Michal Polan
Fakulta matematiky, fyziky a informatiky
Univerzita Komenského v Bratislave
odbor informatika
1.4.2001

1. Zadanie

Pre daný TM M nech $L_M = \{ u\#v^R \mid u, v$ sú konfigurácie v M a zároveň $u \vdash v \}$. Dokážte, že L_M je bezkontextový.

2. Definície a tvrdenia

Možnosti zápisu konfigurácie TM:
1, Konfigurácia so stavom:
(wqv), $v,w \in \Gamma^*$ pričom konfigurácie s hlavou na začiatku, resp. na konci zapisujeme ako (wq), resp (q**B**w)

2, Konfigurácia s bičíkom ⌈:
(q, w⌈ v), $v,w \in \Gamma^*$ pričom konfigurácie s hlavou na začiatku, resp na konci zapisujeme ako (q,w⌈), resp. (q, ⌈**B**w)

3, Konfigurácia s udaním pozície od začiatku slova:
$(q, a_1a_2...a_n,i)$, $w=a_1a_2...a_n$, $w \in \Gamma^*$ pričom je výhodnejšie používať Turingov stroj s jednosmernou páskou a potom sa hlava môže nachádzať na miestach 1 - začiatok, až n+1 - koniec (vtedy čítame Blank)

Krok výpočtu:
Krok výpočtu TS je relácia ⊢ nakonfiguráciách definovaná nasledovne:
$\forall$ a, b, c $\in \Gamma \cup \{\mathbf{B}\}$, p,q $\in$ K
- $\Gamma^*qa\Gamma^* \vdash \Gamma^*pb\Gamma^* \Leftrightarrow (p,b,0) \ni \delta(q,a)$
- $\Gamma^*qa\Gamma^* \vdash \Gamma^*bp\Gamma^* \Leftrightarrow (p,b,1) \ni \delta(q,a)$
- $\Gamma^*cqa\Gamma^* \vdash \Gamma^*pcb\Gamma^* \Leftrightarrow (p,b,-1) \ni \delta(q,a)$

Veta 1. K ľubovoľnej bezkontextovej gramatike G existuje zásobníkový automat A, ktorý akceptuje L(G) prázdnou pamäťou.

Veta 2. Nech A je zásobníkový automat. Potom existuje bezkontextová gramatika G taká, že L(G)=N(A).

Dôkazy jednotlivých viet sú uvedené v skriptách.

Zásobníkový automat:
Nedeterministickým zásobníkovým automatom nazývame 7-ticu $(K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, kde K je konečná množina stavov, Σ je abeceda vstupných symbolov, Γ je abeceda zásobníkových symbolov, $Z_0 \in \Gamma$ je symbol ktorý je v zásobníku na začiatku výpočtu, $F \subseteq K$ je množina akceptačných stavov a je δ prechodová funkcia.
(Definície prechodovej funkcie, konfigurácie a kroku výpočtu sa nachádzajú v skriptách)

3. Riešenie

Úlohou je dokázať, že daný jazyk L_M je bezkontextový. Bezkontextovosť jazyka L_M je najjednoduchšie dokázať zostrojením zásobníkového automatu, ktorý tento jazyk akceptuje.

Podľa vety 1 a 2 (uvedené vyššie) je existencia tohto automatu zároveň aj dôkaz, že daný jazyk L_M je bezkontextový.

Pri riešení úlohy sa preto ďalej budem zaoberať hlavne konštrukciu hľadaného zásobníkového automatu.

Vlastnosti jazyka L_M.

Pre jednoduchšiu konštrukciu automatu je potrebné objasniť si najprv vlastnosti jazyka L_M. Jazyk L_M akceptuje slová, ktoré sú spojením dvoch konfigurácii turingového stroja, oddelených znakom "#", pričom druhá v poradí je zapísaná opačne (revers). Bez ujmy na všeobecnosti budem uvažovať že dané slová u, v sú zapísané v prvom tvare konfigurácie (t.j. konfigurácia so stavom)

Vzhľadom na to že slovo v je odvodené zo slova u (resp. konfigurácia v, u) môžeme rozdeliť situáciu na 3 základné prípady a preskúmať aké slová akceptuje jazyk L_M.

