné sumy, kterou jsme si vysvětlili na konci této části textu. (b) Stejnou sumu vyjádřete pomocí $\sum_{i=0}^{n} a_i$, $\sum_{i=0}^{n} a_i^2$ a $\sum_{i=0}^{n} a_i^3$ [viz (13)].

▸ **30.** [*M23*] (J. Binet, 1812.) Bez indukce dokažte tuto identitu:

$$\left(\sum_{j=1}^{n} a_j x_j\right)\left(\sum_{j=1}^{n} b_j y_j\right) = \left(\sum_{j=1}^{n} a_j y_j\right)\left(\sum_{j=1}^{n} b_j x_j\right) + \sum_{1\le j<k\le n} (a_j b_k - a_k b_j)(x_j y_k - x_k y_j).$$

[Důležitým speciálním případem je, jsou-li $w_1, \ldots, w_n, z_1, \ldots, z_n$ libovolná komplexní čísla a přiřadíme-li $a_j = w_j$, $b_j = \overline{z}_j$, $x_j = \overline{w}_j$, $y_j = z_j$:

$$\left(\sum_{j=1}^{n} |w_j|^2\right)\left(\sum_{j=1}^{n} |z_j|^2\right) = \left|\sum_{j=1}^{n} w_j z_j\right|^2 + \sum_{1\le j<k\le n} |w_j \overline{z}_k - w_k \overline{z}_j|^2.$$

Členy $|w_j \overline{z}_k - w_k \overline{z}_j|^2$ jsou nezáporné, a proto slavná *Cauchy-Schwarzova nerovnost*

$$\left(\sum_{j=1}^{n} |w_j|^2\right)\left(\sum_{j=1}^{n} |z_j|^2\right) \ge \left|\sum_{j=1}^{n} w_j z_j\right|^2$$

je důsledkem Binetova vzorce.]

31. [*M20*] Pomocí Binetova vzorce vyjádřete sumu $\sum_{1\le j<k\le n} (u_j - u_k)(v_j - v_k)$ pomocí $\sum_{j=1}^{n} u_j v_j$, $\sum_{j=1}^{n} u_j$ a $\sum_{j=1}^{n} v_j$.

32. [*M20*] Dokažte, že

$$\prod_{j=1}^{n}\sum_{i=1}^{m}a_{ij}=\sum_{1\le i_1,\ldots,i_n\le m}a_{i_1 1}\ldots a_{i_n n}.$$

▸ **33.** [*M30*] Jednoho dne večer přišel Dr. Matrix na několik vztahů, které jsou snad ještě zajímavější než ve cvičení 20:

$$\frac{1}{(a-b)(a-c)}+\frac{1}{(b-a)(b-c)}+\frac{1}{(c-a)(c-b)}=0,$$

$$\frac{a}{(a-b)(a-c)}+\frac{b}{(b-a)(b-c)}+\frac{c}{(c-a)(c-b)}=0,$$

$$\frac{a^2}{(a-b)(a-c)}+\frac{b^2}{(b-a)(b-c)}+\frac{c^2}{(c-a)(c-b)}=1,$$

$$\frac{a^3}{(a-b)(a-c)}+\frac{b^3}{(b-a)(b-c)}+\frac{c^3}{(c-a)(c-b)}=a+b+c.$$

Dokažte, že výše uvedené vzorce jsou speciálním případem obecnějšího zákona; nechť $x_1, x_2, \ldots, x_n$ jsou různá čísla, a ukažte, že

$$\sum_{j=1}^{n}\left(x_j^r\Big/\prod_{\substack{1\le k\le n\\ k\ne j}}(x_j-x_k)\right)=\begin{cases}0, & \text{pro } 0\le r<n-1;\\ 1, & \text{pro } r=n-1;\\ \sum_{j=1}^{n}x_j, & \text{pro } r=n.\end{cases}$$

34. [*M25*] Dokažte, že

$$\sum_{k=1}^{n}\frac{\prod_{1\le r\le n,\, r\ne m}(x+k-r)}{\prod_{1\le r\le n,\, r\ne k}(k-r)}=1$$

za předpokladu $1\le m\le n$ a libovolného x. Je-li například $n=4$ a $m=2$, pak

$$\frac{x(x-2)(x-3)}{(-1)(-2)(-3)}+\frac{(x+1)(x-1)(x-2)}{(1)(-1)(-2)}+\frac{(x+2)x(x-1)}{(2)(1)(-1)}+\frac{(x+3)(x+1)x}{(3)(2)(1)}=1.$$

35. [*VM20*] Notace $\sup_{R(j)}a_j$ označuje nejmenší horní závoru prvků a_j, přesně stejným způsobem jako u součtových a součinových notací $\sum$ a $\prod$. (Je-li $R(j)$ splněno jen pro konečně mnoho j, vyjadřuje se stejná veličina často zápisem $\max_{R(j)}a_j$.) Ukažte, jak upravit pravidla (a), (b), (c) a (d) pro práci v *této* notaci. Zejména rozeberte následující analogii pravidla (a):

$$(\sup\nolimits_{R(i)}a_i)+(\sup\nolimits_{S(j)}b_j)=\sup\nolimits_{R(i)}(\sup\nolimits_{S(j)}(a_i+b_j))$$

a napište vhodnou definici notace za podmínky, kdy $R(j)$ není splněno pro *žádné* j.

CVIČENÍ – Druhá skupina

Determinanty a matice. Následující zajímavé problémy jsou určeny pro čtenáře, který má za sebou alespoň úvod do determinantů a do základů teorie matic. Determinant je možné vypočítat pomocí důvtipné kombinace následujících operací: (a) vytknutím hodnoty z řádku nebo sloupce, (b) přičtením násobku jednoho řádku (nebo sloupce) k jinému řádku (sloupci), (c) rozvojem algebraických doplňků (kofaktorů). Nejjednodušší a nejčastěji používaná verze operace (c) vypadá tak, že odstraníme celý první řádek

a sloupec, a to za předpokladu, že prvek v levém horním rohu je $+1$ a zbývající prvky v celém prvním řádku nebo v celém prvním sloupci jsou nulové; poté vyhodnotíme výsledný menší determinant. Obecně je algebraický doplněk prvku a_{ij} v determinantu řádu $n\times n$ roven $(-1)^{i+j}$ krát $(n-1)\times(n-1)$ determinant vypočtený odstraněním řádku a sloupce, v němž leží prvek a_{ij}. Hodnota determinantu je rovna $\sum a_{ij}\cdot$alg.doplněk(a_{ij}), přičemž sčítáme při konstantním i nebo j a druhý index necháme probíhat od 1 do n.

Je-li (b_{ij}) *inverzní* of matice (a_{ij}), pak se b_{ij} rovná algebraickému doplňku a_{ji} (*nikoli* a_{ij}), děleno determinantem celé matice.

Speciální význam mají následující typy matic:

Vandermondova matice,

$$a_{ij} = x_j^i$$

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & & & \vdots \\ x_1^n & x_2^n & \dots & x_n^n \end{pmatrix}$$

Kombinatorická matice,

$$a_{ij} = y + \delta_{ij}x$$

$$\begin{pmatrix} x+y & y & \dots & y \\ y & x+y & \dots & y \\ \vdots & & & \vdots \\ y & y & \dots & x+y \end{pmatrix}$$

Cauchyho matice,

$$a_{ij} = 1/(x_i + y_j)$$

$$\begin{pmatrix} 1/(x_1+y_1) & 1/(x_1+y_2) & \dots & 1/(x_1+y_n) \\ 1/(x_2+y_1) & 1/(x_2+y_2) & \dots & 1/(x_2+y_n) \\ \vdots & & & \vdots \\ 1/(x_n+y_1) & 1/(x_n+y_2) & \dots & 1/(x_n+y_n) \end{pmatrix}$$

36. [*M23*] Ukažte, že determinant kombinatorické matice je roven $x^{n-1}(x+ny)$.

▸ **37.** [*M24*] Ukažte, že determinant Vandermondovy matice je roven

$$\prod_{1\le j\le n} x_j \prod_{1\le i<j\le n} (x_j - x_i).$$

▸ **38.** [*M25*] Ukažte, že determinant Cauchyho matice je roven

$$\prod_{1\le i<j\le n} (x_j - x_i)(y_j - y_i) \Big/ \prod_{1\le i,j\le n} (x_i + y_j).$$

39. [*M23*] Ukažte, že inverzní matice ke kombinatorické matici je kombinatorická matice s prvky $b_{ij} = \big(-y + \delta_{ij}(x+ny)\big)/x(x+ny)$.

40. [*M24*] Ukažte, že inverzní matice k Vandermondově matici je dána vztahem

$$b_{ij} = \left(\sum_{\substack{1\le k_1<\dots<k_{n-j}\le n \\ k_1,\dots,k_{n-j}\ne i}} (-1)^{j-1} x_{k_1}\dots x_{k_{n-j}} \right) \Big/ x_i \prod_{\substack{1\le k\le n \\ k\ne i}} (x_k - x_i).$$

Nenechte se odradit komplikovanou sumou v čitateli – je to jen koeficient u x^{j-1} v polynomu $(x_1 - x)\dots(x_n - x)/(x_i - x)$.

41. [*M26*] Ukažte, že inverzní matice ke Cauchyho matici je dána vztahem

$$b_{ij} = \Bigg(\prod_{1\le k\le n}(x_j+y_k)(x_k+y_i)\Bigg)\Bigg/(x_j+y_i)\Bigg(\prod_{\substack{1\le k\le n\\ k\ne j}}(x_j-x_k)\Bigg)\Bigg(\prod_{\substack{1\le k\le n\\ k\ne i}}(y_i-y_k)\Bigg).$$

42. [*M18*] Čemu je roven součet všech n^2 prvků inverzní matice ke kombinatorické matici?

43. [*M24*] Čemu je roven součet všech n^2 prvků inverzní matice k Vandermondově matici? [*Pomůcka:* Využijte cvičení 33.]

▸ **44.** [*M26*] Čemu je roven součet všech n^2 prvků inverzní matice ke Cauchyho matici?

▸ **45.** [*M25*] *Hilbertova matice*, nazývaná někdy segmentem *nekonečné* Hilbertovy matice řádu $n\times n$, je taková matice, pro kterou je $a_{ij} = 1/(i+j-1)$. Ukažte, že se jedná o speciální případ Cauchyho matice, nalezněte k ní matici inverzní, ukažte, že všechny prvky inverzní matice jsou celočíselné a že součet všech prvků inverzní matice je n^2. [*Poznámka:* Pomocí Hilbertových matic se často testují různé algoritmy pro manipulaci s maticemi, protože jsou numericky nestabilní a jsou k nim známy inverzní matice. Je ale chybou porovnávat *známou* inverzní matici z tohoto cvičení s *vypočtenou* inverzní matici k Hilbertově matici, protože invertovanou matici je nutné předem vyjádřit v zaokrouhlených číslech; inverzní matice k přibližné verzi Hilbertovy matice se již bude od inverzní matice k přesné matici mírně lišit, a to díky zmíněné nestabilitě výpočtu. Protože prvky inverzní matice jsou celočíselné a protože inverzní matice je stejně nestabilní jako původní, můžeme inverzní matici vyjádřit přesně a můžeme se pokusit určit také inverzní matici k inverzní. Celočíselné prvky v inverzní matici jsou ale hodně velké.] K řešení tohoto problému jsou potřeba elementární znalosti faktoriálů a binomických koeficientů, které probíráme v částech 1.2.5 a 1.2.6.

▸ **46.** [*M30*] Nechť A je matice typu $m\times n$ a nechť B je matice typu $n\times m$. Pokud je dáno $1\le j_1, j_2, \ldots, j_m \le n$, nechť $A_{j_1 j_2 \ldots j_m}$ označuje matici typu $m\times m$ složenou ze sloupců $j_1, \ldots, j_m$ matice A, a nechť $B_{j_1 j_2 \ldots j_m}$ označuje matici typu $m\times m$ složenou z řádků $j_1, \ldots, j_m$ matice B. Dokažte *Cauchy-Binetovu větu*

$$\det(AB) = \sum_{1\le j_1<j_2<\cdots<j_m\le n} \det(A_{j_1 j_2 \ldots j_m})\det(B_{j_1 j_2 \ldots j_m}).$$

(Všimněte si speciálních případů: (i) $m=n$, (ii) $m=1$, (iii) $B=A^T$, (iv) $m>n$, (v) $m=2$.)

47. [*M27*] (C. Krattenthaler.) Dokažte, že

$$\det\begin{pmatrix}(x+q_2)(x+q_3) & (x+p_1)(x+q_3) & (x+p_1)(x+p_2)\\ (y+q_2)(y+q_3) & (y+p_1)(y+q_3) & (y+p_1)(y+p_2)\\ (z+q_2)(z+q_3) & (z+p_1)(z+q_3) & (z+p_1)(z+p_2)\end{pmatrix}$$
$$= (x-y)(x-z)(y-z)(p_1-q_2)(p_1-q_3)(p_2-q_3),$$

a zobecněte tuto rovnost na identitu pro determinant řádu $n\times n$ v $3n-2$ proměnných $x_1, \ldots, x_n, p_1, \ldots, p_{n-1}, q_2, \ldots, q_n$. Vzorec porovnejte s výsledkem cvičení 38.

1.2.4. Celočíselné funkce a teorie čísel

Pokud je x libovolné reálné číslo, píšeme

$\lfloor x \rfloor$ = největší celé číslo, které je menší nebo rovno x (*dolní celá část* čísla x);
$\lceil x \rceil$ = nejmenší celé číslo, které je větší nebo rovno x (*horní celá část* čísla x).

Před rokem 1970 se zápisem $[x]$ často vyjadřovaly obě z uvedených funkcí, obvykle pak první z nich; tyto notace, které zavedl K. E. Iverson v 60. letech, jsou ale užitečnější, protože výrazy $\lfloor x \rfloor$ a $\lceil x \rceil$ se v praxi vyskytují zhruba stejně často. Funkci $\lfloor x \rfloor$ se někdy jednoduše říká *celá část* neboli *entier*, z francouzského výrazu pro „celý, úplný“ (pod tímto názvem ji zná například programovací jazyk Algol 60).

Snadno si ověříme platnost následujících vzorců a příkladů:

$$\lfloor \sqrt{2} \rfloor = 1, \quad \lceil \sqrt{2} \rceil = 2, \quad \lfloor +\frac{1}{2} \rfloor = 0, \quad \lceil -\frac{1}{2} \rceil = 0, \quad \lfloor -\frac{1}{2} \rfloor = -1 \text{ (\textit{nikoli} nula!)};$$

$$\lceil x \rceil = \lfloor x \rfloor \qquad \text{tehdy a jen tehdy, je-li } x \text{ celé číslo},$$
$$\lceil x \rceil = \lfloor x \rfloor + 1 \quad \text{tehdy a jen tehdy, není-li } x \text{ celé číslo};$$
$$\lfloor -x \rfloor = -\lceil x \rceil; \qquad x - 1 < \lfloor x \rfloor \le x \le \lceil x \rceil < x + 1.$$

Ve cvičeních na konci této části textu najdete další důležité vzorce spojené s operacemi horní a dolní celé části.

Jsou-li x a y libovolná reálná čísla, definujeme následující binární operace:

$$x \bmod y = x - y\lfloor x/y \rfloor, \quad \text{je-li } y \neq 0; \qquad x \bmod 0 = x. \tag{1}$$

Z této definice jasně vidíme, že pokud je $y \neq 0$,

$$0 \le \frac{x}{y} - \left\lfloor \frac{x}{y} \right\rfloor = \frac{x \bmod y}{y} < 1\,. \tag{2}$$

Z toho vyplývá, že

a) je-li $y > 0$, pak $0 \le x \bmod y < y$;
b) je-li $y < 0$, pak $0 \ge x \bmod y > y$;
c) výraz $x - (x \bmod y)$ je celočíselným násobkem y.

Výraz $x \bmod y$ nazýváme *zbytek* po dělení x číslem y; podobně $\lfloor x/y \rfloor$ nazýváme (neúplný) *podíl*.

Jsou-li x a y dvě celá čísla, má operace „mod“ známý význam:

$$5 \bmod 3 = 2, \qquad 18 \bmod 3 = 0, \qquad -2 \bmod 3 = 1. \tag{3}$$

Přitom $x \bmod y = 0$ platí tehdy a jen tehdy, pokud je x násobkem y, neboli tehdy a jen tehdy, pokud je x dělitelné y. Zápis $y \backslash x$, který čteme „y dělí x“, pak znamená, že y je kladné celé číslo a $x \bmod y = 0$.

Operace „mod“ je ale užitečná také při libovolných reálných hodnotách x a y. U trigonometrických funkcí můžeme například psát

$$\tan x = \tan (x \bmod \pi).$$

Výraz $x \bmod 1$ tvoří *zlomkovou část* čísla x; podle (1) je

$$x = \lfloor x \rfloor + (x \bmod 1). \tag{4}$$

Autoři textů z teorie čísel zapisují často zkratku „mod“ v jiném, i když velmi příbuzném významu. Následujícím zápisem tak v teorii čísel vyjádříme pojem takzvané *kongruence*: Výrok

$$x \equiv y \pmod{z} \tag{5}$$

znamená, že $x \bmod z = y \bmod z$, což je totéž ‚jako když $x - y$ je celočíselným násobkem z. Výraz (5) pak čteme: „x je kongruentní s y modulo z“.

Nyní se podíváme na některé základní vlastnosti kongruencí, které budeme v knize používat při výkladu souvisejícím s teorií čísel. Všechny proměnné v následujících vztazích jsou celočíselné. O dvou celých číslech x a y říkáme, že jsou *nesoudělná*, jestliže nemají společného dělitele, a jestliže tedy jejich největší společný dělitel je roven 1; píšeme $x \perp y$. Pojem nesoudělných čísel se používá často a běžně se říká, že je zlomek zkrácen na „základní tvar“, pokud jsou čitatel a jmenovatel nesoudělná čísla.

Věta A. Pokud je $a \equiv b$ a $x \equiv y$, je i $a \pm x \equiv b \pm y$ a $ax \equiv by$ (modulo m).

Věta B. Pokud je $ax \equiv by$ a $a \equiv b$ a pokud dále $a \perp m$, je $x \equiv y$ (modulo m).

Věta C. $a \equiv b$ (modulo m) tehdy a jen tehdy, pokud $an \equiv bn$ (modulo mn), je-li zároveň $n \neq 0$.

Věta D. Pokud je $r \perp s$, pak $a \equiv b$ (modulo rs) tehdy a jen tehdy, je-li $a \equiv b$ (modulo r) a $a \equiv b$ (modulo s).

Věta A tedy říká, že sčítání, odčítání a násobení modulo m je možné provádět úplně stejně jako běžné sčítání, odčítání a násobení. Věta B hovoří o operaci dělení a ukazuje, že pokud je dělitel nesoudělný s modulem, můžeme také krátit společnými činiteli modulo m. Ve větách C a D se pak zabýváme změnou modulu. Všechna tvrzení si dokážeme ve cvičeních.

Důsledkem vět A a B je následující důležitá věta:

Věta F (*Malá Fermatova věta*, 1640). *Pokud p je prvočíslo, platí $a^p \equiv a$* (modulo p) *pro všechna celá a.*

Důkaz. Je-li a násobkem p, je zřejmě $a^p \equiv 0 \equiv a$ (modulo p). Musíme tedy rozebrat jen případ $a \bmod p \neq 0$. Protože p je prvočíslo, znamená to, že $a \perp p$. Uvažujme čísla

$$0 \bmod p, \quad a \bmod p, \quad 2a \bmod p, \quad \ldots, \quad (p-1)a \bmod p. \tag{6}$$

Všech těchto p čísel je *různých*, protože pokud by bylo $ax \bmod p = ay \bmod p$, pak podle definice (5) $ax \equiv ay$ (modulo p), a tudíž podle Věty B $x \equiv y$ (modulo p).

Protože ve vztahu (6) je definováno p různých čísel, která jsou všechna nezáporná a menší než p, vidíme, že první z čísel je nula a zbytek tvoří 1, 2, …, $p-1$ v nějakém pořadí. Podle Věty A je proto

$$(a)(2a) \ldots \big((p-1)a\big) \equiv 1 \cdot 2 \ldots (p-1) \pmod{p}. \tag{7}$$

Vynásobíme-li nyní obě strany této kongruence číslem a, dostaneme

$$a^p\big(1\cdot 2\ldots(p-1)\big)\equiv a\big(1\cdot 2\ldots(p-1)\big)\pmod{p};\tag{8}$$

a tím je věta dokázána, protože všechny dělitelé 1, 2, ..., $p-1$ jsou nesoudělné s p a podle Věty B je můžeme zkrátit. ∎

CVIČENÍ

1. [*00*] Kolik je $\lfloor 1.1\rfloor$, $\lfloor -1.1\rfloor$, $\lceil -1.1\rceil$, $\lfloor 0.99999\rfloor$, a $\lfloor \lg 35\rfloor$?

▸ **2.** [*01*] Kolik je $\lceil\lfloor x\rfloor\rceil$?

3. [*M10*] Nechť n je celé číslo a nechť x je reálné číslo. Dokažte, že:
a) $\lfloor x\rfloor < n$ tehdy a jen tehdy, je-li $x < n$;
b) $n \le \lfloor x\rfloor$ tehdy a jen tehdy, je-li $n \le x$;
c) $\lceil x\rceil \le n$ tehdy a jen tehdy, je-li $x \le n$;
d) $n < \lceil x\rceil$ tehdy a jen tehdy, je-li $n < x$;
e) $\lfloor x\rfloor = n$ tehdy a jen tehdy, je-li $x-1 < n \le x$, a tehdy a jen tehdy, je-li $n \le x < n+1$;
f) $\lceil x\rceil = n$ tehdy a jen tehdy, je-li $x \le n < x+1$, a tehdy a jen tehdy, je-li $n-1 < x \le n$.

[*Tyto vztahy jsou nejdůležitější při dokazování vlastností celých částí* $\lfloor x\rfloor$ *a* $\lceil x\rceil$.]

▸ **4.** [*M10*] Na základě předchozího cvičení dokažte, že $\lfloor -x\rfloor = -\lceil x\rceil$.

5. [*16*] Je-li dáno kladné reálné číslo x, napište jednoduchý vzorec, který vyjádří x *zaokrouhleno na nejbližší celé číslo*. Požadované pravidlo zaokrouhlování by tedy mělo dávat $\lfloor x\rfloor$ pro $x \bmod 1 < \frac12$, respektive $\lceil x\rceil$ pro $x \bmod 1 \ge \frac12$. Jako odpověď uveďte jediný vzorec, který pokryje oba případy. Rozeberte, jaké zaokrouhlení ze vzorce vznikne, bude-li x záporné číslo.

▸ **6.** [*20*] Které z následujících tvrzení platí pro všechna kladná reálná čísla x?
(a) $\lfloor\sqrt{\lfloor x\rfloor}\,\rfloor = \lfloor\sqrt{x}\,\rfloor$; (b) $\lceil\sqrt{\lceil x\rceil}\,\rceil = \lceil\sqrt{x}\,\rceil$; (c) $\lceil\sqrt{\lfloor x\rfloor}\,\rceil = \lceil\sqrt{x}\,\rceil$.

7. [*M15*] Ukažte, že $\lfloor x\rfloor + \lfloor y\rfloor \le \lfloor x+y\rfloor$, přičemž tato rovnost platí tehdy a jen tehdy, je-li $x \bmod 1 + y \bmod 1 < 1$. Platí podobný vztah i pro horní celou část?

8. [*00*] Kolik je 100 mod 3, 100 mod 7, −100 mod 7, −100 mod 0?

9. [*05*] Kolik je 5 mod −3, 18 mod −3, −2 mod −3?

▸ **10.** [*10*] Kolik je 1,1 mod 1; 0,11 mod 0,1; 0,11 mod −0,1?

11. [*00*] Co podle našich dohod znamená „$x \equiv y$ (modulo 0)“?

12. [*00*] Jaká celá čísla jsou nesoudělná s 1?

13. [*M00*] Podle dohody říkáme, že největší společný dělitel 0 a n je $|n|$. Která celá čísla jsou tedy nesoudělná s nulou?

