Závěr

Pokročilé SEO je spíše o technických maličkostech a především o chytrosti. Pokud bude majitel stránky přemýšlet o tom, jak jeho web vnímají uživatelé, a bude se jim snažit vyjít vstříc, výsledky se vždy dostaví a web bude prosperovat. Vždy je potřeba přemýšlet, zda má daná optimalizace přínos pro uživatele. Pokud tomu tak není, je zbytečná. Zatímco základní SEO by mělo být naprostou samozřejmostí, pokročilé SEO je mnohdy považováno za nadstandard. Správné principy SEO jsou součástí kvalitního webdesignu, které by měl dodržovat každý tvůrce webových stránek.

Poděkování. Tento článek byl vytvořen za podpory projektu IGA UP 2016, reg. č. IGA_PrF_2016_027. Děkuji Eduardu Bartlovi za jeho cenné rady a připomínky k textu.

Literatura

[1] Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American, 2001, s. 29–37.

[2] Clark, A.: Search engine optimization 2016: Learn SEO with smart internet marketing strategies. 2015.

[3] Trnecka, M.: Základní optimalizace pro webové vyhledávače. MFI, roč. 25 (2016), č. 3, s. 227–234.

[4] Williams, A.: SEO 2014 & Beyond: Search engine optimization will never be the same again! 2014.

Bobřík učí informatiku
6. díl – Kódování

DANIEL LESSNER – JIŘÍ VANÍČEK

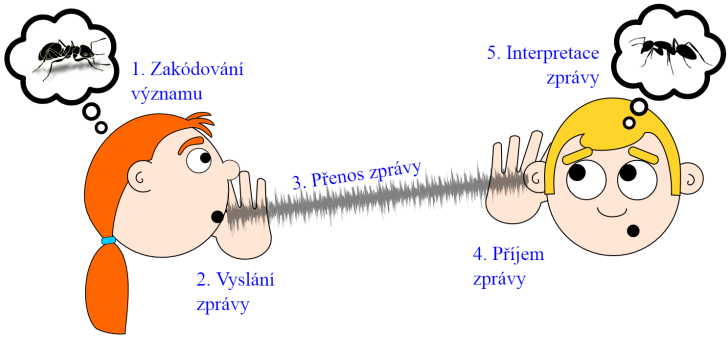
Matematicko-fyzikální fakulta, UK Praha

Pedagogická fakulta, Jihočeská univerzita v Českých Budějovicích

V minulém dílu seriálu o situačních informatických úlohách jsme se seznámili s pojmem informace a s tím, jak informace ubírá možnosti. Příliš jsme ale nezdůrazňovali, že každá informace, pokud ji chceme zpracovat, musí být nějak fyzicky zaznamenána.

Abychom problematiku blíže prozkoumali, změníme úhel pohledu. Informaci lze definovat (nikoliv v matematickém smyslu) také jako zprávu spolu s jejím významem. Zpráva je zde posloupností znaků z předem dané množiny, tzv. abecedy. Do přesnějšího vymezení slova „význam“ se pouštět nebudeme, k tomu mají co říci spíše jiné obory, např. filozofie nebo psychologie.

V následujícím textu budeme používat jednoduchý model komunikační situace ukázaný na obr. 1. Při komunikaci se partnerovi snažíme sdělit naši myšlenku. Přímo (telepaticky) to ale neumíme. Myšlenku proto nějak zaznamenáme a tento záznam předáme partnerovi. Může jít o vyřčená slova nesená vzduchem, obrázek na papíře či pípání telegrafu. Proces vytváření zprávy odpovídající dané myšlence nazýváme kódování.



Obr. 1 Model komunikační situace[1]

Komunikační partner obdrží naši zprávu (sérii změn tlaku vzduchu na ušní bubínek, barevné body na papíře promítnuté na sítnici oka apod.). Pokud partner zná způsob použitého kódování (např. když mluví stejnou řečí nebo když jsme se na způsobu předem domluvili), může ze zprávy dekódovat původní myšlenku.

