Automaty a gramatiky 1

Vojtěch Šára

March 9, 2021

Nejjednodušší automat je vypínač. Deterministický konečný automat (DFA) $A = (Q, \Sigma, \delta, q_0, F)$
Q - konečná neprázdná množina stavů
Σ - konečná neprázdná množina vstupních symbolů abecedy
δ - přechodová funkce $Q \times \Sigma \to Q$ typicky reprezentovaná hranami grafu
$q_0 \in Q$ - počáteční stav, vede do něj šipka odnikud.
F - množina koncových stavů

Pro představu si můžeme představit vyhledávací automat z Knuth Morris Pratta.

Automat mohu reprezentovat diagramem, tabulkou či stavovým stromem.

Slovo je konečná (i prázdná) posloupnost symbolů z Σ, prázdné slovo se značí $\lambda \vee \epsilon$ Množina všech slov v abecedě Σ značíme $\Sigma^\star$
Množina neprázdných slov je Σ^+
Jazyk je množina slov v abecedě Σ - $L \subseteq \Sigma^\star$

Zřetězení - $u.v$ nebo uv
Mocnina (jen přirozenými) - $(u^0 = \lambda, u^1 = u, u^{n+1} = u^n.u)$
Délka slova $|auto| = 4$
Počet výskytů je kolikrát je podřetězec v řetězci.

Rozšířená přechodová funkce $\delta^\star$ je transitivní uzávěr funkce δ

Jazyk je rozpoznatelný konečným automatem jestliže existuje konečný automat A takový, že $L = L(A) = \{w | w \in \Sigma^\star \wedge \delta^\star(q_0, w) \in F\}$

Jazyky rozpoznatelné konečnými automaty nazýváme regulární.

Jsou jazyky, které nejsou regulární.

0.1 Iterační (pumping) lemma

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

1. $y \neq \lambda$
2. $|\text{xy}| \leq n$
3. $\forall k \in \mathbb{N}_0$ slovo xy^kz je také v L.

Důkaz:
Regulární - existuje DFA A s n stavy, že L = L(A).
Vezměme libovolné slovo z L délky $m \geq n$.
Zbytek důkazu neni.

Kongruence

Na konečné abecedě Σ a relaci ekvivalence $\sim$ na $\Sigma^\star$. Potom:

1. $\sim$ je pravá kongruence, jestliže $(\forall u, v, w \in \Sigma^\star) u \sim v \Rightarrow uw \sim vw$
2. $\sim$ je konečného indexu, jestliže rozklad $\Sigma^\star / \sim$ má konečný počet tříd.
3. Třídy kongruence $= [u]_\sim$

Myhill-Nerodova věta

Necht' L je jazyk nad konečnou abecedou Σ. Potom následující je ekvivalentní:

1. L je rozpoznatelný konečným automatem
2. existuje pravá kongruence konečného indexu tak, že L je sjednocením jistých tříd rozkladu $\Sigma^\star / \sim$

Pro důkaz definuji ekvivalenci jako slova, která nás z počátečního stavu dostanou do stejného stavu. Tato ekvivalence je ekvivalencí, je to i pravá kongruence.
Pro druhý směr volíme stavy jako třídy rozkladu $\Sigma^\star / \sim$
Počáteční stav q_0.... . .

Věta Myhil-Nerod je konstruktivní, nabízí nám tedy návod jak sestavit automat. Na přednášce tady byl příklad.

Zajímavost: jazyk $L = \{u | u = a^+ b^i c^i \vee u = b^i c^j\}$ není regulární (Myhill-Nerod) ale vždy lze pumpovat první písmeno.

Dosažitelnost

Stav je dosažitelný, pokud existuje slovo, kterým se do něj z počátečního stavu dostanu. Dosažitelné stavy hledáme iterativně - nějaký průchod stavovým prostorem. Korektnost - nachází pouze dosažitelné stavy, Úplnost - všechny dosažitelné stavy najdu.

Nejednoznačnost

Automat přijímající daný jazyk není určen jednoznačně. Tady byl na přednášce příklad takového jazyka.

Automatový homomorfismus

$h : Q_1 \to Q_2$ je automat. homomorfismus, jestliže:

1. $h(q_{0_1}) = q_{0_2}$
2. $h(\delta_1(q, x)) = \delta_2(h(q), x)$stejné přechodové funkce
3. stejné koncové stavy

Ekvivalence automatů

Dva konečné automaty jsou ekvivalentní, pokud rozpoznávají stejný jazyk. Homomorfismus implikuje ekvivalenci.
Jestli jsou všechny stavy ekvivalentní, pak jsou automaty ekvivalentní. Neekvivalentní = rozlišitelné.

Existuje algoritmus hledání všech rozlišitelných stavů. Postupně vyškrtávám nehodící. Tento algoritmus je korektní.
Algoritmus nalezení reduktu - vyškrtnu nedosažitelné, pak spojím ekvivalentní stavy. Tím dostávám redukovaný DFA.