Automaty a Gramatiky

Poznámky z přednášek

Letní semestr 2020/2021

Viktor Soukup, Lukáš Salak

Verze 0.1
2. března 2021

Obsah

1 První přednáška

Intuice: Automaty a gramatiky - dva způsoby popisu:

Turingovy stroje	$\leftrightarrow$	gramatiky Typu 0
lineárně omezené automaty	$\leftrightarrow$	kontextové gramatiky, monotónní gramatiky
zásobníkové automaty	$\leftrightarrow$	bezkontextové gramatiky
konečné automaty (DFA,NFA, λNFA)	$\leftrightarrow$	regulární jazyky

Poznámka: Nejjednodušší jsou dolu, turingův stroj je nejkomplikovanější.
Každá gramatika odpovídá nějaké skupině automatů.

Proč to řešíme?

1. zpracování přirozeného jazyka,
2. překladače (lexikální, syntaktická analýza...),
3. návrh, popis, verifikace hardware...
4. hledání výskytu slova v textu (grep),
5. verifikace systémů s konečně mnoha stavy

Příklad:
1. konečný automat modelující spínač on/off
   pridať obrázok zo slidov, som lenivý toto vytvárať v latexe.

2. Lexikální analýza
   Konečný automat rozpoznávajíci slovo *then*
   taktiež pridať zo slidov, je to na slide 8

Definice: Deterministický konečný automat (DFA) $A = (Q, \Sigma, \delta, q_0, F)$ sestáva z:

1. konečné množiny **stavů**, zpravidla značíme Q
2. konečné neprázdné množiny **vstupních symbolů (abecedy)**, znažíme Σ
3. **přechodové funkce** zobrazení $Q \times X \to Q$, značíme δ, která bude reprezentovaná hranami grafu
4. **počátečného stavu** $q_0 \in Q$, vede do něj šipka 'odnikud'
5. neprázdné **množiny koncových (přijímajícich) stavů** (final states) $F \subseteq Q$, označených dvojitým kruhem či šipkou 'ven'.

Poznámka: Pokud pro některou dvojici stavu a písmene není definovaný přechod, přidáme nový stav *fail* a přechodovou funkci doplníme na totální přidáním šipek do *fail*.

Pokud je množina F prázdná, přidáme do ní i Q nový stav *final* do kterého vedou jen přechody z něj samotného $\forall s \in \Sigma : \delta(final, s) = final$.

Doplniť obrázok, slide 9

Příklad: Automat A přijímajíci $L = x01y : x, y \in \{0,0\}*$.

Automat $A = (\{q_0, q_1, q_2\}, 0, 1, \delta, q_0, q_1)$
Reprezentace stavovým diagramem (grafem), pomocí tabulky nebo stavovým stromem
doplniť obrázky? imo self explanatory (10. slide)

Definice: Abeceda, slova, jazyky
Mějme neprázdnou množinu symbolů Σ.

1. **Slovo** je konečná (i prázdná) posloupnost symbolů $s \in \Sigma$, **prázdné slovo** se značí λ nebo ϵ.
2. **Množinu všech slov v abecedě** Σ značíme Σ^*,
3. množinu všech neprádzných slov v abecedě značíme Σ^+.
4. **jazyk** $L \subseteq \Sigma^*$ je množina slov v abecedě Σ.

Definice: Operace na Σ^*:

1. **zřetězení slov** $u.v$ nebo uv,
2. **mocnina** (počet opakování) $u^n(u^0 = \lambda, u^1 = u, u^{n+1} = u^n.u)$
3. **délka slova** $|u|(|\lambda| = 0, |auto| = 4)$.
4. **počet výskytů** $s \in \Sigma$ ve slově u značíme $|u|_s(|zmrzlina|_z = 2)$.

Definice: Rozšířená přechodová funkce
Mějme přechodovou funkci $\delta : Q \times \Sigma \to Q$.
Rozšířenou přechodovou funkci δ^*: $Q \times \Sigma^* \to Q$ (tranzitivní uzávěr δ)
definujeme induktivně:

1. $\delta^*(q, \lambda) = q$,
2. $\delta^*(q, wx) = \delta(\delta^*(q, w)x)$ pro $x \in \Sigma, w \in \Sigma^*$.

Poznámka: Pokud se v textu objeví δ aplikované na slova, míní se tím δ^*.

Pridať obrázok a príklad? slide 12

Definice:
1. **Jazykem rozpoznávaným (akceptovaným, přijímaným)** konečným automatem $A = (Q, \Sigma, \delta, q_0, F)$ nazveme jazyk $L(A) = \{w : w \in \Sigma^* \& \delta^*(q_0, w) \in F\}$.
2. Slovo w je *pijmano* automatem A, právě když $w \in L(A)$.
3. Jazyk L je rozpoznatelný konečným automatem, jestliže existuje konečný automat A takový, že $L = L(A)$.
4. Třídu jazyků rozpoznatelných konečnými automaty označíme $\mathcal{F}$, nazveme **regulární jazyky**.
pridať example? slide 13 dolu, pokračuje na 14. slide

Věta: *Iterační (pumping) lemma pro regulární jazyky:*

Mějme regulární jazyk L. Pak existuje konstanta $n \in \mathbb{N}$ (závislá na L) tak, že každé $w \in L; |w| \geq n$ můžeme rozdělit na tři části, $w = xyz$, že:

1. $y \neq \lambda$
2. $|xy| \leq n$

3. *$\forall k \in \mathbb{N}_0$, slovo xy^kz je také v L.*

Důkaz:

1. *Mějme regulární jazyk L, pak existuje DFA A s n stavy, že $L = L(A)$.*
2. *Vezměme libovolné slovo $a_1a_2a_3 \dots a_m = w \in L$ délky $m \geq n, a_i \in \Sigma$.*
3. *Definujeme: $\forall i p_i = \delta^*(q_0, a_1a_2 \dots a_i)$. Platí $p_0 = q_0$.*
4. *Máme $n + 1 p_i$ a n stavů, některý se opakuje, vezměme první takový, t. j. ($\exists i, j : 0 \leq i < jqleqn \& p_i = p_j$).*
5. *Definujeme $x = a_1a_2 \dots a_i, y = a_{i+1}a_{i+2} \dots a_j, z = a_{j+1}a_{j+2} \dots a_m$, t.j. $w = xyz, y \neq \lambda, |xy| \leq n$.*
6. *pak y^k můžeme opakovat libovolněkrát a vstup je také akceptovaný.*

□

Příklad: Doplniť príklady na pumping lemma (slidy 17+)

The End