průměrně dobří studenti. Velmi obtížná je stávající situace pro studenty kombinovaného studia. Tito studenti, většinou zaměstnaní lidé, sice i za normálních okolností studují zčásti distančně, výrazně na ně ale dopadají sekundární problémy současného uzavření: mají rodiny a doma děti, které nemohou chodit do školy.

V obtížných dobách máme být optimisty. Můj optimismus vychází z naděje, že řešení současné situace plošným uzavření společnosti bude opuštěno, školy – základní, střední i vysoké – budou opět otevřeny a společnost se i v jiných ohledech dostane do normálnějšího stavu.

Knuthovy vánoční stromky

EDUARD BARTL

Přírodovědecká fakulta UP, Olomouc

Zájemcům o informatiku je patrně dobře známé jméno amerického informatika, držitele Turingovy ceny, Donalda Ervina Knutha. Tento skromně vystupující vědec s jedinečným smyslem pro humor a renesančním záběrem (Knuth je mimo jiné zdatným varhaníkem) je tvůrcem typografického systému TEX, v němž je vysázen i časopis MFI,[1)] a autorem řady vědeckých článků a knih. Pozoruhodná je například knížka [3], ve které Knuth formou rozhovoru vysvětluje konstrukci takzvaných surreálných čísel. Bezpochyby nejvýznamnějším Knuthovým dílem je však mnohosvazková monografie *The Art of Computer Programming*, kterou začal psát během svých doktorských studií na Kalifornském technologickém institutu[2)] a na které ve svých více než 80 letech, jakožto emeritní profesor, stále pracuje.

The Art of Computer Programming je nepřebernou studnicí algoritmů z mnoha oblastí informatiky vybavených názornými obrázky vypracova-

[1)] Rád bych využil této příležitosti a poděkoval RNDr. Miloslavu Závodnému, který se mnoho let o sazbu časopisu ve zmíněném typografickém systému stará.

[2)] Kalifornský technologický institut (také zvaný Caltech) je soukromá univerzita, která se nachází v Pasadeně v USA. Absolventy jsou například Walter Bright, americký informatik, tvůrce programovacího jazyka D, Stephen Wolfram, britský fyzik a matematik, tvůrce počítačového programu Mathematica, nebo John McCarthy, americký informatik, jeden ze zakladatelů umělé inteligence (jako první začal tento termín používat) a tvůrce programovacího jazyka Lisp.

nými s pečlivostí, která je Knuthovi vlastní. Tuto skutečnost jsem si nedávno opět uvědomil, když jsem hledal způsoby generování a uspořádání všech bitových řetězců[3)] konečné délky. Ve čtvrtém svazku The Art of Computer Programming [4] jsem narazil na jednu vskutku zajímavou metodu, která uspořádává bitové řetězce do schématu Knuthem pojmenovaného jako *Christmas tree pattern*.

Důvodem pro tento poněkud neobvyklý název je zejména fakt, že schéma pro větší délky řetězců skutečně připomíná vánoční stromek. Knuth dále uvádí, že tento název zvolil také proto, že zmíněnou problematiku uspořádání všech bitových řetězců přednesl na tradiční vánoční přednášce „Christmas Tree Lecture“ [7] pořádané v prosinci 2002 na Stanfordově univerzitě.

Donald Ervin Knuth (zdroj: Wikipedia)

1. Jak vypadá Knuthův stromek

Jak bylo řečeno, jde nám o problém generování všech bitových řetězců konečné délky; tuto délku si označíme jako n. Chceme tedy získat a vhodně uspořádat celkem 2^n bitových řetězců. Vyřešit tento problém je snadné; stačí vzít přirozená čísla od 0 do 2^n-1, převést je do binární soustavy a vypsat je v pořadí, které je dáno uspořádáním přirozených čísel $0, \ldots, 2^n-1$ podle jejich velikosti. Například pro $n = 3$ získáme 8 tříprvkových bitových řetězců

$$000\ 001\ 010\ 011\ 100\ 101\ 110\ 111$$

Knuthovy stromky na to ovšem jdou mnohem chytřeji. Ukažme si nejprve, jak vypadá nejjednodušší Knuthův stromek, který uspořádává všechny bitové řetězce délky 1; říkáme, že se jedná o Knuthův stromek řádu 1. Jde o prostý řádek

$$0\ 1$$

[3)] Bitovými řetězci máme na mysli řetězce složené z nul a jedniček. Tento název (v angličtině *bit strings*) používá právě Donald Knuth v [4].