▸ **14.** [*12*] Je-li $x \bmod 3 = 2$ a $x \bmod 5 = 3$, kolik je $x \bmod 15$?

15. [*10*] Dokažte, že $z(x \bmod y) = (zx) \bmod (zy)$. [Věta C je přímým důsledkem tohoto distributivního zákona.]

16. [*M10*] Předpokládejme, že $y > 0$. Ukažte, že pokud $(x-z)/y$ je celé číslo a pokud $0 \le z < y$, pak $z = x \bmod y$.

17. [*M15*] Dokažte větu A přímo z definice kongruence; dále dokažte polovinu věty D: Je-li $a \equiv b$ (modulo rs), pak $a \equiv b$ (modulo r) a $a \equiv b$ (modulo s). (Zde jsou r a s libovolná celá čísla.)

18. [*M15*] Pomocí věty B dokažte druhou polovinu věty D: Je-li $a \equiv b$ (modulo r) a $a \equiv b$ (modulo s), pak $a \equiv b$ (modulo rs) za předpokladu, že $r \perp s$.

▸ **19.** [*M10*] (*Zákon inverzí.*) Je-li $n \perp m$, pak existuje celé číslo n' takové, že $nn' \equiv 1$ (modulo m). Dokažte pomocí rozšířené verze Euklidova algoritmu (Algoritmus 1.2.1E).

20. [*M15*] Pomocí zákona inverzí a věty A dokažte větu B.

21. [*M22*] (*Základní věta aritmetiky.*) Pomocí věty B a cvičení 1.2.1–5 dokažte, že pro každé celé číslo $n > 1$ existuje jeho *jednoznačná* reprezentace jako součin prvočísel (až na pořadí činitelů). Jinými slovy: dokažte, že existuje právě jeden způsob zápisu $n = p_1 p_2 \ldots p_k$, kde každé p_j je prvočíslo a $p_1 \le p_2 \le \cdots \le p_k$.

▸ **22.** [*M10*] Na příkladu ukažte, že věta B nemusí vždy platit, není-li a nesoudělné s m.

23. [*M10*] Na příkladu ukažte, že věta D nemusí vždy platit, není-li r nesoudělné s s.

▸ **24.** [*M20*] Do jaké míry je možné zobecnit věty A, B, C a D, aby platily nejen pro celá čísla, ale i pro libovolná reálná čísla?

25. [*M02*] Ukažte, že podle věty F $a^{p-1} \bmod p = [a$ není násobkem $p]$, pokud je p prvočíslo.

26. [*M15*] Nechť p je liché prvočíslo, nechť a je libovolné celé číslo a nechť $b = a^{(p-1)/2}$. Ukažte, že $b \bmod p$ je buďto 0, nebo 1, anebo $p-1$. [*Nápověda:* Uvažujte $(b+1)(b-1)$.]

27. [*M15*] Je-li n kladné celé, definujeme $\varphi(n)$ jako počet čísel z množiny $\{0, 1, \ldots, n-1\}$, která jsou nesoudělná s n. Tudíž $\varphi(1) = 1$, $\varphi(2) = 1$, $\varphi(3) = 2$, $\varphi(4) = 2$ atd. Ukažte, že $\varphi(p) = p - 1$, pokud je p prvočíslo, a vypočtěte $\varphi(p^e)$, kde e je kladné celé číslo.

▸ **28.** [*M25*] Ukažte, že stejným postupem jako Větu F je možné dokázat i její následující rozšíření nazývané *Eulerova věta*: $a^{\varphi(m)} \equiv 1$ (modulo m) pro *libovolné* kladné celé číslo m takové, že $a \perp m$. (Zejména číslo n' ve cvičení 19 můžeme převést na $n^{\varphi(m)-1} \bmod m$.)

29. [*M22*] Funkci $f(n)$ kladné celočíselné proměnné n nazýváme *multiplikativní*, pokud je $f(rs) = f(r)f(s)$ pro každé $r \perp s$. Ukažte, že všechny následující funkce jsou multiplikativní: (a) $f(n) = n^c$, kde c je libovolná konstanta; (b) $f(n) = [n$ není dělitelné k^2 pro žádné celé číslo $k > 1]$; (c) $f(n) = c^k$, kde k je počet různých prvočísel, která dělí n; (d) součin libovolných dvou multiplikativních funkcí.

30. [*M30*] Dokažte, že funkce $\varphi(n)$ z cvičení 27 je multiplikativní. Na základě tohoto tvrzení vypočtěte $\varphi(1000000)$ a uveďte metodu pro jednoduchý výpočet $\varphi(n)$, při kterém využijeme rozkladu n na prvočísla.

31. [*M22*] Dokažte, že pokud je $f(n)$ multiplikativní, je multiplikativní i $g(n) = \sum_{d \backslash n} f(d)$.

32. [*M18*] Dokažte rovnost dvojitých sum

$$\sum_{d \backslash n} \sum_{c \backslash d} f(c,d) = \sum_{c \backslash n} \sum_{d \backslash (n/c)} f(c, cd)$$

pro libovolnou funkci $f(x, y)$.

33. [*M18*] Jsou dána dvě celá čísla m a n; vypočtěte (a) $\lfloor \frac{1}{2}(n+m) \rfloor + \lfloor \frac{1}{2}(n-m+1) \rfloor$; (b) $\lceil \frac{1}{2}(n+m) \rceil + \lceil \frac{1}{2}(n-m+1) \rceil$. (Zvláštní pozornost si zasluhuje speciální případ $m = 0$.)

▸ **34.** [*M21*] Jakou nutnou a postačující podmínku musí splňovat reálné číslo $b > 1$, aby bylo $\lfloor \log_b x \rfloor = \lfloor \log_b \lfloor x \rfloor \rfloor$ pro všechna reálná $x \ge 1$?

▸ **35.** [*M20*] Jsou-li dána celá čísla m a n, přičemž $n > 0$, dokažte, že

$$\lfloor (x+m)/n \rfloor = \lfloor (\lfloor x \rfloor + m)/n \rfloor$$

pro všechna reálná x. (Pokud je $m = 0$, jedná se o důležitý speciální případ.) Platí analogický výsledek také pro funkci horní celé části?

36. [*M23*] Dokažte, že $\sum_{k=1}^{n} \lfloor k/2 \rfloor = \lfloor n^2/4 \rfloor$; dále vypočtěte $\sum_{k=1}^{n} \lceil k/2 \rceil$.

▸ **37.** [*M30*] Nechť m a n jsou dvě celá čísla, $n > 0$. Ukažte, že

$$\sum_{0 \le k < n} \left\lfloor \frac{mk+x}{n} \right\rfloor = \frac{(m-1)(n-1)}{2} + \frac{d-1}{2} + d\lfloor x/d \rfloor,$$

kde d je největší společný dělitel m a n a kde x je libovolné reálné číslo.

38. [*M26*] (E. Busche, 1909.) Dokažte, že pro všechna reálná x a y, přičemž $y > 0$, platí

$$\sum_{0 \le k < y} \left\lfloor x + \frac{k}{y} \right\rfloor = \lfloor xy + \lfloor x+1 \rfloor (\lceil y \rceil - y) \rfloor.$$

Zejména pokud y je kladné celé číslo n, dostáváme důležitý vzorec

$$\lfloor x \rfloor + \left\lfloor x + \frac{1}{n} \right\rfloor + \cdots + \left\lfloor x + \frac{n-1}{n} \right\rfloor = \lfloor nx \rfloor.$$

39. [*VM35*] Funkci f, definovanou jako $f(x)+f\left(x+\frac{1}{n}\right)+\cdots+f\left(x+\frac{n-1}{n}\right) = f(nx)$, pro libovolné kladné celé číslo n, nazýváme *replikativní funkcí*. V předchozím cvičení jsme si ukázali, že $\lfloor x \rfloor$ je replikativní. Nyní dokažte, že i následující funkce jsou replikativní:

a) $f(x) = x - \frac{1}{2}$;
b) $f(x) = [x$ je celé číslo$]$;
c) $f(x) = [x$ je kladné celé číslo$]$;
d) $f(x) = [$existuje racionální číslo r a celé číslo m takové, že $x = r\pi + m]$;
e) tři další funkce jako v (d), přičemž r a/nebo m jsou omezeny jen na kladné hodnoty;
f) $f(x) = \log|2\sin\pi x|$, pokud je povolena hodnota $f(x) = -\infty$;
g) součet libovolných dvou replikativních funkcí;
h) konstantní násobek replikativní funkce;
i) funkce $g(x) = f(x - \lfloor x \rfloor)$, kde $f(x)$ je replikativní.

40. [*VM46*] Prostudujte třídu replikativních funkcí; vyšetřete všechny replikativní funkce speciálního typu. Je například funkce z části (a) cvičení 39 jedinou spojitou replikativní funkcí? Bylo by zajímavé studovat také obecnější třídu funkcí, pro které

$$f(x) + f\left(x + \frac{1}{n}\right) + \cdots + f\left(x + \frac{n-1}{n}\right) = a_n f(nx) + b_n.$$

Zde jsou a_n a b_n čísla, která závisí na n, ale ne na x. Derivace a (za podmínky $b_n = 0$) integrály těchto funkcí jsou stejného typu. Požadujeme-li $b_n = 0$, dostaneme například Bernoulliho polynomy, trigonometrické funkce $\cot\pi x$ a $\csc^2\pi x$ a také Hurwitzovu zobecněnou funkci zeta $\zeta(s,x) = \sum_{k\ge 0} 1/(k+x)^s$ pro pevné s. Pro $b_n \neq 0$ dostáváme další známé funkce, například funkci ψ.

41. [*M23*] Nechť $a_1, a_2, a_3, \ldots$ je posloupnost 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, ...; najděte výraz pro zjištění a_n z n, přičemž využijete funkce dolní a/nebo horní celé části.

42. [*M24*] (a) Dokažte, že

$$\sum_{k=1}^{n} a_k = na_n - \sum_{k=1}^{n-1} k(a_{k+1} - a_k), \qquad \text{pokud } n > 0.$$

(b) Předchozí vzorec je užitečný při výpočtu jistých sum, ve kterých se používá funkce dolní celé části. Dokažte, že pokud je b celé číslo ≥ 2,

$$\sum_{k=1}^{n} \lfloor \log_b k \rfloor = (n+1)\lfloor \log_b n \rfloor - \big(b^{\lfloor \log_b n \rfloor + 1} - b\big)/(b-1).$$

43. [*M23*] Vypočtěte $\sum_{k=1}^{n} \lfloor \sqrt{k}\, \rfloor$.

44. [*M24*] Ukažte, že $\sum_{k\geq 0} \sum_{1\leq j<b} \lfloor (n + jb^k)/b^{k+1} \rfloor = n$, pokud jsou b a n celá čísla taková, že $n \geq 0$ a $b \geq 2$. Čemu je tato suma rovna pro $n < 0$?

▸ **45.** [*M28*] Výsledek cvičení 37 je poněkud překvapující, protože z něj vyplývá, že

$$\sum_{0\leq k<n} \left\lfloor \frac{mk+x}{n} \right\rfloor = \sum_{0\leq k<m} \left\lfloor \frac{nk+x}{m} \right\rfloor.$$

Tato „relace reciprocity“ je jedním z mnoha podobných vztahů (viz část 3.3.3). Ukažte, že pro libovolnou funkci f platí

$$\sum_{0\leq j<n} f\left(\left\lfloor \frac{mj}{n} \right\rfloor\right) = \sum_{0\leq r<m} \left\lceil \frac{rn}{m} \right\rceil \big(f(r-1) - f(r)\big) + nf(m-1).$$

Zejména pak dokažte, že

$$\sum_{0\leq j<n} \binom{\lfloor mj/n \rfloor + 1}{k} + \sum_{0\leq j<m} \left\lceil \frac{jn}{m} \right\rceil \binom{j}{k-1} = n\binom{m}{k}.$$

[*Nápověda:* Uvažte záměnu proměnné $r = \lfloor mj/n \rfloor$. O binomických koeficientech $\binom{m}{k}$ hovoříme v části 1.2.6.]

46. [*M29*] (*Obecný zákon reciprocity.*) Rozšiřte vzorec ze cvičení 45 na výraz pro výpočet sumy $\sum_{0\leq j<\alpha n} f(\lfloor mj/n \rfloor)$, kde α je *libovolné* kladné reálné číslo.

▸ **47.** [*M31*] Je-li p liché prvočíslo, definujeme *Legendrův symbol* $\left(\frac{q}{p}\right)$ jako $+1$, 0 nebo -1 podle toho, je-li $q^{(p-1)/2} \bmod p$ rovno 1, 0 nebo $p-1$. (Ve cvičení 26 jsme dokázali, že tyto tři hodnoty jsou jediné možné.)

a) Jestliže q není násobkem p, ukažte, že čísla

$$(-1)^{\lfloor 2kq/p \rfloor}(2kq \bmod p), \qquad 0 < k < p/2,$$

jsou v nějakém pořadí kongruentní s čísly 2, 4, ..., $p-1$ (modulo p). Je tudíž $\left(\frac{q}{p}\right) = (-1)^{\sigma}$, kde $\sigma = \sum_{0\leq k<p/2} \lfloor 2kq/p \rfloor$.

b) Pomocí výsledku (a) vypočtěte $\left(\frac{2}{p}\right)$.

c) Je-li dáno liché číslo q, ukažte, že $\sum_{0\leq k<p/2} \lfloor 2kq/p \rfloor \equiv \sum_{0\leq k<p/2} \lfloor kq/p \rfloor \bmod 2$. [*Nápověda:* Uvažujte hodnotu $\lfloor (p-1-2k)q/p \rfloor$.]

d) Podle obecného zákonu reciprocity ze cvičení 46 odvoďte *zákon kvadratické reciprocity*, $\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{(p-1)(q-1)/4}$, pokud jsou p a q dvě různá lichá prvočísla.

48. [*M26*] Dokažte nebo vyvraťte následující rovnosti pro celá čísla m a n:

$$\text{(a)}\ \left\lfloor \frac{m+n-1}{n} \right\rfloor = \left\lceil \frac{m}{n} \right\rceil; \qquad \text{(b)}\ \left\lfloor \frac{n+2-\lfloor n/25 \rfloor}{3} \right\rfloor = \left\lfloor \frac{8n+24}{25} \right\rfloor.$$

49. [*M30*] Předpokládejme, že celočíselná funkce $f(x)$ splňuje dvě jednoduché podmínky: (i) $f(x+1) = f(x)+1$; (ii) $f(x) = f(f(nx)/n)$ pro všechna kladná celá n. Dokažte, že buďto pro všechna racionální x platí $f(x) = \lfloor x \rfloor$, nebo pro všechna racionální x platí $f(x) = \lceil x \rceil$.

1.2.5. Permutace a faktoriály

Permutace n prvků je uspořádání n různých prvků do řady. Tři prvky $\{a, b, c\}$ tak můžeme uspořádat do šesti různých permutací:

$$a\,b\,c, \qquad a\,c\,b, \qquad b\,a\,c, \qquad b\,c\,a, \qquad c\,a\,b, \qquad c\,b\,a. \tag{1}$$

Vlastnosti permutací mají velký význam při analýze algoritmů a později v této knížce o nich odvodíme řadu zajímavých tvrzení.* Naším prvním úkolem bude permutace *spočítat*: Kolik je možných permutací n prvků? První (nejlevější) prvek můžeme vybrat n způsoby, a po provedení tohoto výběru nám na druhé místo zbývá $n-1$ prvků; pro první dvě pozice tak máme $n(n-1)$ možností. Podobně zjistíme, že třetí prvek, má-li být různý od prvních dvou, můžeme vybrat $n-2$ způsoby, takže pro první tři prvky již máme $n(n-1)(n-2)$ možností. Jestliže počet způsobů, kterými vybereme a uspořádáme k prvků z n, označíme jako p_{nk}, dostáváme

$$p_{nk} = n(n-1)\ldots(n-k+1). \tag{2}$$

Celkový počet permutací je tudíž roven $p_{nn} = n(n-1)\ldots(1)$.

Proces *vytvoření* všech permutací n prvků můžeme vést indukcí, přičemž vycházíme z již vytvořených všech permutací $n-1$ prvků; také v našich aplikacích bude tento postup důležitý. Nyní přepíšeme permutace (1) s tím, že namísto písmen $\{a, b, c\}$ použijeme čísla $\{1, 2, 3\}$; dostaneme

$$1\,2\,3, \qquad 1\,3\,2, \qquad 2\,1\,3, \qquad 2\,3\,1, \qquad 3\,1\,2, \qquad 3\,2\,1. \tag{3}$$

Dále se tedy zamyslíme, jak z nich odvodit permutace prvků $\{1, 2, 3, 4\}$. Z $n-1$ prvků na n prvků se přitom můžeme dostat dvěma základním způsoby.

Metoda 1. Pro každou permutaci $a_1\,a_2\,\ldots\,a_{n-1}$ prvků $\{1, 2, \ldots, n-1\}$ vytvoříme n nových tak, že vložíme číslo (prvek) n na všechny možné pozice; dostaneme

$$n\,a_1\,a_2\,\ldots\,a_{n-1}, \quad a_1\,n\,a_2\,\ldots\,a_{n-1}, \quad \ldots, \quad a_1\,a_2\,\ldots\,n\,a_{n-1}, \quad a_1\,a_2\,\ldots\,a_{n-1}\,n.$$

Z permutace 2 3 1 v (3) máme tak například 4 2 3 1, 2 4 3 1, 2 3 4 1, 2 3 1 4. Je zřejmé, že tímto způsobem získáme všechny permutace n prvků a že žádnou z nich neodvodíme více než jednou.

* Permutace jsou natolik důležité, že Vaughan Pratt navrhl, aby se jim říkalo stručně „perms“. Kdyby se tento jeho návrh prosadil, mohly by se učebnice informatiky o něco zkrátit (a možná i zlevnit).

Metoda 2. Pro každou permutaci $a_1\,a_2\,\ldots\,a_{n-1}$ prvků $\{1,2,\ldots,n{-}1\}$ vytvoříme n nových následujícím způsobem: Nejprve vytvoříme pole

$$a_1\,a_2\,\ldots\,a_{n-1}\,\tfrac{1}{2},\quad a_1\,a_2\,\ldots\,a_{n-1}\,\tfrac{3}{2},\quad \ldots,\quad a_1\,a_2\,\ldots\,a_{n-1}\,\bigl(n-\tfrac{1}{2}\bigr).$$

Poté přejmenujeme prvky každé permutace na čísla $\{1,2,\ldots,n\}$, a to *se zachováním pořadí*. Z permutace 2 3 1 v (3) dostaneme například

$$2\,3\,1\,\tfrac{1}{2},\quad 2\,3\,1\,\tfrac{3}{2},\quad 2\,3\,1\,\tfrac{5}{2},\quad 2\,3\,1\,\tfrac{7}{2}$$

a dále po přejmenování

$$3\,4\,2\,1,\quad 3\,4\,1\,2,\quad 2\,4\,1\,3,\quad 2\,3\,1\,4.$$

Postup můžeme popsat ještě jinak: vezmeme permutaci $a_1a_2\ldots a_{n-1}$ a číslo k, $1\le k\le n$, a dále přičteme jedničku ke každému a_j, jehož hodnota je $\ge$ $\ge k$; dostaneme tudíž permutaci $b_1b_2\ldots b_{n-1}$ prvků $\{1,\ldots,k-1,k+1,\ldots,n\}$ a konečně $b_1b_2\ldots b_{n-1}k$ je permutací prvků $\{1,\ldots,n\}$.

Opět je zřejmé, že pomocí této konstrukce dostaneme každou permutaci n prvků právě jednou. Stejně bychom mohli číslo k umístit nikoli vpravo, ale vlevo nebo na jinou pevnou pozici.

Je-li p_n počet permutací n prvků, pak z obou uvedených metod vidíme, že $p_n = np_{n-1}$; to znamená další dva důkazy rovnosti $p_n = n(n-1)\ldots(1)$, kterou jsme si již ukázali ve vztahu (2).

Počet permutací p_n je důležitá veličina; označuje ji takzvaný n *faktoriál* a zapisujeme

$$n! = 1\cdot 2\cdot\ldots\cdot n = \prod_{k=1}^{n} k. \tag{4}$$

Podle dohody o prázdném součinu (viz část 1.2.3) musí být

$$0! = 1 \tag{5}$$

a s touto dohodou již základní rovnost

$$n! = (n-1)!\,n \tag{6}$$

platí pro všechna kladná celá n.

Faktoriály se při práci s počítači vyskytují tak často, že je vhodné si zapamatovat hodnoty alespoň prvních několika:

$$0! = 1,\quad 1! = 1,\quad 2! = 2,\quad 3! = 6,\quad 4! = 24,\quad 5! = 120.$$

Hodnoty faktoriálů velmi rychle rostou; například 1000! je celé číslo s více než 2500 desítkovými číslicemi.

Je dobré si zapamatovat také hodnotu 10! = 3 628 800; stačí vědět, že 10! je zhruba $3\frac{1}{2}$ miliónu. Toto číslo v jistém slova smyslu přibližně rozděluje problémy, které se dají prakticky spočítat a které se spočítat nedají. Jestliže algoritmus zkouší více než 10! případů, znamená to větší spotřebu času, než jaká je prakticky rozumná. Na druhé straně pokud se rozhodneme vyzkoušet těchto 10! případů a každý z nich zabere například jednu milisekundu strojového času,

poběží výpočet zhruba hodinu. Tyto úvahy jsou samozřejmě velmi vágní, ale získáme z nich alespoň nějakou představu, co je a není výpočetně proveditelné.

Je přirozené se ptát, jaký vztah má $n!$ k jiným matematickým veličinám. Dá se nějak určit třeba 1000!, aniž bychom pracně prováděli všechna násobení podle vztahu (4)? Odpověď nalezl James Stirling a uvedl ji ve svém slavném díle *Methodus Differentialis* (1730), strana 137; vzorec je

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n. \tag{7}$$

Znaménko „$\approx$“ přitom znamená „je přibližně rovno“ a „e“ je základ přirozených logaritmů, které jsme si zavedli v části 1.2.2. Stirlingovu aproximaci (7) dokážeme v části 1.2.11.2; ve cvičení 24 nás čeká jednoduchý důkaz méně přesného vztahu.

Jako příklad využití tohoto vzorce můžeme nyní vypočítat

$$40320 = 8! \approx 4\sqrt{\pi}\left(\frac{8}{e}\right)^8 = 2^{26}\sqrt{\pi}\,e^{-8} \approx 67108864 \cdot 1,77245 \cdot 0,00033546 \approx 39902.$$

V tomto případě je chyba zhruba 1 %; později uvidíme, že relativní chyba se pohybuje okolo $1/(12n)$.

Kromě přibližné hodnoty uvedené ve vztahu (7) můžeme také poměrně snadno zjistit přesnou hodnotu $n!$ rozloženou na prvočísla. Prvočíslo p je ve skutečnosti dělitelem $n!$ s násobností

$$\mu = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots = \sum_{k>0} \left\lfloor \frac{n}{p^k} \right\rfloor. \tag{8}$$

Pokud je například $n = 1000$ a $p = 3$, dostáváme

$$\begin{aligned}\mu &= \left\lfloor \frac{1000}{3} \right\rfloor + \left\lfloor \frac{1000}{9} \right\rfloor + \left\lfloor \frac{1000}{27} \right\rfloor + \left\lfloor \frac{1000}{81} \right\rfloor + \left\lfloor \frac{1000}{243} \right\rfloor + \left\lfloor \frac{1000}{729} \right\rfloor \\ &= 333 + 111 + 37 + 12 + 4 + 1 = 498,\end{aligned}$$

takže 1000! je dělitelné 3^{498}, ale ne 3^{499}. I když vzorec (8) je zapsán jako nekonečná suma, je ve skutečnosti pro každou hodnotu n a p konečný, protože všechny členy jsou od jisté pozice rovny nule. Ze cvičení 1.2.4–35 vyplývá, že $\lfloor n/p^{k+1} \rfloor = \lfloor \lfloor n/p^k \rfloor / p \rfloor$; tato rovnost nám usnadní výpočet ve vztahu (8), protože můžeme dělit hodnotu předchozího členu číslem p a zbytek zahodit.

