

# 1. DOMÁCÍ ÚKOL Z ADS II.

Vyhledávání v textu

Odevzdávejte emailem na [tarken@kam.mff.cuni.cz](mailto:tarken@kam.mff.cuni.cz) do 10.11.2017 9:00.

Při vymýšlení úkolů můžete spolupracovat, chtěl bych ale, abyste řešení sepsali každý sám.

**PŘÍKLAD PRVNÍ** Nalezněte příklad slov a textu v němž je asymptoticky více než lineární počet výskytů. Tedy ukažte, že pro každé  $n$ , kde  $n$  je počet všech písmen na vstupu (součet všech slov i textu) existuje vstup takový, že počet výskytů je asymptoticky větší než lineární.

[3 body]

**PŘÍKLAD DRUHÝ** *Cenzor* dostane text a množinu zakázaných řetězců. Cenzor nalezne ne-jevější výskyt zakázaného řetězce ten smaže a postup opakuje dokud existuje zakázaný vzor v textu. Navrhněte algoritmus, který usnadní cenzorovu práci.

[3 body]

**PŘÍKLAD TŘETÍ** *Hlášení vybraných výskytů*: Vymyslete algoritmy, které dostanou na vstupu slova  $s_1, s_2, \dots, s_k$  a text  $T$ . Výstupem algoritmu pak je pro každou pozici,

1. na které **končí** nějaké slovo: index pozice a **nejdelší** slovo, které na ní **končí**,
2. na které **končí** nějaké slovo: index pozice a **nejkratší** slovo, které na ní **končí**,
3. na které **začíná** nějaké slovo: index pozice a **nejdelší** slovo, které na ní **začíná**,
4. na které **začíná** nějaké slovo: index pozice a **nejkratší** slovo, které na ní **začíná**.

[4 body]

---

Nezapomeňte, že se hodnotí i srozumitelnost řešení a že správné řešení (pro příklady 2 a 3) obsahuje:

1. slovně dostatečně detailně popsaný algoritmus (volitelně pseudokód),
2. důkaz správnosti,
3. časovou složitost se zdůvodněním,
4. prostorovou složitost se zdůvodněním.