Knuthův stromek řádu 2 uspořádává všechny řetězce délky 2 do dvou řádků, jejichž sloupce jsou specificky zarovnané:

$$\begin{array}{c} 10 \\ 00\ \ 01\ \ 11 \end{array}$$

Obecný případ, tedy Knuthův stromek řádu $n+1$, je pak definovaný tak, že přepíšeme každý řádek Knuthova stromku řádu n do jednoho, nebo dvou řádků podle následujícího pravidla:

(i) pokud má řádek pouze jeden prvek (tento prvek si označíme jako σ_1), pak ho přepíšeme do jediného řádku ve tvaru

$$\sigma_1 0\ \sigma_1 1$$

(ii) pokud má řádek více než jeden prvek (tyto prvky si označíme jako $\sigma_1, \ldots, \sigma_s$, kde $s > 1$), pak ho přepíšeme do dvou řádků v tomto tvaru

$$\begin{array}{c} \sigma_2 0\ \ldots\ \ \sigma_s 0 \\ \sigma_1 0\ \sigma_1 1\ \ldots\ \sigma_{s-1} 1\ \sigma_s 1 \end{array}$$

Knuthův stromek řádu 3, který řeší výše uvedený příklad s uspořádáním všech bitových řetězců délky 3, je tedy vytvořen ze stromku řádu 2 následujícím způsobem. První řádek stromku řádu 2 má pouze jeden prvek $\sigma_1 = 10$; podle části (i) zmíněného pravidla přepíšeme tento řádek do jediného řádku o dvou prvcích (červeně je zvýrazněna nula a jednička, která byla podle pravidla připojena k bitovému řetězci σ_1):

$$100\ 101$$

Druhý řádek stromku řádu 2 má tři prvky $\sigma_1 = 00$, $\sigma_2 = 01$ a $\sigma_3 = 11$. Tento řádek tedy přepíšeme podle části (ii) pravidla do dvou řádků (červeně jsou opět zvýrazněny nuly a jedničky připojené k σ_1, σ_2 a σ_3):

$$\begin{array}{c} 010\ 110 \\ 000\ 001\ 011\ 111 \end{array}$$

Celý Knuthův stromek řádu 3 tudíž vypadá následovně:

$$\begin{array}{c} 100\ 101 \\ 010\ 110 \\ 000\ 001\ 011\ 111 \end{array}$$

Knuthův stromek řádu 4 má tuto podobu:

$$
\begin{array}{ccccc}
 & & 1010 & & \\
 & 1000 & 1001 & 1011 & \\
 & & 1100 & & \\
 & 0100 & 0101 & 1101 & \\
 & 0010 & 0110 & 1110 & \\
0000 & 0001 & 0011 & 0111 & 1111
\end{array}
$$

Analogickým způsobem pak můžeme konstruovat stromky vyšších řádů. Pro ilustraci si ještě předveďme Knuthův stromek řádu 8 (zobrazen je na konci článku). Tento stromek ukazuje uspořádání všech bitových řetězců délky 8 a můžeme tedy na něj nahlížet jako na schéma zachycující všechny bajty.

Vidíme, že definice Knuthova stromku řádu n je rekurzivní a všimněme si dále, že nezáleží pouze na uspořádání řetězců v jednotlivých řádcích, ale také že roli hraje uspořádání celých řádků a zarovnání jednotlivých sloupců. Z uvedených obrázků je také patrná jistá estetika – Knuthovy stromky vyšších řádů mohou svým tvarem vzdáleně připomínat fraktály. To, co je však skutečně zajímavé, není jejich vzhled, nýbrž vlastnosti ukazující, jak se v těchto schématech potkávají klasické výsledky z různých oblastí informatiky a matematiky.

Podívejme se nejprve na tři základní vlastnosti, které se dají snadno dokázat přímo z definice Knuthových stromků užitím matematické indukce:

(a) Každý bitový řetězec délky n se v Knuthově stromku řádu n vyskytuje právě jednou. Stromek řádu n má tedy 2^n bitových řetězců délky n.

(b) Jestliže očíslujeme sloupce Knuthova stromku zleva doprava, přičemž číslovat budeme od nuly,[4)] pak k-tý sloupec obsahuje všechny bitové řetězce s k jedničkami.

(c) Pro každý řádek Knuthova stromku platí, že bitový řetězec v $(k+1)$-tém sloupci vznikne z bitového řetězce v k-tém sloupci přepsáním jedné nuly na jedničku.

V Knuthových stromcích se toho ovšem skrývá mnohem více. Podívejme se tedy na ně podrobněji.

[4)] Toto číslování sloupců budeme používat i ve zbytku článku.

2. Spernerův teorém

Jedna z nejzajímavějších vlastností je ukryta v prostředním sloupci,[5] který tvoří, můžeme říct, *kmen* Knuthova stromku. Než se dostaneme k popisu zmíněné vlastnosti, uděláme poměrně podrobnou, zdánlivě nesouvisející odbočku k jednomu klasickému problému takzvané extremální teorie množin.

