V.Forner 2i1

Prezentácia úlohy č.3

Zadanie:

Nech L je regulárny jazyk. Pomocou nedeterministických dvojsmerných konečných automatov dokážte, že jazyk $L' = \{ uv \mid \exists (w \in \Sigma^*) : uvw \in L ,\ |u|=|w|=|v| \}$ je regulárny.

Definícia:

Dvojsmerným (konečným nedeterministickým) automatom nazývame 5-ticu $A=(K, \Sigma,\delta,q0,F)$, kde K, Σ,δ,q_0 sú ako u KA a

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

$\delta(q,\$) \subseteq \{0,1\}$,

$\delta(q,¢) \subseteq \{-1,0\}\ \forall q\in K.$

Definícia:

Nech $\lceil \notin \Sigma$. Konfiguráciou konečného deterministického automatu A nazývame dvojicu $(q,¢\ \Sigma^*\lceil\Sigma^*\$) \cup (q,\lceil¢\ \Sigma^*\$)$.

Definícia:

Krokom výpočtu $\vdash$ dvojsmerného konečného nedeterministického automatu A nazývame binárnu relácie na nmožine konfigurácií definovanú následovne:

$(p,u\lceil av) \vdash (q,ua\lceil v)$ ak $\{(q,1)\} \in \delta(p,a)$

$(p,u\lceil av) \vdash (q,u\lceil av)$ ak $\{(q,0)\} \in \delta(p,a)$

$(p,ua\lceil v) \vdash (q,u\lceil\#av)$ ak $\{(q,-1)\} \in \delta(p,a)$.

Definícia:

Jazyk akceptovaný konečným dvojsmerným nedeterministickým automatom A je

$L(A)=\{w \mid (q_0,¢\lceil w\$) \vdash^* (q_f,¢w\lceil\$) , q_f \in K\}$

Veta:

Ku každej regulárnej gramatike L existuje ekvivalentná gramatika L' v normálnom tvare, kde $P\subseteq N\times TN\cup\{\varepsilon\}$ a L=L'.

Riešenie:

Nech G je gramatika generujúca jazyk L a nech G je v normálnom tvare tj.

$P \subseteq N\times TN\cup\{\varepsilon\}$.

Nech N je množina neterminálov a T množina terminálov gramatiky G, σ je začiatočný neterminál. Zostrojíme automat A, ktorý akceptuje jazyk L' takto:

$G=(N,T,P,\sigma)$

$N = \{ N_i \}\ i=0..n$

$T = \Sigma_G$

$P \subseteq N \times TN \cup \{\varepsilon\}$

σ – počiatočný stav

Zostrojíme automat A, ktorý akceptuje jazyk L' takto:

$\Sigma_A = \Sigma_G$
$K = \{q_{0,Ni}, q_{1,Ni}, q_{2,Ni}, q_{3,Ni}, q_f \mid N_i \in N\}$ N – množina neterminálov gramatiky G
$F = \{q_f\}$
Počiatočný stav – $q_{0,\sigma}$

$\delta(q_{0,Ni},x) = \{ (q_{1,Nj},1) \mid N_i \Rightarrow_G x\,N_j \}, x \in \Sigma$ } automat simuluje generovanie
$\delta(q_{1,Ni},x) = \{ (q_{0,Nj},1) \mid N_i \Rightarrow_G x\,N_j \}, x \in \Sigma$ } uv v gramatike G a zároveň
$\delta(q_{0,Ni},\$) = \{ (q_{2,Ni},-1) \}$ } kontroluje či má 2*k znakov
$\delta(q_{2,Ni},x) = \{ (q_{3,Nj},-1) \mid N_i \Rightarrow_G x\,N_j \}, x \in \Sigma$
$\delta(q_{3,Ni},x) = \{ (q_{2,Ni},-1) \}, x \in \Sigma$ } automat sa snaží vygenerovať
$\delta(q_{2,Ni},¢) = \{ (q_f,1) \mid N_i \Rightarrow_G \varepsilon \}$ } slovo w
$\delta(q_f,x) = \{(q_f,1)\}, x \in \Sigma$

$A = (K, \Sigma, \delta, q_{0,\sigma}, F)$

Dôkaz:

Potrebujeme ukázať, že platí L(A)=L'.Nato potrebujeme 2 inklúzie:

L' ⊆ L(A) :

Nech $x \in L'$. Potom x=uv pre vhodné $u,v \in \Sigma^*$ také že |u|=|v| a existuje $w \in \Sigma^*$ také, že |w|=|u| a $uvw \in L$.Treba nájsť akceptačný výpočet slova uv na automate A.

1°.Nech $\varepsilon \in L'$
Nech |u|=|v|=0, $uv \in L'$. Existuje $w \in \Sigma^*$ také, že |w|=|u|=|v| a $uvw \in L$. Teda u=v=w=ε, uvw=ε. Existuje odvodenie slova uvw v gramatike G : $\sigma \Rightarrow \varepsilon$
Potom akceptačný výpočet uv vyzerá takto:
$(q_{0,\sigma}, \$) \vdash (q_{2,\sigma}, ¢) \vdash (q_f,\$)$
a preto $uv \in L'$.

2°. Nech tvrdenie platí pre všetky |u|=|v|=k. Ukážem, že platí aj pre |u|=|v|=k+1.

