PREZENTÁCIA Z FORMÁLNYCH JAZYKOV A AUTOMATOV, ÚLOHA 3.2

DANIEL ŠIMKO

Pomocou LBA (lineárne ohraničených automatov) dokážte uzavretosť triedy kontextových jazykov na inverzný homomorfizmus. Teda, pre daný LBA $\mathcal{A}$ a homorfizmus h zostrojte LBA $\mathcal{A}'$, taký že $L(\mathcal{A}') = h^{-1}(L(\mathcal{A})) = \{w \mid h(w) \in L(\mathcal{A})\}$.

Predpokladáme :

$$h(a) = w \ \forall a \in \Sigma_1; w \in \Sigma_2^*$$

kde n je dĺžka najdlhšieho w, n¿0 (n=0 je triviálny prípad)

$$\mathcal{A} = (K, \Sigma_2, \Gamma, \delta, q_0, F)$$

kde $K = \{q_n\}$

Z poslednej rovnosti zadania nám hneď napadá intuitívne riešenie : preložiť slovo zo vstupnej pásky daným homomorfizmom a zistiť, či toto slovo akceptuje automat $\mathcal{A}$. Ale pri preklade nastáva problém : výsledné slovo môže mať rôznu dĺžku ako slovo pôvodné, v najhoršom prípade môže byť až n-krát také dlhé. Keďže pásku nemôžeme ľubovoľne naťahovať, musíme sa s týmto problémom nejako vysporiadať. Jedným z možných riešení je rozdeliť si každé políčko na páske na n políčok - tým dosiahneme, že slovo zo vstupu dokážeme preložiť a výsledok sa zmestí na pásku. Skomplikuje sa nám ale výpočet - pri každom načítaní dostaneme n znakov. Budeme teda musieť primerane upraviť delta funkciu pôvodného automatu. Vytvorme teda nový automat :

$$\mathcal{A}' = (K', \Sigma_1, \Gamma', \delta', q_0', F')$$

a zaoberajme sa najprv delta funkciou. Na preklad použijeme postupne tieto štyri pravidlá :

$$\begin{aligned}
(I) \quad & \delta'(q_0', x) = (q_0', v, 1) \\
(II) \quad & \delta'(q_0', \$) = (q_0'', \$, -1) \\
(III) \quad & \delta'(q_0'', v) = (q_0'', v, -1) \\
(IV) \quad & \delta'(q_0'', \cent) = ((?), \cent, 1)
\end{aligned}$$

kde v reprezentuje špeciálny znak pracovnej abecedy nového automatu získaný z homomorfizmu ako vektor dĺžky n, ktorý obsahuje zľava písmená z pravej strany homomrfizmu h(x) doplnené epsilonmi. Teda napr. ak h(x)=abc a n=5 tak potom $v = (a, b, c, \varepsilon, \varepsilon)$. V prvidle (IV) máme zatiaľ neoznačený stav, ale je jasné, že to bude vlastne modifikovaný q_0.

Poznámka. Ak sa teraz pozrieme na pásku, odmyslíme si zátvorky, čiarky a epsilony, máme slovo zobrazené homomorfizmom. To sa ľahko dokáže pomocou indukcie vzhľadom na dĺžku slova a využijeme to neskôr pri dôkaze.

Teraz teda musíme prepísať delta funkciu pôvodného automatu tak, aby dokázal pracovať s našou momentálnou verziou pásky. Musíme sa teda vedieť pohybovať vo vnútri jedného vektora ako po páske. V stave musíme mať teda aj miesto vo vektore, kde sa práve nachádzame. Okrem toho (pre neskoršie potreby) v ňom potrebujeme mať aj smer posledného posunu. Hrôzostrašné pravidlo (V) našej delta funkcie odvodenej od pôvodnej teda vyzerá :

$(q_l^{i+d,d}, u, 0) \in \delta'(q_k^{i,j}, v); u = (v_1, ..., v_{i-1}, u_0, v_{i+1}, ..., v_n); (q_l, u_0, d) \in \delta(q_k, v_i); i \in 1...n$

všimnime si ešte, že $v_i \neq \varepsilon,\ v \neq \$\ a\ v \neq ¢$.