Budeme uvažovat n prvkovou množinu U; dejme tomu, že tato množina obsahuje čísla $1, 2, \ldots, n$ jako svoje prvky. Symbolem $\mathcal{P}(U)$ budeme označovat její *potenční množinu*, tedy množinu všech podmnožin množiny U; symbolem $|A|$ pak označíme počet prvků nějaké množiny A. Zajímat nás budou množiny podmnožin $\mathcal{F} \subseteq \mathcal{P}(U)$, které tvoří takzvaný *antiřetězec*. V antiřetězci jsou libovolné dvě různé množiny neporovnatelné vzhledem k množinové inkluzi $\subseteq$; to znamená, platí-li $A, B \in \mathcal{F}$ a $A \neq B$, pak $A \nsubseteq B$. Příkladem antiřetězce na množině $U = \{1, 2, 3, 4, 5\}$ je tato množina

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

K tomuto antiřetězci však můžeme přidat ještě další prvky tak, aby nebyla podmínka neporovnatelnosti porušena; následující množina je také antiřetězcem:

$$\{\{1\}, \{2, 3\}, \{3, 4\}, \{2, 4, 5\}\}. \tag{1}$$

Zajímat nás bude, kolik prvků budeme ještě schopni přidat tak, aby nově vzniklá množina byla stále antiřetězcem. Jinými slovy, pokusíme se odpovědět na otázku, kolik prvků má největší antiřetězec.

Abychom byli schopni odpovědět na tuto otázku, podíváme se nejprve na jeden jednoduchý způsob, jak systematicky vytvářet antiřetězce určitého typu: je totiž jasné, že stačí vzít všechny (nebo jen některé) podmnožiny se stejným počtem prvků. Například množina složená ze všech jednoprvkových podmnožin

$$\{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$$

je zjevně antiřetězcem obsahujícím $\binom{5}{1} = 5$ prvků. Množina dvouprvkových podmnožin

$$\{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{2, 3\}, \{2, 4\}, \{2, 5\}, \{3, 4\}, \{3, 5\}, \{4, 5\}\}$$

---

[5] Případně v prostředních dvou sloupcích. Z příkladů v předchozí sekci jsme totiž mohli vypozorovat, že pro sudé n je uprostřed jeden sloupec, pro liché n jsou však uprostřed sloupce dva.

je taktéž antiřetězec, který je větší než předcházející antiřetězec; má celkem $\binom{5}{2} = 10$ prvků. Obecně tedy množina složená ze všech k prvkových podmnožin je antiřetězcem o velikosti $\binom{n}{k}$.[6)] Zamyslíme-li se nyní nad tím, jak fungují kombinační čísla $\binom{n}{k}$, záhy zjistíme, že největší antiřetězec získaný tímto způsobem bude složen ze všech podmnožin o $k = \lfloor \frac{n}{2} \rfloor$ prvcích, kde $\lfloor \cdot \rfloor$ značí operaci odřezávající desetinnou část daného čísla. Velikost tohoto antiřetězce je proto rovna číslu

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}. \tag{2}$$

Je poměrně překvapujícím zjištěním, že není možné zkonstruovat antiřetězec, který by měl větší počet prvků než (2) a to ani v případě, že se budeme snažit antiřetězec sestavit tak, že bude obsahovat podmnožiny různých velikostí.[7)] Platí tedy tvrzení, že je-li množina $\mathcal{F} \subseteq \mathcal{P}(U)$ antiřetězcem, pak $|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. Tomuto tvrzení se říká *Spernerův teorém*; publikoval ho roku 1928 německý matematik Emanuel Sperner (1905–1980). Číslu (2) se pak často říká *Spernerovo číslo*.

Nahlédneme-li do původního Spernerova článku [6], zjistíme, že toto jednoduše vypadající tvrzení má dosti složitý a dlouhý důkaz. Časem se však objevilo několik prací, které ke Spernerovu teorému přistupují zcela odlišně a dokazují ho výrazně kratším způsobem. Ukážeme si jeden takový důkaz, který byl publikován roku 1966 pod příznačným názvem „A short proof of Sperner's lemma" [5] americkým matematikem Davidem Lubellem. Tento důkaz je založen na jednoduchých úvahách, ke kterým využijeme kombinatoriku ze střední školy. Poté se podíváme na to, jak souvisí s Knuthovými stromky (méně trpělivý čtenář tedy může důkaz přeskočit).

Pro Lubellův důkaz jsou klíčové posloupnosti podmnožin $C_0, \ldots, C_n$ množiny U, pro které platí tyto dvě podmínky:

$$\emptyset = C_0 \subset C_1 \subset \cdots \subset C_{n-1} \subset C_n = U,$$
$$|C_i| = i \text{ pro } i = 0, \ldots, n.$$

Posloupnost tedy začíná prázdnou množinou, množina C_1 obsahuje právě jeden prvek z U, C_2 vznikne z C_1 přidáním dalšího prvku z U a stejným způsobem pokračujeme až k množině C_n, která proto musí obsahovat

[6)] Připomeňme, že $\binom{n}{k}$ je takzvané *kombinační číslo*, které je rovno $\frac{n!}{k! \cdot (n-k)!}$. $\binom{n}{k}$ je počet možností, kterými lze vybrat k prvků z n prvků, přičemž nezáleží na pořadí vybraných prvků; jinými slovy, je to počet k prvkových podmnožin n prvkové množiny.