Každá informace, kterou kdy přijmeme, je nějakým způsobem reprezentována, zakódována. Možnosti kódování zpráv jsou ovšem různé. Abychom se v těchto možnostech zorientovali a mohli rozhodnout, které kódování je lepší, potřebujeme se vyznat v podstatných vlastnostech kódování. Některé z nich nyní projdeme.

[1]Obrázek převzat z učebnice Lessner: Učebnice informatiky pro každého, ksvi.mff.cuni.cz/ucebnice/index.php/Učebnice/Informace/Úvod_a_komunikace se svolením autora.

Jako první se v informatice nabízí úspornost kódování. Chceme, aby byly zakódované zprávy co možná nejkratší, aby šly rychle předávat. Neměly by používat zbytečně velkou abecedu, zároveň kódování a dekódování nemá být příliš složité. Čtenáře jistě napadne, že tyto požadavky jsou vlastně protichůdné; v praxi proto hledáme optimální řešení pro danou situaci. Např. internetový provoz běžně komprimován není. První mobilní připojení ale na komprimaci přenášených dat spoléhala, neboť telefony dekomprimaci stíhaly v reálném čase.

Další protichůdné požadavky najdeme ve vztahu k chybám při přenosu. V praxi k chybám dochází a dobré kódování si s nimi musí umět poradit. Např. ve zprávě se použije několik kontrolních znaků navíc a příjemce díky nim případné poškození zprávy odhalí, nebo dokonce opraví (aniž by musel žádat o opakovaný přenos zprávy).

S otázkou chybovosti souvisí otázka důvěrnosti zprávy. Odesilatel chce mít jistotu, že nezměněnou zprávu přečte právě jen příjemce, a příjemce chce jistotu, že zprávu opravdu odeslal odesilatel. Dalšími důležitými vlastnostmi kódování jsou snadnost použití (nikoliv jako u úlohy Kódy měst) a robustnost (v úloze Chybějící část zprávy).

V této části představujeme několik soutěžních úloh z Bobříka informatiky, při nichž soutěžící musí vzít v úvahu vlastnosti kódování. Soutěžící musí prokázat schopnost daný kód přečíst či rozluštit, zakódovat zprávu, ale také posoudit, zda zvolený způsob kódování není z hlediska některé své vlastnosti problematický nebo nevhodný.

Kódy měst

Kategorie Kadet, autor Willem van der Vegt.

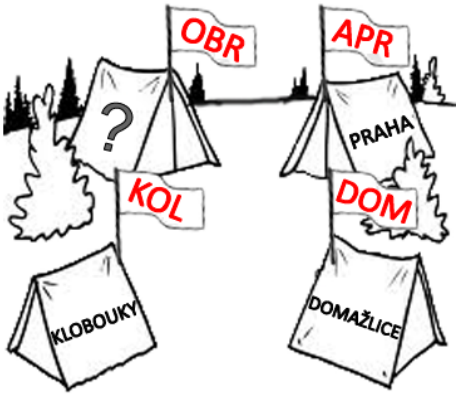
Zadání

Na skautském táboře se sjeli účastníci z celé republiky. Na svůj stan si každý vyvěsil praporek s kódem svého bydliště. Kód byl vytvořen následovně:

- Písmena názvu obce se seřadila podle četnosti výskytu.
- Nejčastěji se vyskytující znak se vzal jako první.
- Písmena se stejným počtem výskytů se seřadila v tom pořadí, v jakém se vyskytla v názvu města.
- Použila se první tři seřazená písmena.

Např. kód pro město HODONÍN byl ONH (písmena O a N se objevují dvakrát; O se objevuje první; z ostatních písmen použitých jednou je první

na řadě H: ONHDÍ). Několik dalších příkladů je vidět na obr. 2.



Obr. 2 Kódování měst

Která z následujících obcí nemá kód OBR?