Rovnost (8) vyplývá z toho, že $\lfloor n/p^k \rfloor$ je počet celých čísel v množině $\{1, 2, \ldots, n\}$, které jsou násobkem p^k. Jestliže se nyní podíváme na celá čísla v součinu (4), počítá se zde každé celé číslo, které je dělitelné p^j, ale už ne p^{j+1}, právě j-krát: jednou v $\lfloor n/p \rfloor$, jednou v $\lfloor n/p^2 \rfloor$, ..., jednou v $\lfloor n/p^j \rfloor$. Tím postihneme všechny výskyty prvočísla p v rozkladu $n!$. [Viz A. M. Legendre, *Essai sur la Théorie des Nombres*, druhé vydání (Paříž: 1808), strana 8.]

Nyní přichází další přirozená otázka: jestliže jsme definovali $n!$ pro všechna nezáporná celá n, nemá smysl funkce faktoriálu také pro racionální n a pro reálná čísla? Kolik je třeba $\left(\frac{1}{2}\right)!$? Tuto myšlenku si ilustrujeme zavedením „termiálové“

funkce

$$n? = 1 + 2 + \cdots + n = \sum_{k=1}^{n} k, \tag{9}$$

která je analogií k faktoriálu, pouze jsme násobení zaměnili za sčítání. Součet této aritmetické řady již známe; podle vztahu 1.2.3–(15) je to:

$$n? = \tfrac{1}{2}n(n+1). \tag{10}$$

Odtud můžeme „termiálovou“ funkci snadno zobecnit na libovolné n, přičemž namísto vztahu (9) použijeme (10). Máme tak například $(\frac{1}{2})? = \frac{3}{8}$.

Sám Stirling učinil několik pokusů o zobecnění $n!$ na necelá n. Aproximaci (7) rozšířil na nekonečný součet, který ale bohužel pro žádnou hodnotu n nekonvergoval; z jeho metody vznikla mimořádně dobrá aproximace, ale nedala se rozšířit na *přesnou* hodnotu. [Podrobný výklad této poněkud neobvyklé situace viz K. Knopp, *Theory and Application of Infinite Series*, 2. vydání (Glasgow: Blackie, 1951), 518–520, 527, 534.]

Stirling pracoval dále a všiml si, že

$$\begin{aligned} n! = 1 &+ \Big(1 - \frac{1}{1!}\Big)n + \Big(1 - \frac{1}{1!} + \frac{1}{2!}\Big)n(n-1) \\ &+ \Big(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!}\Big)n(n-1)(n-2) + \cdots \end{aligned} \tag{11}$$

(Tento vzorec dokážeme v další části textu.) Zdánlivě nekonečný součet ve vztahu (11) je ve skutečnosti pro každé nezáporné celé n konečný; nedostáváme z něj ale kýžené zobecnění $n!$, protože nekonečný součet neexistuje, právě *kromě* případu, kdy n je nezáporné celé číslo. (Viz cvičení 16.)

Stále odhodlaný Stirling nalezl nakonec takovou posloupnost $a_1, a_2, \ldots$, pro kterou

$$\ln n! = a_1 n + a_2 n(n-1) + \cdots = \sum_{k\ge 0} a_{k+1} \prod_{0\le j\le k} (n-j). \tag{12}$$

Neuměl již ale *dokázat*, že tato suma definuje $n!$ pro všechny racionální hodnoty n, i když odvodil hodnotu $(\frac{1}{2})! = \sqrt{\pi}/2$.

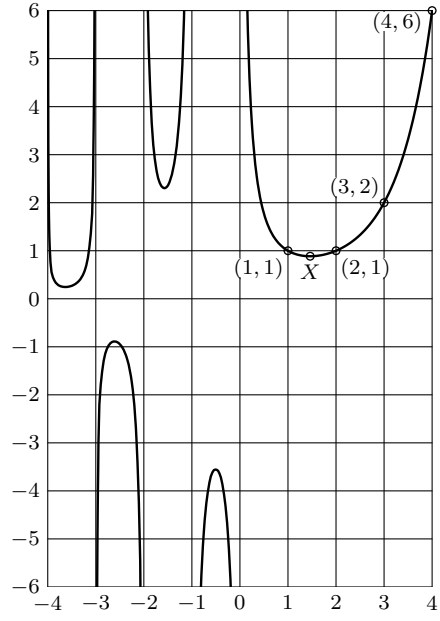
Zhruba ve stejné době se stejným problémem zabýval také Leonhard Euler, který jako první nalezl vhodné zobecnění:

$$n! = \lim_{m\to\infty} \frac{m^n m!}{(n+1)(n+2)\ldots(n+m)}. \tag{13}$$

Tuto myšlenku sdělil Euler dne 13. října 1729 v dopise Christianu Goldbachovi. Jeho vzorec definuje $n!$ pro libovolnou hodnotu n s výjimkou záporných celých čísel (kdy je jmenovatel nulový); v takovém případě se $n!$ klade roven nekonečnu. Ve cvičeních 8 a 22 si vysvětlíme, proč je definice podle vztahu (13) rozumná.

Téměř o dvě století později, v roce 1900, C. Hermite dokázal, že Stirlingův vzorec (12) skutečně definuje $n!$ pro necelá čísla n a že Eulerovo a Stirlingovo zobecnění jsou fakticky totožné.

Kdysi se faktoriály zapisovaly mnoha různými notacemi. Euler je zapisoval $[n]$, Gauss jako Πn a v Anglii a Itálii se rozšířily symboly $\underline{\lfloor n}$ a $\underline{n\rfloor}$. Notaci $n!$,



Obr. 7. Funkce $\Gamma(x) = (x-1)!$. Lokální minimum funkce v bodě X má souřadnice $(1,46163\,21449\,68362\,34126\,26595,\ 0,88560\,31944\,10888\,70027\,88159)$.

která se dnes pro celočíselná n používá, zavedl poměrně málo známý matematik Christian Kramp v algebraickém spisu [*Élémens d'Arithmétique Universelle* (Köln: 1808), strana 219].

Jestliže ale n *není* celé, je notace $n!$ méně obvyklá; spíše používáme jinou notaci, kterou zavedl A. M. Legendre:

$$n! = \Gamma(n+1) = n\Gamma(n). \tag{14}$$

Této funkci $\Gamma(x)$ se říká *funkce gama*; podle (13) tak můžeme zapsat definici

$$\Gamma(x) = \frac{x!}{x} = \lim_{m\to\infty} \frac{m^x m!}{x(x+1)(x+2)\ldots(x+m)}. \tag{15}$$

Graf funkce $\Gamma(x)$ je zachycen na obrázku 7.

Vztahy (13) a (15) definují faktoriály a funkci gama nejen pro reálné, ale i pro komplexní hodnoty; proměnnou, která má reálnou i imaginární část, ale zpravidla namísto písmene n či x označujeme jako z. Faktoriál a funkce gama jsou spojeny nejen pravidlem $z! = \Gamma(z+1)$, ale také vztahem

$$(-z)!\,\Gamma(z) = \frac{\pi}{\sin \pi z}, \tag{16}$$

který platí pro každé necelé z. (Viz cvičení 23.)

I když funkce $\Gamma(z)$ je pro nulu a pro záporná celá z nekonečná, funkce $1/\Gamma(z)$ je dobře definována pro všechna komplexní z. (Viz cvičení 1.2.7–24.)

V pokročilých aplikacích funkce gama se často používá důležitý vzorec pro křivkový integrál, který odvodil Hermann Hankel:

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i}\oint \frac{e^t\,dt}{t^z}; \tag{17}$$

přičemž křivka pro komplexní integraci začíná v $-\infty$, poté vede okolo počátku proti směru hodinových ručiček a nakonec se vrací k $-\infty$. [*Zeitschrift für Math. und Physik* **9** (1864), 1–21.]

V mnoha vzorcích diskrétní matematiky se používají podobné součiny jako faktoriály, kterým se říká *faktoriálové mocniny* neboli Pochhammerovy symboly. Veličiny $x^{\underline{k}}$ a $x^{\overline{k}}$ (čteme „x na klesající k“ a „x na rostoucí k“) jsou definovány takto (kde k je kladné celé číslo):

$$x^{\underline{k}} = x(x-1)\ldots(x-k+1) = \prod_{j=0}^{k-1}(x-j); \tag{18}$$

$$x^{\overline{k}} = x(x+1)\ldots(x+k-1) = \prod_{j=0}^{k-1}(x+j). \tag{19}$$

Číslo p_{nk} ze vztahu (2) je tedy například rovno $n^{\underline{k}}$. Všimněte si, že máme

$$x^{\overline{k}} \;=\; (x+k-1)^{\underline{k}} \;=\; (-1)^k(-x)^{\underline{k}}. \tag{20}$$

Z obecných vztahů

$$x^{\underline{k}} = \frac{x!}{(x-k)!}\,, \qquad\qquad x^{\overline{k}} = \frac{\Gamma(x+k)}{\Gamma(x)} \tag{21}$$

je pak možné definovat faktoriálové mocniny pro jiné hodnoty k. [Notace $x^{\overline{k}}$ a $x^{\underline{k}}$ zavedli postupně A. Capelli, *Giornale di Mat. di Battaglini* **31** (1893), 291–313, a L. Toscano, *Comment. Accademia della Scienze* **3** (1939), 721–757.]

Zajímavou historii faktoriálů od dob Stirlinga až do současnosti shrnuje P. J. Davis v článku „Leonhard Euler's integral: A historical profile of the gamma function“, *AMM* **66** (1959), 849–869. Viz též J. Dutka, *Archive for History of Exact Sciences* **31** (1984), 15–34.

CVIČENÍ

1. [*00*] Kolika způsoby je možné zamíchat balíček 52 karet?

2. [*10*] V notaci podle vztahu (2) ukažte, že je $p_{n(n-1)} = p_{nn}$, a vysvětlete proč.

3. [*10*] Jaké permutace čísel $\{1,2,3,4,5\}$ vytvoříme z permutace 3 1 2 4 pomocí Metody 1, respektive Metody 2?

▸ **4.** [*13*] Jestliže známe $\log_{10} 1000! = 2567,60464\ldots$, zjistěte, kolik přesně desítkových číslic je potřeba k zápisu 1000!. Jaká je *nejvýznamnější* číslice? A jaká je *nejméně významná* číslice?

5. [*15*] Odhadněte 8! s pomocí následující přesnější varianty Stirlingovy aproximace :

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1+\frac{1}{12n}\right).$$

▸ **6.** [*17*] Podle vztahu (8) napište 20! jako součin prvočinitelů.

7. [*M10*] Ukažte, že „zobecněná termiálová“ funkce ve vztahu (10) splňuje rovnost $x? = x + (x-1)?$ pro všechna reálná čísla x.

8. [*VM15*] Ukažte, že limita ve vztahu (13) se pro nezáporné celé číslo n rovná $n!$.

9. [*M10*] Vypočtěte hodnotu $\Gamma(\frac{1}{2})$ a $\Gamma(-\frac{1}{2})$, pokud je dáno $(\frac{1}{2})! = \sqrt{\pi}/2$.

▸ **10.** [*VM20*] Platí rovnost $\Gamma(x+1) = x\Gamma(x)$ pro všechna reálná čísla x? (Viz cvičení 7.)

11. [*M15*] Nechť reprezentace čísla n ve dvojkové soustavě je $n = 2^{e_1} + 2^{e_2} + \cdots + 2^{e_r}$, kde $e_1 > e_2 > \cdots > e_r \geq 0$. Ukažte, že $n!$ je dělitelné 2^{n-r}, ale ne 2^{n-r+1}.

▸ **12.** [*M22*] (A. Legendre, 1808.) Zobecníme-li výsledek předchozího cvičení, nechť p je prvočíslo a nechť reprezentace čísla n v p-ární číselné soustavě je $n = a_k p^k + a_{k-1} p^{k-1} + \cdots + a_1 p + a_0$. Vyjádřete číslo μ ze vztahu (8) pomocí jednoduchého vzorce s n, p a čísel a.

13. [*M23*] (*Wilsonova věta*, ve skutečnosti zformuloval Leibniz, 1682.) Je-li p prvočíslo, pak $(p-1)! \bmod p = p-1$. Dokažte větu tím, že z množiny $\{1, 2, \ldots, p-1\}$ spárujete čísla, jejichž součin modulo p je roven 1.

▸ **14.** [*M28*] (L. Stickelberger, 1890.) V notaci podle cvičení 12 můžeme stanovit $n! \bmod p$ vyjádřené v p-ární reprezentaci pro *libovolné* kladné celé číslo n a tím zobecnit Wilsonovu větu. Dokažte tedy, že $n!/p^\mu \equiv (-1)^\mu a_0! \, a_1! \ldots a_k!$ (modulo p).

15. [*VM15*] *Permanent* čtvercové matice je definován stejným rozvojem jako determinant, ale u permanentu mají všechny členy znaménko plus, zatímco v determinantu se střídají znaménka plus a minus. Permanent matice

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

je tudíž $aei + bfg + cdh + gec + hfa + idb$. Jaký je permanent matice

$$\begin{pmatrix} 1\times 1 & 1\times 2 & \ldots & 1\times n \\ 2\times 1 & 2\times 2 & \ldots & 2\times n \\ \vdots & \vdots & \ddots & \vdots \\ n\times 1 & n\times 2 & \ldots & n\times n \end{pmatrix}?$$

16. [*VM15*] Ukažte, že nekonečný součet ve vztahu (11) nekonverguje, kromě případu, kdy je n nezáporné celé číslo.

17. [*VM20*] Dokažte, že nekonečný součin

$$\prod_{n\geq 1} \frac{(n+\alpha_1)\ldots(n+\alpha_k)}{(n+\beta_1)\ldots(n+\beta_k)}$$

se rovná $\Gamma(1+\beta_1)\ldots\Gamma(1+\beta_k)/\Gamma(1+\alpha_1)\ldots\Gamma(1+\alpha_k)$, pokud je $\alpha_1 + \cdots + \alpha_k = \beta_1 + \cdots + \beta_k$ a pokud žádné z β není záporné celé číslo.

18. [*M20*] Předpokládejme, že $\pi/2 = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdot\cdots$. (Je to takzvaný „Wallisův součin“, který zjistil J. Wallis v roce 1655; dokážeme jej ve cvičení 1.2.6–43.) Na základě předchozího cvičení dokažte, že $(\frac{1}{2})! = \sqrt{\pi}/2$.

19. [*VM22*] Vyjádřete veličinu za symbolem „$\lim_{m\to\infty}$“ ve vztahu (15) pomocí $\Gamma_m(x)$. Ukažte, že

$$\Gamma_m(x) = \int_0^m \Big(1-\frac{t}{m}\Big)^m t^{x-1}\,dt = m^x \int_0^1 (1-t)^m t^{x-1}\,dt, \quad \text{pokud } x>0.$$

20. [*VM21*] Na základě faktu, že $0 \le e^{-t} - (1-t/m)^m \le t^2 e^{-t}/m$, je-li $0 \le t \le m$, a na základě předchozího cvičení ukažte, že $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}\,dt$, je-li $x>0$.

21. [*VM25*] (L. F. A. Arbogast, 1800.) Nechť $D_x^k u$ vyjadřuje k-tou derivaci funkce u podle x. Podle pravidla pro derivaci složené funkce je $D_x^1 w = D_u^1 w\, D_x^1 u$. Jestliže toto pravidlo aplikujeme na druhé derivace, dostaneme $D_x^2 w = D_u^2 w (D_x^1 u)^2 + D_u^1 w\, D_x^2 u$. Ukažte, že *obecný vzorec* je

$$D_x^n w = \sum_{j=0}^{n} \sum_{\substack{k_1+k_2+\cdots+k_n=j\\ k_1+2k_2+\cdots+nk_n=n\\ k_1,k_2,\ldots,k_n\ge 0}} D_u^j w \frac{n!}{k_1!\,(1!)^{k_1}\ldots k_n!\,(n!)^{k_n}} (D_x^1 u)^{k_1}\ldots(D_x^n u)^{k_n}.$$

▸ **22.** [*VM20*] Pokuste se vžít do Eulerovy role a najít způsob zobecnění $n!$ na neceločíselné hodnoty n. Protože $(n+\frac12)!/n!$ krát $((n+\frac12)+\frac12)!/(n+\frac12)!$ se rovná $(n+1)!/n! = n+1$, zdálo by se přirozené, že $(n+\frac12)!/n!$ bude přibližně rovno $\sqrt{n}$. Podobně $(n+\frac13)!/n!$ by mohlo být $\approx \sqrt[3]{n}$. Zformulujte hypotézu o hodnotě poměru $(n+x)!/n!$ pro n blížící se k nekonečnu. Je hypotéza správná pro celé číslo x? Říká něco o odpovídající hodnotě $x!$, pokud x není celé číslo?

23. [*VM20*] Dokažte vztah (16), pokud je $\pi z \prod_{n=1}^\infty (1-z^2/n^2) = \sin \pi z$.

▸ **24.** [*VM21*] Dokažte užitečné nerovnosti

$$\frac{n^n}{e^{n-1}} \le n! \le \frac{n^{n+1}}{e^{n-1}}, \qquad \text{pro celé } n \ge 1.$$

[*Nápověda:* $1+x \le e^x$ pro všechna reálná x; tudíž $(k+1)/k \le e^{1/k} \le k/(k-1)$.]

25. [*M20*] Platí pro faktoriálové mocniny pravidlo podobné běžným pravidlům umocňování, tedy $x^{m+n} = x^m x^n$?

1.2.6. Binomické koeficienty

Kombinace k-té třídy z n prvků jsou všechny možné varianty výběru k různých prvků z množiny n prvků, přičemž nezáleží na pořadí. Kombinace 3. třídy z pěti prvků $\{a,b,c,d,e\}$ jsou tudíž

$$abc,\quad abd,\quad abe,\quad acd,\quad ace,\quad ade,\quad bcd,\quad bce,\quad bde,\quad cde. \tag{1}$$

Zjistit celkový počet kombinací k-té třídy z n prvků je jednoduché: podle vztahu (2) z předchozí části víme, že permutací prvních k prvků existuje $n(n-1)\ldots(n-k+1)$; mezi nimi se každá kombinace k-té třídy opakuje právě $k!$-krát, protože je zde uvedena ve všech svých permutacích (ve všech pořadích). Počet kombinací, který označujeme jako $\binom{n}{k}$, je proto roven

$$\binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots(1)}. \tag{2}$$

Tabulka 1
TABULKA BINOMICKÝCH KOEFICIENTŮ (PASCALŮV TROJÚHELNÍK)

r	$\binom{r}{0}$	$\binom{r}{1}$	$\binom{r}{2}$	$\binom{r}{3}$	$\binom{r}{4}$	$\binom{r}{5}$	$\binom{r}{6}$	$\binom{r}{7}$	$\binom{r}{8}$	$\binom{r}{9}$
0	1	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0	0
2	1	2	1	0	0	0	0	0	0	0
3	1	3	3	1	0	0	0	0	0	0
4	1	4	6	4	1	0	0	0	0	0
5	1	5	10	10	5	1	0	0	0	0
6	1	6	15	20	15	6	1	0	0	0
7	1	7	21	35	35	21	7	1	0	0
8	1	8	28	56	70	56	28	8	1	0
9	1	9	36	84	126	126	84	36	9	1

Například

$$\binom{5}{3} = \frac{5 \cdot 4 \cdot 3}{3 \cdot 2 \cdot 1} = 10,$$

což je právě počet kombinací uvedených v (1).

Číslo $\binom{n}{k}$, které čteme „n nad k“, se nazývá *binomický koeficient*; tyto koeficienty mají neobyčejně široké spektrum uplatnění. Také při analýze algoritmů jsou to snad nejdůležitější veličiny, a proto by se s nimi měl čtenář velmi dobře seznámit.

Pomocí vztahu (2) můžeme definovat $\binom{n}{k}$ i pro necelé číslo n. Přesněji řečeno definujeme symbol $\binom{r}{k}$ pro všechna reálná čísla r a všechna celá čísla k následovně:

$$\begin{aligned} \binom{r}{k} &= \frac{r(r-1)\ldots(r-k+1)}{k(k-1)\ldots(1)} = \frac{r^{\underline{k}}}{k!} = \prod_{j=1}^{k} \frac{r+1-j}{j}, && \text{pro celé } k \geq 0; \\ \binom{r}{k} &= 0, && \text{pro celé } k < 0. \end{aligned} \tag{3}$$

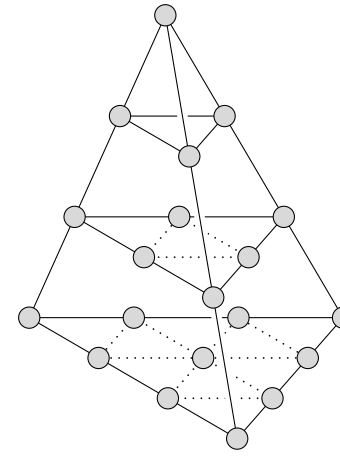
Jako speciální případy můžeme uvést

$$\binom{r}{0} = 1, \qquad \binom{r}{1} = r, \qquad \binom{r}{2} = \frac{r(r-1)}{2}. \tag{4}$$

V tabulce 1 jsou pak uvedeny hodnoty binomických koeficientů pro malé celočíselné hodnoty r a k; pro $0 \leq r \leq 4$ je dobré se je naučit zpaměti.

Binomické koeficienty mají dlouhou a zajímavou historii. Tabulka 1 se nazývá „Pascalův trojúhelník“, protože ji poprvé uvedl Blaise Pascal ve svém *Traité du Triangle Arithmétique* roku 1653. Toto velmi významné pojednání se jako jedna z prvních prací zabývalo teorií pravděpodobnosti, ale samotné binomické koeficienty Pascal nevynalezl (v té době již byly v Evropě dobře známé). Tabulka 1 se objevila také ve spise *Szu-yüan Yü-chien* („Drahocenné zrcadlo čtyř živlů“), jež vydal čínský matematik Chu Shih-Chieh roku 1303; ten zde tvrdil, že se jedná již o starou myšlenku. Yang Hui je roku 1261 připisoval Chia

Hsienovi (kolem 1100), jehož dílo je dnes ztraceno. Nejstarší známý podrobný výklad binomických koeficientů pochází z 10. století a napsal jej Halāyudha formou poznámek ke klasickému hindskému spisu Piṅgala: *Chandaḥśāstra*. [Viz G. Chakravarti, *Bull. Calcutta Math. Soc.* **24** (1932), 79–88.] Jiný indický matematik, Mahāvīra, vysvětlil pravidlo (3) pro výpočet $\binom{r}{k}$ v kapitole 6 své *Gaṇita Sāra Saṅgraha*, napsané okolo roku 850; v roce 1150 pak jeho pravidlo zopakoval Bhāskara ke konci své slavné knihy *Līlāvatī*. Pro malé hodnoty k byly binomické koeficienty známy ještě mnohem dříve; objevily se v řeckých a římských dílech s geometrickou interpretací podle obrázku 8. Notaci $\binom{r}{k}$ zavedl Andreas von Ettingshausen v § 31 své knihy *Die combinatorische Analysis* (Vídeň: 1826).



Obr. 8. Geometrická interpretace $\binom{n+2}{3}$ pro $n = 4$

Čtenář si v tabulce 1 zajisté všiml několika zajímavých vlastností. Binomické koeficienty vystupují doslova v tisících různých rovností a jejich pozoruhodné vlastnosti byly ustavičně zkoumány po staletí. Ve skutečnosti je již o nich známo tolik relací, že když se někomu podaří objevit novou rovnost, bývá z ní zpravidla nadšený toliko sám objevitel. Pro snazší manipulaci se vztahy, které vznikají při analýze algoritmů, je určitý mechanismus pro práci s binomickými koeficienty přímo nutností, a proto se v této části textu jeden takový jednoduchý způsob pokusíme vysvětlit. Mark Twain se kdysi pokusil zredukovat všechny vtipy zhruba na desítku jakýchsi primitivních typů (sedlákova dcera, tchýně atd.); podobně i my se pokusíme zhustit všechny ty tisíce rovností do malé množiny základních operací a s nimi již zvládneme vyřešit téměř každý problém s binomickými koeficienty, na který narazíme.