[7)] Podobně, jako jsme to učinili na začátku této sekce v příkladu s antiřetězcem (1).

všechny prvky z U. Tuto posloupnost podmnožin proto můžeme nazvat *řetězcem*.[8] Položme si otázku, kolik existuje takových různých řetězců. Odpověď je jednoduchá: řetězce se liší tím, v jakém pořadí přidáváme prvky z U; takových pořadí je přesně tolik, kolik je permutací n prvkové množiny U. Různých řetězců je tedy $n!$.

Uvažujme nyní libovolný antiřetězec $\mathcal{F} \subseteq \mathcal{P}(U)$ a libovolnou k prvkovou množinu $A \in \mathcal{F}$ tohoto antiřetězce. Zkusme vypočítat, kolik výše definovaných řetězců $C_0, \dots, C_n$ obsahuje A jako svůj prvek. Opět se jedná o jednoduchou kombinatorickou úvahu: jestliže je A jedním z prvků řetězce $C_0, \dots, C_n$, pak musí být na k-tém místě tohoto řetězce (protože A má k prvků), neboli $A = C_k$. Navíc tento řetězec musel vzniknout tak, že se **nejprve** postupně k prázdné množině přidalo v nějakém pořadí všech k prvků z A a teprve **poté** se k množině A přidalo v nějakém pořadí všech zbývajících $n - k$ prvků z U. Takových možností je celkem $k! \cdot (n-k)!$; počet řetězců obsahujících A je tedy roven právě tomuto číslu.

Poslední úvaha Lubellova důkazu, kterou je potřeba učinit, je následující. Označme si symbolem m_k počet k prvkových množin antiřetězce $\mathcal{F}$.[9] Je jasné, že počet všech prvků antiřetězce $\mathcal{F}$ je možné zapsat pomocí čísel m_k takto:

$$|\mathcal{F}| = \sum_{k=0}^{n} m_k. \tag{3}$$

Víme, že každá z k prvkových množin antiřetězce je obsažena v $k! \cdot (n-k)!$ řetězcích; těchto k prvkových množin je m_k, takže počet všech řetězců, které obsahují tyto k prvkové množiny je $m_k \cdot k! \cdot (n-k)!$. Všechny množiny antiřetězce $\mathcal{F}$ jsou tak obsaženy celkem v $\sum_{k=0}^{n} m_k \cdot k! \cdot (n-k)!$ navzájem různých řetězcích.[10] Celkový počet řetězců, jak jsme ukázali výše, je $n!$, takže právě uvedený součet nemůže přesáhnout tento faktoriál; platí tedy

---

[8] Jedná se o řetězec v tom smyslu, že je každý prvek posloupnosti porovnatelný s libovolným jiným prvkem této posloupnosti. Řetězec podmnožin je proto jistým opakem antiřetězce podmnožin. Pro jistotu dodejme, že termín „řetězec podmnožin" nijak nesouvisí s termínem „bitový řetězec". Anglicky psané texty jsou v tomto ohledu uvážlivější, pro řetězec podmnožin se používá slovo „chain" pro bitový řetězec slovo „string".

[9] Ve výše uvedeném příkladu antiřetězce (1) je $m_0 = 0$ (antiřetězec neobsahuje prázdnou množinu), $m_1 = 1$ (antiřetězec obsahuje jednu jednoprvkovou množinu) a dále pak $m_2 = 2$, $m_3 = 1$ a $m_4 = 0$.

[10] Bude dobré si uvědomit, že neexistuje řetězec, který by obsahoval dvě různé množiny $A, B \in \mathcal{F}$ daného antiřetězce. Důvod je nasnadě; A a B jsou neporovnatelné množiny (jelikož to jsou prvky antiřetězce) a řetězce z definice obsahují pouze porovnatelné množiny.

$\sum_{k=0}^{n} m_k \cdot k! \cdot (n-k)! \leq n!$. Uvedenou nerovnost můžeme vydělit $n!$ a dále provést několik jednoduchých úprav, čímž dostaneme

$$\begin{aligned}
\sum_{k=0}^{n} m_k \cdot \frac{k! \cdot (n-k)!}{n!} &\leq 1, \\
\sum_{k=0}^{n} \frac{m_k}{\frac{n!}{k! \cdot (n-k)!}} &\leq 1, \\
\sum_{k=0}^{n} \frac{m_k}{\binom{n}{k}} &\leq 1. \qquad (4)
\end{aligned}$$

Jak jsme si již řekli, největší hodnoty dosahují kombinační čísla $\binom{n}{k}$ pro $k = \lfloor \frac{n}{2} \rfloor$; levou stranu nerovnosti (4) tak můžeme zdola omezit:

$$\sum_{k=0}^{n} \frac{m_k}{\binom{n}{k}} \geq \sum_{k=0}^{n} \frac{m_k}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} = \frac{\sum_{k=0}^{n} m_k}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}.$$

Z tohoto omezení a nerovnosti (4) plyne:

$$\frac{\sum_{k=0}^{n} m_k}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq 1.$$

Využitím vztahu (3) konečně dostáváme

$$\frac{|\mathcal{F}|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq 1,$$

neboli

$$|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor},$$

což je přesně nerovnost, o které mluví Spernerův teorém.

