Týden 2

Přednáška - první část

(viz také slidy k této přednášce ...)
Po rekapitulaci některých předchozích partií jsme si udělali soutěž (o pivo Primus) v poznání regulárních jazyků (viz slidy).

Redukce konečného automatu

Podívali jsme se na následující automat (který by vznikl po dotažení modulární konstrukce z první přednášky: má $3 \cdot 4 = 12$ stavů, jež vznikly přejmenováním původních dvojic).

	0	1
$\leftrightarrow s_1$	s_1	s_2
s_2	s_3	s_4
s_3	s_5	s_6
$\leftarrow s_4$	s_7	s_2
s_5	s_8	s_4
s_6	s_9	s_6
$\leftarrow s_7$	s_1	s_9
s_8	s_5	s_{10}
s_9	s_6	s_{11}
s_{10}	s_{12}	s_{10}
s_{11}	s_{11}	s_9
s_{12}	s_8	s_{11}

Rychle jsme zjistili, že všechny stavy jsou dosažitelné a promýšleli jsme, jak zjistit, zda existují dva různé stavy $q, q' \in Q = \{s_1, s_2, \ldots, s_{12}\}$ pro něž $L_q^{toAcc} = L_{q'}^{toAcc}$. Jde tedy o zjištění, zda pro q, q' existuje (či neexistuje) rozlišující slovo.
Uvědomili jsme si, že se nabízí induktivní postup, při němž konstruujeme ekvivalence $\sim_0, \sim_1, \sim_2, \ldots$ na množině Q definované takto:

$$q \sim_i q' \Leftrightarrow \text{pro stavy } q, q' \text{ neexistuje rozlišující slovo délky} \leq i.$$

Každé ekvivalenci $\sim_i$ odpovídá příslušný rozklad R_i na množině Q.
Výchozí R_0 jsme hravě sestrojili; má dvě třídy, přijímající a nepřijímající stavy.

$$R_0\text{:}\quad \mathrm{I} = \{s_1, s_4, s_7\},\ \ \mathrm{II} = \{s_2, s_3, s_5, s_6, s_8, s_9, s_{10}, s_{11}, s_{12}\}$$

Pak jsme sestrojili R_1 (postupem popsaným ve studijním textu).

$$R_1\text{:}\quad \mathrm{I} = \{s_1, s_4, s_7\},\ \ \mathrm{II} = \{s_2, s_5\},\quad \mathrm{III} = \{s_3, s_6, s_8, s_9, s_{10}, s_{11}, s_{12}\}.$$

Uvědomili jsme si, že postup se opírá o toto pozorování:

> pokud pro q, q' existuje rozlišující slovo délky $i + 1$, tak nutně existuje $b \in \Sigma$ (v našem případě $\Sigma = \{0, 1\}$) tak, že $q \xrightarrow{b} q''$, $q' \xrightarrow{b} q'''$ a pro q'', q''' existuje rozlišující slovo délky i.

(Pozn.: použili jsme symbol b pro proměnnou, ať si uvědomíme, že nemusíme vždy stereotypně používat a.)
Kdybychom pokračovali, zjistili bychom, že

$$R_2: \quad \mathrm{I} = \{s_1, s_4\}, \ \ \mathrm{II} = \{s_7\}, \quad \mathrm{III} = \{s_2, s_5\}, \quad \mathrm{IV} = \{s_3, s_8\}, \quad \mathrm{V} = \{s_6, s_9, s_{10}, s_{11}, s_{12}\},$$

atd., až bychom dospěli k pevnému bodu, tedy k rozkladu, který se již nezjemní. Ten má 3-prvkovou třídu $\{s_6, s_9, s_{11}\}$ a pak už jen jednoprvkové třídy.
Redukcí (původně 12-stavového) automatu tedy vznikne ekvivalentní automat s 10 stavy (jeho stavy odpovídají třídám závěrečného rozkladu).

Přednáška - druhá část

Mj. jsme se zamýšleli nad důkazem věty 3.18 (z části 3.8., str. 87)

Věta. Jazyk $L \subseteq \Sigma^*$ je regulární (tzn. přijímaný konečným automatem) právě tehdy, když je množina kvocientů $\{w \backslash L \mid w \in \Sigma^*\}$ konečná.

