Týden 8

Přednáška-pondělí

Na začátku jsme se stručně vrátili k vybraným věcem z předchozích týdnů.

Nebezkontextové jazyky

Metodou obrázků derivačních stromů z části 5.3. jsme si ukázali, proč jazyk

$$L = \{a^n b^n c^n \mid n \geq 0\}$$

není bezkontextový. Porozuměli jsme tak pumping lemmatu (neboli $uvwxy$-teorému) pro bezkontextové jazyky:

> *Věta.* Nechť L je bezkontextový jazyk. Pak existuje přirozené číslo n tž. každé slovo $z \in L$, $|z| \geq n$, (tedy každé 'dlouhé' slovo z jazyka L) lze psát ve tvaru $z = uvwxy$, přičemž platí
>
> - $vx \neq \varepsilon$,
> - $|vwx| \leq n$,
> - pro vš. $i \geq 0$ je $uv^iwx^iy \in L$.

Intuici k poznání nebezkontextových jazyků jsme si posílili zvážením jazyků $L_1 = \{ww \mid w \in \{0,1\}^*\}$, $L_2 = \{\, 0^m 1^n 0^m \mid m = 2n \,\}$, u nichž obou jsme usoudili, že ne všechna jejich dlouhá slova lze napsat v patřičném tvaru $uvwxy$, a že tedy nejsou bezkontextové. (U jazyka L_1 jde např. o slova tvaru $0^n1^n0^n1^n$.)

Uzávěrové vlastnosti třídy CFL

Připomněli jsme si, že snadno umíme ukázat, že CFL je uzavřena vůči sjednocení.
Pak jsme si všimli, že jazyky $L_1 = \{a^ib^jc^k \mid i = j\}$ a $L_2 = \{a^ib^jc^k \mid j = k\}$ jsou bezkontextové, ale $L_1 \cap L_2 = \{a^nb^nc^n \mid n \geq 0\}$ je jazyk, o němž víme, že bezkontextový není.
Tedy CFL není uzavřena na průnik. Díky de Morganovým pravidlům ihned vidíme, že CFL není uzavřena ani na doplněk.
(Jelikož $(L_1 \cap L_2 = \overline{(\overline{L_1} \cup \overline{L_2})})$, tak z uzavřenosti na doplněk by díky uzavřenosti na sjednocení vyplynula uzavřenost na průnik.)
Poznámka. Vezměme alespoň na vědomí, že třída DCFL, tj. třída jazyků rozpoznatelných deterministickými zásobníkovými automaty (koncovým stavem), je vlastní podtřídou CFL. Ta je mimochodem uzavřena vůči doplňku, ne ovšem vůči průniku, a tedy ani ne vůči sjednocení.

Turingovy stroje, (výpočetní) problémy

Připomněli jsme si, co bylo cílem (třiadvacetiletého) Alana Turinga, když v roce 1936 zaváděl tzv. Turingův stroj, a nastínili jsme význam tzv. *Church-Turingovy teze*, která se stručně dá vyjádřit sloganem

$$\textit{algoritmus} = \textit{Turingův stroj}.$$

Sestrojili jsme konkrétní Turingův stroj M_1 s dvěma koncovými stavy q_{accept} a q_{reject}, který přijímá (nebezkontextový) jazyk

$$\{a^n b^n c^n \mid n \geq 1\}.$$

Snažili jsme se tedy přehledně zkonstruovat množinu instrukcí

$(q_0, a) \rightarrow (q_1, \bar{a}, +1)$
$(q_1, x) \rightarrow (q_1, x, +1)$ pro $x \in \{a, \bar{b}\}$
$(q_1, b) \rightarrow (q_2, \bar{b}, +1)$
...

Zároveň jsme připomněli definici Turingova stroje jako struktury $M = (Q, \Sigma, \Gamma, \delta, q_0, F)$; speciálně jsme si uvědomili, že přechodová funkce

$$\delta : (Q - F) \times \Gamma \rightarrow Q \times \Gamma \times \{-1, 0, +1\}$$

je, de facto, příslušnou množinou instrukcí ...
Připomněli jsme si, že na výpočet Turingova stroje M se dá pohlížet jako na posloupnost konfigurací $K_0 \vdash_M K_1 \vdash_M K_2 \vdash_M \cdots$.
(V našem konkrétním případě stroje M_1 např. platí $(q_0 aaabbbccc) \vdash (\bar{a} q_1 aabbbccc) \vdash (\bar{a} a q_1 abbbccc) \vdash (\bar{a} a a q_1 bbbccc) \vdash (\bar{a} a a \bar{b} q_2 bbccc) \vdash \cdots \vdash (\bar{a}\bar{a}\bar{a}\bar{b}\bar{b}\bar{b}\bar{c}\bar{c}\bar{c} q_{accept} \square)$; ale např. $(q_0 aabbbccc) \vdash^* \bar{a}\bar{a}\bar{b}\bar{b} q_{reject} b\bar{c}\bar{c}c$).)
Uvědomili jsme si, že náš *stroj* (tedy algoritmus) *řeší problém*

NÁZEV: *Příslušnost k jazyku* $L = \{a^n b^n c^n \mid n \geq 1\}$

VSTUP: $w \in \{a, b, c\}^*$

VÝSTUP: ANO, když $w \in L$, NE jinak.