Jak to všechno souvisí s Knuthovými stromky? Souvislost není obtížná. Na jednotlivé prvky Knuthova stromku řádu n se totiž můžeme dívat jako na *charakteristické vektory* podmnožin n prvkové množiny. Například ve výše uvedeném příkladu Knuthova stromku řádu 4 je prvek 0000 charakteristickým vektorem prázdné množiny $\emptyset$, prvek 0010 charakteristickým vektorem množiny $\{3\}$, prvek 0101 charakteristickým vektorem množiny $\{2, 4\}$

a podobně. Knuthův stromek nám při tomto množinovém pohledu poskytuje velmi jednoduchý důkaz Spernerova teorému; stačí pouze vhodně uchopit základní vlastnosti stromků uvedených v předchozí sekci.

Nejprve si uvědomíme, že díky základní vlastnosti (a) jsou v Knuthově stromku řádu n všechny možné podmnožiny n prvkové množiny, jsme proto schopni najít v něm všechny antiřetězce, o jejichž délky se zajímá Spernerův teorém. Ze základní vlastnosti (c) dále plyne, že pro dva bezprostředně sousedící prvky σ_k a σ_{k+1} na stejném řádku platí $\sigma_k \subseteq \sigma_{k+1}$. Každý řádek tak tvoří *řetězec podmnožin*.[11)] Pokud budeme chtít vytvořit antiřetězec, tak to můžeme udělat pouze tak, že z daného řádku Knuthova stromku vybereme nejvýše jeden prvek; více než jeden prvek vybrat nemůžeme prostě proto, že každý řádek je řetězcem, obsahuje tedy vzájemně porovnatelné prvky, které však v antiřetězci nemohou být. Tímto způsobem tedy budeme schopni vytvořit antiřetězec, který bude mít nejvýše tolik prvků, kolik je řádků v Knuthově stromku, přičemž počet řádků odpovídá výšce sloupce (případně sloupců pro n liché), které jsou uprostřed. Ze základní vlastnosti (b) ovšem víme, že k-tý sloupec obsahuje všechny podmnožiny s k prvky; počet prvků k-tého sloupce je tedy $\binom{n}{k}$. Prostřední sloupec má tedy výšku rovnou Spernerovu číslu, tedy $\binom{n}{\lfloor \frac{n}{2} \rfloor}$. Každý antiřetězec (který jsme vytvořili výběrem nejvýše jednoho prvku z každého řádku Knuthova stromku) proto může mít maximálně $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ prvků.

Výše prezentovaný Lubellův důkaz je velmi elegantní a zpravidla je uváděn jako nejjednodušší ze všech prozatím objevených důkazů Spernerova teorému. Vidíme však, že jedinečný způsob, jakým jsou v Knuthově stromku uspořádány všechny podmnožiny n prvkové množiny, umožňuje nerovnost uvedenou ve Spernerově teorému takřka odečíst z obrázku stromku – není třeba žádných složitých úvah a žádného složitého odvozování.

3. Systémy závorek

V této sekci se podíváme na vztah Knuthova stromku a takzvaných *systémů závorek*. Víme, že libovolný prvek na řádku vznikne tak, že se jedna nula z předcházejícího prvku přepíše na jedničku; každý řádek tak tvoří řetězec podmnožin. Přepisování nul na jedničky samozřejmě není nahodilé. Podrobnější pohled ukazuje, že se toto přepisování řídí jistým pra-

[11)] V tomto řetězci se navíc sousedící množiny liší pouze v jednom prvku; to ovšem není pro důkaz Spernerova teorému podstatné. Tuto skutečnost však využijeme později.

vidlem, které je dobře popsatelné právě prostřednictvím souvislosti Knuthova stromku se systémy závorek.

Těmito systémy máme na mysli libovolné (tedy i prázdné) posloupnosti levých a pravých kulatých závorek.[12)] Následuje několik příkladů systémů závorek: `()))`, `(())`, `(`, `()()`, `))((` a `(()())`. Tyto systémy mohou být *korektně uzávorkované* nebo *nekorektně uzávorkované*. Pro korektně uzávorkované systémy platí, volně řečeno, že každá závorka je v páru s právě jednou opačnou závorkou, přičemž jako první je uvedena levá závorka, jako druhá je potom uvedena pravá závorka (prázdný systém také považujeme za korektně uzávorkovaný). Pokud nějaká závorka nemá svého „partnera“ do páru, pak říkáme, že je *volná*. Uvedené systémy `(())`, `(()())` a `()()` jsou proto korektně uzávorkované,[13)] naproti tomu systémy `()))`, `(` a `))((` jsou nekorektně uzávorkované (v prvním případě jsou poslední dvě pravé závorky volné, ve druhém a třetím případě jsou volné všechny závorky).

