

Longest Shortest Words in Regular Languages

Bc. Matúš Juran
RNDr. Peter Kostolányi, PhD.

- Konečný popis jazyka - napr. automat, gramatika, regulárny výraz
- Ak je neprázdny, obsahuje najkratšie slovo
- Má zmysel pri zohľadnení popisnej zložitosti
- Dolný a horný odhad jeho dĺžky
- Operácie na jazykoch - prienik, komplement

- Nedeterministické automaty
 - Alpoge, Ang, Schaeffer, Shallit (2011)
- Dvojsmerné automaty
 - E. Dobronravov, N. Dobronravov, Okhotin (2019)
 - Krymski, Okhotin (2020)
- Regulárne výrazy - komplement
 - Ellul, Krawetz, Shallit, Wang (2005)

- Rotujúce automaty - môžu sa vrátiť z konca pásky na začiatok
- Alternujúce automaty - zovšeobecnenie nedeterminizmu, všeobecné a existenčné stavy
- Dokážu popisovať len regulárne jazyky

- \mathcal{L} - postupnosť konečných množín jazykov
- \mathcal{I} - indexová množina
- $lsw_{\mathcal{L}} : \mathcal{I} \rightarrow \mathbb{N}$
- $lsw_{\mathcal{L}}(i)$ - maximálna dĺžka najkratšieho slova

- Výsledky založené na výskume dvojsmerných automatov
- Landauova funkcia -
$$g(n) = \max\{\text{lcm}(p_1, \dots, p_k) \mid p_1 + \dots + p_k \leq n\} = e^{(1+o(1))\sqrt{n \ln n}}$$
- Vieme postupne simulovať niekoľko NFA
- $\text{lsw}_{\mathcal{L}}(n) \geq g(n) - 1$

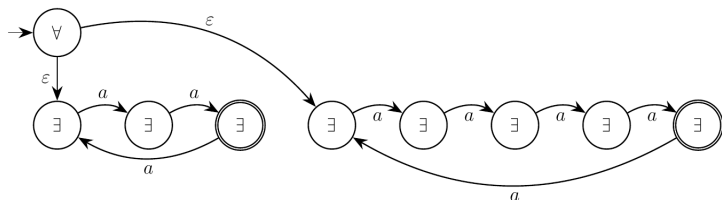
- Veľkosť abecedy lineárna od počtu stavov
- $w_0 = \varepsilon, w_i = w_{i-1}a_iw_{i-1}a_iw_{i-1}$
- $lsw_{\mathcal{L}}(n) \geq 3^{\lfloor \frac{n}{3} \rfloor} - 1 \approx 1.44^n$
- Väčšie ohraničené abecedy - nemôžeme priamo využiť techniku pre 2DFA

- $L \subseteq \Sigma_1^*$
- $L(\Sigma_2 L)^*$
- $L = L_1 \cap \dots \cap L_k$
- $L(\Sigma_2 L)^* = L_1(\Sigma_2 L_1)^* \cap \dots \cap L_k(\Sigma_2 L_k)^*$
- $L' \subseteq \Sigma_2^*$
- $lsw_{\mathcal{L}}(2n) \geq g(n)^2 - 1$ pre 2-znakovú abecedu
- $lsw_{\mathcal{L}}(kn) \geq g(n)^k - 1$ pre k -znakovú abecedu

- Prienik dvoch jazykov
- $lsw_{\mathcal{L}}((n, m)) \geq g(n)g(m) - 1$ pre binárne abecedy
- $lsw_{\mathcal{L}}((n, m)) \geq 1.44^{n+m} - 1$ pre abecedu s $n + m$ znakmi

Alternujúce automaty

- Podobné výsledky ako pre rotujúce automaty
- Viacnásobné čítanie vstupu (postupne/paralelne)
- Stavovo efektívne konštrukcie pre prienik
- Akceptovanie daných jazykov



- Regulárny výraz \rightarrow NFA \rightarrow AFA
- K AFA jednoduchý komplement
- Dlhé najkratšie slovo v komplemente $L(E) \rightarrow$ dlhé najkratšie slovo akceptované AFA

- $lsw_{\mathcal{L}}(25n + 112) \geq (2^n - 1)(n + 1) + 1$ pre $|\Sigma| = 5$
- $lsw_{\mathcal{L}}(75n + 363) \geq 3(2^n - 1)(n + 1) + 3$ pre $|\Sigma| = 2$
- Rastie pomalšie ako 1.44^n pre n stavov
- Stačí konštantne veľká abeceda

- Odhady pre prieniky dvoch jazykov
- $lsw_{\mathcal{L}}((n+1, m+1)) \geq g(m)(g(n)-1)$ pre $|\Sigma| = 2$
- $lsw_{\mathcal{L}}((25n+112, 25m+113)) \geq ((2^n-1)(n+1)+1)((2^m-1)(m+1)+2)$ pre $|\Sigma| = 10$
- $lsw_{\mathcal{L}}((3n+3, 3m+5)) \geq 3^n(1+3^m)$ vo všeobecnosti

- Odhady cez transformácie medzi modelmi
- Horšie výsledky

- Modely mali dodatočné obmedzenia
- Možné pokračovanie - uvoľnenie obmedzení

- Lema 7
- Ďalšie operácie
- Silnejšie modely