A) BOROHRÁDEK B) BOROVNICE C) BROUMOV D) OBORA

Co má tato úloha společného s informatikou

V úloze se sleduje, který z kódů je vytvořen podle popsaných pravidel. Můžeme v ní vidět nevhodně navržené kódování. Mnoho různých měst totiž dostává stejný kód, který tím pádem nelze jednoznačně dekódovat. S podobnou situací se setkáváme ve školním rozvrhu hodin. Na prvním stupni může vystačit jednopísmenná zkratka předmětu, později by ale nebylo možné odlišit zeměpis od základů společenských věd nebo historii od hudební výchovy. Proto jsou později zkratky předmětů aspoň dvoupísmenné. Ještě zapeklitější je pak situace na školách vysokých s množstvím podobných předmětů. Podobně se tvůrce rozvrhu musí rozhodovat v případě zkratek jmen učitelů. Zvolený systém musí všechny učitele jednoznačně rozlišit, musí umožnit snadné odvození zkratky učitele podle jména (i obráceně) a nesmí selhat, ani když na školu přijde učitel se stejnými iniciálami.

Zdůvodnění správné odpovědi

Správná odpověď je BOROHRÁDEK: jeho kód není OBR, ale ORB (dvakrát je přítomno O a R, proto musí být nejprve R a po něm B).

Velikosti kalhot

Kategorie Junior, autor Juha Vartiainen.

Zadání

Radek pracuje ve firmě, která vyrábí pánské kalhoty různých velikostí – S, M, L a X. Jeho úkolem je vložit kalhoty do krabice a na víko přidat razítko podle velikosti kalhot. Při razítkování se nedělají žádné mezery, vytváří se jedna řada znaků, např. SMMXLS.

Firma se rozhodla snížit počet razítek na dvě: O a ◇. Radek tedy musí vymyslet jiný způsob značení vík beden, aby dvěma znaky nahradil čtyři písmena.

Který z následujících způsobů značení může bez obav použít?

A) S: ◇, M: ◇◇, L: ◇◇◇, X: ◇◇◇◇
B) S: ◇, M: O, L: ◇◇, X: O◇
C) S: O, M: O◇, L: OO, X: O◇O
D) S: ◇, M: O◇, L: OO◇, X: OOO

Co má tato úloha společného s informatikou

Častou metodou kódování je kódování s proměnnou délkou kódu pro jednotlivé znaky. Je používáno především tam, kde je požadována co nejkratší délka výsledného kódu (např. Huffmanova komprese, podrobně popsaná na https://cs.wikipedia.org/wiki/Huffmanovo_kódování). Protože ale potom není hned jasné, kde jednotlivé kódy končí, je potřeba zajistit, aby kód žádného znaku netvořil začátek jiného kódu. Příklady nevhodných návrhů ukazuje právě tato úloha.

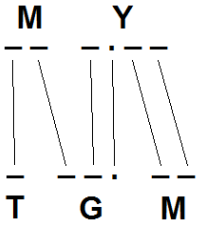
Podobný problém kdysi řešili výrobci telegrafu při používání Morseovy abecedy. Operátoři byli cvičeni v tom, aby dělali mezery mezi slovy třikrát delší než mezi písmeny. Pokud by je nedělali správně, mohlo docházet k omylům, např. řada znaků −−−.−− může být mezerami rozdělena různými způsoby, jak je vidět na obr. 3. Podívejte se na video na https://goo.gl/3O3z6y.

Zdůvodnění správné odpovědi

Protože mezi značkami pro jednotlivé kalhoty nejsou žádné mezery, musí být jednotlivé kódy jednoznačné. To platí pouze u odpovědi D.

- V kódu A) S: ◇, M: ◇◇, L: ◇◇◇, X: ◇◇◇◇ by například X (◇◇◇◇) mohlo znamenat LS, SL i MM.
- V kódu B) S: ◇, M: O, L: ◇◇, X: O◇ by např. X (O◇) mohlo znamenat i MS.

