

1. Najděte v textu  $T$  všechny výskyty řetězců, které mají od vzorku  $P$  Hammingovu vzdálenost rovnou nejvýše  $k$ . Použijte metodu dynamického programování ([TSA] str. 199 – 202).

a)  $T = ccacbaabccaccbcabccc$ ,  $P = abcba$ ,  $k = 2$

b)  $T = 000111011000101010111110$ ,  $P = 110010$ ,  $k = 3$

2. Ukažte, že Levenshteinova vzdálenost splňuje axiomy metriky.

3. Napište všechna slova, která mají od vzorku  $aba$  nad abecedou  $\{a, b, c\}$  Levenshteinovu vzdálenost rovnu a) 1 b) 2

4. Abeceda  $A$  obsahuje  $n$  symbolů, slovo  $w$  má délku  $m$ . Odhadněte shora počet slov nad abecedou  $A$ , která mají od  $w$  Hammingovu vzdálenost rovnu  $k$  ( $0 \leq k \leq m$ ).

5. Abeceda  $A$  obsahuje  $n$  symbolů, slovo  $w$  má délku  $m$ . Odhadněte shora počet slov nad abecedou  $A$ , která mají od  $w$  Levenshteinovu vzdálenost rovnu  $k$  ( $0 \leq k \leq m$ ).

6. Najděte v textu  $T$  všechny výskyty řetězců, které mají od vzorku  $P$  Levenshteinovu vzdálenost rovnou nejvýše  $k$ . Použijte metodu dynamického programování ([TSA] str. 202 – 205).

a)  $T = aacacacbaabbbcbccacc$ ,  $P = abbcba$ ,  $k = 2$

b)  $T = 010011101000010101011100$ ,  $P = 11100$ ,  $k = 1$

7. Sestrojte deterministický automat nad abecedou  $A$ , který přijímá právě množinu  $M$  slov nad touto abecedou.

a)  $A = \{a, b, c\}$ ,  $M = \{a, b, ba, bc, aaa, bab, ccc, abbc, abcc\}$ .

b)  $A = \{0, 1\}$ ,  $M = \{10, 11, 101, 111, 1011, 1101, 10001, 10011, 10111, 11101, 11111\}$ .

8. Sestrojte deterministický automat, který v textu nad abecedou  $A$  vyhledá právě každé slovo množiny  $M$  z předchozí úlohy.

9. Sestavte automat, který v textu nad abecedou  $A$  vyhledává všechna slova popsaná regulárním výrazem  $R$ .

a)  $A = \{a, b, c\}$ ,  $R = c^*(ac + bb)^*$

b)  $A = \{0, 1\}$ ,  $R = 0^*(101 + 11)^*0$

10. Demonstrujte fakt, že paměťová složitost deterministického automatu pro vyhledávání slov odpovídajících regulárnímu výrazu může růst exponenciálně s počtem přijímaných slov. Použijte regulární výraz  $R = a(a+b)^{m-1}$  ( $m \geq 1$ ).

11. Sestavte tabulky pro simulaci činnosti vyhledávacího automatu metodou bitového paralelismu pro daný text  $t$ , vzorek  $p$  a Hammingovu vzdálenost  $k$

a)  $t = abcbcaaccbbaa$ ,  $p = bbac$ ,  $k = 2$

b)  $t = accbbaaabcba$ ,  $p = acbb$ ,  $k = 2$