Riešenia prvej sady domácich úloh

Peter Kostolányi

24. marca 2025

Úloha 1. Nech $m, n \in \mathbb{N} - \{0\}$. Nájdite tesný horný odhad na počet stavov *minimálneho* deterministického konečného automatu akceptujúceho jazyk typu $L_m \cup L_n$, kde L_m je jazyk akceptovaný deterministickým konečným automatom s m stavmi a L_n je jazyk akceptovaný deterministickým konečným automatom s n stavmi.[1] Svoje tvrdenie dokážte.

Riešenie. Dokážeme, že pre jazyky $L_m, L_n \subseteq \Sigma^*$ nad *spoločnou* abecedou Σ je hľadané tesné horné ohraničenie $f(m, n)$ dané ako $f(m, n) = mn$.

Uvažujme najprv *ľubovoľnú* dvojicu jazykov $L_m, L_n \subseteq \Sigma^*$ takých, že $L_m = L(A_m)$ a $L_n = L(A_n)$ pre nejaké deterministické konečné automaty $A_m = (K_m, \Sigma, \delta_m, q_{0,m}, F_m)$ a $A_n = (K_n, \Sigma, \delta_n, q_{0,n}, F_n)$ spĺňajúce $|K_m| = m$ a $|K_n| = n$. Deterministický konečný automat $A = (K, \Sigma, \delta, q_0, F)$ taký, že $|K| = mn$ a $L(A) = L_m \cup L_n$ potom môže byť daný ako $K = K_m \times K_n$, $q_0 = [q_{0,m}, q_{0,n}]$, $F = F_m \times K_n \cup K_m \times F_n$ a $\delta([p, q], c) = [\delta_m(p, c), \delta_n(q, c)]$ pre všetky $p \in K_m$, $q \in K_n$ a $c \in \Sigma$. V dôsledku toho je $f(m, n) \leq mn$.

Na dôkaz obrátenej nerovnosti uvažujme abecedu $\Sigma = \{a, b\}$ a jazyky

$$L_m = \{w \in \Sigma^* \mid \#_a(w) \not\equiv 0 \pmod{m}\}$$

a

$$L_n = \{w \in \Sigma^* \mid \#_b(w) \not\equiv 0 \pmod{n}\}.$$

Tieto jazyky sú evidentne rozoznávané deterministickými konečnými automatmi nad abecedou Σ o m resp. n stavoch. Dokážeme, že pravá syntaktická ekvivalencia $\equiv_{(L_m \cup L_n)}$ indukovaná jazykom $L_m \cup L_n$ nad abecedou Σ je indexu mn, v dôsledku čoho musí mať aj minimálny automat pre jazyk $L_m \cup L_n$ nad Σ presne mn stavov. Skutočne: pre ľubovoľné $p_1, p_2 \in \{0, \ldots, m-1\}$ a $r_1, r_2 \in \{0, \ldots, n-1\}$ také, že $[p_1, r_1] \neq [p_2, r_2]$ uvažujme slová $u = a^{p_1} b^{r_1}$ a $v = a^{p_2} b^{r_2}$. Vezmime $z = a^{m-p_1} b^{n-r_1}$. Potom $\#_a(uz) \equiv 0 \pmod{m}$ a $\#_b(uz) \equiv 0 \pmod{n}$, kým na druhej strane z predpokladu $[p_1, r_1] \neq [p_2, r_2]$ dostávame $\#_a(vz) \not\equiv 0 \pmod{m}$ alebo $\#_b(vz) \not\equiv 0 \pmod{n}$ – teda $uz \notin L_m \cup L_n$ a $vz \in L_m \cup L_n$, z čoho $u \not\equiv_{(L_m \cup L_n)} v$. Pravá syntaktická ekvivalencia $\equiv_{(L_m \cup L_n)}$ je teda indexu aspoň mn, pričom z dokázaného odhadu $f(m, n) \leq mn$ vyplýva, že jej index musí byť *presne* mn. $\square$

Úloha 2. Nech Σ, Γ sú abecedy, $L \subseteq \Sigma^*$ je (rozšírený) kontextový jazyk a $\tau \colon \Sigma^* \to 2^{\Gamma^*}$ je substitúcia taká, že pre všetky $c \in \Sigma$ je $\tau(c)$ kontextový jazyk neobsahujúci prázdne slovo ε. Musí byť v takom prípade $\tau(L) \in \mathscr{L}_{ECS}$?

Riešenie. Trieda $\mathscr{L}_{ECS}$ *je uzavretá* na takéto substitúcie. Nech A je lineárne ohraničený automat akceptujúci L a pre všetky $c \in \Sigma$ je A_c lineárne ohraničený automat taký, že $L(A_c) = \tau(c)$. Lineárne ohraničený automat B akceptujúci $\tau(L)$ môže pracovať nasledovne:

1. Vstupnú pásku so slovom $w \in \Gamma^*$ upraví na dvojstopú; na jednej stope ponechá vstupné slovo w a na druhej vstupe nedeterministicky vypíše nejaké slovo $x \in \Sigma^*$ také, že $|x| \leq |w|$. Je zrejmé, že uvedená nerovnosť musí byť splnená kedykoľvek $w \in \tau(x)$. Nech $x = a_1 \ldots a_k$ pre $k = |x|$ a nejaké $a_1, \ldots, a_k \in \Sigma$.

2. Simuláciou automatu A sa na príslušnej stope pásky presvedčí o tom, že $x \in L$.

[1] Je teda potrebné nájsť funkciu $f \colon \mathbb{N}^2 \to \mathbb{N}$ s nasledujúcimi dvoma vlastnosťami:

(i) Ak A je deterministický konečný automat s m stavmi a B je deterministický konečný automat s n stavmi, má minimálny deterministický konečný automat pre jazyk $L(A) \cup L(B)$ vždy *nanajvýš* $f(m, n)$ stavov.

(ii) Pre všetky $m, n \in \mathbb{N} - \{0\}$ existujú deterministické konečné automaty A, B s m resp. n stavmi také, že minimálny deterministický konečný automat pre jazyk $L(A) \cup L(B)$ má *presne* $f(m, n)$ stavov.

3. Symboly vstupného slova w nedeterministicky označí tak, aby týmto označením bola daná faktorizácia $w = w_1 \dots w_k$ slova w na podslová $w_1, \dots, w_k \in \Gamma^*$ (pre $x = \varepsilon$ je toto možné iba v prípade, že $w = \varepsilon$; pre $x = \varepsilon$ a $w \neq \varepsilon$ sa automat zasekne).

4. Pre $j = 1, \dots, k$ sa automat simuláciou automatu A_{a_j} presvedčí o tom, že $w_j \in \tau(a_j)$. V prípade, že všetky tieto simulované automaty akceptujú, akceptuje aj automat B.

Je zrejmé, že automat B akceptuje slovo $w \in \Gamma^*$ práve vtedy, keď existuje $x \in L$ také, že $x = a_1 \dots a_k$ pre nejaké $k \in [|w|]$ a $a_1, \dots, a_k \in \Sigma$, pričom $w \in \tau(a_1) \dots \tau(a_k)$. Keďže pre všetky $c \in \Sigma$ je $\tau(c) \subseteq \Gamma^+$, nastane táto situácia práve vtedy, keď existuje $x \in L$ také, že $w \in \tau(x)$ – čiže práve vtedy, keď $w \in \tau(L)$. Naozaj teda $L(B) = \tau(L)$. $\square$