Ve většině aplikací jsou r i k ve vzorci $\binom{r}{k}$ celá čísla a některé z níže popisovaných technik je možné uplatnit jen pro ně. U každého z uvedených vztahů si proto dávejte pozor na omezující podmínky, které jsou zapsány vpravo. Podíváme-li se například na vztah (3), musí být k celé číslo, zatímco pro r žádné omezení neplatí. Dá se říci, že čím méně omezení je zapsáno, tím je daný vztah užitečnější.

Nyní se podívejme na základní postupy pro manipulaci s binomickými koeficienty:

A. Reprezentace pomocí faktoriálů. Ze vztahu (3) okamžitě dostáváme

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}, \quad \text{pro celé } n \geq \text{celé } k \geq 0. \tag{5}$$

Takto můžeme popsané kombinace faktoriálů vyjádřit pomocí binomických koeficientů a naopak.

B. Podmínka symetrie. Ze vztahů (3) a (5) odvodíme

$$\binom{n}{k} = \binom{n}{n-k}, \quad \text{pro celé } n \geq 0, \text{ celé } k. \tag{6}$$

Uvedený vzorec platí pro všechna celá k. *Pokud je k záporné nebo větší než n, je binomický koeficient nulový* (za předpokladu, že n je nezáporné celé číslo).

C. Vytýkání a roznásobení závorek. Podle definice (3) platí

$$\binom{r}{k} = \frac{r}{k}\binom{r-1}{k-1}, \quad \text{pro celé } k \neq 0. \tag{7}$$

Tento vztah je velmi užitečný pro kombinování binomických koeficientů s ostatními částmi výrazu. Jednoduchou transformací dostaneme pravidla

$$k\binom{r}{k} = r\binom{r-1}{k-1}, \quad \frac{1}{r}\binom{r}{k} = \frac{1}{k}\binom{r-1}{k-1},$$

z nichž první platí pro všechna celá k a druhá za podmínky, že nebylo provedeno dělení nulou. Můžeme odvodit také podobnou relaci:

$$\binom{r}{k} = \frac{r}{r-k}\binom{r-1}{k}, \quad \text{pro celé } k \neq r. \tag{8}$$

Nyní si ilustrujeme popsané transformace na důkazu vztahu (8) pomocí vztahů (6) a (7):

$$\binom{r}{k} = \binom{r}{r-k} = \frac{r}{r-k}\binom{r-1}{r-1-k} = \frac{r}{r-k}\binom{r-1}{k}.$$

[*Poznámka:* Toto odvození je vzhledem k podmínkám v (6) a (7) platné jen v případě, že je r kladné celé číslo $\neq k$; přesto (8) je platné pro *libovolné* $r \neq k$. Vztah můžeme dokázat jednoduchým a důležitým postupem: Ověřili jsme, že

$$r\binom{r-1}{k} = (r-k)\binom{r}{k}$$

pro *nekonečně mnoho hodnot* r. Obě strany této rovnosti jsou *polynomy* v r. Nenulový polynom stupně n může mít nejvýše n různých nulových bodů; po odečtení to znamená, že *pokud se dva polynomy stupně* $\leq n$ *shodují* v $n+1$ *a více různých bodech, pak jsou identické.* Pomocí tohoto principu je možné rozšířit platnost celé řady rovností z celých čísel na všechna reálná.]

D. Součtový vzorec. Základní relace

$$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}, \quad \text{pro celé } k, \tag{9}$$

je v tabulce 1 zjevně platná (každé číslo je součtem dvou čísel nad sebou a vlevo) a snadno ji ověříme v obecném tvaru z (3). Druhá možnost je vyjít ze vztahů (7) a (8), podle kterých je

$$r\binom{r-1}{k} + r\binom{r-1}{k-1} = (r-k)\binom{r}{k} + k\binom{r}{k} = r\binom{r}{k}.$$

Vztah (9) se často používá při důkazech indukcí podle r (pokud je r celé číslo).

E. Sumační vzorce. Opakovanou aplikací vztahu (9) dostáváme

$$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1} = \binom{r-1}{k} + \binom{r-2}{k-1} + \binom{r-2}{k-2} = \cdots$$

nebo

$$\binom{r}{k} = \binom{r-1}{k-1} + \binom{r-1}{k} = \binom{r-1}{k-1} + \binom{r-2}{k-1} + \binom{r-2}{k} = \cdots$$

Dostali jsme se tedy ke dvěma důležitým sumačním vzorcům, které můžeme vyjádřit následovně:

$$\sum_{k=0}^{n}\binom{r+k}{k} = \binom{r}{0} + \binom{r+1}{1} + \cdots + \binom{r+n}{n} = \binom{r+n+1}{n},$$

pro celé $n \geq 0$. (10)

$$\sum_{k=0}^{n}\binom{k}{m} = \binom{0}{m} + \binom{1}{m} + \cdots + \binom{n}{m} = \binom{n+1}{m+1},$$

pro celé $m \geq 0$, celé $n \geq 0$. (11)

Rovnost (11) můžeme snadno dokázat indukcí podle n, je ale zajímavé se podívat také na možnost jejího odvození z (10) při dvojím uplatnění vztahu (6):

$$\sum_{0\leq k\leq n}\binom{k}{m} = \sum_{0\leq m+k\leq n}\binom{m+k}{m} = \sum_{-m\leq k<0}\binom{m+k}{m} + \sum_{0\leq k\leq n-m}\binom{m+k}{k}$$
$$= 0 + \binom{m+(n-m)+1}{n-m} = \binom{n+1}{m+1}$$

za předpokladu, že $n \geq m$. Pokud je $n < m$, je (11) zřejmé.

Se vztahem (11) se budeme v různých aplikacích setkávat velmi často; již v předcházející části textu jsme odvodili jeho speciální případy. Pro $m = 1$ máme například náš starý známý vzorec pro součet aritmetické posloupnosti: :

$$\binom{0}{1} + \binom{1}{1} + \cdots + \binom{n}{1} = 0 + 1 + \cdots + n = \binom{n+1}{2} = \frac{(n+1)n}{2}.$$

Předpokládejme nyní, že chceme odvodit jednoduchý vzorec pro součet $1^2 + 2^2 + \cdots + n^2$. K tomu nejprve nahlédneme, že $k^2 = 2\binom{k}{2} + \binom{k}{1}$, a tudíž

$$\sum_{k=0}^{n} k^2 = \sum_{k=0}^{n}\left(2\binom{k}{2} + \binom{k}{1}\right) = 2\binom{n+1}{3} + \binom{n+1}{2}.$$

Tato odpověď je vyjádřena pomocí binomických koeficientů; podle potřeby ji můžeme vrátit do polynomické notace:

$$1^2 + 2^2 + \cdots + n^2 = 2\frac{(n+1)n(n-1)}{6} + \frac{(n+1)n}{2} = \tfrac{1}{3}n(n+\tfrac{1}{2})(n+1). \quad (12)$$

Podobným způsobem můžeme získat také součet $1^3+2^3+\cdots+n^3$; dokonce *libovolný* polynom $a_0+a_1k+a_2k^2+\cdots+a_mk^m$ je možné vyjádřit jako $b_0\binom{k}{0}+$ $+b_1\binom{k}{1}+\cdots+b_m\binom{k}{m}$, pro vhodně zvolené koeficienty $b_0,\ldots,b_m$. K tématu se ještě později vrátíme.

F. Binomická věta. Binomická věta bude samozřejmě pro nás jedním z nejdůležitějších nástrojů:

$$(x+y)^r=\sum_k\binom{r}{k}x^ky^{r-k},\quad\text{pro celé } r\ge 0. \tag{13}$$

Například $(x+y)^4=x^4+4x^3y+6x^2y^2+4xy^3+y^4$. (Alespoň tak můžeme zdůvodnit, proč jsme čísla $\binom{r}{k}$ nazvali „binomickými koeficienty“.)

Na tomto místě je důležité si všimnout, že jsme ve vztahu (13) napsali $\sum_k$, nikoli $\sum_{k=0}^r$, jak by čtenář mohl očekávat. Jestliže pro k neurčíme žádné omezení, sčítáme přes *všechna* celá čísla, tj. $-\infty<k<+\infty$; obě notace jsou ale přesto ekvivalentní, protože pro $k<0$ a pro $k>r$ jsou příslušné členy (13) nulové. Vhodnější je používat prostší zápis $\sum_k$, protože zjednodušením podmínek si usnadníme i veškeré manipulace se sumami. Jestliže si nemusíme hlídat horní a dolní hranice sumy, ušetříme si tím obrovské množství práce, takže je nejlépe pokud možno všude ponechávat hranice sum neurčené. Naše notace má ještě jednu výhodu: pokud je r nezáporné číslo, stane se ze vztahu (13) *nekonečná* suma a podle *binomické věty* z diferenciálního počtu je *vztah* (13) *platný pro všechna* r, pokud $|x/y|<1$.

Ještě je třeba poznamenat, že ze vztahu (13) vyplývá

$$0^0=1. \tag{14}$$

Tuto dohodu budeme důsledně používat.

Speciální případ $y=1$ je ve vztahu (13) natolik důležitý, že jej vyslovíme samostatně:

$$\sum_k\binom{r}{k}x^k=(1+x)^r,\qquad\text{pro celé } r\ge 0 \text{ nebo } |x|<1. \tag{15}$$

Objev binomické věty oznámil Isaac Newton ve svých dopisech Oldenburgovi ze 13. června a 24. října 1676. [Viz D. Struik, *Source Book in Mathematics* (Harvard Univ. Press, 1969), 284–291.] Tento vztah nemohl ale zjevně doložit žádným důkazem; v té době si potřebu řádného důkazu nikdo plně neuvědomoval. První pokus o důkaz provedl L. Euler v roce 1774, i když jej nedokončil. S prvním skutečným důkazem přišel nakonec roku 1812 C. F. Gauss. V Gaussově díle se fakticky objevily vůbec první uspokojivé důkazy *jakýchkoli* vlastností nekonečných součtů.

Začátkem 19. století objevil N. H. Abel překvapivé zobecnění binomické věty (13):

$$(x+y)^n=\sum_k\binom{n}{k}x(x-kz)^{k-1}(y+kz)^{n-k},\quad\text{pro celé } n\ge 0,\ x\ne 0. \tag{16}$$

Tuto rovnost ve *třech* proměnných, x, y a z (viz cvičení 50 až 52), Abel publikoval a dokázal ve svazku 1 později slavného díla A. L. Crelle, *Journal für die reine und angewandte Mathematik* (1826), na str. 159–160. Je zajímavé poznamenat, že Abel přispěl do stejného svazku 1 ještě mnoha dalšími věcmi, mimo jiné slavným pojednáním o neřešitelnosti algebraických rovnic stupně 5 a více pomocí radikálů a poznámkami k binomické větě. Řada odkazů na (16) viz H. W. Gould, *AMM* **69** (1962), 572.

G. Změna znaménka horního indexu. Základní rovnost

$$\binom{r}{k} = (-1)^k \binom{k-r-1}{k}, \quad \text{pro celé } k, \tag{17}$$

vyplývá bezprostředně z definice (3), v níž změníme znaménko každého člena čitatele. Je to častá transformace horního indexu.

Jedním ze snadných důsledků vztahu (17) je sumační vzorec

$$\sum_{k\le n} \binom{r}{k}(-1)^k = \binom{r}{0} - \binom{r}{1} + \cdots + (-1)^n \binom{r}{n} = (-1)^n \binom{r-1}{n}, \quad \text{pro celé } n. \tag{18}$$

Tuto rovnost dokážeme snadno indukcí s využitím (9), ale můžeme ji také odvodit přímo ze vztahů (17) a (10):

$$\sum_{k\le n} \binom{r}{k}(-1)^k = \sum_{k\le n} \binom{k-r-1}{k} = \binom{-r+n}{n} = (-1)^n \binom{r-1}{n}.$$

Další důležitá aplikace vztahu (17) platí pro celočíselné r:

$$\binom{n}{m} = (-1)^{n-m} \binom{-(m+1)}{n-m}, \quad \text{pro celé } n \ge 0, \text{ celé } m. \tag{19}$$

(Ve vztahu (17) položte $r = n$ a $k = n - m$; dále využijte (6).) Číslo n jsme přesunuli z horní pozice do dolní.

H. Zjednodušení součinů. Součiny binomických koeficientů je obvykle možné přeformulovat několika různými způsoby, a to pomocí rozvoje do faktoriálů podle (5). Například

$$\binom{r}{m}\binom{m}{k} = \binom{r}{k}\binom{r-k}{m-k}, \quad \text{pro celé } m, \text{ celé } k. \tag{20}$$

Vztah (20) stačí dokázat pro případ, kdy r je celé $\ge m$ (viz poznámky za vztahem (8)), a pro $0 \le k \le m$. Potom je

$$\binom{r}{m}\binom{m}{k} = \frac{r!\,m!}{m!\,(r-m)!\,k!\,(m-k)!} = \frac{r!\,(r-k)!}{k!\,(r-k)!\,(m-k)!\,(r-m)!} = \binom{r}{k}\binom{r-k}{m-k}.$$

Vztah (20) je velice užitečný v případech, kdy je stejný index (konkrétně m) zapsán nahoře i dole a kdy potřebujeme, aby byl uveden jen na jednom místě, nikoli na dvou. Všimněte si, že (7) je speciálním případem (20) pro $k = 1$.

I. Součet součinů. Sérii manipulací s binomickými koeficienty završíme v následujících velmi obecných rovnostech, které dokážeme ve cvičeních na konci této části textu. Tyto vztahy ukazují sčítání součinů dvou binomických koeficientů, přičemž uvažujeme s různými místy zápisu proměnné k:

$$\sum_k \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}, \quad \text{pro celé } n. \tag{21}$$

$$\sum_k \binom{r}{m+k}\binom{s}{n+k} = \binom{r+s}{r-m+n},$$
$$\text{pro celé } m,\ \text{celé } n,\ \text{celé } r \geq 0. \tag{22}$$

$$\sum_k \binom{r}{k}\binom{s+k}{n}(-1)^{r-k} = \binom{s}{n-r}, \ \text{pro celé } n,\ \text{celé } r \geq 0. \tag{23}$$

$$\sum_{k=0}^{r} \binom{r-k}{m}\binom{s}{k-t}(-1)^{k-t} = \binom{r-t-s}{r-t-m},$$
$$\text{pro celé } t \geq 0,\ \text{celé } r \geq 0,\ \text{celé } m \geq 0. \tag{24}$$

$$\sum_{k=0}^{r} \binom{r-k}{m}\binom{s+k}{n} = \binom{r+s+1}{m+n+1},$$
$$\text{pro celé } n \geq \text{celé } s \geq 0,\ \text{celé } m \geq 0,\ \text{celé } r \geq 0. \tag{25}$$

$$\sum_{k\geq 0} \binom{r-tk}{k}\binom{s-t(n-k)}{n-k}\frac{r}{r-tk} = \binom{r+s-tn}{n}, \ \text{pro celé } n. \tag{26}$$

Z těchto rovností je zdaleka nejdůležitější vztah (21), který je vhodné se naučit zpaměti. Řeknu vám, jak si jej zapamatovat. Zkuste pravou stranu přečíst jako počet způsobů, kterými vybereme n osob z r mužů a s žen; každý člen na levé straně pak tvoří počet způsobů, kterými vybereme k mužů a $n-k$ žen. Vztahu (21) se obvykle říká Vandermondova konvoluce, protože jej publikoval A. Vandermonde v *Mém. Acad. Roy. Sciences* (Paříž, 1772), část 1, 489–498. Objevila se nicméně mnohem dříve, v již zmíněném pojednání Chu Shih- Chieha z roku 1303 [viz J. Needham, *Science and Civilisation in China* **3** (Cambridge University Press, 1959), 138–139].

Je-li ve vztahu (26) rovno $r = tk$, zkrátíme stejný činitel i v čitateli a odstraníme tím nulového jmenovatele; vztah (26) tak přejde na polynomickou rovnici v proměnných r, s, t. Zřejmě (21) je speciálním případem (26) pro $t = 0$.

Dále poukážeme na jedno netriviální uplatnění vztahů (23) a (25): Často je totiž vhodné nahradit jednoduchý binomický koeficient na pravé straně složitějším výrazem vlevo, zaměnit pořadí sumace a výraz zjednodušit. Levé strany těchto vztahů můžeme považovat za rozvoj výrazů

$$\binom{s}{n+a} \quad \text{ve vyjádření} \quad \binom{s+k}{n}.$$

Vztah (23) se používá pro záporná a, vztah (25) pro kladná a.

Tím končí naše exkurze do světa binomických koeficientů. Doporučuji čtenáři naučit se zejména vztahy (5), (6), (7), (9), (13), (17), (20) a (21) – zatrhněte si je barevnou tužkou!

Máme-li k dispozici všechny tyto metody, můžeme již vyřešit téměř libovolný problém nejméně třemi různými způsoby. Uvedené techniky si nyní ilustrujeme na následujících příkladech.

Příklad 1. Je-li r kladné celé číslo, čemu se rovná $\sum_k \binom{r}{k}\binom{s}{k}k$?

Řešení. S využitím vztahu (7) se zbavíme samostatného k:

$$\sum_k \binom{r}{k}\binom{s}{k}k = \sum_k \binom{r}{k}\binom{s-1}{k-1}s = s\sum_k \binom{r}{k}\binom{s-1}{k-1}.$$

Nyní aplikujeme vztah (22), přičemž $m=0$ a $n=-1$. Odpověď je proto

$$\sum_k \binom{r}{k}\binom{s}{k}k = \binom{r+s-1}{r-1}s, \quad \text{pro celé } r \ge 0.$$

Příklad 2. Čemu se rovná $\sum_k \binom{n+k}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1}$, je-li n nezáporné celé číslo?

Řešení. Tento problém je o něco obtížnější; sumační index k je uveden hned na šesti místech! Nejprve aplikujeme vztah (20) a dostaneme

$$\sum_k \binom{n+k}{k}\binom{n}{k}\frac{(-1)^k}{k+1}.$$

Nyní se nám již dýchá lehčeji, protože jsme se z původního vztahu zbavili několika nepříjemných vlastností. Další krok by už měl být zřejmý: podobně jako v Příkladu 1 aplikujeme vztah (7):

$$\sum_k \binom{n+k}{k}\binom{n+1}{k+1}\frac{(-1)^k}{n+1}. \tag{27}$$

Výborně, zmizelo další k. Před námi jsou dvě zhruba stejně slibné linie útoku. Můžeme nahradit $\binom{n+k}{k}$ za $\binom{n+k}{n}$, přičemž předpokládáme $k \ge 0$, a poté vypočítat sumu podle (23):

$$\begin{aligned}
&\sum_{k\ge 0} \binom{n+k}{n}\binom{n+1}{k+1}\frac{(-1)^k}{n+1} \\
&\quad = -\frac{1}{n+1}\sum_{k\ge 1} \binom{n-1+k}{n}\binom{n+1}{k}(-1)^k \\
&\quad = -\frac{1}{n+1}\sum_{k\ge 0} \binom{n-1+k}{n}\binom{n+1}{k}(-1)^k + \frac{1}{n+1}\binom{n-1}{n} \\
&\quad = -\frac{1}{n+1}(-1)^{n+1}\binom{n-1}{-1} + \frac{1}{n+1}\binom{n-1}{n} = \frac{1}{n+1}\binom{n-1}{n}.
\end{aligned}$$

Binomický koeficient $\binom{n-1}{n}$ se rovná nule s výjimkou případu $n = 0$, kdy je roven jedné. Odpověď na náš problém můžeme tedy pohodlně formulovat podle Iversonovy konvence (viz 1.2.3–(16)) jako $[n=0]$ nebo pomocí Kroneckerova delta jako δ_{n0} (viz 1.2.3–(19)).

Další možnost řešení (27) spočívá ve využití vztahu (17), kdy dostaneme

$$\sum_k \binom{-(n+1)}{k} \binom{n+1}{k+1} \frac{1}{n+1}.$$

Nyní můžeme aplikovat vztah (22), kterým přejdeme k sumě

$$\binom{n+1-(n+1)}{n+1-1+0} \frac{1}{n+1} = \binom{0}{n} \frac{1}{n+1}.$$

A opět jsme dostali stejnou odpověď:

$$\sum_k \binom{n+k}{2k} \binom{2k}{k} \frac{(-1)^k}{k+1} = \delta_{n0}, \qquad \text{pro celé } n \geq 0. \tag{28}$$

Příklad 3. Čemu se rovná $\sum_k \binom{n+k}{m+2k} \binom{2k}{k} \frac{(-1)^k}{k+1}$ pro kladná celá čísla m a n?

Řešení. Pokud by m bylo rovno nule, dostali bychom se ke stejnému vzorci jako v Příkladu 2. Díky proměnné m nemůžeme ale tuto metodu ani zahájit, protože v prvním kroku jsme použili vztah (20) – a ten zde již neplatí. V této situaci se vyplatí problém ještě více zkomplikovat a nahradit nežádoucí výraz $\binom{n+k}{m+2k}$ sumou členů ve tvaru $\binom{x+k}{2k}$, protože z celého problému vytvoříme součet podproblémů, které již vyřešit umíme. Využijeme tedy (25) a dosadíme

$$r = n + k - 1, \quad m = 2k, \quad s = 0, \quad n = m - 1,$$

odkud dostaneme

$$\sum_k \sum_{0 \leq j \leq n+k-1} \binom{n+k-1-j}{2k} \binom{2k}{k} \binom{j}{m-1} \frac{(-1)^k}{k+1}. \tag{29}$$

Jako první chceme provést sumaci podle k, ale pro záměnu pořadí sumace musíme sčítat jen podle hodnot k, které jsou ≥ 0 a $\geq j - n + 1$. Druhá z podmínek vede bohužel k problémům, protože pro $j \geq n$ *neznáme* požadovanou sumu. Situaci ale zachráníme zjištěním, že při $n \leq j \leq n + k - 1$ jsou členy ve vztahu (29) nulové. Z této podmínky vyplývá, že $k \geq 1$, a tudíž $0 \leq n+k-1-j \leq k-1 < 2k$; první z binomických koeficientů ve vztahu (29) tím pádem zmizí. Podmínku ve druhé sumě můžeme tedy nahradit za $0 \leq j < n$; záměna pořadí sumace je již triviální. Po sečtení (28) podle k dostáváme

$$\sum_{0 \leq j < n} \binom{j}{m-1} \delta_{(n-1-j)0},$$

jejíž všechny členy jsou nulové, kromě $j = n - 1$. Naše konečná odpověď je tudíž rovna

$$\binom{n-1}{m-1}.$$

Řešení tohoto problému bylo poměrně komplikované, ale přitom nijak tajuplné; ke každému kroku jsme měli rozumný důvod. Celé odvození si ovšem důkladně prostudujte, protože jsme si na něm ukázali jisté drobné operace s podmínkami ve zkoumaných vztazích. Ve skutečnosti existuje ještě lepší způsob řešení tohoto problému; transformace zadané sumy do podoby, v níž můžeme aplikovat vztah (26), je ponechána bystrému čtenáři (viz cvičení 30).

Příklad 4. Dokažte, že

$$\sum_k A_k(r,t)A_{n-k}(s,t) = A_n(r+s,t), \quad \text{pro celé } n \ge 0, \tag{30}$$

kde $A_n(x,t)$ je polynom n-tého stupně v proměnné x, který splňuje podmínku

$$A_n(x,t) = \binom{x-nt}{n}\frac{x}{x-nt}, \quad \text{pro } x \ne nt.$$

Řešení. Můžeme předpokládat, že pro $0 \le k \le n$ je $r \ne kt \ne s$, protože obě strany vztahu (30) jsou polynomy v proměnných r, s, t. Naším úkolem je vypočítat hodnotu vzorce

$$\sum_k \binom{r-kt}{k}\binom{s-(n-k)t}{n-k}\frac{r}{r-kt}\,\frac{s}{s-(n-k)t},$$

který, když nic jiného, vypadá mnohem hůře než všechny předešlé! Všimněte si ale výrazné podobnosti se vztahem (26) a také si všimněte případu $t = 0$.

