Formálne jazyky a automaty
Prezentácia úlohy 1.3

Nalevajko Tomáš
Fakulta matematiky, fyziky a informatiky
Univerzita Komenského v Bratislave
odbor informatika
4. skupina
e-mail:9nalevajko@st.fmph.uniba.sk

February 28, 2001

1 Problém

Nech jazyk L je regulárny. Pomocou dvojsmerných nedeterministických konečných automatov dokážeme, že jazyk

$$L' = \{uv \mid (\exists w \in \Sigma^*)(uvw \in L \wedge |u| = |w| = |v|)\}$$

je regulárny.

2 Definície a tvrdenia

Definícia 1 *Triedou $\mathcal{R}$ regulárnych jazykov nazývame množinu jazykov L, pre ktoré existuje regulárna gramatika G, ktorá generuje jazyk L, tj.*

$$\mathcal{R} = \{L \mid (\exists G)(L = L(G))\}$$

Definícia 2 *Dvojsmerným konečným nedeterministickým automatom (2NKA) nazývame 5-ticu $A = (K, \Sigma, \delta, q_0, F)$, kde K je konečná množina stavov, Σ je vstupná abeceda, $q_0 \in K$ je počiatočný stav, $F \subseteq K$ je množina akceptačných stavov,*

$$\delta : K \times \Sigma \cup \{¢, \$\} \longrightarrow 2^{K \times \{-1,0,1\}}$$

pričom platí $\delta(q, \$) \subseteq \{0, 1\}$ a $\delta(q, ¢) \subseteq K \times \{-1, 0\}$ pre všetky $q \in K$.

Definícia 3 *Nech $\# \notin \Sigma$. Dvojicu $(q, ¢\Sigma^*\#\Sigma^*\$)$ nazývame konfigurácia 2NKA A.*

Definícia 4 *Krokom výpočtu $\vdash_A$ 2NKA A nazývame binárnu reláciu na množine konfigurácií definovanú nasledovne:*

$$(p, u\#av) \vdash_A (q, ua\#v) \Leftrightarrow (q, 1) \in \delta(p, a)$$

$$(p, u\#av) \vdash_A (q, u\#av) \Leftrightarrow (q, 0) \in \delta(p, a)$$

$$(p, ua\#v) \vdash_A (q, u\#av) \Leftrightarrow (q, -1) \in \delta(p, a)$$

Definícia 5 *Jazyk akceptovaný 2NKA A definujeme nasledovne:*

$$L(A) = \{w \mid (q_0, ¢\#w\$) \vdash_A^* (q_f, ¢w\#\$), q_f \in F\}$$

Veta 1 *K ľubovolnému 2NKA A existuje regulárna gramatika G, pre ktorú platí $L(G) = L(A)$.*

Veta 2 *K ľubovolnej regulárnej gramatike G existuje 2NKA A taký, že $L(A) = L(G)$.*

3 Riešenie

Kedže jazyk L je regulárny, podľa definície regulárnych jazykov platí, že existuje regulárna gramatika G taká, že $L(G) = L$. Potom z vety 2 vyplýva, že ku gramatike G existuje 2NKA A taký, že $L(A) = L(G)$. Pri dokazovaní nášho tvrdenia najprv zostrojíme 2NKA A' pomocou automatu A. Popíšeme konštrukciu automatu A' v kapitole 3.1 a v kapitole 3.2 následne dokážeme, že naša konštrukcia je správna. Kedže vieme zostrojiť 2NKA A' taký, že $L(A') = L'$, tak podľa vety 1 môžme povedať, že jazyk L' je regulárny.

3.1 Konštrukcia

Nech $A = (K, \Sigma, \delta, q_0, F)$, pričom $K = \{q_0, q_1, \dots, q_n\}$, $F = \{q_{k_1}, q_{k_2}, \dots, q_{k_l} \mid k_1, k_2, \dots, k_l \in \{0, 1, \dots, n\}, 0 \leq l \leq n\}$. Zostrojíme $A' = (K', \Sigma, \delta', q_0, F')$ nasledovne:

Množina stavov K' bude obsahovať tri množiny stavov, pričom každá množina stavov akoby zodpovedala množine stavov K, tj. stav $q_i, q'_i, p_i \in K'$ zodpovedá stavu $q_i \in K$.

$$K' = Q \cap Q' \cap P \cap \{q_f\}$$

$$Q = K$$

$$Q' = \{q'_0, q'_1, \dots, q'_n\}$$

$$P = \{p_0, p_1, \dots, p_n\}$$

Množina stavov Q bude slúžiť na výpočet slova uv, množinu Q' využijeme na výpočet slova w, vlastne iba na odsimulovanie jeho výpočtu a v stavoch množiny P budeme iba posúvať hlavu. Na koniec stav q_f bude akceptačný, tj.

