Konštrukcia LBA ku kontextovej gramatike

Juraj Vaško

7. marca 2001

Úloha

Majme danú kontextovú gramatiku G.
Zostrojme lineárne ohraničený automat (LBA) A, taký že $L(G) = L(A)$.
Pri simulácii použime spätné odvodenie.

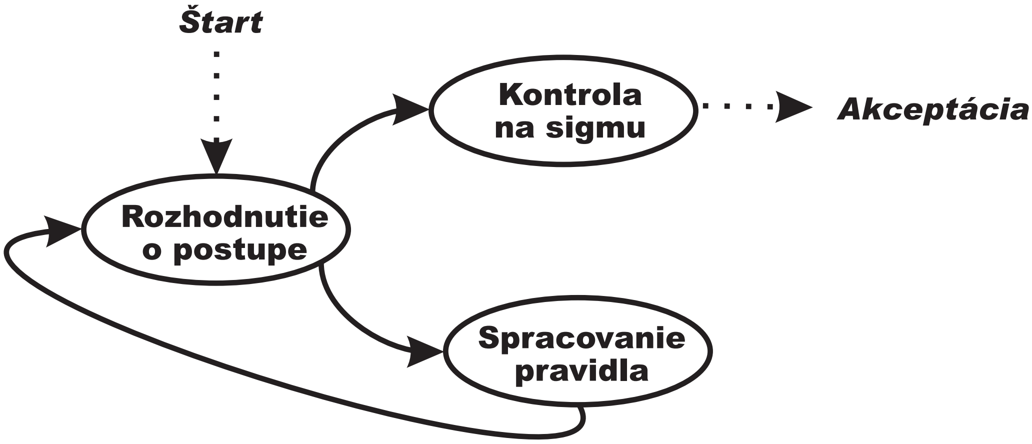
Vstupná gramatika

$G = (N, T, P, \sigma)$
$P = \{P_1, P_2, \ldots, P_m\} \wedge P_i = L_i \rightarrow R_i$
$L_i = l_{i,1} l_{i,2} \ldots l_{i,|L_i|}$
$R_i = r_{i,1} r_{i,2} \ldots r_{i,|R_i|}$

Lineárne ohraničený automat

Konečnostavové zariadenie s prepisovateľnou páskou dĺžky vstupu pracujúce v krokovom režime. V každom kroku hlava automatu načíta a zapíše jeden znak na pásku pričom sa môže pohnúť o jeden krok vpravo alebo vľavo, prípadne zmeniť stav automatu.

Myšlienka konštrukcie automatu



- Automat pracuje v cykle, v každej iterácii sa nedeterministicky rozhoduje, či skončí alebo spracuje ďalšie pravidlo.

- Pri spracúvaní pravidla automat nedeterministicky vyberie pravidlo, jeho pozíciu na páske, skontroluje výskyt jeho pravej strany a prepíše ho jeho ľavou stranou, v prípade nenájdenia očakávaných symbolov sa zasekne

- Pri ukončení cyklu automat skontroluje, či sa na páske nachádza len symbol σ, ak áno skončí ak nie zasekne sa

Problém voľného miesta

V priebehu výpočtu sa vetná forma postupne skracuje, páska však zostáva stále rovnako dlhá.

Riešením je pridanie nového symbolu podobného symbolu ϵ zastupujúceho prázdne miesto ďalej označovaného ako $\emptyset$.

Na pásku (bez symbolov $\cent$ a \$) sa budeme nahliadať cez funkciu h, ktorá na výstupe odstráni symboly $\emptyset$
$h : (N \cup T \cup \{\emptyset\})^* \to (N \cup T)^*$

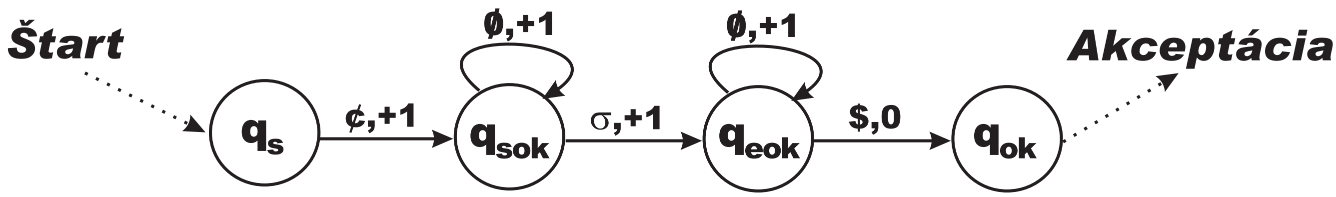
$$\begin{aligned} h(aw) &= a\, h(w) \leftarrow a \in N \cup T \\ h(aw) &= h(w) \leftarrow a = \emptyset \\ h(\epsilon) &= \epsilon \end{aligned}$$

Zápis výsledného automatu

$A = (K, T, \Gamma, \delta, q_s, \{q_{ok}\})$
$\Gamma = \Sigma \cup N \cup \{\cent, \$, \emptyset\}$
$K = \{q_s, q_{sok}, q_{eok}, q_{ok}, q_p, q_r\} \cup \{p_{i,j} | j \in \{1 \dots |R_i|\}\} \cup \{s_{i,j} | j \in \{1 \dots |L_i|\}\}$

Ukončovacia vetva

Výpočet môže skončiť akceptáciou, ak sme slovo spätne prepísali až po σ a teda treba overiť, že $h(páska) = \sigma$