Jinými slovy, náš *stroj* (algoritmus) *rozhoduje (rozhodovací, neboli* ANO/NE*) problém*

NÁZEV: *Příslušnost k jazyku* $L = \{a^n b^n c^n \mid n \geq 1\}$

VSTUP: $w \in \{a, b, c\}^*$

OTÁZKA: Je $w \in L$?

Pak jsme si načrtli, jak by pracoval Turingův stroj M_2, který řeší následující problém (jenž není typu ANO/NE):

NÁZEV: *Zdvojení slova v abecedě* $\{a, b\}$.

VSTUP: $w \in \{a, b\}^*$.

VÝSTUP: ww.

Uvědomili jsme si, že každý Turingův stroj $M = (Q, \Sigma, \Gamma, \delta, q_0, F)$ přirozeně definuje *částečné zobrazení*

$$f_M : \Sigma^* \to \Gamma^* .$$

Zobrazení je obecně částečné (tedy ne nutně totální) proto, že pro některé vstupy $w \in \Sigma^*$ se výpočet M nemusí zastavit a příslušný výstup tedy není definován; výrazem $!M(w)$ se zpravidla označuje fakt, že M se pro vstup w zastaví (a hodnota výstupu $f_M(w)$, někdy označovaná přímo jako $M(w)$, je definována).

Připomněli jsme si další konkrétní problémy (jako např. SAT [problém splnitelnosti booleovských formulí] a problém minimální kostry grafu) a přirozené způsoby kódování vstupů a výstupů slovy ve vhodné abecedě (která lze nakonec zakódovat „binárně“, tedy pomocí slov v abecedě $\{0, 1\}$).

Přednáška-čtvrtek

Na přednášce se psala druhá zápočtová písemka.

Partie textu k prostudování

Uzávěrové vlastnosti CFL (část 5.2.). Nebezkontextové jazyky (část 5.3.).

Problémy a algoritmy (část 6.1.), Turingovy stroje (část 6.2.).

(Máte si udělat přinejmenším dobrou první představu a zamyslet se nad příklady, speciálně těmi plánovanými na cvičení, ať se můžete na cvičení aktivně účastnit a případné problémy si tam objasnit.)

Cvičení

Prezentace referátů

Referát č. 13 (Pumping lemma pro regulární jazyky)

Vysvětlete tzv. pumping lemma pro regulární jazyky
(můžete vyjít např. z `http://www.cs.vsb.cz/jancar/TEORET-INF/teoret-inf.pdf`)
a naznačte na příkladu, jak jej lze využít pro důkaz neregularity nějakého jazyka.

Referát č. 14 (Pumping lemma pro bezkontextové jazyky)

Připomeňte pumping lemma pro bezkontextové jazyky (např. ze studijního textu) a vysvětlete souvislost se hrou dvou hráčů popsanou např. v
`http://www.cs.vsb.cz/jancar/TEORET-INF/teoret-inf.pdf`.

Příklad 8.1

Na přednášce jsme mj. ukázali, že třída CFL není uzavřena na doplněk. Existuje tedy jazyk $L \subseteq \Sigma^*$ (pro nějakou abecedu Σ), který je bezkontextový, přičemž jeho doplněk $\overline{L} = \Sigma^* - L$ bezkontextový není.
Víme, že jazyk $L_1 = \{ww \mid w \in \{a,b\}^*\}$ bezkontextový není. Zkuste prokázat, že $L_2 = \overline{L_1}$ bezkontextový je a představuje tak konkrétní příklad jazyka z CFL, jehož doplněk v CFL není.
Nápověda. Vystihněte co nejjednodušeji, jak vypadá slovo z L_2, které má sudou délku, tedy $2d$ pro nějaké $d \geq 1$. Pak se zkuste zamyslet nad následujícím vztahem:

$$2d = d_1 + 1 + d_2 + d_1 + 1 + d_2 = d_1 + 1 + d_1 + d_2 + 1 + d_2\,.$$

Nakonec si tipněte, jestli se může podařit navrhnout *deterministický* zásobníkový automat přijímající L_2.
(Pomůže vám nějak u předchozí otázky, když se dozvíte, že třída DCFL je uzavřena vůči doplňku?)

Příklad 8.2

(Na základě alespoň intuitivních argumentů) zjistěte, které z daných jazyků
jsou regulární:
jsou bezkontextové, ale ne regulární:
nejsou bezkontextové:

$L_1 = \{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$
$L_2 = \{w \in \{a,b\}^* \mid |w|_a \text{ je sudé }\}$
$L_3 = \{w \in \{a,b\}^* \mid w \text{ obsahuje podslovo } abba\}$
$L_4 = \{w \in \{a,b,c\}^* \mid |w|_a = |w|_b = |w|_c\}$
$L_5 = \{w \in \{a,b\}^* \mid |w|_a \text{ je prvočíslo}\}$
$L_6 = \{\, 0^m 1^n \mid m \leq 2n \,\}$
$L_7 = \{\, 0^m 1^n 0^m \mid m = 2n \,\}$

Příklad 8.3

Navrhněte (co nejjednodušší) Turingův stroj M řešící problém příslušnosti k jazyku (palindromů) $L = \{w \in \{a,b\}^* \mid w = w^R\}$.

Diskuse témat druhé zápočtové písemky

Připomenutí a diskuse témat zřejmých z ukázkové druhé zápočtové písemky a případných problémů, se kterými přijdou studenti.