Mohli bychom se pokusit zaměnit

$$\binom{r-kt}{k}\frac{r}{r-kt} \qquad \text{za} \qquad \binom{r-kt-1}{k-1}\frac{r}{k},$$

ale ve druhém z uvedených výrazů již ztrácíme analogii se vztahem (26) a navíc jej nemůžeme použít pro $k = 0$. Vhodnější je využít metodu *parciálních zlomků*, při které zlomek se složitým jmenovatelem nahradíme součtem zlomků s jednoduššími jmenovateli. Skutečně dostáváme

$$\frac{1}{r-kt}\,\frac{1}{s-(n-k)t} = \frac{1}{r+s-nt}\left(\frac{1}{r-kt}+\frac{1}{s-(n-k)t}\right).$$

Dosadíme-li tuto rovnost do naší sumy, máme

$$\frac{s}{r+s-nt}\sum_k \binom{r-kt}{k}\binom{s-(n-k)t}{n-k}\frac{r}{r-kt} + \frac{r}{r+s-nt}\sum_k \binom{r-kt}{k}\binom{s-(n-k)t}{n-k}\frac{s}{s-(n-k)t},$$

a jestliže ve druhém ze vzorců zaměníme k za $n-k$, vypočteme pomocí vztahu (26) oba; požadovaný výsledek již odtud vyplyne okamžitě. Rovnosti (26) a (30) publikoval H. A. Rothe, *Formulæ de Serierum Reversione* (Lipsko: 1793); dodnes jsou často „objevovány“ jejich speciální případy. Zajímavější historie těchto rovností a některá jejich zobecnění viz H. W. Gould a J. Kaucký, *Journal of Combinatorial Theory* **1** (1966), 233–247.

Příklad 5. Vypočtěte hodnoty a_0, a_1, a_2, ... takové, že

$$n! = a_0 + a_1 n + a_2 n(n-1) + a_3 n(n-1)(n-2) + \cdots \tag{31}$$

pro všechna nezáporná celá čísla n.

Řešení. Odpověď nám dává vztah 1.2.5–(11), který jsme v předcházející části textu uvedli bez důkazu. Nyní ale řekněme, že jej zatím neznáme. Problém zřejmě má řešení, protože můžeme položit $n = 0$ a vypočítat a_0, potom položit $n = 1$ a vypočítat a_1 atd.

Nejprve musíme přepsat vztah (31) pomocí binomických koeficientů:

$$n! = \sum_k \binom{n}{k} k!\, a_k. \tag{32}$$

Problém řešení podobných rovnic, jako je tato, v proměnné a_k se nazývá *inverzní problém*, přičemž u nich můžeme uplatnit níže uvedený postup.

Myšlenka řešení vychází ze speciálního případu $s = 0$ v rovnici (23):

$$\sum_k \binom{r}{k}\binom{k}{n}(-1)^{r-k} = \binom{0}{n-r} = \delta_{nr}, \quad \text{pro celé } n, \text{ celé } r \geq 0. \tag{33}$$

Význam tohoto vztahu spočívá v tom, že pro $n \neq r$ je suma nulová; problém tak vyřešíme snáze, protože velká část členů se vyruší stejně jako v Příkladu 3:

$$\begin{aligned}\sum_n n! \binom{m}{n}(-1)^{m-n} &= \sum_n \sum_k \binom{n}{k} k!\, a_k \binom{m}{n}(-1)^{m-n} \\ &= \sum_k k!\, a_k \sum_n \binom{n}{k}\binom{m}{n}(-1)^{m-n} \\ &= \sum_k k!\, a_k \delta_{km} = m!\, a_m.\end{aligned}$$

Všimněte si, jak jsme se sečtením vhodných násobků vztahu (32) pro $n = 0, 1, \ldots$ dostali k rovnici, v níž se vyskytuje jen hodnota a_m. Máme tedy

$$a_m = \sum_{n\geq 0}(-1)^{m-n}\frac{n!}{m!}\binom{m}{n} = \sum_{0\leq n\leq m}\frac{(-1)^{m-n}}{(m-n)!} = \sum_{0\leq n\leq m}\frac{(-1)^n}{n!}.$$

Tím jsme dokončili řešení Příkladu 5. Podívejme se nyní blíže na důsledky vztahu (33): Pokud jsou r a m nezáporná celá čísla, máme

$$\sum_k \binom{r}{k}(-1)^{r-k}\left(c_0\binom{k}{0} + c_1\binom{k}{1} + \cdots + c_m\binom{k}{m}\right) = c_r,$$

protože ostatní členy se po sečtení vyruší. Pečlivou volbou koeficientů c_i můžeme nyní *libovolný* polynom v k vyjádřit jako součet binomických koeficientů s horním indexem k. Zjistili jsme tedy, že

$$\sum_k \binom{r}{k}(-1)^{r-k}(b_0 + b_1 k + \cdots + b_r k^r) = r!\, b_r, \text{pro celé } r \geq 0, \tag{34}$$

kde $b_0 + \cdots + b_r k^r$ reprezentuje libovolný polynom libovolného stupně r nebo menšího. (Pro studenty numerické analýzy nebude tento vzorec žádným překvapením, protože $\sum_k \binom{r}{k}(-1)^{r-k} f(x+k)$ je „r-tá derivace“ funkce $f(x)$.)

Ze vztahu (34) odvodíme okamžitě mnoho jiných relací, které jsou na pohled komplikované a které se často dokazují jen velmi zdlouhavě, například

$$\sum_k \binom{r}{k}\binom{s-kt}{r}(-1)^k = t^r, \quad \text{pro celé } r \geq 0. \tag{35}$$

V podobných učebnicích, jako je tato, se setkáváme s množstvím působivých příkladů různých triků apod., ale texty se často nezmiňují o problémech, které jsou na pohled jednoduché, ale u nichž běžné postupy selhávají. Z výše uvedených příkladů by se mohlo zdát, že s binomickými koeficienty zvládneme snad všechno; musíme se ale zmínit, že navzdory vztahům (10), (11) a (18) podle všeho neexistuje jednoduchý vzorec pro analogickou sumu

$$\sum_{k=0}^{n}\binom{m}{k} = \binom{m}{0} + \binom{m}{1} + \cdots + \binom{m}{n}, \tag{36}$$

je-li $n < m$. (Pro $n = m$ je odpověď jednoduchá; víte ji? Viz cvičení 36.)

Pokud je však m pevně dané záporné číslo, má tato suma uzavřenou formu funkce v proměnné n, například

$$\sum_{k=0}^{n}\binom{-2}{k} = (-1)^n\left\lceil\frac{n+1}{2}\right\rceil. \tag{37}$$

Existuje také jednoduchý vzorec

$$\sum_{k=0}^{n}\binom{m}{k}\left(k-\frac{m}{2}\right) = -\frac{m}{2}\binom{m-1}{n}, \tag{38}$$

kterým snadno vypočteme zdánlivě obtížnější sumu.

Jak zjistíme, kdy máme přestat v pokusech o zjednodušování vzpurné sumy? Dnes již naštěstí existuje způsob, kterým uvedený problém v mnoha důležitých případech vyřešíme: Je to algoritmus, jehož autory jsou R. W. Gosper a D. Zeilberger a který odhalí uzavřené formy v binomických koeficientech, pokud existují, a naopak dokáže nemožnost řešení, pokud neexistují. Algoritmus Gosper-Zeilberger je mimo rámec této knihy, ale je vysvětlen v *CMath* § 5.8. Viz též kniha $A = B$ od autorů Petkovšek, Wilf a Zeilberger (Wellesley, Mass.: A. K. Peters, 1996).

Nejdůležitějším nástrojem pro systematickou, mechanickou práci se součty binomických koeficientů je využití vlastností *hypergeometrických funkcí*, což jsou nekonečné řady definované v rostoucích mocninách faktoriálů takto:

$$F\left(\begin{matrix}a_1, \ldots, a_m\\ b_1, \ldots, b_n\end{matrix}\,\middle|\, z\right) = \sum_{k\geq 0}\frac{a_1^{\bar{k}}\ldots a_m^{\bar{k}}}{b_1^{\bar{k}}\ldots b_n^{\bar{k}}}\,\frac{z^k}{k!}\,. \tag{39}$$

Úvod do těchto důležitých funkcí najdete v částech 5.5 a 5.6 *CMath*. Historické odkazy viz též J. Dutka, *Archive for History of Exact Sciences* **31** (1984), 15–34.

Myšlenka binomických koeficientů má několik důležitých zobecnění, která nyní stručně probereme. Nejprve pro dolní index k výrazu $\binom{r}{k}$ uvažujme libovolné reálné hodnoty; podrobněji viz cvičení 40 až 45. Máme také zobecnění

$$\binom{r}{k}_q = \frac{(1-q^r)(1-q^{r-1})\dots(1-q^{r-k+1})}{(1-q^k)(1-q^{k-1})\dots(1-q^1)}, \tag{40}$$

které pro q limitně se blížící k 1 přechází v obyčejné binomické koeficienty $\binom{r}{k}_1 = \binom{r}{k}$; tuto vlastnost snadno nahlédneme, pokud každý člen v čitateli a jmenovateli zkrátíme $1-q$. Základní vlastnosti těchto „q-nomických koeficientů“ probíráme ve cvičení 58.

Pro naše účely je ale nejdůležitějším zobecněním tento *multinomický koeficient*:

$$\binom{k_1+k_2+\cdots+k_m}{k_1, k_2, \dots, k_m} = \frac{(k_1+k_2+\cdots+k_m)!}{k_1!\,k_2!\dots k_m!}, \quad \text{pro celé } k_i \ge 0. \tag{41}$$

Nejdůležitější vlastností multinomických koeficientů je zobecnění vztahu (13):

$$(x_1+x_2+\cdots+x_m)^n = \sum_{k_1+k_2+\cdots+k_m=n} \binom{n}{k_1, k_2, \dots, k_m} x_1^{k_1} x_2^{k_2} \dots x_m^{k_m}. \tag{42}$$

Je také důležité poznamenat, že každý multinomický koeficient je možné vyjádřit pomocí binomických koeficientů:

$$\binom{k_1+k_2+\cdots+k_m}{k_1, k_2, \dots, k_m} = \binom{k_1+k_2}{k_1}\binom{k_1+k_2+k_3}{k_1+k_2}\cdots\binom{k_1+k_2+\cdots+k_m}{k_1+\cdots+k_{m-1}}, \tag{43}$$

takže zde opět můžeme uplatnit postupy, které známe pro manipulaci s binomickými koeficienty. Obě strany vztahu (20) tak tvoří trinomický koeficient

$$\binom{r}{k,\ m-k,\ r-m}.$$

V závěru této části se podíváme na stručnou analýzu transformace z polynomu vyjádřeného v mocninách x na polynom vyjádřený pomocí binomických koeficientů. Koeficientům v této transformaci se říká *Stirlingova čísla* a používají se při studiu různých algoritmů.

Stirlingova čísla existují přitom dvojího typu: jsou to Stirlingova čísla prvního druhu, která označujeme $\left[{n \atop k}\right]$, a druhého druhu, označovaná $\left\{{n \atop k}\right\}$. Tyto notace, jež zavedl Jovan Karamata [*Mathematica* (Cluj) **9** (1935), 164–178], mají oproti jiným navrhovaným symbolikám velké výhody [viz D. E. Knuth, *AMM* **99** (1992), 403–422]. Složené závorky ve výrazu $\left\{{n \atop k}\right\}$ si zapamatujeme snadno, protože ty obvykle označují množiny, a číslo $\left\{{n \atop k}\right\}$ je počet způsobů, kterými je možné rozdělit množinu n prvků do k disjunktních podmnožin neboli tříd rozkladu (viz cvičení 64). Také Stirlingova čísla prvního druhu, $\left[{n \atop k}\right]$, mají kombinatorickou interpretaci a budeme ji studovat v části 1.3.3: $\left[{n \atop k}\right]$ je totiž počet permutací n písmen při k cyklech.

Tabulka 2

STIRLINGOVA ČÍSLA PRVNÍHO A DRUHÉHO DRUHU

n	$\begin{bmatrix} n \\ 0 \end{bmatrix}$	$\begin{bmatrix} n \\ 1 \end{bmatrix}$	$\begin{bmatrix} n \\ 2 \end{bmatrix}$	$\begin{bmatrix} n \\ 3 \end{bmatrix}$	$\begin{bmatrix} n \\ 4 \end{bmatrix}$	$\begin{bmatrix} n \\ 5 \end{bmatrix}$	$\begin{bmatrix} n \\ 6 \end{bmatrix}$	$\begin{bmatrix} n \\ 7 \end{bmatrix}$	$\begin{bmatrix} n \\ 8 \end{bmatrix}$
0	1	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0	0
3	0	2	3	1	0	0	0	0	0
4	0	6	11	6	1	0	0	0	0
5	0	24	50	35	10	1	0	0	0
6	0	120	274	225	85	15	1	0	0
7	0	720	1764	1624	735	175	21	1	0
8	0	5040	13068	13132	6769	1960	322	28	1

n	$\left\{ n \atop 0 \right\}$	$\left\{ n \atop 1 \right\}$	$\left\{ n \atop 2 \right\}$	$\left\{ n \atop 3 \right\}$	$\left\{ n \atop 4 \right\}$	$\left\{ n \atop 5 \right\}$	$\left\{ n \atop 6 \right\}$	$\left\{ n \atop 7 \right\}$	$\left\{ n \atop 8 \right\}$
0	1	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0	0
3	0	1	3	1	0	0	0	0	0
4	0	1	7	6	1	0	0	0	0
5	0	1	15	25	10	1	0	0	0
6	0	1	31	90	65	15	1	0	0
7	0	1	63	301	350	140	21	1	0
8	0	1	127	966	1701	1050	266	28	1

Aproximace pro velké hodnoty n viz L. Moser a M. Wyman, *J. London Math. Soc.* **33** (1958), 133–146; *Duke Math. J.* **25** (1958), 29–43; D. E. Barton, F. N. David, a M. Merrington, *Biometrika* **47** (1960), 439–445; **50** (1963), 169–176; N. M. Temme, *Studies in Applied Math.* **89** (1993), 233–243; H. S. Wilf, *J. Combinatorial Theory* **A64** (1993), 344–349; H.-K. Hwang, *J. Combinatorial Theory* **A71** (1995), 343–351.

V Tabulce 2 jsou znázorněny Stirlingovy trojúhelníky, které jsou v jistém slova smyslu analogií Pascalova trojúhelníku.

Stirlingova čísla prvního druhu slouží pro převod faktoriálových mocnin na obyčejné mocniny:

$$\begin{aligned} x^{\underline{n}} &= x(x-1)\dots(x-n+1) \\ &= \begin{bmatrix} n \\ n \end{bmatrix} x^n - \begin{bmatrix} n \\ n-1 \end{bmatrix} x^{n-1} + \cdots + (-1)^n \begin{bmatrix} n \\ 0 \end{bmatrix} \\ &= \sum_k (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} x^k. \end{aligned} \tag{44}$$

Například z Tabulky 2 vyplývá:

$$\binom{x}{5} = \frac{x^{\underline{5}}}{5!} = \frac{1}{120}(x^5 - 10x^4 + 35x^3 - 50x^2 + 24x).$$

Stirlingova čísla druhého druhu se používají pro převod obyčejných mocnin na faktoriálové mocniny:

$$x^n = \left\{ {n \atop n} \right\} x^{\underline{n}} + \cdots + \left\{ {n \atop 1} \right\} x^{\underline{1}} + \left\{ {n \atop 0} \right\} x^{\underline{0}} = \sum_k \left\{ {n \atop k} \right\} x^{\underline{k}}. \tag{45}$$

Tento vztah byl ve skutečnosti důvodem, pro který vůbec Stirling začal ve svém *Methodus Differentialis* (Londýn: 1730) čísla $\left\{ {n \atop k} \right\}$ studovat. Podle Tabulky 2 máme například

$$\begin{aligned} x^5 &= x^{\underline{5}} + 10x^{\underline{4}} + 25x^{\underline{3}} + 15x^{\underline{2}} + x^{\underline{1}} \\ &= 120\binom{x}{5} + 240\binom{x}{4} + 150\binom{x}{3} + 30\binom{x}{2} + \binom{x}{1}. \end{aligned}$$

Nyní uvedeme nejdůležitější rovnosti se Stirlingovými čísly. Proměnné m a n v nich vždy označují nezáporná celá čísla.

Součtové vzorce:

$$\begin{aligned} \left[{n+1 \atop m} \right] &= n\left[{n \atop m} \right] + \left[{n \atop m-1} \right]; \\ \left\{ {n+1 \atop m} \right\} &= m\left\{ {n \atop m} \right\} + \left\{ {n \atop m-1} \right\}. \end{aligned} \tag{46}$$

Inverzní vzorce (srovnej se vztahem (33)):

$$\sum_k \left[{n \atop k} \right] \left\{ {k \atop m} \right\} (-1)^{n-k} = \delta_{mn}, \quad \sum_k \left\{ {n \atop k} \right\} \left[{k \atop m} \right] (-1)^{n-k} = \delta_{mn}. \tag{47}$$

Speciální hodnoty:

$$\binom{0}{n} = \left[{0 \atop n} \right] = \left\{ {0 \atop n} \right\} = \delta_{n0}, \qquad \binom{n}{n} = \left[{n \atop n} \right] = \left\{ {n \atop n} \right\} = 1; \tag{48}$$

$$\left[{n \atop n-1} \right] = \left\{ {n \atop n-1} \right\} = \binom{n}{2}; \tag{49}$$

$$\left[{n+1 \atop 0} \right] = \left\{ {n+1 \atop 0} \right\} = 0, \quad \left[{n+1 \atop 1} \right] = n!, \quad \left\{ {n+1 \atop 1} \right\} = 1, \quad \left\{ {n+1 \atop 2} \right\} = 2^n - 1. \tag{50}$$

Rozvojové vzorce:

$$\sum_k \left[{n \atop k} \right] \binom{k}{m} = \left[{n+1 \atop m+1} \right], \quad \sum_k \left[{n+1 \atop k+1} \right] \binom{k}{m} (-1)^{k-m} = \left[{n \atop m} \right]; \tag{51}$$

$$\sum_k \left\{ {k \atop m} \right\} \binom{n}{k} = \left\{ {n+1 \atop m+1} \right\}, \quad \sum_k \left\{ {k+1 \atop m+1} \right\} \binom{n}{k} (-1)^{n-k} = \left\{ {n \atop m} \right\}; \tag{52}$$

$$\sum_k \binom{m}{k} (-1)^{m-k} k^n = m! \left\{ {n \atop m} \right\}; \tag{53}$$

$$\sum_k \binom{m-n}{m+k}\binom{m+n}{n+k}\left\{{m+k \atop k}\right\} = \left[{n \atop n-m}\right],$$
$$\sum_k \binom{m-n}{m+k}\binom{m+n}{n+k}\left[{m+k \atop k}\right] = \left\{{n \atop n-m}\right\}; \tag{54}$$

$$\sum_k \left\{{n+1 \atop k+1}\right\}\left[{k \atop m}\right](-1)^{k-m} = \binom{n}{m}; \tag{55}$$

$$\sum_{k\le n} \left[{k \atop m}\right]\frac{n!}{k!} = \left[{n+1 \atop m+1}\right], \qquad \sum_{k\le n} \left\{{k \atop m}\right\}(m+1)^{n-k} = \left\{{n+1 \atop m+1}\right\}. \tag{56}$$

Několik dalších základních rovností se Stirlingovými čísly je uvedeno ve cvičeních 1.2.6–61 a 1.2.7–6, a dále ve vztazích (23), (26), (27) a (28) z části 1.2.9.

Rovnost (49) je jen jedním případem obecnějšího jevu: Oba druhy Stirlingových čísel $\left[{n \atop n-m}\right]$ a $\left\{{n \atop n-m}\right\}$ jsou polynomy stupně $2m$ v proměnné n, a to pro každé nezáporné celé m. Pro $m = 2$ a $m = 3$ dostáváme například vztahy

$$\left[{n \atop n-2}\right] = \binom{n}{4} + 2\binom{n+1}{4}, \qquad \left\{{n \atop n-2}\right\} = \binom{n+1}{4} + 2\binom{n}{4},$$
$$\left[{n \atop n-3}\right] = \binom{n}{6} + 8\binom{n+1}{6} + 6\binom{n+2}{6}; \quad \left\{{n \atop n-3}\right\} = \binom{n+2}{6} + 8\binom{n+1}{6} + 6\binom{n}{6}. \tag{57}$$

Má proto smysl definovat čísla $\left[{r \atop r-m}\right]$ a $\left\{{r \atop r-m}\right\}$ pro libovolné reálné (nebo i komplexní) hodnoty r. Při tomto zobecnění jsou oba druhy Stirlingových čísel svázány zajímavým zákonem duality

$$\left\{{n \atop m}\right\} = \left[{-m \atop -n}\right], \tag{58}$$

jenž vyplýval z původního Stirlingova výkladu. Navíc rovnost (45) platí i obecně, a to v tom slova smyslu, že nekonečná řada

$$z^r = \sum_k \left\{{r \atop r-k}\right\} z^{\underline{r-k}} \tag{59}$$

konverguje, pokud je reálná část čísla z kladná. Související vzorec (44) je podobně možné zobecnit na asymptotickou (nikoli však konvergentní) řadu:

$$z^{\underline{r}} = \sum_{k=0}^{m} \left[{r \atop r-k}\right](-1)^k z^{r-k} + O(z^{r-m-1}). \tag{60}$$

(Viz cvičení 65.) V částech 6.1, 6.2 a 6.5 *CMath* najdete další informace o Stirlingových číslech a o jejich úpravách v rovnicích. Viz též cvičení 4.7–21, kde je popsána obecná množina trojúhelníků, jejichž velmi speciálním případem jsou právě Stirlingova čísla.

CVIČENÍ

1. [*00*] Kolik je možných kombinací třídy $n-1$ z n prvků?

2. [*00*] Kolik je $\binom{0}{0}$?

3. [*00*] Kolik je možností rozdání bridžových karet pro jednoho hráče (13 karet z balíčku 52)?

4. [*10*] Uveďte odpověď na cvičení 3 jako součin prvočinitelů.

▶ **5.** [*05*] Pomocí Pascalova trojúhelníku vysvětlete, proč je $11^4 = 14641$.

▶ **6.** [*10*] Pascalův trojúhelník (viz Tabulka 1) můžeme pomocí součtového vzorce (9) rozšířit na všechny strany. Uveďte tři řádky, které Tabulku 1 rozšíří směrem *nahoru* (tedy pro $r = -1$, -2 a -3).

7. [*12*] Je-li n pevně dané celé číslo, při které hodnotě k má výraz $\binom{n}{k}$ nejvyšší hodnotu?

8. [*00*] Která vlastnost Pascalova trojúhelníku se odráží v „podmínce symetrie“ (6)?

9. [*01*] Čemu se rovná $\binom{n}{n}$? (Uvažujte všechna celá čísla n.)

▶ **10.** [*M25*] Pokud je p prvočíslo, ukažte, že:

a) $\binom{n}{p} \equiv \left\lfloor \frac{n}{p} \right\rfloor$ (modulo p).

b) $\binom{p}{k} \equiv 0$ (modulo p), pro $1 \le k \le p-1$.

c) $\binom{p-1}{k} \equiv (-1)^k$ (modulo p), pro $0 \le k \le p-1$.

d) $\binom{p+1}{k} \equiv 0$ (modulo p), pro $2 \le k \le p-1$.

e) (É. Lucas, 1877.)

$$\binom{n}{k} \equiv \binom{\lfloor n/p \rfloor}{\lfloor k/p \rfloor} \binom{n \bmod p}{k \bmod p} \quad \text{(modulo } p\text{)}.$$

f) Jestliže vyjádření čísel n a k v p-ární číselné soustavě můžeme zapsat jako

$$n = a_r p^r + \cdots + a_1 p + a_0, \qquad k = b_r p^r + \cdots + b_1 p + b_0,$$

$$\text{pak} \quad \binom{n}{k} \equiv \binom{a_r}{b_r} \cdots \binom{a_1}{b_1} \binom{a_0}{b_0} \quad \text{(modulo } p\text{)}.$$

▶ **11.** [*M20*] (E. Kummer, 1852.) Nechť p je prvočíslo. Ukažte, že pokud p^n dělí

$$\binom{a+b}{a}$$

ale p^{n+1} je nedělí, pak n je rovno počtu *přenosů*, které vzniknou při přičtení a k b v p-ární číselné soustavě. [*Nápověda:* Viz cvičení 1.2.5–12.]