Nech $u=a_1a_2...a_npb_1b_2...b_n$, a automat krokom odvodenia zapísal písmeno x na pásku

1, hlava automatu ostala stáť - c=0

$v= a_1a_2...a_npxb_2...b_n$

2, hlava automatu sa posunula doprava c=1

$v=a_1a_2..a_nxpb_2...b_n$

3, hlava automatu sa posunula doľava c=-1

$v=a_1a_2...a_{n-1}pa_nxb_2..b_n$

Podstatné je, že zmena oproti pôvodnému slovu nastala maximálne v troch znakoch. Slová akceptované jazykom L_M budú teda tvaru

$$\mathbf{u\#v^R \text{ teda } a_1a_2...a_npb_1b_2...b_n\#b_n...b_2z_1z_2z_3a_{n-1}...a_2a_1,}$$

kde $z_1z_2z_3$ sú tri zmenené písmená podľa použitého pravidla.

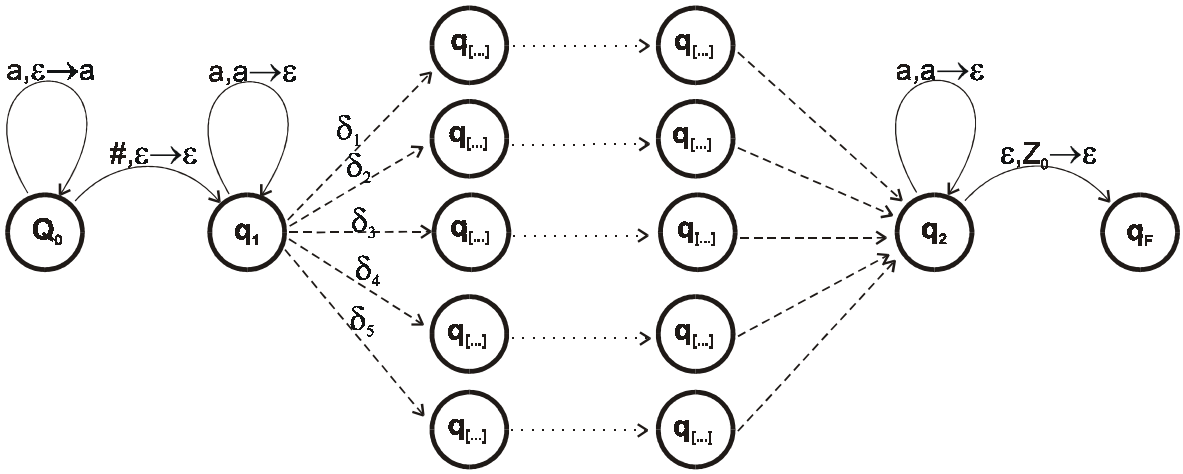
4. Konštrukcia

Konštrukcia automatu bude pomerne jednoduchá. Automat má akceptovať slova $w = a_1a_2...a_npb_1b_2...b_n\#b_n...b_2\mathbf{z_1z_2z_3}a_{n-1}...a_2a_1$, kde $\mathbf{z_1z_2z_3}$ sú tri konkrétne zmenené písmena.

Činnosť automatu:

Automat bude čítať slovo zo vstupnej pásky a zapisovať ho do zásobníka až kým nenačíta znak #. Po načítaní tohto znaku bude automat načítané písmená vyhadzovať zo zásobníka (v prípade, že je to požadované slovo znaky sú rovnaké, inak automat slovo neakceptuje). V určitej časti slova sa automat nedeterministicky rozhodne, že práve teraz prišli hľadané zmenené písmená a predá riadenie stavu, ktorý nedeterministicky vyberie aká časť delta funkcie bola použitá v kroku odvodenia a overí, či zmenené písmená mohli byť naozaj následkom vybranej časti delta funkcie. Po overení zmeny automat porovná zvyšné písmená vstupného slova so zásobníkom.