$$F' = \{q_f\}$$

$$\delta'(q_i, z) = \{(q_j, 1)\} \Leftrightarrow q_j \in \delta(q_i, z) \quad z \in \Sigma, 0 \leq i, j \leq n \tag{1}$$

$$\delta'(q_i, \$) = \{(q'_i, -1)\} \tag{2}$$

$$\delta'(q'_i, z_1) = \{(p_j, -1)\} \Leftrightarrow q_j \in \delta(q_j, z_2) \quad z_1, z_2 \in \Sigma, 0 \leq i, j \leq n \tag{3}$$

$$\delta'(p_i, z) = \{(q'_i, -1\} \quad z \in \Sigma \tag{4}$$

$$\delta'(q'_i, \cent) = \{(q_f, 1)\} \Leftrightarrow q_i \in F' \tag{5}$$

$$\delta'(q_f, z) = \{(q_f, 1)\} \quad z \in \Sigma \tag{6}$$

V (1) presne simulujeme automat A, čo je jasne vidieť. Pri tejto simulácii spravíme $2k$ krokov. V (2) sa iba prepneme do inej množiny stavov. Na to aby sme odsimulovali výpočet slova w, musíme spraviť ešte k krokov, v ktorých simulujeme jeho výpočet. Avšak páska má $2k$ symbolov, preto v každom druhom kroku, iba posunieme pásku (4). Teda krokov (3) je len k a v nich sa simuluje výpočet slova w. V (5) prechádzame do akceptačného stavu za daného predpokladu. V (6) iba presunieme pásku na koniec, aby sme slovo mohli akceptovať.

3.2 Formálny dôkaz

$L' \subseteq L(A')$:

Chceme ukázať, že pre každé slovo $uv \in L'$ existuje akceptačný výpočet v automate A'. Nech $uv \in L' \Rightarrow |u| = |v|$. Označme $|u| = k = |v|, u = u_1 u_2 \dots u_k, v = v_1 v_2 \dots v_k$. Keďže $uv \in L'$, tak $(\exists w \in \Sigma^*)(uvw \in L(A))$, tj. existuje akceptačný výpočet v A pre slovo uvw. Tento výpočet vyzerá nasledovne:

$$(q_0, uvw) \vdash_A^* (q_i, w) \vdash_A^* (q_j, \epsilon)$$

Výpočet slova uv v A' vyplýva z konštrukcie δ-funkcie. Kým čítame slovo uv, výpočet je nasledovný

$$(q_0, \cent\#u_1 \dots u_k v_1 \dots v_k\$) \vdash_{A'}^* (q_i, \cent u_1 \dots u_k v_1 \dots v_k\#\$)$$

Potom v $2k$ krokoch robíme k presunov, pomocou ktorých simulujeme výpočet slova w. Keďže v automate A sme sa na k krokov dostali k akceptovaniu slova, tak aj v automate A' sa dostaneme do stavu, z ktorého môžeme prejsť do stavu akceptačného. Simulovanie výpočtu slova w je nasledovné:

$$(q_i', ¢u_1 \dots u_k v_1 \dots v_{k-1}\#v_k\$) \vdash_{A'} (p_j, ¢u_1 \dots u_k v_1 \dots v_{k-2}\#v_{k-1}v_k\$) \vdash_{A'}$$

$$(q_j', ¢u_1 \dots u_k v_1 \dots v_{k-3}\#v_{k-2}v_{k-1}v_k\$) \vdash_{A'}^* (q_m', \#¢u_1 \dots u_k v_1 \dots v_k\$)$$

Čiže ak automat A akceptoval slovo uvw, tak aj automat A' môže prejsť do stavu akceptačného.

$$(q_f, ¢\#u_1 \dots u_k v_1 \dots v_k\$) \vdash_{A'}^* (q_f, ¢u_1 \dots u_k v_1 \dots v_k\#\$)$$

$L(A') \subseteq L'$:
Nech existuje akceptačný výpočet pre slovo uv v A'. Tento akceptačný výpočet musí obsahovať konfiguráciu $(q_i, ¢uv\#\$)$, pretože musí podľa δ-funkcie prečítať celú pásku. Z konštrukcie δ'-funkcie vyplýva, že v automate A existuje výpočet pre slovo uvw a to, $(q_0, uvw) \vdash_A^* (q_i, w)$. Následne automat A' nedeterministicky generuje slovo w. Podľa δ-funkcie máme zabezpečené, že sa vykoná práve k-krokov, v ktorých sa simuluje generovanie slova w. Nezaujíma nás aké je to slovo, ktorého prefixom je uv, ale zaujíma nás či také slovo existuje. Preto sa teraz bude automat iba presúvať medzi stavmi, tak aby presne po k-krokoch skončil v stave, z ktorého môže prejsť do akceptačného stavu. Ak slovo uv je prefixom nejakého slova z L, tak nedeterminizmus nám zaručí, že automat sa rohodne tak, aby po k krokoch generujúcich w skončil automat v stave, z ktorého môže prejsť do akceptačného stavu.

Ak sa automat dostane do stavu q_f, slovo w existuje a hlava sa presunie koniec a automat skončí v konfigurácii $(q_f, ¢uv\#\$)$.