

ADS2 Jiří Kalvoda: Cvičení 2

Informace k cv. jsou na <https://kam.mff.cuni.cz/~jirikalkvoda/vyuka/25z/ads2>.

Vyhledávání v textu – algoritmus KMP

Definice: Pro řetězec $\alpha = a_0a_1a_2 \dots a_{n-1}$ definujeme:

- prefix řetězce α : $\alpha[:l] = a_0a_1a_2 \dots a_{l-1}$.
- suffix řetězce α : $\alpha[k:] = a_ka_{k+1}a_{k+2} \dots a_{n-1}$.
- souvislý podřetězec řetězce α : $\alpha[k:l] = a_ka_{k+1}a_{k+2} \dots a_{l-1}$.

Úloha 1 (Konstrukce): Sestavte KMP automaty pro následující řetězce:

- kocka
- aaaaaaaaaa
- toronto

Úloha 2 (Timeout): Naivní algoritmus, který zkouší všechny možné začátky jehly v seně a vždy porovnává řetězce, má časovou složitost $\mathcal{O}(JS)$. Může být opravdu tak pomalý, uvážíme-li, že porovnávání řetězců skončí, jakmile najde první neshodu? Sestrojte vstup, na kterém algoritmus poběží $\Theta(JS)$ kroků, přestože nic nenajde.

Úloha 3 (Cenzura): Chceme otestovat jestli se v seně S nachází skrytý text $T = \{J_1, \dots, J_N\}$ (v tomto pořadí). Nicméně text je skrytý, takže mezi každými dvěma jehlami může být výplň – libovolný řetězec.

Definice: Máme-li řetězec α , nazvěme jeho *rotaci* o k pozic řetězec $\alpha[k:] \alpha[:k]$.

Úloha 4 (Rotace): Jak o dvou řetězcích zjistit, zda je jeden rotací druhého?

Úloha 5 (Minimální rotace*): Navrhněte algoritmus, který v lineárním čase nalezne tu z rotací zadaného řetězce, jež je lexikograficky minimální.

Úloha 6 (Prefixy a suffixy): Pro dané slovo najděte nejdelší vlastní prefix, který je současně suffixem.

Úloha 7 (Fibonacci): Fibonacciho slova definujeme $f_0 = a$, $f_1 = b$, $f_n = f_{n-2} + f_{n-1}$ pro nějaké znaky a, b . Najděte nejdelší fibonacciho slovo v seně (pro libovolné znaky a, b).