Důkaz lze nalézt v průběhu výuky v loňském roce …
Končili jsme úvahami o mohutnosti množin …

Partie textu k prostudování

V návaznosti na část 3.3. (modulární návrh) se jedná zejména o část 3.4. (dosažitelné stavy, normovaný tvar) a část 3.6. (minimalizace konečných automatů) (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 1 a 2

Příklad 2.1

Definujte jazyk
$L = \{w \in \{a,b\}^* \mid w$ začíná a a končí b nebo začíná b a končí $a\}$
jako (přirozené) sjednocení $L = L_1 \cup L_2$, kde
$L_1 = \{w \in \{a,b\}^* \mid \dots\}$
$L_2 = \{w \in \{a,b\}^* \mid \dots\}$
Zkonstruujte (co nejmenší) automat A_1 tak, že $L(A_1) = L_1$, a automat A_2 tak, že $L(A_2) = L_2$.
Aplikujte algoritmickou konstrukci vedoucí k automatu pro sjednocení $L_1 \cup L_2$. Přitom ovšem konstruujte jen dosažitelné stavy, podle definice
1/ $q_0 \in \mathit{Reach}(A)$,
2/ $(q \in \mathit{Reach}(A) \wedge q \longrightarrow q'\ ) \Longrightarrow q' \in \mathit{Reach}(A)$.
Definujte pak *normovaný tvar* automatu, při němž jsou stavy označeny $1, 2, 3, \dots$ a jsou uspořádány podle „vzdálenosti“ od počátečního stavu. Jak přesně vyjádříte onu vzdálenost, aby poskytla úplné uspořádání (dosažitelných) stavů?
Uveďte sestrojený automat do normovaného tvaru.

Příklad 2.2

O konečném automatu řekneme, že je *redukovaný*, jestliže

- všechny jeho stavy jsou dosažitelné
- a pro každé dva různé stavy q, q' platí $L_q^{toAcc} \neq L_{q'}^{toAcc}$.

Zkuste u vašeho (5-stavového?) automatu z předchozího příkladu nalézt pro každou dvojici stavů $q \neq q'$ slovo w tak, že $q \xrightarrow{w} F$ a $q' \not\xrightarrow{w} F$ (tedy $q' \xrightarrow{w} (Q{-}F)$) či naopak $q \not\xrightarrow{w} F$ a $q' \xrightarrow{w} F$; takovému slovu w budeme říkat

rozlišující slovo pro stavy q, q'.

Povedlo se? Pak je váš automat redukovaný.

Příklad 2.3

Dokončete redukci automatu z přednášky (tam jsme jen začali)

	0	1
$\leftrightarrow s_1$	s_1	s_2
s_2	s_3	s_4
s_3	s_5	s_6
$\leftarrow s_4$	s_7	s_2
s_5	s_8	s_4
s_6	s_9	s_6
$\leftarrow s_7$	s_1	s_9
s_8	s_5	s_{10}
s_9	s_6	s_{11}
s_{10}	s_{12}	s_{10}
s_{11}	s_{11}	s_9
s_{12}	s_8	s_{11}

Příklad 2.4

Připomeňme si větu:
Věta. Jazyk $L \subseteq \Sigma^*$ je regulární (tzn. přijímaný konečným automatem) právě tehdy, když je množina kvocientů $\{w\backslash L \mid w \in \Sigma^*\}$ konečná.
Vysvětlete, proč pro jazyk

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

jsou jazyky (kvocienty) $a\backslash L$, $aa\backslash L$, $aaa\backslash L$, … navzájem různé. Vyvoďte, že L není regulární.

Příklad 2.5

Vybudujte si intuici o (ne)regulárních jazycích promyšleným zodpovězením otázek na s. 88 a 89. Ty se týkají (ne)regularity jazyků
$\{w \in \{a,b\}^* \mid |w|_a \bmod 2 = 0\}$
$\{w \in \{a,b\}^* \mid w$ začíná nebo končí dvojicí stejných písmen $\}$
$\{w \in \{a,b\}^* \mid |w|_a < |w|_b\}$
$\{w \in \{a,b,c\}^* \mid$ jestliže w neobsahuje podřetězec abc, pak končí $bca\}$
$\{w \in \{a,b\}^* \mid |w|_a > |w|_b$ nebo w končí $baa\}$
$\{w \in \{a,b\}^* \mid |w|_a > |w|_b$ nebo $|w|_b \geq 2\}$
$\{u \mid$ ex. $w \in \{a,b\}^*$ tak, že $u = ww^R\}$, stručněji také psáno $\{ww^R \mid w \in \{a,b\}^*\}$,
$\{w \in \{a,b\}^* \mid w = w^R\}$
$\{w \in \{a\}^* \mid w = w^R\}$
$\{ww \mid w \in \{a,b\}^*\}$
$\{ww \mid w \in \{a\}^*\}$
$\{w \in \{a,b\}^* \mid$ rozdíl počtů znaků a a znaků b ve w je větší než $100\}$
$\{w \in \{a,b\}^* \mid$ součin $|w|_a$ a $|w|_b$ je větší nebo roven $100\}$
$\{w \in \{a\}^* \mid |w|$ je prvočíslo $\}$

Příklad 2.6

Pro následující vztahy vysvětlete, proč obecně platí či neplatí.

- $L_1 \cdot L_2 = L_2 \cdot L_1$
- $L_1(L_2 \cup L_3) = L_1L_2 \cup L_1L_3$
- $(L_1 \cup L_2)^* = L_1^*(L_2 \cdot L_1^*)^*$
- $(L_1 \cap L_2)^* = L_1^* \cap L_2^*$
- $w \cdot (w\backslash L) = L$