12. [*M22*] Existuje nějaké kladné celé číslo n, pro které jsou všechny nenulové hodnoty v n-tém řádku Pascalova trojúhelníku *liché*? Pokud ano, najděte všechna taková n.

13. [*M13*] Dokažte sumační vzorec (10).

14. [*M21*] Vypočtěte hodnotu $\sum_{k=0}^{n} k^4$.

15. [*M15*] Dokažte binomickou větu (13).

16. [*M15*] Jsou-li dána kladná celá n a k, dokažte symetrickou rovnost

$$(-1)^n\binom{-n}{k-1} = (-1)^k\binom{-k}{n-1}.$$

▸ **17.** [*M18*] Dokažte Chu-Vandermondův vzorec (21) ze vztahu (15) a s využitím rovnosti $(1+x)^{r+s} = (1+x)^r(1+x)^s$.

18. [*M15*] Dokažte vztah (22) pomocí (21) a (6).

19. [*M18*] Dokažte vztah (23) indukcí.

20. [*M20*] Dokažte vztah (24) s použitím vztahů (21) a (19) a poté ukažte, že při jiném využití (19) dostanete (25).

▸ **21.** [*M05*] Obě strany vztahu (25) jsou polynomy v proměnné s; proč není tato rovnice identitou v s?

22. [*M20*] Dokažte vztah (26) pro speciální případ $s = n - 1 - r + nt$.

23. [*M13*] Předpokládejme, že (26) platí pro (r, s, t, n) a pro $(r,\ s-t,\ t,\ n-1)$; dokažte jej také pro $(r,\ s+1,\ t,\ n)$.

24. [*M15*] Vysvětlete, proč kombinací výsledků předchozích dvou cvičení dostaneme důkaz vztahu (26).

25. [*VM30*] Nechť je polynom $A_n(x,t)$ definován podle vztahu (30). Dále nechť $z = x^{t+1} - x^t$. Dokažte, že při dostatečně malé hodnotě z platí $\sum_k A_k(r,t)z^k = x^r$. [*Poznámka:* Pro $t = 0$ znamená tento výsledek v podstatě binomickou větu, přičemž uvedená rovnost je jejím důležitým zobecněním. V důkazu můžete využít platnost obyčejné binomické věty (15).] *Nápověda:* Začněte rovností

$$\sum_j (-1)^j\binom{k}{j}\binom{r-jt}{k}\frac{r}{r-jt} = \delta_{k0}.$$

26. [*VM25*] Z předpokladů v předcházejícím cvičení dokažte, že

$$\sum_k \binom{r-tk}{k} z^k = \frac{x^{r+1}}{(t+1)x - t}.$$

27. [*VM21*] Vyřešte Příklad 4 z textu pomocí výsledků cvičení 25 a dokažte vztah (26) z předchozích dvou cvičení. [*Nápověda:* Viz cvičení 17.]

28. [*M25*] Dokažte, že

$$\sum_k \binom{r+tk}{k}\binom{s-tk}{n-k} = \sum_{k\ge 0}\binom{r+s-k}{n-k}t^k,$$

pro nezáporné celé číslo n.

29. [*M20*] Ukažte, že vztah (34) je speciálním případem obecné rovnosti dokázané ve cvičení 1.2.3–33.

▸ **30.** [*M24*] Ukažte, že je možné Příklad 3 vyřešit lepším způsobem, než jaký byl uveden v textu; upravte sumu tak, že na ni budete moci aplikovat vztah (26).

▸ **31.** [*M20*] Vypočtěte

$$\sum_k \binom{m-r+s}{k}\binom{n+r-s}{n-k}\binom{r+k}{m+n}$$

a výsledek vyjádřete pomocí r, s, m a n, pokud m a n jsou celá čísla. Začněte náhradou

$$\binom{r+k}{m+n} \qquad \text{za} \qquad \sum_j \binom{r}{m+n-j}\binom{k}{j}.$$

32. [*M20*] Ukažte, že $\sum_k \left[{n \atop k}\right] x^k = x^{\overline{n}}$, kde $x^{\overline{n}}$ je rostoucí faktoriálová mocnina definovaná vztahem 1.2.5–(19).

33. [*M20*] (A. Vandermonde, 1772.) Ukažte, že binomická věta platí také v případě, že do ní namísto obyčejných mocnin zapíšeme faktoriálové mocniny. Jinými slovy: dokažte

$$(x+y)^{\underline{n}} = \sum_k \binom{n}{k} x^{\underline{k}} y^{\underline{n-k}}; \qquad (x+y)^{\overline{n}} = \sum_k \binom{n}{k} x^{\overline{k}} y^{\overline{n-k}}.$$

34. [*M23*] (*Torelliho součet.*) Ve světle předchozího cvičení ukažte, že Abelovo zobecnění binomické věty (16) platí také pro rostoucí mocniny:

$$(x+y)^{\overline{n}} = \sum_k \binom{n}{k} x(x-kz+1)^{\overline{k-1}}(y+kz)^{\overline{n-k}}.$$

35. [*M23*] Dokažte součtové vzorce (46) pro Stirlingova čísla přímo z definic, tedy ze vztahů (44) a (45).

36. [*M10*] Čemu se rovná součet $\sum_k \binom{n}{k}$ čísel v každém řádku Pascalova trojúhelníku? Jaký je součet těchto čísel se střídavými (alternujícími) znaménky, $\sum_k \binom{n}{k}(-1)^k$?

37. [*M10*] Z odpovědí na předchozí cvičení odvoďte hodnotu součtu každého druhého čísla v řádku, tedy $\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots$

38. [*VM30*] (C. Ramus, 1834.) Zobecněte výsledek předchozího cvičení a ukažte, že pro $0 \le k < m$ platí následující vzorec:

$$\binom{n}{k} + \binom{n}{m+k} + \binom{n}{2m+k} + \cdots = \frac{1}{m} \sum_{0 \le j < m} \left(2\cos\frac{j\pi}{m}\right)^n \cos\frac{j(n-2k)\pi}{m}.$$

Například:

$$\binom{n}{1} + \binom{n}{4} + \binom{n}{7} + \cdots = \frac{1}{3}\left(2^n + 2\cos\frac{(n-2)\pi}{3}\right).$$

[*Nápověda:* Najděte správnou kombinaci těchto koeficientů, násobených m-tou odmocninou jedničky.] Tato identita je pozoruhodná zejména pro $m \ge n$.

39. [*M10*] Čemu je roven součet $\sum_k \left[{n \atop k}\right]$ čísel v každém řádku Stirlingova prvního trojúhelníku? Jaký je součet těchto čísel se střídavými (alternujícími) znaménky? (Viz cvičení 36.)

40. [*VM17*] *Funkce* β $\mathrm{B}(x,y)$ je pro kladná reálná čísla x, y definována vztahem $\mathrm{B}(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$.

a) Ukažte, že $\mathrm{B}(x,1) = \mathrm{B}(1,x) = 1/x$.
b) Ukažte, že $\mathrm{B}(x+1,\, y) + \mathrm{B}(x,\, y+1) = \mathrm{B}(x,y)$.
c) Ukažte, že $\mathrm{B}(x,y) = \big((x+y)/y\big)\,\mathrm{B}(x,\, y+1)$.

41. [*VM22*] Ve cvičení 1.2.5–19 jsme dokázali relaci mezi funkcí gama a funkcí beta; pro kladné celé m jsme tak ukázali, že $\Gamma_m(x) = m^x\,\mathrm{B}(x, m+1)$.

a) Dokažte, že

$$\mathrm{B}(x,y) = \frac{\Gamma_m(y)m^x}{\Gamma_m(x+y)}\mathrm{B}(x,\, y+m+1).$$

b) Ukažte, že

$$\mathrm{B}(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

42. [*VM10*] Vyjádřete binomický koeficient $\binom{r}{k}$ pomocí výše definované funkce beta. (Tímto způsobem můžeme definici rozšířit na všechny reálné hodnoty k.)

43. [*VM20*] Ukažte, že $\mathrm{B}(1/2, 1/2) = \pi$. (Ze cvičení 41 můžeme nyní odvodit, že $\Gamma(1/2) = \sqrt{\pi}$.)

44. [*VM20*] Pomocí zobecněného binomického koeficientu, definovaného ve cvičení 42, ukažte

$$\binom{r}{1/2} = 2^{2r+1} \Big/ \binom{2r}{r}\, \pi.$$

45. [*VM21*] Pomocí zobecněného binomického koeficientu, definovaného ve cvičení 42, zjistěte $\lim_{r\to\infty} \binom{r}{k}/r^k$.

▸ **46.** [*M21*] Pomocí Stirlingovy aproximace, vztahu 1.2.5–(7), najděte přibližnou hodnotu $\binom{x+y}{y}$ při velkých hodnotách x a y. Zejména určete přibližnou velikost čísla $\binom{2n}{n}$ pro velké hodnoty n.

47. [*M21*] Je-li k celé číslo, ukažte

$$\binom{r}{k}\binom{r-1/2}{k} = \binom{2r}{k}\binom{2r-k}{k} \Big/ 4^k = \binom{2r}{2k}\binom{2k}{k} \Big/ 4^k.$$

Uveďte také jednodušší vzorec pro speciální případ $r = -1/2$.

▸ **48.** [*M25*] Ukažte, že

$$\sum_{k\ge 0} \binom{n}{k} \frac{(-1)^k}{k+x} = \frac{n!}{x(x+1)\dots(x+n)} = \frac{1}{x\binom{n+x}{n}},$$

pokud jmenovatele nejsou nulové. [Všimněte si, že tento vztah definuje převrácenou hodnotu binomického koeficientu a také rozvoj výrazu $1/x(x+1)\dots(x+n)$ do parciálních zlomků.]

49. [*M20*] Ukažte, že z rovnosti $(1+x)^r = (1-x^2)^r(1-x)^{-r}$ vyplývá relace nad binomickými koeficienty.

50. [*M20*] Dokažte Abelův vzorec (16) ve speciálním případě $x + y = 0$.

51. [*M21*] Dokažte Abelův vzorec (16) tak, že dosadíte $y = (x+y) - x$, rozvinete pravou stranu v mocninách $(x+y)$ a aplikujete výsledky předchozího cvičení.

52. [*VM11*] Dokažte, že pro nezáporné celé číslo n nemusí Abelova binomická věta (16) vždy platit; vypočtěte pravou stranu pro $n = x = -1$, $y = z = 1$.

53. [*M25*] (a) Dokažte následující rovnost indukcí podle m, pokud m a n jsou celá čísla:

$$\sum_{k=0}^{m} \binom{r}{k}\binom{s}{n-k}(nr-(r+s)k) = (m+1)(n-m)\binom{r}{m+1}\binom{s}{n-m}.$$

(b) Využijte důležité relace ze cvičení 47,

$$\binom{-1/2}{n} = \frac{(-1)^n}{2^{2n}}\binom{2n}{n},\quad \binom{1/2}{n} = \frac{(-1)^{n-1}}{2^{2n}(2n-1)}\binom{2n}{n} = \frac{(-1)^{n-1}}{2^{2n-1}(2n-1)}\binom{2n-1}{n} - \delta_{n0},$$

a ukažte, že je možné získat následující vzorec jako speciální případ rovnosti z části (a):

$$\sum_{k=0}^{m}\binom{2k-1}{k}\binom{2n-2k}{n-k}\frac{-1}{2k-1}=\frac{n-m}{2n}\binom{2m}{m}\binom{2n-2m}{n-m}+\frac{1}{2}\binom{2n}{n}.$$

(Tento výsledek je výrazně obecnější než (26) pro případ $r=-1$, $s=0$, $t=-2$.)

54. [*M21*] Uvažujte Pascalův trojúhelník (který je uveden v Tabulce 1) jako matici. Jak k ní vypadá *inverzní* matice?

55. [*M21*] Uvažujte každý ze Stirlingových trojúhelníků (viz Tabulka 2) jako matici a opět k nim určete matici inverzní.

56. [*20*] (*Kombinatorická číselná soustava.*) Pro každé celé číslo $n=0,\,1,\,2,\,\ldots,\,20$, najděte tři celá čísla a, b, c, pro která je $n=\binom{a}{3}+\binom{b}{2}+\binom{c}{1}$ a $a>b>c\geq 0$. Víte, jak bude tato konstrukce pokračovat pro vyšší hodnoty n?

▸ **57.** [*M22*] Ukažte, že koeficient a_m ve Stirlingově pokusu o zobecnění faktoriálové funkce, tedy vztahu 1.2.5–(12), je

$$\frac{(-1)^m}{m!}\sum_{k\geq 1}(-1)^k\binom{m-1}{k-1}\ln k.$$

58. [*M23*] (H. A. Rothe, 1811.) V notaci podle vztahu (40) dokažte „q-nomickou větu“:

$$(1+x)(1+qx)\ldots(1+q^{n-1}x)=\sum_k\binom{n}{k}_q q^{k(k-1)/2}x^k.$$

Najděte také q-nomické zobecnění základních identit (17) a (21).

59. [*M25*] Posloupnost čísel A_{nk}, $n\geq 0$, $k\geq 0$, splňuje relace $A_{n0}=1$, $A_{0k}=\delta_{0k}$, $A_{nk}=A_{(n-1)k}+A_{(n-1)(k-1)}+\binom{n}{k}$ pro $nk>0$. Určete A_{nk}.

▸ **60.** [*M23*] Viděli jsme, že $\binom{n}{k}$ je počet kombinací k-té třídy z n, tedy počet způsobů, jakými je možné vybrat k různých prvků z množiny n prvků. Obyčejným kombinacím se podobají také *kombinace s opakováním*, ve kterých můžeme každý prvek vybrat i vícekrát. Do seznamu (1) tak můžeme u kombinací s opakováním navíc zapsat *aaa*, *aab*, *aac*, *aad*, *aae*, *abb* atd. Kolik je kombinací k-té třídy z n prvků s opakováním?

61. [*M25*] Vypočtěte sumu

$$\sum_k\begin{bmatrix}n+1\\k+1\end{bmatrix}\begin{Bmatrix}k\\m\end{Bmatrix}(-1)^{k-m},$$

a získejte tak podobný vztah jako v (55).

▸ **62.** [*M23*] V textu jsou uvedeny vzorce pro součty se součinem dvou binomických koeficientů. Mezi sumami se součinem tří binomických koeficientů se zdá být nejužitečnější následující vztah a identita ze cvičení 31:

$$\sum_k(-1)^k\binom{l+m}{l+k}\binom{m+n}{m+k}\binom{n+l}{n+k}=\frac{(l+m+n)!}{l!\,m!\,n!},\quad\text{pro celé } l,m,n\geq 0.$$

(Suma zahrnuje kladné i záporné hodnoty k.) Dokažte tuto rovnost.
[*Nápověda:* Existuje velmi krátký důkaz, na jehož začátku aplikujeme výsledek cvičení 31.]

63. [*M30*] Jsou-li l, m a n celá čísla a je-li $n \geq 0$, dokažte že

$$\sum_{j,k}(-1)^{j+k}\binom{j+k}{k+l}\binom{r}{j}\binom{n}{k}\binom{s+n-j-k}{m-j} = (-1)^l\binom{n+r}{n+l}\binom{s-r}{m-n-l}.$$

▸ **64.** [*M20*] Ukažte, že $\{{n \atop m}\}$ je počet způsobů rozkladu množiny n prvků do m neprázdných disjunktních podmnožin. Množinu $\{1,2,3,4\}$ můžeme například rozdělit na dvě podmnožiny $\{{4 \atop 2}\} = 7$ způsoby: $\{1,2,3\}\{4\}$; $\{1,2,4\}\{3\}$; $\{1,3,4\}\{2\}$; $\{2,3,4\}\{1\}$; $\{1,2\}\{3,4\}$; $\{1,3\}\{2,4\}$; $\{1,4\}\{2,3\}$. *Nápověda:* Využijte vztah (46).

65. [*VM35*] (B. F. Logan.) Dokažte vztahy (59) a (60).

66. [*VM30*] Předpokládejme, že x, y a z jsou reálná čísla, která splňují vztah

$$\binom{x}{n} = \binom{y}{n} + \binom{z}{n-1},$$

kde $x \geq n-1$, $y \geq n-1$, $z > n-2$ a n je celé číslo ≥ 2. Dokažte, že

$$\binom{x}{n-1} \leq \binom{y}{n-1} + \binom{z}{n-2} \qquad \text{tehdy a jen tehdy, je- li} \qquad y \geq z;$$
$$\binom{x}{n+1} \leq \binom{y}{n+1} + \binom{z}{n} \qquad \text{tehdy a jen tehdy, je- li} \qquad y \leq z.$$

▸ **67.** [*M20*] Často potřebujeme zjistit, jestli binomické koeficienty nejsou příliš velké. Dokažte tuto snadno zapamatovatelnou horní hranici

$$\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k, \qquad \text{je-li } n \geq k \geq 0.$$

68. [*M25*] (A. de Moivre.) Dokažte, že pro nezáporné celé n platí

$$\sum_k \binom{n}{k} p^k (1-p)^{n-k} |k - np| = 2\lceil np \rceil \binom{n}{\lceil np \rceil} p^{\lceil np \rceil} (1-p)^{n+1-\lceil np \rceil}.$$

1.2.7. Harmonická čísla

Následující suma bude mít pro naši další práci velký význam:

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = \sum_{k=1}^{n} \frac{1}{k}, \quad n \geq 0. \tag{1}$$

V klasické matematice se tato suma příliš často neobjevuje, a nemá proto zavedenu ani žádnou standardní notaci; při analýze algoritmů se s ní ale setkáváme každou chvíli, a proto ji budeme důsledně označovat H_n. (Kromě H_n najdeme v matematické literatuře také notace h_n a S_n, případně $\psi(n+1)+\gamma$. Písmeno H znamená „harmonická“ a číslo H_n nazýváme *harmonické číslo*, protože rozvoj (1) se běžně nazývá harmonickou řadou.)

Na první pohled by se mohlo zdát, že čísla H_n nijak příliš nerostou ani při velkých hodnotách n, protože do řady přičítáme stále menší a menší čísla. Ve skutečnosti ale není těžké nahlédnout, že H_n při vhodných velkých hodnotách n roste nade všechny meze, protože

$$H_{2^m} \geq 1 + \frac{m}{2}. \tag{2}$$

Tato dolní hranice vychází z pozorování, že pro $m \geq 0$ máme

$$\begin{aligned} H_{2^{m+1}} &= H_{2^m} + \frac{1}{2^m+1} + \frac{1}{2^m+2} + \cdots + \frac{1}{2^{m+1}} \\ &\geq H_{2^m} + \frac{1}{2^{m+1}} + \frac{1}{2^{m+1}} + \cdots + \frac{1}{2^{m+1}} = H_{2^m} + \tfrac{1}{2}. \end{aligned}$$

Jestliže se tedy m zvýší o jedničku, zvýší se levá strana vztahu (2) nejméně o $\frac{1}{2}$.

Hodnotu čísel H_n je ale důležité zkoumat detailněji než jen pomocí nerovnosti (2). Přibližná velikost čísla H_n je dobře známou veličinou (přinejmenším v matematických kruzích) a můžeme ji vyjádřit

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \epsilon, \quad 0 < \epsilon < \frac{1}{252n^6}. \tag{3}$$

Číslo $\gamma = 0,5772156649\ldots$ je zde *Eulerova konstanta*, kterou zavedl Leonhard Euler v *Commentarii Acad. Sci. Imp. Pet.* **7** (1734), 150–161. Přesné hodnoty H_n pro malá n jsou spolu s hodnotou γ na 40 desetinných míst uvedeny v tabulkách přílohy A. Vztah (3) si odvodíme v části 1.2.11.2.

H_n je tedy poměrně rozumně blízké přirozenému logaritmu n. Ve cvičení 7(a) si jednoduše ukážeme, že má H_n skutečně poněkud logaritmické chování.

V jistém slova smyslu se ale H_n pro rostoucí n pouze blíží nekonečnu, protože podobná suma

$$1 + \frac{1}{2^r} + \frac{1}{3^r} + \cdots + \frac{1}{n^r} \tag{4}$$

je pro všechna n ohraničená, pokud je r libovolný reálný exponent *větší* než jedna. (Viz cvičení 3.) Sumu ze vztahu (4) označujeme jako $H_n^{(r)}$.

Pokud je exponent r ve vztahu (4) roven nejméně 2, je hodnota $H_n^{(r)}$ dosti blízká své maximální hodnotě $H_\infty^{(r)}$, s výjimkou velmi malých n. Veličina $H_\infty^{(r)}$ je v matematice velmi dobře známá a říká se jí *Riemannova funkce zeta*:

$$H_\infty^{(r)} = \zeta(r) = \sum_{k\geq 1} \frac{1}{k^r}. \tag{5}$$

Je známo, že pokud je r sudé celé číslo, je hodnota $\zeta(r)$ rovna

$$H_\infty^{(r)} = \frac{1}{2}\,|B_r|\,\frac{(2\pi)^r}{r!}, \qquad \text{pro celé } r/2 \geq 1, \tag{6}$$

kde B_r je Bernoulliho číslo (viz část 1.2.11.2 a Příloha A). Zejména je:

$$H_\infty^{(2)} = \frac{\pi^2}{6}, \quad H_\infty^{(4)} = \frac{\pi^4}{90}, \quad H_\infty^{(6)} = \frac{\pi^6}{945}, \quad H_\infty^{(8)} = \frac{\pi^8}{9450}. \tag{7}$$

Těchto výsledků dosáhl Euler; podrobnější výklad a důkaz viz *CMath*, § 6.5.

Nyní se podíváme na několik důležitých sum s harmonickými čísly. Nejprve

$$\sum_{k=1}^{n} H_k = (n+1)H_n - n. \tag{8}$$

Tento vztah vyplývá z jednoduché záměny pořadí sumace:

$$\sum_{k=1}^{n}\sum_{j=1}^{k}\frac{1}{j}=\sum_{j=1}^{n}\sum_{k=j}^{n}\frac{1}{j}=\sum_{j=1}^{n}\frac{n+1-j}{j}.$$

Vztah (8) je speciálním případem sumy $\sum_{k=1}^{n}\binom{k}{m}H_k$, kterou nyní vypočteme pomocí důležité metody nazývané sumace po částech (viz cvičení 10). Sumace po částech je užitečnou technikou výpočtu sumy $\sum a_k b_k$ a můžeme ji použít, pokud existuje jednoduchá forma veličin $\sum a_k$ a $(b_{k+1}-b_k)$. V tomto případě je

$$\binom{k}{m}=\binom{k+1}{m+1}-\binom{k}{m+1},$$

a proto

$$\binom{k}{m}H_k=\binom{k+1}{m+1}\Big(H_{k+1}-\frac{1}{k+1}\Big)-\binom{k}{m+1}H_k;$$

tudíž

$$\begin{aligned}\sum_{k=1}^{n}\binom{k}{m}H_k&=\Big(\binom{2}{m+1}H_2-\binom{1}{m+1}H_1\Big)+\cdots\\&\quad+\Big(\binom{n+1}{m+1}H_{n+1}-\binom{n}{m+1}H_n\Big)-\sum_{k=1}^{n}\binom{k+1}{m+1}\frac{1}{k+1}\\&=\binom{n+1}{m+1}H_{n+1}-\binom{1}{m+1}H_1-\frac{1}{m+1}\sum_{k=0}^{n}\binom{k}{m}+\frac{1}{m+1}\binom{0}{m}.\end{aligned}$$

Aplikací vztahu 1.2.6–(11) již dostaneme požadovaný vzorec:

$$\sum_{k=1}^{n}\binom{k}{m}H_k=\binom{n+1}{m+1}\Big(H_{n+1}-\frac{1}{m+1}\Big).\tag{9}$$

(Toto odvození a jeho konečný výsledek jsou analogií k výpočtu

$$\int_1^n x^m\ln x\,dx=\frac{n^{m+1}}{m+1}\Big(\ln n-\frac{1}{m+1}\Big)+\frac{1}{(m+1)^2}$$

metodou, která se v učebnicích integrálního počtu nazývá integrace per partes.)