Nech $u=a_1 \ldots a_{k+1}$, $v=a_{k+2} \ldots a_{2k+2}$, $uv \in L'$ a existuje $w=a_{2k+3} \ldots a_{3k+3}$, $a_i \in \Sigma, \forall i \in \{1..3k+3\}$ a $uvw \in L$, |u|=|v|=|w|. Teda existuje odvodenie slova uvw v gramatike G
$\sigma \Rightarrow a_1 N_{k1} \Rightarrow a_1 a_2 N_{k2} \Rightarrow \ldots \Rightarrow a_1..a_{3k+2} N_{3k+2} \Rightarrow a_1..a_{3k+3} N_{3k+3} \Rightarrow a_1..a_{3k+3} = uvw$
$k_i \in \{1..n\}$

Potom akceptačný výpočet slova uv na automate A vyzerá takto:
$(q_{0,\sigma}, a_1) \vdash (q_{1,N1}, a_2) \vdash \ldots \vdash (q_{1,N2k+1}, a_{2k+2}) \vdash (q_{0,N2k+2}, \$) \vdash$
$(q_{2,N2k+3}, a_{2k+2}) \vdash (q_{3,N2k+3}, a_{2k+1}) \vdash (q_{2,N2k+4}, a_{2k}) \vdash \ldots \vdash (q_{3,N3k+3}, a_1) \vdash (q_{2,N3k+3}, ¢) \vdash$
$(q_f, a_1) \vdash \ldots \vdash (q_f,\$)$
preto $uv \in L(A)$.

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

L(A)⊆L':

Treba ukázať, že ak je slovo uv akceptované automatom A, tak patrí do L'. Keďže každé slovo uv patriace do L' má párny počet znakov a existuje k nemu sufix w tak, že uvw∈L a |u|=|v|=|w|. Preto stačí ukázať, že vieme nájsť w∈Σ* tak, že existuje odvodenie slova uvw v gramatike G a že w má polovičný počet znakov ako uv.

Nech existuje akceptačný výpočet slova uv na automate A. Nech
$(q_{0,\sigma},a_1) \vdash (q_{1,N1},a2) \vdash \dots \vdash (q_{1,N2k-1},a_{2k}) \vdash (q_{0,N2k},\$) \vdash$
$(q_{2,N2k+1},a_{2k}) \vdash (q_{3,N2k+1},a_{2k-1}) \vdash \dots \vdash (q_{3,N3k-1},a_1) \vdash (q_{2,N3k}, ¢) \vdash$
$(q_f,a1) \vdash \dots \vdash (q_f,\$)$ je akceptačný výpočet.

Keďže slovo uv sa kontroluje v stavoch $q_{0,Ni}$ a $q_{1,Ni}$, Výpočet uv sa dostane do stavu $q_{0,Ni}$ po prečítaní uv, len ak má uv párny počet znakov. Toto dokážem matematickou indukciou

1°.|uv| = 0 $(q_{0,\sigma},\$)$ niet čo dokazovať.

2°.Nech tvrdenie platí pre všetky uv také, že |uv|=2k. Dokážem, že platí aj pre |uv|=2k+2
$uv=a_1 \dots a_{2k+2}$
$(q_{0,\sigma},a_1) \vdash (q_{1,N1},a_2) \vdash^{IP}. \vdash (q_{0,N2k},a_{2k+1}) \vdash (q_{1,N2k+1},a_{2k+2}) \vdash (q_{0,N2k+2},\$)$

Keďže výpočet sa dostane ďalej do stavu $q_{2,Ni}$ len zo stavu $q_{0,Nj}$ ak automat číta znak \$, ak automat akceptuje slovo uv, musí prejsť stavom $q_{2,Ni}$ a preto musí mať párny počet znakov. Keďže má párny počet znakov dá sa rozdeliť na dve polovice tak, že |u|=|v|.

V stavoch $q_{2,Ni}$ a $q_{3,Ni}$ sa kontroluje,či existuje w∈Σ* také, že uvw∈L. Ak také w existuje, automat sa dostane do stavu q2,Ni pri čítaní znaku ¢. Keďže simuluje prácu gramatiky G len v stavoch $q_{2,Ni}$, dostane sa do konfigurácie $(q_{2,Ni}, ¢)$ po k presunoch zo stavu $q_{2,Ni}$ do $q_{3,Ni}$, ak |uv|=2k. Toto dokážem matematickou indukciou na dľžku uv

1°.|uv|=0 – niet čo dokazovať
$|uv|=2$, $uv=a_1a_2$
$(q_{2,Ni}a_2) \vdash (q_{3,Nj}a_1) \vdash (q_{2,Nj}\ ¢)$

2°.Nech tvrdenie platí pre |uv|=2k. Dokážem, že platí pre |uv|=2k+2.
$uv=a_1..a_{2k+2}$
$(q_{2,Ni},a_{2k+2}) \vdash (q_{3,Nj},a2k+1) \vdash^{IP}. \vdash (q_{2,Nj},a_2) \vdash (q_{3,Nj},a_1) \vdash (q_{2,Nj}\ ¢)$

Preto 2|w|=|uv| a keďže uv má párny počet znakov a dá sa rozdeliť na polovicu platí |u|=|v|=|w|. Z konštrukcie automatu, ktorá kopíruje generovanie slova uvw v gramatike G vyplýva, že existuje odvodenie slova uvw v G.
⇒ uvw∈L ⇒uv∈L'.
⇒L(A)⊆L'.

⇒L(A)=L'.
Keďže jazyk je akceptovaný konečným dvojsmerným automatom je regulárny.