Obecně platí, že libovolný systém závorek (ať už je korektně uzávorkovaný nebo není) je možné napsat ve formě následující posloupnosti:

$$\alpha_0\texttt{)}\ldots\alpha_{p-1}\texttt{)}\alpha_p\texttt{(}\alpha_{p+1}\ldots\texttt{(}\alpha_q, \tag{5}$$

kde $0 \leq p \leq q$ a $\alpha_0, \ldots, \alpha_q$ jsou korektně uzávorkované systémy závorek. Vidíme, že v tomto obecném tvaru je p volných pravých závorek a $q - p$ volných levých závorek. Uveďme si dva příklady: korektně uzávorkovaný systém `(()())` se ve tvaru (5) zapíše tak, že p a q nastavíme na nulu, platí tedy

$$\alpha_0 = \texttt{(()())};$$

pro nekorektně uzávorkovaný systém `))(()))(()` platí

$$\alpha_0\texttt{)}\alpha_1\texttt{)}\alpha_2\texttt{)}\alpha_3\texttt{(}\alpha_4 = \texttt{))(()))(()},$$

to znamená, že $p = 3$, $q = 4$, systémy závorek α_0, α_1 a α_3 jsou prázdné a pro zbývající systémy platí $\alpha_2 = \texttt{(())}$, $\alpha_4 = \texttt{()}$.[14)]

[12)] Bylo by jistě vhodnější říkat těmto posloupnostem *řetězce závorek*, rádi bychom se však vyhnuli použití slova „řetěze“ v dalším (již třetím) významu.

[13)] Korektně uzávorkované systémy jsou nám dobře známy z programovacích jazyků. Například programy napsané ve funkcionálních jazycích obsahují korektně uzávorkované systémy (stejně tak běžně psané texty (jako je kupříkladu tento) by měly obsahovat korektně uzávorkované systémy)).

[14)] Všimněme si, že všechny systémy α_0 až α_4 jsou korektně uzávorkované.

Dále platí, že libovolný systém závorek je prvkem následující posloupnosti systémů:

$$\begin{aligned}&\alpha_0\texttt{)}\ldots\alpha_{q-1}\texttt{)}\alpha_{q-1}\texttt{)}\alpha_q,\\&\alpha_0\texttt{)}\ldots\alpha_{q-1}\texttt{)}\alpha_{q-1}\texttt{(}\alpha_q,\\&\alpha_0\texttt{)}\ldots\alpha_{q-1}\texttt{(}\alpha_{q-1}\texttt{(}\alpha_q,\\&\ldots\\&\alpha_0\texttt{(}\ldots\alpha_{q-1}\texttt{(}\alpha_{q-1}\texttt{(}\alpha_q.\end{aligned}\tag{6}$$

Na první pohled to vypadá komplikovaně, jde však pouze o to, že první prvek posloupnosti (6) je systémem, který obsahuje pouze pravé volné závorky, druhý prvek vznikne z prvního tak, že poslední pravou volnou závorku přepíšeme na levou závorku (ta je proto také volná) a stejným způsobem pokračujeme i nadále. Poslední prvek posloupnosti (6) je pak systémem, který obsahuje pouze levé volné závorky. Libovolný systém, ať už korektně či nekorektně uzávorkovaný, tak musí být členem posloupnosti (6).

Vraťme se nyní ke Knuthovým stromkům a podívejme se na to, jak souvisí se systémy závorek. Spojitost se vyjeví v okamžiku, když přepíšeme v Knuthově stromku každou nulu na pravou závorku a každou jedničku na levou závorku. Stromky jsou pak kolekcí systémů závorek. Například Knuthův stromek řádu 3 bude vypadat takto:

```
    ()) ()(
    )() (()
))) ))( )(( (((
```

Přepsaný Knuthův stromek řádu 4 vypadá následovně:

```
           ()()
      ())) ())( ()((
           (())
      )()) )()( (()(
      ))() )(() ((()
)))) )))( ))(( )((( ((((
```

Můžeme tedy vidět, že každý řádek začíná systémem, ve kterém jsou volné pouze pravé závorky a každý další systém závorek na tomto řádku vznikne z předchozího systému přepsáním poslední volné pravé závorky na levou (tato vlastnost se dá opět dokázat matematickou indukcí). Řádky stromku jsou tedy přesně ve tvaru posloupností systémů závorek (6).