V závěru této části textu se podíváme na jiný typ sumy, $\sum_k\binom{n}{k}x^kH_k$, kterou si na chvíli stručněji označíme S_n. Vidíme, že

$$\begin{aligned}S_{n+1}=\sum_k\left(\binom{n}{k}+\binom{n}{k-1}\right)x^kH_k&=S_n+x\sum_{k\ge1}\binom{n}{k-1}x^{k-1}\Big(H_{k-1}+\frac{1}{k}\Big)\\&=S_n+x\,S_n+\frac{1}{n+1}\sum_{k\ge1}\binom{n+1}{k}x^k.\end{aligned}$$

Je tudíž $S_{n+1} = (x+1)S_n + \big((x+1)^{n+1} - 1\big)/(n+1)$, přičemž máme

$$\frac{S_{n+1}}{(x+1)^{n+1}} = \frac{S_n}{(x+1)^n} + \frac{1}{n+1} - \frac{1}{(n+1)(x+1)^{n+1}}.$$

Z této rovnice a z faktu, že $S_1 = x$, pak dostáváme

$$\frac{S_n}{(x+1)^n} = H_n - \sum_{k=1}^{n} \frac{1}{k(x+1)^k}. \tag{10}$$

Nová suma je součástí nekonečné řady 1.2.9–(17) pro $\ln\big(1/(1-1/(x+1))\big) = \ln(1+1/x)$, přičemž pro $x > 0$ je řada konvergentní; rozdíl je

$$\sum_{k>n} \frac{1}{k(x+1)^k} < \frac{1}{(n+1)(x+1)^{n+1}} \sum_{k\ge 0} \frac{1}{(x+1)^k} = \frac{1}{(n+1)(x+1)^n x}.$$

Tím jsme dokázali následující větu:

Věta A. *Je-li $x > 0$, pak*

$$\sum_{k=1}^{n} \binom{n}{k} x^k H_k = (x+1)^n \left(H_n - \ln\Big(1 + \frac{1}{x}\Big) \right) + \epsilon,$$

kde $0 < \epsilon < 1/\big(x(n+1)\big)$. ∎

CVIČENÍ

1. [*01*] Čemu se rovná H_0, H_1 a H_2?

2. [*13*] Ukažte, že mírnou úpravou argumentu, pomocí něhož jsme v textu dokázali $H_{2^m} \ge 1 + m/2$, můžeme dokázat také $H_{2^m} \le 1 + m$.

3. [*M21*] Zobecněním argumentu z předchozího cvičení ukažte, že pro $r > 1$ zůstává suma $H_n^{(r)}$ ohraničená pro všechna n. Najděte příslušnou horní hranici.

▸ **4.** [*10*] Rozhodněte, které z následujících výroků platí pro všechna kladná celá n: (a) $H_n < \ln n$. (b) $H_n > \ln n$. (c) $H_n > \ln n + \gamma$.

5. [*15*] S pomocí tabulek v Příloze A uveďte hodnotu H_{10000} na 15 desetinných míst.

6. [*M15*] Dokažte, že harmonická čísla přímo souvisí se Stirlingovými čísly, která jsme si zavedli v předcházející části textu; fakticky je

$$H_n = \begin{bmatrix} n+1 \\ 2 \end{bmatrix} \Big/ n!.$$

7. [*M21*] Nechť $T(m,n) = H_m + H_n - H_{mn}$. (a) Ukažte, že s rostoucím m nebo n se $T(m,n)$ nikdy nezvyšuje (pokud jsou m a n kladná čísla). (b) Vypočtěte minimální a maximální hodnoty $T(m,n)$ pro $m, n > 0$.

8. [*VM18*] Porovnejte vztah (8) se sumou $\sum_{k=1}^{n} \ln k$; odhadněte rozdíl jako funkci proměnné n.

▸ **9.** [*M18*] Věta A platí jen pro $x > 0$; čemu se rovná zmíněná suma pro $x = -1$?

10. [*M20*] (*Sumace po částech.*) Speciální případy obecné metody sumace po částech jsme využili ve cvičení 1.2.4–42 a v odvození vztahu (9). Dokažte obecný vzorec

$$\sum_{1\le k<n}(a_{k+1}-a_k)b_k = a_nb_n - a_1b_1 - \sum_{1\le k<n} a_{k+1}(b_{k+1}-b_k).$$

▸ **11.** [*M21*] Metodou sumace po částech vypočtěte

$$\sum_{1<k\le n}\frac{1}{k(k-1)}H_k.$$

▸ **12.** [*M10*] Vypočtěte $H_\infty^{(1000)}$ s přesností nejméně na 100 desetinných míst.

13. [*M22*] Dokažte rovnost

$$\sum_{k=1}^{n}\frac{x^k}{k} = H_n + \sum_{k=1}^{n}\binom{n}{k}\frac{(x-1)^k}{k}.$$

(Všimněte si zejména speciálního případu $x=0$, ze kterého vyplývá identita související se cvičením 1.2.6–48.)

14. [*M22*] Ukažte, že $\sum_{k=1}^n H_k/k = \frac12(H_n^2+H_n^{(2)})$, a vypočtěte $\sum_{k=1}^n H_k/(k+1)$.

▸ **15.** [*M23*] Vyjádřete $\sum_{k=1}^n H_k^2$ pomocí n a H_n.

16. [*18*] Vyjádřete sumu $1+\frac13+\cdots+\frac{1}{2n-1}$ pomocí harmonických čísel.

17. [*M24*] (E. Waring, 1782.) Nechť p je liché prvočíslo. Ukažte, že čitatel H_{p-1} je dělitelný p.

18. [*M33*] (J. Selfridge.) Jaká nejvyšší mocnina 2 dělí čitatele zlomku $1+\frac13+\cdots+\frac{1}{2n-1}$?

▸ **19.** [*M30*] Uveďte všechna nezáporná celá čísla n, pro která je H_n celé číslo. [*Nápověda:* Pokud má H_n lichého čitatele a sudého jmenovatele, nemůže být celým číslem.]

20. [*VM22*] Sumační problémy, jako je například problém vedoucí v této části textu k Větě A, je možné řešit také analytickým způsobem: Pokud je $f(x)=\sum_{k\ge0}a_kx^k$ a pokud tato řada pro $x=x_0$ konverguje, dokažte, že

$$\sum_{k\ge0}a_kx_0^kH_k = \int_0^1\frac{f(x_0)-f(x_0y)}{1-y}\,dy.$$

21. [*M24*] Vypočtěte $\sum_{k=1}^n H_k/(n+1-k)$.

22. [*M28*] Vypočtěte $\sum_{k=0}^n H_kH_{n-k}$.

▸ **23.** [*VM20*] Uvažujte nyní funkci $\Gamma'(x)/\Gamma(x)$ a ukažte, jak je možné vytvořit přirozené zobecnění funkce H_n na neceločíselné hodnoty n. Můžete využít rovnosti $\Gamma'(1)=-\gamma$ a předběhnout tak následující cvičení.

24. [*VM21*] Ukažte, že

$$xe^{\gamma x}\prod_{k\ge1}\Big(\Big(1+\frac{x}{k}\Big)e^{-x/k}\Big) = \frac{1}{\Gamma(x)}.$$

(Uvažujte částečné součiny tohoto nekonečného součinu.)

1.2.8. Fibonacciho čísla

Posloupnost

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ \ldots, \tag{1}$$

ve které je každé z čísel součtem dvou předcházejících, hraje významnou roli v nejméně desítce zdánlivě nesouvisejících algoritmů, které budeme později studovat. Čísla v posloupnosti se označují F_n a formálně je definujeme

$$F_0 = 0; \qquad F_1 = 1; \qquad F_{n+2} = F_{n+1} + F_n, \quad n \geq 0. \tag{2}$$

Tuto slavnou posloupnost zveřejnil roku 1202 Leonardo Pisánský (Leonardo z Pisy), který bývá nazýván Leonardo Fibonacci (*Filius Bonaccii*, syn Bonaccia). Jeho *Liber Abaci* (Kniha o abaku) obsahuje následující cvičení: „Kolik párů králíků se rozmnoží z jediného páru za jeden rok?“ Pro řešení tohoto problému jsou dále uvedeny zjednodušující předpoklady, podle nichž každý pár zplodí každý měsíc nový párek potomků a tento nový pár začne rodit ve věku jednoho měsíce. Králíci navíc neumírají. Za jeden měsíc budeme mít 2 páry králíků, po dvou měsících jsou již 3, příštího měsíce původní pár i pár narozený v prvním měsíci přivedou na svět nový pár a celkem jich bude 5 a tak dále.

Fibonacci byl zdaleka největším evropským matematikem středověku. Studoval dílo al-Chwārizmīho (po němž jsou pojmenovány „algoritmy“, viz část 1.1) a přispěl celou řadou originálních poznatků do aritmetiky i geometrie. Jeho spisy byly přetištěny ve dvou svazcích roku 1857 [B. Boncompagni, *Scritti di Leonardo Pisano* (Řím, 1857–1862); posloupnost F_n je uvedena ve svazku 1, 283–285]. Problém králíků nebyl samozřejmě předložen jako praktická aplikace z biologie nebo z růstu populace; bylo to cvičení ze sčítání. Také na počítačích je dobrým cvičením ze sčítání (viz cvičení 3); sám Fibonacci napsal: „Toto [sčítání] je možné provádět po nekonečně mnoho měsíců.“

Než posloupnost $\langle F_n \rangle$ popsal ve svém díle Fibonacci, hovořili o ní již dříve indičtí učenci, kteří se dlouho zajímali o rytmické vzorky tvořené z jednodobých a dvoudobých not či slabik. Počet takovýchto rytmů s n dobami je právě F_{n+1}; proto se již Gopāla (před 1135) a Hemacandra (okolo 1150) výslovně zmiňují o posloupnosti čísel 1, 2, 3, 5, 8, 13, 21, 34, ... [Viz P. Singh, *Historia Math.* **12** (1985), 229–244; viz též cvičení 4.5.3–32.]

Stejná posloupnost se objevuje také v díle Johannese Keplera z roku 1611, který přemítal o číslech okolo sebe [J. Kepler, *The Six-Cornered Snowflake* (Oxford: Clarendon Press, 1966), 21]. Kepler podle všeho nevěděl, že se o této posloupnosti již zmínil Fibonacci. Fibonacciho čísla byla mnohokrát pozorována v přírodě, zřejmě z důvodů podobných původním úvahám k problému s králíky. [Neobyčejně přehledné vysvětlení viz Conway a Guy, *The Book of Numbers* (New York: Copernicus, 1996), 113–126.]

První náznaky úzké spojitosti mezi čísly F_n a algoritmy přišly na světlo roku 1837, kdy É. Léger využil Fibonacciho posloupnost ke studiu efektivity Euklidova algoritmu. Zjistil, že pokud čísla m a n v Algoritmu 1.1E nejsou větší než F_k, proběhne krok E2 nejvýše $k+1$-krát. To byl první případ praktického uplatnění Fibonacciho posloupnosti. (Viz Věta 4.5.3F.) Během 70. let 19. století získal matematik É. Lucas velmi přesné výsledky o Fibonacciho číslech a zejména pomocí

nich dokázal, že číslo $2^{127}-1$ se 39 číslicemi je prvočíslem. Lucas pojmenoval posloupnost $\langle F_n \rangle$ „Fibonacciho čísla“ a takto se jí říká dodnes.

Fibonacciho čísla jsme již krátce zkoumali v části 1.2.1 (viz vztah (3) a cvičení 4), kde jsme zjistili, že $\phi^{n-2} \le F_n \le \phi^{n-1}$, pokud je n kladné celé číslo a pokud

$$\phi = \tfrac{1}{2}(1+\sqrt{5}\,). \tag{3}$$

Za chvilku uvidíme, že tato veličina ϕ je s Fibonacciho čísly velmi úzce spjata.

Samotné číslo ϕ má velmi zajímavou historii. Euklides je nazýval „poměrem krajnosti a středu“; poměr A ku B je totiž stejný jako $A+B$ ku A, pokud je poměr A ku B roven ϕ. Renesanční spisovatelé jej nazývali „božským poměrem“ (divina proportio) a v 19. století začal být běžně nazýván „zlatý řez“. Mnoho umělců a spisovatelů tvrdí, že poměr ϕ ku 1 je ze všech esteticky nejpříjemnější proporcí, a jejich názor se potvrzuje i z pohledu estetiky programování počítačů. Celý příběh čísla ϕ si můžete přečíst ve vynikajícím článku “The Golden Section, Phyllotaxis, and Wythoff’s Game”, který napsal H. S. M. Coxeter, *Scripta Math.* **19** (1953), 135–143; viz též kapitola 8 knihy *The 2nd Scientific American Book of Mathematical Puzzles and Diversions*, jejímž autorem je Martin Gardner (New York: Simon and Schuster, 1961). Několik hojně rozšířených mýtů o čísle ϕ rozebral George Markowsky v *College Math. J.* **23** (1992), 2–19. Jeden z prvních evropských počtářských mistrů Simon Jacob, který zemřel roku 1564, již věděl, že se poměr F_{n+1}/F_n blíží k ϕ [viz P. Schreiber, *Historia Math.* **22** (1995), 422–424].

Notace, které používáme v této části textu, jsou poněkud zapomenuté. Ve výrazné části vyšší matematické literatury se F_n označuje jako u_n a číslo ϕ se pro změnu nazývá τ. Naše notace se používá téměř výhradně v rekreační matematice (a také v některé brakové literatuře!) a rychle se rozšiřuje. Označení ϕ pochází ze jména řeckého umělce Feidiáse, který údajně používal zlatý řez na svých sochách. Zápis F_n je pak v souladu se zvyklostí používanou ve *Fibonacci Quarterly*, kde čtenář najde řadu zajímavých faktů o Fibonacciho posloupnosti. Dobrým zdrojem z klasické literatury k F_n je kapitola 17 knihy L. E. Dickson, *History of the Theory of Numbers* **1** (Carnegie Inst. of Washington, 1919).

Fibonacciho čísla splňují množství zajímavých rovností, z nichž některé jsou uvedeny ve cvičeních na konci této části textu. Jedna z nejčastěji uváděných relací, o které se zmiňuje Kepler v dopise z roku 1608, ale kterou poprvé publikoval až J. D. Cassini [*Histoire Acad. Roy. Paříž* **1** (1680), 201], je

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n, \tag{4}$$

což snadno dokážeme indukcí. Zajímavější způsob důkazu stejného vzorce začíná jednoduchým induktivním důkazem maticové rovnosti

$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n. \tag{5}$$

Poté z obou stran rovnosti vypočteme determinant.

Z relace (4) vidíme, že F_n a F_{n+1} jsou nesoudělné, protože jakýkoli společný dělitel by musel být také dělitelem čísla $(-1)^n$.

Podle definice (2) okamžitě zjistíme, že

$$F_{n+3} = F_{n+2} + F_{n+1} = 2F_{n+1} + F_n\,; \quad F_{n+4} = 3F_{n+1} + 2F_n;$$

a dále obecně indukcí

$$F_{n+m} = F_m F_{n+1} + F_{m-1} F_n \tag{6}$$

pro každé kladné celé číslo m.

Jestliže ve vztahu (6) položíme m za násobek n, indukcí zjistíme, že

$$F_{nk} \textit{ je násobkem } F_n.$$

Každé třetí číslo je tak sudé, každé čtvrté je násobkem 3, každé páté je násobkem 5 atd.

Ve skutečnosti platí ale mnohem více tvrzení. Jestliže největšího společného dělitele m a n označíme jako $\mathrm{nsd}(m, n)$, platí dosti překvapivá věta:

Věta A (É. Lucas, 1876). *Číslo dělí F_m a F_n tehdy a jen tehdy, je-li dělitelem F_d, kde $d = \mathrm{nsd}(m, n)$; zejména,*

$$\mathrm{nsd}(F_m, F_n) = F_{\mathrm{nsd}(m,n)}. \tag{7}$$

Důkaz. Tento výsledek dokážeme pomocí Euklidova algoritmu. Vidíme, že vzhledem k (6) je každý společný dělitel F_m a F_n také dělitelem F_{n+m}, a naopak, každý společný dělitel F_{n+m} a F_n je také dělitelem $F_m F_{n+1}$. Protože F_{n+1} je s F_n nesoudělné, společný dělitel F_{n+m} a F_n dělí také F_m. Dokázali jsme tudíž, že pro libovolné číslo d

$$d \textit{ dělí } F_m \textit{ a } F_n \textit{ tehdy a jen tehdy, pokud } d \textit{ dělí } F_{m+n} \textit{ a } F_n. \tag{8}$$

Nyní ukážeme, že *každá* posloupnost $\langle F_n \rangle$, pro kterou platí tvrzení (8) a pro kterou je $F_0 = 0$, splňuje také Větu A.

Nejprve si ujasníme, že výrok (8) můžeme indukcí podle k rozšířit do tvaru

$$d \textit{ dělí } F_m \textit{ a } F_n \textit{ tehdy a jen tehdy, pokud } d \textit{ dělí } F_{m+kn} \textit{ a } F_n,$$

kde k je libovolné nezáporné celé číslo. Tento výsledek můžeme formulovat stručněji:

$$d \textit{ dělí } F_{m \bmod n} \textit{ a } F_n \textit{ tehdy a jen tehdy, pokud } d \textit{ dělí } F_m \textit{ a } F_n. \tag{9}$$

Jestliže nyní r je zbytek po dělení m číslem n, tedy jestliže $r = m \bmod n$, pak každý společný dělitel $\{F_m, F_n\}$ je také společným dělitelem $\{F_n, F_r\}$. Z toho vyplývá, že při manipulaci v Algoritmu 1.1E zůstává množina společných dělitelů $\{F_m, F_n\}$ nezměněna i při měnícím se m a n; konečně pokud je $r = 0$, jsou společné dělitele jednoduše děliteli $F_0 = 0$ a $F_{\mathrm{nsd}(m,n)}$. ∎

Většinu důležitých výsledků ohledně Fibonacciho čísel je možné odvodit z reprezentace F_n pomocí ϕ, kterou si nyní ukážeme. Metoda, kterou povedeme následující odvození, je neobyčejně důležitá, a matematicky orientovaný čtenář

by ji měl pečlivě prostudovat; stejnou metodu budeme podrobněji zkoumat v následující části textu.

Začneme vytvořením nekonečné řady

$$\begin{aligned} G(z) &= F_0 + F_1 z + F_2 z^2 + F_3 z^3 + F_4 z^4 + \cdots \\ &= z + z^2 + 2z^3 + 3z^4 + \cdots \end{aligned} \tag{10}$$

Zde nemáme *a priori* žádný důvod očekávat, že popsaná nekonečná suma existuje nebo že je funkce $G(z)$ vůbec nějak zajímavá – ale nepodléhejme skepsi a podívejme se, jaké závěry o funkci $G(z)$ odvodíme, pokud existuje. Výhodou takového postupu je, že $G(z)$ je jediná veličina, která reprezentuje *celou* Fibonacciho posloupnost najednou; pokud navíc zjistíme, že $G(z)$ je nějaká „známá“ funkce, můžeme stanovit i její koeficienty. Funkci $G(z)$ nazveme *generující funkcí* pro posloupnost $\langle F_n \rangle$.

Můžeme tedy pokračovat v šetření funkce $G(z)$:

$$\begin{aligned} zG(z) &= F_0 z + F_1 z^2 + F_2 z^3 + F_3 z^4 + \cdots, \\ z^2 G(z) &= \qquad F_0 z^2 + F_1 z^3 + F_2 z^4 + \cdots; \end{aligned}$$

a tudíž po odečtení

$$\begin{aligned} (1 - z - z^2)G(z) = F_0 &+ (F_1 - F_0)z + (F_2 - F_1 - F_0)z^2 \\ &+ (F_3 - F_2 - F_1)z^3 + (F_4 - F_3 - F_2)z^4 + \cdots \end{aligned}$$

Všechny členy až na druhý se vzhledem k definici F_n vyruší, takže celý výraz se rovná z. Vidíme tedy, že pokud funkce $G(z)$ existuje, je

$$G(z) = z/(1 - z - z^2). \tag{11}$$

Tuto funkci *můžeme* ve skutečnosti rozvinout do nekonečné řady v proměnné z (Taylorovy řady); při zpětném postupu zjistíme, že koeficienty rozvoje vztahu (11) do mocninné řady musí být Fibonacciho čísla.

Nyní budeme pracovat s funkcí $G(z)$ a zjistíme další vlastnosti Fibonacciho posloupnosti. Jmenovatel $1 - z - z^2$ je tvořen kvadratickou rovnicí se dvěma kořeny, $\frac{1}{2}(-1 \pm \sqrt{5}\,)$; po menším výpočtu zjistíme, že funkci $G(z)$ můžeme metodou parciálních zlomků rozvinout do tvaru

$$G(z) = \frac{1}{\sqrt{5}} \left(\frac{1}{1 - \phi z} - \frac{1}{1 - \widehat{\phi}\, z} \right), \tag{12}$$

kde

$$\widehat{\phi} = 1 - \phi = \tfrac{1}{2}(1 - \sqrt{5}\,). \tag{13}$$

Veličina $1/(1 - \phi z)$ je součtem nekonečné geometrické řady $1 + \phi z + \phi^2 z^2 + \cdots$, takže máme

$$G(z) = \frac{1}{\sqrt{5}} \left(1 + \phi z + \phi^2 z^2 + \cdots - 1 - \widehat{\phi}\, z - \widehat{\phi}^{\,2} z^2 - \cdots \right).$$

Nyní se podíváme na koeficient při z^n, který musí být roven F_n, a tudíž

$$F_n = \frac{1}{\sqrt{5}} \left(\phi^n - \widehat{\phi}^{\,n} \right). \tag{14}$$

To je důležitý výraz pro Fibonacciho čísla v uzavřené formě, který byl poprvé objeven začátkem 18. století. (Viz D. Bernoulli, *Comment. Acad. Sci. Petrop.* **3** (1728), 85–100, §7; viz též A. de Moivre, *Philos. Trans.* **32** (1722), 162–178, který ukázal, jak řešit obecné lineárně rekurentní posloupnosti v podstatě stejným způsobem, jaký jsme zde odvodili ve vztahu (14).)

Mohli bychom jednoduše vyslovit vztah (14) a dokázat jej indukcí. Podstatou celého poměrně dlouhého odvození bylo ale ukázat, jak je vůbec možné tuto rovnici *objevit*, a to pomocí důležité metody generujících funkcí, která je důležitým postupem pro řešení širokého spektra problémů.

Ze vztahu (14) můžeme dokázat spoustu věcí. Nejprve si všimneme, že $\hat\phi$ je *záporné* číslo $(-0,61803\ldots)$ s absolutní hodnotou menší než 1, takže při zvyšujícím se n je $\hat\phi^n$ velmi malé. Veličina $\hat\phi^n/\sqrt5$ je dokonce vždy tak malá, že je

$$F_n = \phi^n/\sqrt5 \text{ } \textit{zaokrouhleno na nejbližší celé číslo.} \tag{15}$$

Další výsledky dostaneme přímo z $G(z)$; například:

$$G(z)^2 = \frac15\left(\frac{1}{(1-\phi z)^2} + \frac{1}{(1-\hat\phi z)^2} - \frac{2}{1-z-z^2}\right); \tag{16}$$

koeficient při z^n ve vztahu $G(z)^2$ je $\sum_{k=0}^n F_k F_{n-k}$. Odvodíme tedy, že

$$\begin{aligned}\sum_{k=0}^n F_k F_{n-k} &= \tfrac15\big((n+1)(\phi^n+\hat\phi^n) - 2F_{n+1}\big)\\ &= \tfrac15\big((n+1)(F_n+2F_{n-1}) - 2F_{n+1}\big)\\ &= \tfrac15(n-1)F_n + \tfrac25 nF_{n-1}.\end{aligned} \tag{17}$$

(Druhý krok v tomto odvození vyplývá z výsledků cvičení 11.)

CVIČENÍ

1. [*10*] Jak zní odpověď na původní problém Leonarda Fibonacciho: Kolik párů králíků budeme mít v chovu po jednom roce?

▶ **2.** [*20*] Podíváte-li se na vztah (15), jaká je přibližná hodnota F_{1000}? (Využijte logaritmy z Přílohy A.)

3. [*25*] Napište počítačový program, který vypočte a vypíše hodnoty F_1 až F_{1000} v desítkové notaci. (Velikost zpracovávaných čísel jsme určili v předchozím cvičení.)