Prvý horný index obsahuje údaj, na ktorej pozícii vektora sa práve nachádzame. ten sa zmení v závislosti od posunu v pôvodnej delta funkcii. zmení sa i vektor, pričom sa zmení práve jedno písmeno z pracovnej abecedy Γ v súlade s pôvodnou δ-funkciou. Na páske však zostaneme na tom istom mieste. Kedy sa však presunúť aj na reálnej páske? Všimnime si, že prvý horný index sa môže dostať i na hodnoty 0 a n+1. A to je práve čas na zmenu pozície na páske :

$$(VI) \quad (q_k^{n,-1}, v, -1) \in \delta'(q_k^{0,-1}, v)$$

$$(VII) \quad (q_k^{1,1}, v, 1) \in \delta'(q_k^{n+1,1}, v)$$

Význam indexov je myslím si intuitívne zrejmý rovnako ako i celkový zmysel týchto pravidiel.

Ďalším problémom sú epsilony vo vektoroch. Zatiaľ nemáme žiadne pravidlá na ich spracovanie. Tieto ale nebudú vôbec náročné - epsilony stačí jednoducho preplávať. Teraz vidíme význam druhého horného indexu :

$$(VIII) \quad (q_k^{i+d,d}, v, 0) \in \delta'(q_k^{i,d}, v); i \in 1...n; v_i = \varepsilon$$

Už stačí ošetriť len znaky $\$\ a\ ¢$:

$$(IX) \quad (q_l^{1+d,d}, \$, 0) \in \delta'(q_k^{1,1}, \$); (q_l, \$, d) \in \delta(q_k, \$)$$

$$(X) \quad (q_l^{n+d,d}, ¢, 0) \in \delta'(q_k^{n,-1}, ¢); (q_l, ¢, d) \in \delta(q_k, ¢)$$

Teraz môžeme doplniť zostávajúce parametre. V pravidle (IV) bude namiesto (?) stav $q_0^{1,1}$. Ďalšie množiny by sa dali podľa potreby zúžiť, ale vo všeobecnosti vyzerajú takto :

$$K' = \{q_k^{i,j} \mid \forall q_k \in K; \forall i \in 0...n+1; \forall j \in \{-1, 0, 1\}\}$$

$$\Gamma' = \{v \mid v = (v_1, ..., v_n); v_1...v_n \in \Gamma \cup \{\varepsilon\}\} \cup \{¢, \$\}$$

$$F' = \{q_k^{i,j} \mid q_k \in F\}$$

Máme teda slovo preložené homomorfizmom, ktoré sa skladá z písmen patriacich Σ_2 alebo epsilonov. Matematickou indukciou vzhľadom na dĺžku výpočtu ukážeme, že výpočet v $\mathcal{A}'$ je ekvivalentný výpočtu v $\mathcal{A}$. V prvom kroku výpočtu (stav $q_0^{1,1}$) môžeme načítať \$ alebo v. V prípade \$ máme podľa (IX) ekvivalentný výpočet (ev). V prípade v uvažujeme v_1 : ak je to písmeno, máme ev podľa (V). V prípade epsilonu sa podľa pravidiel (VIII) resp. (VII) dostaneme k najbližšiemu písmenu (resp \$), pričom zostaneme v stave q_0 pôvodného automatu a podľa predchádzajúceho dostaneme ev.

Predpokladajme teraz, že máme ev po nejaké písmeno. Dokážeme, že aj pre ďalší krok výpočtu existuje ekvivalent v $\mathcal{A}'$. Som v stave $q_k^{i,j}$ a na vstupe mám v. Ak i nie je z $1...n$, tak mám jedinú možnosť - pravidlá (VI) alebo (VII). Teraz už je. Ak v je vektor a spracovávaný znak je epsilon, tak ho preskočím pomocou (VIII), kým neprídem k inému znaku. Ak to nie je ¢ alebo \$, tak pomocou (V) získam ev. Až teraz mením dolný index stavu (ekvivalentne s $\mathcal{A}$) a prepisujem niečo na páske! Ak je to ¢ alebo \$, použijem (IX) resp. (X) a zas mám ev.

Poznámka. To, že pri použití (V), (IX) či (X) získam ev vyplíva z ich definície. Pravidlá (VI), (VII) a (VIII) sú oproti $\mathcal{A}$ navyše a slúžia len na potrebný prechod po páske. V skutočnosti na nej žiadne údaje nemenia a rovnako nemenia ani pseudostav automatu (dolný index).