Princíp automatu:



Formálny zápis:

Nech M=(K', Σ', Γ', δ', q_0', F')

Hľadaný automat A definujeme ako:
A=(K, Σ, Γ, δ, q_0, Z_0, F), kde
K={q_0,q_1,q_2,q_F}∪{stavy ošetrujúce jednotlivé pravidlá TM M}
t.j
∀q_1,q_2∈K', ∀a,b∈Γ', ∀c ∈ {-1,0,1} (q_2,b,c) ∈ δ'(q_1,a) ⇒ $q_{[q1,q2,a,b,c]}$ ∈K, $q_{x[q1,q2,a,b,c]}$ ∈K (∀x∈Γ')
Σ=Γ'∪{#}∪K' (označenia stavov); (Γ'∩{#}∩K'=θ)
Γ=Σ∪{ Z_0}
F={q_F}

delta funkcia δ je definovaná nasledovne:

(1) δ(q_0,a,b)=(q_0,ab); a∈Σ, b∈Γ //čítané znaky zapisujem do zásobníka
(2) δ(q_0,#,b)=(q_1,b);
(3) δ(q_1,b,b)=(q_1,ε); //čítané znaky vyhadzujem zo zásobníka (porovnávam zásobnik so vstupnou páskou)

nedeterministický krok (vyberiem pravidlo, kt. zmenilo slovo v):
∀q_1,q_2∈K', ∀a,b∈Γ', ∀c ∈ {-1,0,1}

(4) (q_2,b,c) ∈ δ'(q_1,a) ⇒ δ (q_1, ε,b) = ($q_{[q1,q2,a,b,c]}$,b)
ak c=0
δ ($q_{[q1,q2,a,b,0]}$,b,a) = ($q'_{[q1,q2,a,b,0]}$, ε)
δ ($q'_{[q1,q2,a,b,0]}$,q_2,q_1) = (q_2, ε)
ak c=1
δ ($q_{[q1,q2,a,b,0]}$,q_2,a) = ($q'_{[q1,q2,a,b,0]}$, ε)
δ ($q'_{[q1,q2,a,b,0]}$,b,q_1) = (q_2, ε)
ak c=-1
δ ($q_{[q1,q2,a,b,0]}$,b,a) = ($q'_{[q1,q2,a,b,0]}$, ε)

$\delta\,(q'_{[q1,q2,a,b,0]},c,q_1) = \quad (q_{c[q1,q2,a,b,0]},\ \varepsilon)\ \forall c \in \Gamma$
$\delta\,(q_{c[q1,q2,a,b,0]},q_2,c) = \quad (q_2,\ \varepsilon)$

(5) $\delta(q_2,b,b)=(q_2,\varepsilon)$; //podobne ako (3)
(6) $\delta(q_2,\ \varepsilon,Z)=(q_F,\varepsilon)$; //na vstupe ani v zásobníku nie je nič - končím a akceptujem

5. Dôkaz

$L_M \subset L(A)$

Táto inklúzia je pomerne zrejmá, pretože automat bol konštruovaný tak, aby všetky slova jazyka L_M akceptoval. To znamená, že ak nejaké $w \in L_M$, musí byť zrkadlovo súmerné (podľa znaku "#") s výnimkou maximálne 3 znakov, ktoré zmenil krok odvodenia druhého slova. Pri tomto kroku muselo byť použité nejaké pravidlo z delta funkcie δ'. Teda v automate existujú stavy kt. túto zmenu testujú (4).

$L(A) \subset L_M$

Táto inklúzia je tiež pomerne jednoduchá, pretože aj konštruovaný automat je jednoduchý. Cieľom je vlastne dokázať, že všetky slová, ktoré automat akceptuje patria jazyku L_M. Tento dôkaz môžeme rozdeliť na dve časti (poddôkazy), ktoré sa navzájom neovplyvňujú.