▶ **4.** [*14*] Najděte všechna n, pro která je $F_n = n$.

5. [*20*] Najděte všechna n, pro která je $F_n = n^2$.

6. [*VM10*] Dokažte vztah (5).

▶ **7.** [*15*] Není-li n prvočíslem, pak ani F_n není prvočíslem (s jednou výjimkou). Tvrzení dokažte a najděte výjimku.

8. [*15*] V řadě případů je vhodné definovat F_n i pro *záporná* n, a to podle předpokladu, že $F_{n+2} = F_{n+1} + F_n$ pro *všechna* celá čísla n. Prozkoumejte tuto definici. Kolik je F_{-1}? Kolik je F_{-2}? Je možné vyjádřit F_{-n} jednoduše pomocí F_n?

9. [*M20*] Na základě konvencí ze cvičení 8 stanovte, jestli platí vztahy (4), (6), (14) a (15), pokud indexy mohou nabývat *libovolných* celých čísel.

10. [*15*] Je $\phi^n/\sqrt{5}$ větší než F_n, nebo menší než F_n?

11. [*M20*] Ukažte, že $\phi^n = F_n\phi + F_{n-1}$ a $\hat\phi^n = F_n\hat\phi + F_{n-1}$ pro *všechna* celá čísla n.

▸ **12.** [*M26*] Fibonacciho posloupnost „druhého řádu“ je definována pravidlem

$$\mathcal{F}_0 = 0, \quad \mathcal{F}_1 = 1, \quad \mathcal{F}_{n+2} = \mathcal{F}_{n+1} + \mathcal{F}_n + F_n.$$

Vyjádřete $\mathcal{F}_n$ pomocí F_n a F_{n+1}. [*Nápověda:* Využijte generující funkce.]

▸ **13.** [*M22*] Vyjádřete následující posloupnosti pomocí Fibonacciho čísel. když r, s a c jsou dané konstanty:

a) $a_0 = r$, $a_1 = s$; $a_{n+2} = a_{n+1} + a_n$, pro $n \ge 0$.
b) $b_0 = 0$, $b_1 = 1$; $b_{n+2} = b_{n+1} + b_n + c$, pro $n \ge 0$.

14. [*M28*] Nechť m je pevně dané kladné celé číslo. Najděte a_n, pokud je

$$a_0 = 0, \quad a_1 = 1; \qquad a_{n+2} = a_{n+1} + a_n + \binom{n}{m}, \quad \text{pro } n \ge 0.$$

15. [*M22*] Nechť $f(n)$ a $g(n)$ jsou libovolné funkce a pro $n \ge 0$ nechť

$$\begin{aligned} a_0 &= 0, \quad a_1 = 1, \quad a_{n+2} = a_{n+1} + a_n + f(n);\\ b_0 &= 0, \quad b_1 = 1, \quad b_{n+2} = b_{n+1} + b_n + g(n);\\ c_0 &= 0, \quad c_1 = 1, \quad c_{n+2} = c_{n+1} + c_n + xf(n) + yg(n). \end{aligned}$$

Vyjádřete c_n pomocí x, y, a_n, b_n a F_n.

▸ **16.** [*M20*] Fibonacciho čísla jsou skryta v Pascalově trojúhelníku, pokud se na něj podíváme ze správného úhlu. Ukažte, že součet následujících binomických koeficientů je Fibonacciho číslo:

$$\sum_{k=0}^{n} \binom{n-k}{k}.$$

17. [*M24*] Pomocí konvence ze cvičení 8 dokažte následující zobecnění vztahu (4): $F_{n+k}F_{m-k} - F_nF_m = (-1)^n F_{m-n-k}F_k$.

18. [*20*] Je $F_n^2 + F_{n+1}^2$ vždy Fibonacciho číslem?

▸ **19.** [*M27*] Čemu se rovná $\cos 36^\circ$?

20. [*M16*] Vyjádřete $\sum_{k=0}^{n} F_k$ pomocí Fibonacciho čísel.

21. [*M25*] Čemu se rovná $\sum_{k=0}^{n} F_k x^k$?

▸ **22.** [*M20*] Ukažte, že $\sum_k \binom{n}{k} F_{m+k}$ je Fibonacciho číslo.

23. [*M23*] Zobecněte předcházející cvičení a ukažte, že $\sum_k \binom{n}{k} F_t^k F_{t-1}^{n-k} F_{m+k}$ je vždy Fibonacciho číslo.

24. [*VM20*] Vypočtěte determinant řádu $n \times n$

$$\begin{pmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & 1 \end{pmatrix}.$$

25. [*M21*] Ukažte, že

$$2^n F_n = 2 \sum_{k \text{ liché}} \binom{n}{k} 5^{(k-1)/2}.$$

▸ **26.** [*M20*] Na základě předchozího cvičení ukažte, že $F_p \equiv 5^{(p-1)/2}$ (modulo p), pokud je p liché prvočíslo.

27. [*M20*] Na základě předchozího cvičení ukažte, že pokud je p prvočíslo různé od 5, pak buďto F_{p-1}, nebo F_{p+1} (nikoli obojí) je násobkem p.

28. [*M21*] Kolik je $F_{n+1} - \phi F_n$?

▸ **29.** [*M23*] (*Fibonomické koeficienty.*) Édouard Lucas definoval veličiny

$$\binom{n}{k}_{\mathcal{F}} = \frac{F_n F_{n-1} \ldots F_{n-k+1}}{F_k F_{k-1} \ldots F_1} = \prod_{j=1}^{k} \left(\frac{F_{n-k+j}}{F_j} \right)$$

podobným způsobem jako binomické koeficienty. (a) Vytvořte tabulku čísel $\binom{n}{k}_{\mathcal{F}}$ pro $0 \le k \le n \le 6$. (b) Ukažte, že $\binom{n}{k}_{\mathcal{F}}$ je vždy celé číslo, protože máme

$$\binom{n}{k}_{\mathcal{F}} = F_{k-1} \binom{n-1}{k}_{\mathcal{F}} + F_{n-k+1} \binom{n-1}{k-1}_{\mathcal{F}}.$$

▸ **30.** [*M38*] (D. Jarden, T. Motzkin.) Posloupnost m-tých mocnin Fibonacciho čísel splňuje rekurentní relaci, ve které je každý člen závislý na předchozích $m+1$ členech. Ukažte, že

$$\sum_k \binom{m}{k}_{\mathcal{F}} (-1)^{\lceil (m-k)/2 \rceil} F_{n+k}^{m-1} = 0, \quad \text{je-li } m > 0.$$

Pro $m = 3$ dostáváme například rovnost $F_n^2 - 2F_{n+1}^2 - 2F_{n+2}^2 + F_{n+3}^2 = 0$.

31. [*M20*] Ukažte, že $F_{2n}\phi \bmod 1 = 1 - \phi^{-2n}$ a $F_{2n+1}\phi \bmod 1 = \phi^{-2n-1}$.

32. [*M24*] Zbytek po dělení Fibonacciho čísla jiným je ± Fibonacciho číslo. Ukažte, že modulo F_n platí:

$$F_{mn+r} \equiv \begin{cases} F_r, & \text{pokud } m \bmod 4 = 0; \\ (-1)^{r+1} F_{n-r}, & \text{pokud } m \bmod 4 = 1; \\ (-1)^n F_r, & \text{pokud } m \bmod 4 = 2; \\ (-1)^{r+1+n} F_{n-r}, & \text{pokud } m \bmod 4 = 3. \end{cases}$$

33. [*VM24*] Je-li dáno $z = \pi/2 + i \ln \phi$, ukažte, že $\sin nz / \sin z = i^{1-n} F_n$.

▸ **34.** [*M24*] (*Fibonacciho číselná soustava.*) Nechť notace $k \gg m$ znamená, že $k \ge \ge m+2$. Ukažte, že každé kladné celé číslo n má *jednoznačnou* reprezentaci ve tvaru $n = F_{k_1} + F_{k_2} + \cdots + F_{k_r}$, kde $k_1 \gg k_2 \gg \cdots \gg k_r \gg 0$.

35. [*M24*] (*Číselná soustava ϕ.*) Uvažujme zápis reálných čísel pomocí číslic 0 a 1 a základu ϕ; konkrétně tak $(100,1)_\phi = \phi^2 + \phi^{-1}$. Ukažte, že číslo 1 je možné reprezentovat nekonečně mnoha způsoby, například $1 = (0,11)_\phi = (0,011111\ldots)_\phi$. Jestliže ale navíc vyslovíme požadavek, že se v rozvoji nevyskytují dvě po sobě jdoucí jedničky a že celá reprezentace nekončí v nekonečné posloupnosti 01010101…, pak každé nezáporné číslo má jedinečnou reprezentaci. Jak vypadá reprezentace celých čísel?

▸ **36.** [*M32*] (*Fibonacciho řetězce.*) Nechť $S_1 =$ „a“, $S_2 =$ „b“ a $S_{n+2} = S_{n+1}S_n$, $n > > 0$; jinými slovy S_{n+2} vytvoříme připojením S_n na pravý konec S_{n+1}. Máme $S_3 = =$ „ba“, $S_4 =$ „bab“, $S_5 =$ „$babba$“ atd. Je zřejmé, že S_n má F_n písmen. Zkoumejte nyní vlastnosti S_n. (Kde se vyskytují zdvojená písmena? Dokážete předpovědět hodnotu k-tého písmene S_n? Jaká je hustota výskytu písmen b? A tak dále.)

▸ **37.** [*M35*] (R. E. Gaskell, M. J. Whinihan.) Dva hráči hrají následující hru: Na hromádce je n kamenů; první hráč odebere libovolně mnoho kamenů, nesmí ale odebrat celou hromádku. Dále se hráči v tazích střídají a každý odebere jeden nebo více kamenů, ale odebrat *nesmí více než dvakrát tolik kamenů jako předchozí hráč*. Vítězí hráč, který odebere poslední kámen. (Uvažujme například $n = 11$; hráč A odebere 3 kameny, hráč B může odebrat nejvýše 6 kamenů a odebere jen 1. Zbývá 7 kamenů; hráč A může odebrat 1 nebo 2 a vezme 2; poté hráč B může odebrat nejvýše 4 a vezme 1. Nyní zbývají 4 kameny; hráč A vezme 1, hráč B musí odebrat nejméně jeden kámen a v příštím tahu hráč A vyhraje.)

Jaký nejlepší tah může udělat první hráč, pokud je na začátku hry 1000 kamenů?

38. [*35*] Napište počítačový program, který bude hrát hru popsanou v předchozím cvičení a který bude realizovat vítěznou strategii.

39. [*M24*] Najděte vyjádření a_n v uzavřené formě, pokud je $a_0 = 0$, $a_1 = 1$, a $a_{n+2} = a_{n+1} + 6a_n$ pro $n \geq 0$.

40. [*M25*] Vyřešte rekurentní vztah

$$f(1) = 0; \quad f(n) = \min_{0<k<n} \max\bigl(1 + f(k), 2 + f(n-k)\bigr), \quad \text{pro } n > 1.$$

▸ **41.** [*M25*] (Jurij Matijaševič, 1990.) Nechť $f(x) = \lfloor x + \phi^{-1} \rfloor$. Dokažte, že pokud $n = F_{k_1} + \cdots + F_{k_r}$ je reprezentace čísla n ve Fibonacciho číselné soustavě ze cvičení 34, pak $F_{k_1+1} + \cdots + F_{k_r+1} = f(\phi n)$. Najděte podobný vzorec pro $F_{k_1-1} + \cdots + F_{k_r-1}$.

42. [*M26*] (D. A. Klarner.) Ukažte, že pokud m a n jsou nezáporná celá čísla, existuje jednoznačná posloupnost indexů $k_1 \gg k_2 \gg \cdots \gg k_r$ taková že,

$$m = F_{k_1} + F_{k_2} + \cdots + F_{k_r}, \qquad n = F_{k_1+1} + F_{k_2+1} + \cdots + F_{k_r+1}.$$

(Viz cvičení 34. Čísla k mohou být záporná a r může být nulové.)

1.2.9. Generující funkce

Potřebujeme-li zjistit určité informace o posloupnosti čísel $\langle a_n \rangle = a_0, a_1, a_2, \ldots$, můžeme definovat nekonečnou sumu s „parametrem“ z¿

$$G(z) = a_0 + a_1 z + a_2 z^2 + \cdots = \sum_{n \geq 0} a_n z^n. \tag{1}$$

Poté již můžeme zkoumat vlastnosti funkce G. Tato funkce je jediná veličina, která reprezentuje celou posloupnost; jestliže posloupnost $\langle a_n \rangle$ byla definována induktivně, má podobná funkce velké výhody. Navíc jestliže předpokládáme, že pro nějaké nenulové hodnoty z existuje nekonečná suma uvedená v (1), můžeme zjistit i jednotlivé hodnoty a_0, a_1, ... z funkce $G(z)$, a to metodami diferenciálního počtu.

Funkci $G(z)$ nazýváme *generující funkcí* posloupnosti a_0, a_1, a_2, ... Generující funkce (někdy v české literatuře nazývané také „vytvořující funkce“) před námi otevírají celý nový svět různých metod a rozšiřují náš potenciál pro řešení problémů. Jak jsme se zmínili v předchozí části textu, zavedl generující funkce A. de Moivre, a to pro řešení obecného problému lineární rekurentní posloupnosti. De Moivreovu teorii později rozšířil James Stirling na o něco složitější rekurentní posloupnosti a ukázal, jak kromě aritmetických operací aplikovat také derivování a integrování [*Methodus Differentialis* (Londýn: 1730), Tvrzení 15].

O několik let později začal generující funkce používat několika novými způsoby také L. Euler, například ve svých materiálech o rozkladech množin [*Commentarii Acad. Sci. Pet.* **13** (1741), 64–93; *Novi Comment. Acad. Sci. Pet.* **3** (1750), 125–169]. Metody dále rozvinul Pierre S. Laplace ve svém klasickém díle *Théorie Analytique des Probabilités* (Paříž: 1812).

Jistý význam má otázka konvergence nekonečné sumy (1). V každé učebnici teorii nekonečných řad se dokazuje, že:

a) Jestliže řada konverguje pro určitou hodnotu $z = z_0$, pak konverguje pro všechny hodnoty z takové, že $|z| < |z_0|$.

b) Řada konverguje pro nějaké $z \neq 0$ tehdy a jen tehdy, pokud je posloupnost $\langle \sqrt[n]{|a_n|} \rangle$ ohraničená. (Není-li tato podmínka splněna, můžeme získat konvergentní řadu z posloupnosti $\langle a_n/n! \rangle$ nebo z nějaké jiné příbuzné posloupnosti.)

Při práci s generujícími funkcemi se ale na druhé straně často nevyplatí se příliš zabývat konvergencí řad, protože zkoumáme jenom možné přístupy k řešení jistého problému. Ať už řešení najdeme *jakýmikoli* prostředky, třeba těžkopádnými, zdůvodníme je pak třeba nezávisle. V předcházející části textu jsme například pomocí generující funkce odvodili vztah (14); po nalezení takovéto rovnosti je již ale jednoduché dokázat ji indukcí a už vůbec nemusíme uvažovat, že jsme ji odhalili prostřednictvím generující funkce. Navíc se dá ukázat, že většinu operací (ne-li všechny) s generujícími funkcemi je možné řádně zdůvodnit bez ohledu na konvergenci příslušné řady. Viz například E. T. Bell, *Trans. Amer. Math. Soc.* **25** (1923), 135–154; Ivan Niven, *AMM* **76** (1969), 871–889; Peter Henrici, *Applied and Computational Complex Analysis* **1** (Wiley, 1974), kapitola 1.

Podívejme se nyní na základní techniky práce s generujícími funkcemi.

A. Sčítání. Pokud je $G(z)$ generující funkce k posloupnosti $\langle a_n \rangle = a_0, a_1, \ldots$ a $H(z)$ je generující funkce k $\langle b_n \rangle = b_0, b_1, \ldots$, pak $\alpha G(z) + \beta H(z)$ je generující funkce k posloupnosti $\langle \alpha a_n + \beta b_n \rangle = \alpha a_0 + \beta b_0,\ \alpha a_1 + \beta b_1,\ \ldots$:

$$\alpha \sum_{n\geq 0} a_n z^n + \beta \sum_{n\geq 0} b_n z^n = \sum_{n\geq 0} (\alpha a_n + \beta b_n) z^n. \tag{2}$$

B. Posuv. Pokud je $G(z)$ generující funkce k posloupnosti $\langle a_n \rangle = a_0, a_1, \ldots$, pak $z^m G(z)$ je generující funkce k $\langle a_{n-m} \rangle = 0, \ldots, 0, a_0, a_1, \ldots$:

$$z^m \sum_{n\geq 0} a_n z^n = \sum_{n\geq m} a_{n-m} z^n. \tag{3}$$

Jestliže pro záporné hodnoty n položíme $a_n = 0$, můžeme poslední sumu rozšířit přes všechna $n \geq 0$.

Podobně $\big(G(z) - a_0 - a_1 z - \cdots - a_{m-1} z^{m-1}\big)/z^m$ je generující funkcí k posloupnosti $\langle a_{n+m} \rangle = a_m, a_{m+1}, \ldots$:

$$z^{-m} \sum_{n\geq m} a_n z^n = \sum_{n\geq 0} a_{n+m} z^n. \tag{4}$$

Spojením operací A a B můžeme nyní vyřešit Fibonacciho problém z předchozí části: Zde byla $G(z)$ generující funkcí pro $\langle F_n\rangle$, $zG(z)$ pro $\langle F_{n-1}\rangle$, $z^2G(z)$ pro $\langle F_{n-2}\rangle$, a $(1-z-z^2)G(z)$ pro $\langle F_n - F_{n-1} - F_{n-2}\rangle$. Protože $F_n - F_{n-1} - F_{n-2}$ je pro $n \geq 2$ rovno nule, zjistíme odtud, že $(1-z-z^2)G(z)$ je polynom. Podobně pro každou danou *lineárně rekurentní* posloupnost, tedy pro takovou posloupnost, kde $a_n = c_1a_{n-1} + \cdots + c_ma_{n-m}$, bude generující funkcí polynom dělený $(1 - c_1z - \cdots - c_mz^m)$.

Uvažujme nyní nejjednodušší příklad ze všech: Pokud je $G(z)$ generující funkce ke *konstantní* posloupnosti $1, 1, 1, \ldots$, pak $zG(z)$ generuje posloupnost $0, 1, 1, \ldots$, takže $(1-z)G(z) = 1$. Tím dostáváme jednoduchý, ale důležitý vztah

$$\frac{1}{1-z} = 1 + z + z^2 + \cdots \tag{5}$$

C. Násobení. Pokud je $G(z)$ generující funkce k posloupnosti $a_0, a_1, \ldots$ a $H(z)$ je generující funkce k $b_0, b_1, \ldots$, pak

$$\begin{aligned} G(z)H(z) &= (a_0 + a_1z + a_2z^2 + \cdots)(b_0 + b_1z + b_2z^2 + \cdots) \\ &= (a_0b_0) + (a_0b_1 + a_1b_0)z + (a_0b_2 + a_1b_1 + a_2b_0)z^2 + \cdots; \end{aligned}$$

a tudíž $G(z)H(z)$ je generující funkcí k posloupnosti $c_0, c_1, \ldots$, kde

$$c_n = \sum_{k=0}^{n} a_kb_{n-k}. \tag{6}$$

Vztah (3) je velmi speciálním případem této rovnosti. Další důležitý speciální případ vzniká, pokud jsou všechny koeficienty b_n jedničkové:

$$\frac{1}{1-z}G(z) = a_0 + (a_0 + a_1)z + (a_0 + a_1 + a_2)z^2 + \cdots \tag{7}$$

Dostáváme tak generující funkci k součtu původní posloupnosti.

Pravidlo součinu *tří* funkcí vyplývá ze vztahu (6); $F(z)G(z)H(z)$ generuje $d_0, d_1, d_2, \ldots$, kde

$$d_n = \sum_{\substack{i,j,k\geq 0 \\ i+j+k=n}} a_ib_jc_k. \tag{8}$$

Generující pravidlo součinu *libovolného počtu* funkcí (pokud má tato operace smysl) je

$$\prod_{j\geq 0}\sum_{k\geq 0} a_{jk}z^k = \sum_{n\geq 0} z^n \sum_{\substack{k_0,k_1,\ldots\geq 0 \\ k_0+k_1+\cdots=n}} a_{0k_0}a_{1k_1}\ldots \tag{9}$$

Pokud součástí rekurentní relace pro nějakou posloupnost jsou binomické koeficienty, budeme často potřebovat generující funkci k posloupnosti $c_0, c_1, \ldots$, definovanou vztahem

$$c_n = \sum_k \binom{n}{k} a_kb_{n-k}. \tag{10}$$

V tomto případě je obvykle lepší využít generující funkce k posloupnostem $\langle a_n/n!\rangle$, $\langle b_n/n!\rangle$, $\langle c_n/n!\rangle$, protože máme

$$\Big(\frac{a_0}{0!}+\frac{a_1}{1!}z+\frac{a_2}{2!}z^2+\cdots\Big)\Big(\frac{b_0}{0!}+\frac{b_1}{1!}z+\frac{b_2}{2!}z^2+\cdots\Big)=\Big(\frac{c_0}{0!}+\frac{c_1}{1!}z+\frac{c_2}{2!}z^2+\cdots\Big), \tag{11}$$

kde c_n je dáno vztahem (10).

D. Změna proměnné z. Funkce $G(cz)$ je zřejmě generující funkcí k posloupnosti $a_0, ca_1, c^2a_2, \ldots$ V konkrétním případě je generující funkce k $1, c, c^2, c^3, \ldots$ rovna $1/(1-cz)$.

Alternující (střídavé) členy řady zjistíme pomocí známého triku:

$$\begin{aligned}\tfrac12\big(G(z)+G(-z)\big)&=a_0\ +\ a_2z^2\ +\ a_4z^4\ +\ \cdots,\\ \tfrac12\big(G(z)-G(-z)\big)&=\ \ a_1z\ +\ a_3z^3\ +\ a_5z^5\ +\ \cdots\end{aligned} \tag{12}$$

Pomocí komplexních odmocnin jedné můžeme tuto myšlenku rozšířit a extrahovat každý m-tý člen: Nechť $\omega = e^{2\pi i/m} = \cos(2\pi/m) + i\sin(2\pi/m)$; potom máme

$$\sum_{n \bmod m=r} a_nz^n=\frac1m\sum_{0\le k<m}\omega^{-kr}G(\omega^kz),\quad 0\le r<m. \tag{13}$$

(Viz cvičení 14.) Pokud je například $m=3$ a $r=1$, dostáváme $\omega=-\frac12+\frac{\sqrt3}{2}i$, což je komplexní třetí odmocnina z jedné; odtud vyplývá, že

$$a_1z+a_4z^4+a_7z^7+\cdots=\tfrac13\big(G(z)+\omega^{-1}G(\omega z)+\omega^{-2}G(\omega^2z)\big).$$

E. Derivování a integrování. Metody infinitesimálního počtu nám přinášejí další operace. Pokud je $G(z)$ dána vztahem (1), její derivace je rovna

$$G'(z)=a_1+2a_2z+3a_3z^2+\cdots=\sum_{k\ge0}(k+1)a_{k+1}z^k. \tag{14}$$

Generující funkce k posloupnosti $\langle na_n\rangle$ je $zG'(z)$. Můžeme tudíž kombinovat n-tý člen posloupnosti s polynomem v n pouhými úpravami generující funkce.

Pokud celý postup obrátíme, dostaneme integrací jinou užitečnou operaci:

$$\int_0^z G(t)\,dt=a_0z+\frac12a_1z^2+\frac13a_2z^3+\cdots=\sum_{k\ge1}\frac1ka_{k-1}z^k. \tag{15}$$

Jako speciální případy pak dostáváme derivaci a integrál vztahu (5):

$$\frac{1}{(1-z)^2}=1+2z+3z^2+\cdots=\sum_{k\ge0}(k+1)z^k. \tag{16}$$

$$\ln\frac{1}{1-z}=z+\frac12z^2+\frac13z^3+\cdots=\sum_{k\ge1}\frac1kz^k. \tag{17}$$