- V kódu C) S: O, M: O◇, L: OO, X: O◇O by např. L (OO) mohlo znamenat i SS.
- Kód D) S: ◇, M: O◇, L: OO◇, X: OOO je správný, zde nemůže dojít k žádné záměně.



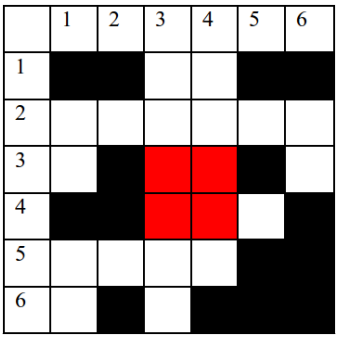
Obr. 3 Mezery na různých místech telegrafického záznamu mohou vytvářet ze stejného sledu teček a čárek různé zprávy

Chybějící část zprávy

Kategorie Senior, autoři Paul Miotti, Ivo Blöchliger.

Zadání

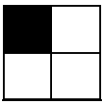
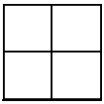
Honza dostal od Jirky tajnou zprávu, jejíž část někdo bohužel zamazal červenou barvou, jak můžeme vidět na obr. 4.



Obr. 4 Chybějící informace

Chlapci možnost poškození zprávy předpokládali. Proto ke vzkazu přidali řádek a sloupec navíc. Každý čtverec v pravém sloupci (6) nebo v posledním řádku (6) je začerněný, pokud je v daném sloupci nebo řádku vzkazu lichý počet černých čtverců.

Která možnost patří místo červeného čtverce?

A)  B) C) D) 

Co má tato úloha společného s informatikou

Úloha se týká zabezpečení přenosu signálu proti chybám, což je klasická partie informatiky. Když se na dálku (třeba po Internetu) přenáší signál, může dojít k jeho poškození. Kvůli poklesu napětí, změně světelných podmínek nebo změně magnetického pole může příjemce signálu obdržet 0 namísto 1 nebo naopak. To změní význam celé zprávy a celou ji tak znehodnotí.

Abychom tomuto nebezpečí předešli, zavádíme kontrolní mechanismy. Například součástí našeho rodného čísla je čtyřčíslí vybrané tak, aby bylo rodné číslo dělitelné jedenácti. Pomocí dělení pak může být snadno odhalena chyba způsobená např. překlepem. Další metodou jsou kontrolní součty (https://cs.wikipedia.org/wiki/Kontrolní_součet), které byly použity v soutěžní úloze. Kontrolní součty v této úloze umožňují nejen zjistit, že k chybě došlo, ale také ve které části zprávy se tak stalo. Čím více kontrolních informací zašleme, tím rozsáhlejšímu poškození zprávy můžeme předejít.

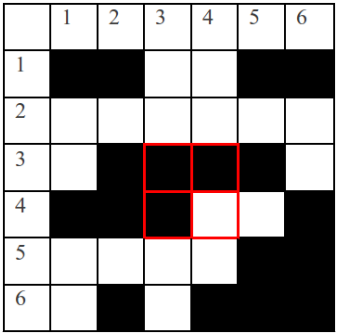
Související informatické poznatky používáme intuitivně např. při diktování čísla bankovního účtu nebo telefonního čísla, kdy jej jednou sdělíme po dvojčíslích a jednou po trojčíslích, nebo necháme příjemce číslo pro kontrolu zopakovat. Podívejte se též na videa se školními aktivitami na detekování chyby na http://csunplugged.org/error-detection.

Zdůvodnění správné odpovědi

Jestliže poslední řádek a sloupec přidávají černý čtverec k lichému počtu černých čtverců a k sudému počtu nepřidává, musí být počet černých čtverců sudý ve všech řádcích a sloupcích (kromě posledního), viz obr. 5.

