

Lineární verze Holubova algoritmu

David Tvrđý

22. března 2020

- 1 Cíle bakalářské práce
- 2 Úvod do problému
- 3 Dosavadní výsledky

Historie

- V roce 2009 představil Š. Holub první polynomiální algoritmus pro testování morfické primitivnosti slova.
- Později byla prezentována lineární verze tohoto algoritmu.

Cíle práce

- Popsat lineární verzi algoritmu.
- Porovnat lineární verzi s původní verzí.
- Implementovat lineární verzi algoritmu.

Slovo

Definice (Slovo)

Mějme konečnou abecedu Σ . Konečné posloupnosti prvků této abecedy, včetně prázdné posloupnosti, budeme nazývat *slova*. Prázdné slovo budeme značit ϵ .

Pevný bod

Definice (Pevný bod)

Mějme homomorfismus $h: \Sigma^* \rightarrow \Sigma^*$. Slovo $w \in \Sigma^*$ je *pevným bodem* homomorfismu h , pokud $h(w) = w$.

Definice (Triviální pevný bod)

Slovo $w \in \Sigma^*$ je *triviálním* pevným bodem, pokud $h(a) = a$ pro každý znak $a \in \Sigma$, který se ve slově w vyskytuje.

Morfická primitivnost

Definice (Morficky primitivní slovo)

Pokud pro každý homomorfismus, že slovo w je jeho pevným bodem, je slovo w i jeho triviálním pevným bodem, pak řekneme, že slovo w je *morficky primitivní*.

Jinak řekneme, že slovo w je *morficky složené*.

- Slovo $acacabab$ je morficky složené, jelikož je netriviálním pevným bodem pro h : $h(c) = ac$, $h(b) = ab$, $h(a) = \epsilon$.
- Naopak slovo $acacababa$ je morficky primitivní.

Morfický rozklad

Definice (Klíčový znak)

Bud' (w_1, \dots, w_k) rozklad slova w , tj. $w = w_1 \dots w_k$. Znak $e \in \Sigma$ je *klíčový znak* (slova w), pokud se ve slově w vyskytuje a pokud pro každé $i, j \in \{1, \dots, k\}$ platí, že

$$(|w_i|_e \leq 1) \wedge (|w_i|_e = |w_j|_e = 1 \implies w_i = w_j),$$

kde $|w_i|_e$ je počet výskytů znaku e ve faktoru w_i .

Definice (Morfický rozklad)

Bud' $F = (w_1, \dots, w_k)$ rozklad slova w . Pokud každý faktor w_i obsahuje klíčový znak, pak řekneme, že F je *morfický rozklad* slova w .

Morfický rozklad

Věta

Slovo w je morficky složené právě tehdy, když existuje nějaký jeho netriviální morfický rozklad F , tedy když existuje nějaký morfický rozklad, který má méně než $|w|$ faktorů.

Aktuální stav

- Aktuálně mám hotovou implementaci algoritmu a vizualizaci jeho chodu na konkrétních vstupech.
- Dále budu pracovat na testování a optimalizaci implementace.

Díky za pozornost!