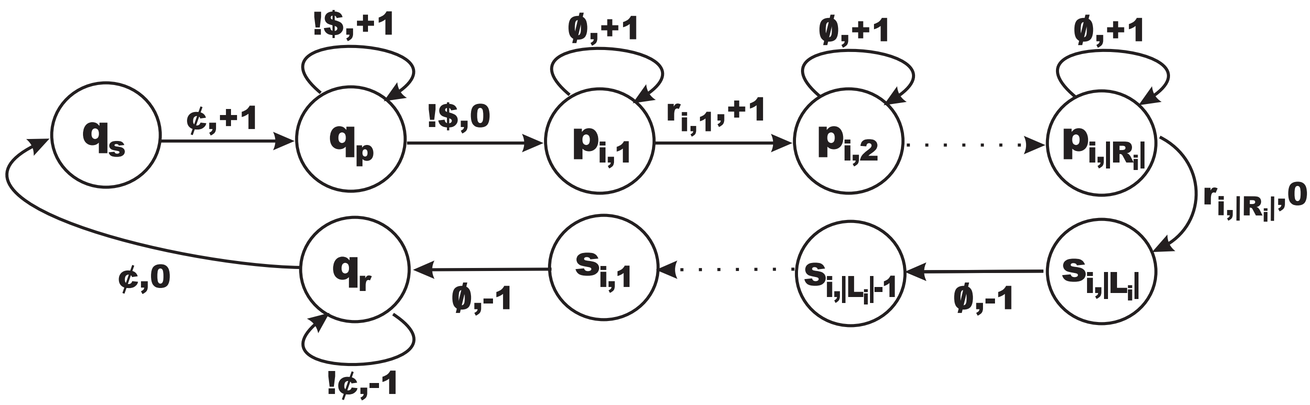
$$\begin{aligned}
\delta(q_s, \cent) &= \{(q_{sok}, \cent, 1)\} \\
\delta(q_s, \cent) &= \{(q_p, \cent, 1)\} \\
\delta(q_{sok}, \emptyset) &= \{(q_{sok}, \emptyset, 1)\} \\
\delta(q_{sok}, \sigma) &= \{(q_{eok}, \sigma, 1)\} \\
\delta(q_{eok}, \emptyset) &= \{(q_{eok}, \emptyset, 1)\} \\
\delta(q_{eok}, \$) &= \{(q_{ok}, \$, 0)\}
\end{aligned}$$

V prípade, že na páske sa nachádza niečo iné ako práve jeden symbol σ a symboly $\emptyset$ automat sa zasekne.

Automat musí vychádzať z konifigurácie $(q_S, \cent, 0)$ avšak po vstupe do tejto vetvy sa už do q_S nemôže vrátiť.

Prepísanie pravidla

Táto vetva vyberie polohu, pravidlo a ak sa trafila prepíše výskyt pravej strany pravidla na jeho ľavú stranu a voľné miesto prepíše symbolmi $\emptyset$.



$$
\begin{aligned}
\delta(q_p, a) &= \{(q_p, a, 1)\} \leftarrow a \in \Gamma/\{\$\} \\
\delta(q_p, a) &= \{(p_{i,1}, a, 0)\} \leftarrow a \in \Gamma/\{\$\} \\
\delta(q_r, a) &= \{(q_r, a, -1)\} \leftarrow a \in \Gamma/\{\not{c}\} \\
\delta(q_r, \not{c}) &= \{(q_s, \not{c}, 0)\} \\
\delta(p_{i,j}, \emptyset) &= \{(p_{i,j}, \emptyset, 1)\} \leftarrow j \in \{1 \ldots |R_i|\} \\
\delta(p_{i,j}, r_{i,j}) &= \{(p_{i,j+1}, \emptyset, 1)\} \leftarrow j \in \{1 \ldots |R_i| - 1\} \\
\delta(p_{i,|R_i|}, r_{i,|R_i|}) &= \{(s_{i,|L_i|}, \emptyset, 0)\} \\
\delta(s_{i,j}, \emptyset) &= \{(s_{i,j-1}, l_{i,j}, -1)\} \leftarrow j \in \{2 \ldots |L_i|\} \\
\delta(s_{i,1}, \emptyset) &= \{(q_r, l_{i,1}, -1)\}
\end{aligned}
$$

Nájdenie výpočtu pre slovo odvodené v G

Smerom od konca odvodenia pre každý krok odvodenia urobíme nasledovné

- Posunieme hlavu na začiatok pravej strany pravidla
- Prečítame pravú stanu pravidla, a zároveň ju tým prepíšeme symbolmi $\emptyset$
- Spätne zapíšeme ľavú stranu pravidla
- Vrátime sa na začiatok pásky prejdeme do q_s

Nakoniec kedže na páske je už len σ prejdeme bez problémov cez ukončovaciu vetvu.

Formálny dôkaz sleduje túto ideu a je založený na matematickej indukcii na dĺžku odvodenia slova v G.

Nájdenie odvodenia pre slovo akceptované automatom

Koniec výpočtu musel prejsť cez ukončovaciu vetvu a na páske bola teda σ a môžeme sa vrátiť do stavu q_s z ktorého vetva začala.

Pre každý prechod cez q_s sme museli vykonať výpočet ktorý

- Našiel a prečítal pravú stranu nejakého pravidla

- Nahradil ju ľavou stranou a vrátil sa do q_s

Potom však môžme toto pravidlo použiť na ďalší krok v odvodení, čím nakoniec odvodíme slovo, ktoré bolo na vstupe automatu.

Formálny dôkaz sleduje túto ideu a je založený na matematickej indukcii na počet prechodov cez q_s.