Správně je odpověď B). Na rozdíl od ostatních odpovědí mají všechny řádky i sloupce (kromě posledních s č. 6) sudý počet černých čtverců. Ostatní odpovědi nejsou správné: odpovědi A) a C) mají ve sloupci 3 a odpověď D) ve sloupci 4 lichý počet černých čtverců.

Doplňující otázka: Mohla zamazaná část zprávy vypadat ještě jinak, než nabízí naše čtyři odpovědi?



Obr. 5 Doplnění chybějící informace

Agent Tajný

Kategorie Senior.

Zadání

Abeceda agenta Tajného má pouze 5 písmen. Tabulka ukazuje, jak se kódují.

S	O	N	E	T
10	01	11	001	000

Dostali jsme čtyři tajné zprávy. Která je od agenta Tajného?

A) 110100001111

B) 110100000100

C) 110100000110

D) 110100011111

Co má tato úloha společného s informatikou

Šifrování je jednou z důležitých součástí informatiky. Např. data přenášená po internetu ve skutečnosti prochází mnoha cizími zařízeními, která je tím pádem mohou číst. Proto je důležité mít možnost posílaná data zašifrovat. Z uvedeného plyne i rozdíl mezi kódováním a šifrováním: smyslem šifrování je skrýt informace v (jinak viditelné) zprávě před nepovolanými,

nebo jim aspoň přečtení zprávy náležitě ztížit. Naopak pro zamýšlené příjemce má být přečtení zprávy snadné a rychlé. Typickým řešením je použití šifrovacích klíčů známých jen omezenému okruhu lidí. Jiným (starším a nebezpečnějším) je zatajení samotného šifrovacího a dešifrovacího algoritmu.

Připomeňme slavné „šifrovací“ příběhy z 2. světové války. Nacistické Německo používalo šifrovací stroj Enigma, který i přes svoji vyspělost Britové dokázali po ukořistění jednoho exempláře „prolomit“ a tajné zprávy dešifrovat již z kraje války.

Američané používali systém šifrování založený na indiánském jazyce Navaho. Zvlášť vycvičení příslušníci tohoto nepočetného kmene byli součástí vojenských jednotek a zajišťovali komunikaci s ostatními. Jejich systém nebyl nikdy prolomen. Šifrování tak výrazně přispělo k úspěchu či porážce ve válce.

Zdůvodnění správné odpovědi

Správně je odpověď C) 11|01|000|001|10. Dešifrovanou zprávou je slovo NOTES. Když budeme ostatní zprávy rozdělovat na bloky nul a jedniček, uvidíme, že některé tyto bloky neodpovídají zašifrovaným písmenům abecedy agenta Tajného (tyto bloky jsou podtrženy):

11|01|000|01|11|1

11|01|000|11|11|1

11|01|000|001|00

Ověřte sami, že ani jednu z těchto tří zpráv nelze rozdělit tak, aby vznikala pouze písmena abecedy agenta Tajného.

Kódování obrázku

Kategorie Senior, autor Ilja Posov.

Zadání

Tomáš vymyslel způsob, jak zakódovat černobílý obrázek. Každý obrázek je popsán výpisem barev v jeho buňkách. Kóduje se zleva doprava a shora dolů.

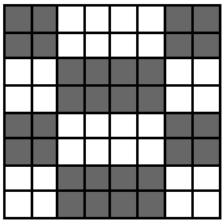
Například BBBBČČČBB zakóduje obrázek .

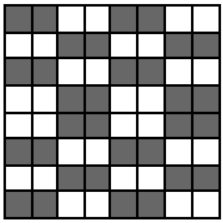
Pak Tomáš svůj kód vylepšil. Stejná písmena za sebou napsal pouze jednou a uvedl, kolikrát se opakují. Takto zapsal předchozí obrázek pouze šesti symboly: 4B3Č2B.