Z uvedeného je pak zřejmé, jak vypadá na začátku zmíněné pravidlo popisující přepisování nul na jedničky v libovolném řádku: postupně se přepisují zprava doleva nuly na jedničky, přeskakují se však takové nuly, ke kterým existuje jednička do páru (ve smyslu korektně uzávorkovaných systémů). Demonstrujeme to na řádku

$$0100 \quad 0101 \quad 1101$$

Knuthova stromku řádu 4, jehož celou podobu jsme si ukázali v předchozí sekci. Druhý prvek vznikne z prvního tak, že nula zcela napravo v tomto prvním prvku je přepsána na jedničku:

$$010\textcolor{red}{0} \mapsto 010\textcolor{red}{1}.$$

Třetí prvek z druhého však nevznikne přepsáním předposlední nuly, protože tato nula je v páru s jedničkou, přepisuje se tedy první nula:

$$\textcolor{red}{0}101 \mapsto \textcolor{red}{1}101.$$

Tento řádek již nebude mít žádný další prvek jednoduše proto, že nemáme další nuly k přepisování.

Toto zjištění pak vede k další zajímavé vlastnosti Knuthových stromků, ke které se dostaneme snadnou úvahou o jednoprvkových řádcích. Pokud se totiž na jednoprvkové řádky podíváme z pohledu právě uvedeného přepisování nul, ke kterým neexistují jedničky do páru, ihned zjistíme, že tyto řádky jsou jednoprvkové proto, že již není co přepisovat. Jednoprvkové řádky proto musí představovat korektně uzávorkované systémy! Ukážeme si to opět na příkladu Knuthova stromku řádu 4. Ten obsahuje pouze dva jednoprvkové řádky:

$$\begin{array}{c}
\textcolor{blue}{1010}\\
1000\ 1001\ 1011\\
\textcolor{blue}{1100}\\
0100\ 0101\ 1101\\
0010\ 0110\ 1110\\
0000\ 0001\ 0011\ 0111\ 1111
\end{array}$$

Tyto řádky odpovídají závorkovým systémům `()()` a `(())`, které jsou, jak vidno, korektně uzávorkované. Žádné další korektně uzávorkované systémy složené ze 4 závorek neexistují. Z obrázku Knuthova stromku řádu 8 (na s. 73) tak může vyčíst, že existuje celkem 14 korektně uzávorkovaných systémů.

Tím však problematika závorkových systémů kódovaných pomocí Knuthových stromků nekončí. Dá se například ukázat, že prvky, které jsou na konci řádků, jsou *lexikograficky uspořádané.*[15] Knuthův stromek proto ve svém kmenu obsahuje korektně uzávorkované systémy v lexikografickém uspořádání.

4. Heslovitě další souvislosti

Spektrum souvislostí Knuthových stromků s různými informatickými a matematickými výsledky nebylo v předchozích sekcích zdaleka vyčerpáno. V Knuthových stromcích je také zajímavým způsobem kódován *Catalanův trojúhelník* zavedený švýcarským matematikem Leonhardem Eulerem (1707–1783),[16] který má rozličné aplikace zejména v kombinatorice. Dále se pomocí Knuthových stromků řádu n a pomocí takzvané *Hanselovy vlastnosti* [2] dají velmi elegantním způsobem vyčíslovat monotonní booleovské funkce o n proměnných.

Podrobný popis těchto témat by však činil článek příliš dlouhým. Zvídavého čtenáře tedy odkazuji na příslušnou kapitolu čtvrtého svazku The Art of Computer Programming [4]. Zároveň s tím vyjadřuji závazek vrátit se někdy v budoucnosti ke zmíněným tématům ve volném pokračování tohoto textu.

Závěr

Čtenáři dlužím jedno vysvětlení. Autorem popsaného uspořádání všech bitových řetězců konečné délky není Donald Knuth, ale nizozemští matematici N. G. de Bruijn, C. van Ebbenhorst Tengbergen a D. Kruyswijk.

[15] Definici lexikografického uspořádání uvádět nebudeme. Pouze poznamenáme, že je použito například pro uspořádání slov ve slovnících. Prvky, které jsou nejvíce napravo v jednotlivých řádcích Knuthova stromku řádu 4, to znamená 1010, 1011, 1100, 1101, 1110, 1111, jsou proto uspořádané stejně jako by byly ve slovníku uspořádaná slova BABA, BABB, BBAA, BBAB, BBBA, BBBB (nulu jsme nahradili písmenem A, jedničku písmenem B).

[16] Ač definovaný Eulerem, je Catalanův trojúhelník pojmenovaný po belgickém matematikovi Eug?novi Catalanovi (1814–1894).

Když se však podíváme do jejich krátké, třístránkové publikace [1], ze které Knuth vychází, zjistíme, že v ní není řeč o všech bitových řetězcích, nýbrž o všech dělitelích daného přirozeného čísla. Vztah mezi bitovými řetězci a děliteli je celkem zjevný, byl to ovšem až Knuth, který si ho uvědomil. Také uspořádání do schématu podobného vánočnímu stromku je Knuthovým nápadem (i když také do určité míry vychází z práce de Bruijna a jeho dvou kolegů). Knuth sám, jak jsme již uvedli v úvodu, nazývá tato schémata Christmas tree patterns. Pojmenování „Knuthovy stromky" jsem zvolil právě kvůli přínosu Donalda Knutha do této oblasti informatiky.