Prvá časť:

Prvá časť bude zjednodušený automat ktorý vznikne vynechaním nedeterministického kroku (t.j. všetkých stavov $q_{[...]}$). Všetky vynechané stavy vždy vyberú zo zásobníka a načítajú rovnaký počet znakov, čiže nemajú principiálny vplyv na túto časť automatu. Táto zjednodušená časť automatu vlastne akceptuje len slová tvaru ww^R. Dôkaz tohto tvrdenia je jednoduchý, dá sa previesť matematickou indukciu vzhľadom na dĺžku slova w.

Druhá časť:

Druhú časť automatu tvorí nedeterministicky krok, spolu so stavmi testujúcimi pravidlo. Táto časť automatu akceptuje práve jedno slovo (zmenené písmena) podľa toho, pre aké pravidlo sa automat v nedeterministickom kroku rozhodol. Teda je to akýsi jednoduchý automat ktorý testuje či na vstupe a v zásobníku je správna postupnosť znakov. Funkčnosť automatu je dosť zrejmá (Je to podobne typ automatu ktorý akceptuje práve jedno slovo "abc").

Teda automat bude akceptovať iba a len slovo ktoré spĺňa podmienky (resp. vlastnosti) daného jazyka LM)

Teoreticky by problém mohol nastať iba vtedy ak by sa automat v určitom kroku rozhodol nesprávne. To je:

1, Automat vybral nesprávne pravidlo

Automat nebude slovo akceptovať, pretože v zásobníku nebudú správne symboly, teda automat sa zasekne niekde v postupnosti stavov q[...].

2, Automat sa začal testovať pravidlo skôr

Automat sa zasekne, lebo nenájde znak pre stav v pôvodnom TM M

3, Automat začal testovať pravidlo neskôr

To isté ako 2.

Absolútne formálne by sa to dalo dokázať cez MI vzhľadom na dĺžku konfigurácie, pričom predpoklad by platilo |w|=5 alebo |w|=5 podľa typu pravidla (prázdna páska alebo jeden znak). Pridaním znaku by sa však funkčnosť automatu nezmenila, pretože, tieto znaky sa len uložia do zásobníka a porovnajú s koncom, teda ak platí predpoklad bude akceptované aj dlhšie slovo.

Pre jednoduchosť to dokážem pre c=0, ostatné prípady by sa dokazovali podobne, prakticky by sa zmenil len predpoklad indukcie.

(0) Automat je v stave q1, na páske sa nachádza nejaský znak a (a môže byť aj **B**), ktorý sa prepíše pomocou delta funkcie na znak b, a automat sa prejde do stavu q2.

Formálne zapísane:

Nech $w=q_1a\#bq_2$ a nech $\delta(q_1,a)=(q_2,b,0)$

Potom v automate budú existovať takéto časti delta funkcie (z konštrukcie)

$\delta\ (q_{[q1,q2,a,b,0]},b,a) = \ (q'_{[q1,q2,a,b,0]}, \varepsilon)$

$\delta\ (q'_{[q1,q2,a,b,0]},q_2,q_1) = \ (q_2, \varepsilon)$

keď sa pozrieme na zvyšok delta funkcie zistíme, že predpoklad triviálne platí.

(1) Nech teda platí pre všetky slová dĺžky 5+2*n. Slovo dĺžky 5+2*(n+1) sa líši iba tým že pred alebo za stavom sa v každom podslove pridalo jedno písmeno. Teda automat v stave q_0 načíta do zásobníka o jedno písmeno viac a v stave q_1 alebo q_2 o jedno písmeno viac zo zásobníka vyhodí.

6. Záver

Úlohou bolo dokázať, že daný jazyk LM je bezkontextový. Bezkontextovosť jazyka L_M som dokázal konštrukciou zásobníkového automatu A, ktorý akceptuje jazyk L_M. Základná myšlienka automatu spočíva v nedeterministickom zvolení si pravidla, ktoré bolo použité v TM na odvodenie druhej časti slova a overenie či toto slovo má požadovaný tvar.