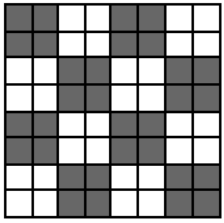
Při dalším vylepšení nepsal opakování jen pro jednotlivá písmena, ale pro celé bloky, které dával do závorek.

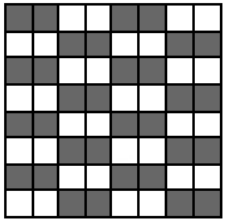
Například 4(ČB)Č zakóduje obrázek .

Který z následujících obrázků odpovídá kódu 2 (4 (2Č2B) 4 (2B2Č))?

A) 

B) 

C) 

D) 

Co má tato úloha společného s informatikou

Úloha úzce souvisí s kódováním, tedy s domluveným způsobem záznamu informace pomocí předem dané sady znaků a pravidel. I text tohoto článku je zakódováním informací, zde pomocí znaků latinské abecedy a pravidel českého jazyka. Jiným příkladem kódování je Morseova abeceda.

Různá kódování jsou různě efektivní – podobně, jako se někdo vyjadřuje stručně a někdo jiný sdělí totéž velmi rozvláčným způsobem. Kódování z našeho příkladu je úsporné ve srovnání s jednoduchou maticí nul a jedniček (které bychom přidělili podle barvy příslušného pixelu).

Pokud chce čtenář přemýšlet nad důsledky, může si rozmyslet, na jakých obrázcích je kódování z naší úlohy nejúspornější. Patrně odhalí, že půjde o obrázky obsahující pravidelné vzory. Kódování uchovává veškerou informaci obsaženou v obrázku a nepravidelný obrázek nelze zakódovat příliš

úsporně. Lze také říci, že takový obrázek obsahuje mnoho informace – na rozdíl od obrázků typu bílá plocha, šachovnice atp. To je pro leckoho překvapivé: více informace zachycuje chaos a nahodilost, nikoliv řád. Je to nakonec známo i ze školy: zapamatování látky založené na několika dobře zorganizovaných pravidlech (počítání se zlomky, rozpoznání románského slohu) je snazší, než zapamatování látky obsahující mnoho výjimek (anglická slovesa v minulém čase).

Zdůvodnění správného řešení

Správný obrázek je C). Ostatní řešení jsou nesprávná:

- Když se podíváme na kód 2 (4 (2Č2B) 4 (2B2Č)), zjistíme, že se celá vnější závorka opakuje dvakrát. Takže správný obrázek musí mít horní a dolní polovinu stejnou, což obrázek D) nesplňuje – nemůže být správným řešením.
- Uvnitř velké závorky jsou další dvě závorky, které se čtyřikrát opakují. Každá z nich je jiná. Znamená to, že první dva řádky se musí lišit od druhých dvou řádků tabulky. Toto nesplňuje obrázek A).
- Kód (2Č2B) v první závorce se opakuje čtyřikrát a můžeme jej rozepsat jako ČČBBČČBB. Tomu neodpovídá první řádek obrázku B), který je souměrný podle svislé osy.

Závěrem

Úlohy ze soutěže jsou k dispozici na webu soutěže Bobřík informatiky http://www.ibobr.cz. Zde jsou stále přístupné archívní soutěžní testy. Učitel, který chce použít takové úlohy ve své výuce informatiky, může žáky nechat test vyplnit a poté nad odpověďmi žáků diskutovat. Může také test spustit při projekci před třídou a hromadně řešit pouze vybrané úlohy. Touto formou se žáci mohou seznámit s informatickými problémy a udělat si lepší představu o tom, jaké problémy informatika řeší a jaké myšlení vyžaduje. Budou pak ke studiu oboru připraveni lépe, než absolvováním pouhého praktického kurzu programování nebo naučením se technickým faktům o počítačích. I ti žáci, kteří obor studovat nepůjdou, se tu mohou naučit užitečné přístupy k řešení problémů, se kterými se ve škole nesetkali.