Rychlé čtení tohoto textu by mohlo v leckterém čtenáři navodit představu, že je článek pouze strohým (místy možná i poněkud nudným) popisem Knuthových stromků a jeho vlastností. Mojí motivací však bylo především ukázat provázanost Knuthových stromků s několika klasickými výsledky z poměrně vzdálených oblastí informatiky a matematiky. Právě provázanost různých oblastí – mnohdy zcela nečekaná, a proto i překvapivá – je pro zmíněné vědní obory příznačná a činí tak jejich studium neobyčejně poutavým.

Literatura

[1] *Bruijn, N. G., Tengbergen, C. E., Kruyswijk, D.*: On the set of divisors of a number. Nieuw Archief Wiskunde, roč. 23 (1951), č. 2, s. 191–193.

[2] *Hansel, P.*: Sur le nombre des fonctions booléennes monotones de n variables. Comptes Rendus Acad Sci, roč. 262 (A) (1966), s. 1088–1090.

[3] *Knuth, D. E.*: Surreal numbers: how two ex-students turned on to pure mathematics and found total happiness: a mathematical novelette. Addison-Wesley, NJ, 1974.

[4] *Knuth, D. E.*: *The Art of Computer Programming.* Vol. 4, Fasc. 4 of *Generating All Trees; History of Combinatorial Generation.* Upper Saddle River, Addison-Wesley, NJ, 2006.

[5] *Lubell, D.*: A short proof of Sperner's lemma. Journal of Combinatorial Theory, roč. 1 (1966), č. 2, s. 299

[6] *Sperner, E.*: Ein Satz über Untermengen einer endlichen Menge. Mathematische Zeitschrift, roč. 27 (1928), č. 1, s. 544–548.

[7] Christmass tree lecture (YouTube).

				10101010				
			10101000	10101001	10101011			
				10101100				
			10100100	10100101	10101101			
			10100010	10100110	10101110			
		10100000	10100001	10100011	10100111	10101111		
				10110010				
			10110000	10110001	10110011			
				10110100				
			10010100	10010101	10110101			
			10010010	10010110	10110110			
		10010000	10010001	10010011	10010111	10110111		
				10111000				
			10011000	10011001	10111001			
			10001010	10011010	10111010			
		10001000	10001001	10001011	10011011	10111011		
			10001100	10011100	10111100			
		10000100	10000101	10001101	10011101	10111101		
		10000010	10000110	10001110	10011110	10111110		
	10000000	10000001	10000011	10000111	10001111	10011111	10111111	
				11001010				
			11001000	11001001	11001011			
				11001100				
			11000100	11000101	11001101			
			11000010	11000110	11001110			
		11000000	11000001	11000011	11000111	11001111		
				11010010				
			11010000	11010001	11010011			
				11010100				
			01010100	01010101	11010101			
			01010010	01010110	11010110			
		01010000	01010001	01010011	01010111	11010111		
				11011000				
			01011000	01011001	11011001			
			01001010	01011010	11011010			
		01001000	01001001	01001011	01011011	11011011		
			01001100	01011100	11011100			
		01000100	01000101	01001101	01011101	11011101		
		01000010	01000110	01001110	01011110	11011110		
	01000000	01000001	01000011	01000111	01001111	01011111	11011111	
				11100010				
			11100000	11100001	11100011			
				11100100				
			01100100	01100101	11100101			
			01100010	01100110	11100110			
		01100000	01100001	01100011	01100111	11100111		
				11101000				
			01101000	01101001	11101001			
			00101010	01101010	11101010			
		00101000	00101001	00101011	01101011	11101011		
			00101100	01101100	11101100			
		00100100	00100101	00101101	01101101	11101101		
		00100010	00100110	00101110	01101110	11101110		
	00100000	00100001	00100011	00100111	00101111	01101111	11101111	
				11110000				
			01110000	01110001	11110001			
			00110010	01110010	11110010			
		00110000	00110001	00110011	01110011	11110011		
			00110100	01110100	11110100			
		00010100	00010101	00110101	01110101	11110101		
		00010010	00010110	00110110	01110110	11110110		
	00010000	00010001	00010011	00010111	00110111	01110111	11110111	
			00111000	01111000	11111000			
		00011000	00011001	00111001	01111001	11111001		
		00001010	00011010	00111010	01111010	11111010		
	00001000	00001001	00001011	00011011	00111011	01111011	11111011	
		00001100	00011100	00111100	01111100	11111100		
	00000100	00000101	00001101	00011101	00111101	01111101	11111101	
	00000010	00000110	00001110	00011110	00111110	01111110	11111110	
00000000	00000001	00000011	00000111	00001111	00011111	00111111	01111111	11111111

Knuthův stromek řádu 8