

A string pattern algorithm

Henrik Andersen Sveinsson

November 10, 2013

1 Algoritmen

Vi ser på et mønster $M = (m_1 \dots m_n)$ av lengde n , og leter etter dette i en string S , der den i 'te bokstaven i S er s_i . Match tilsvarende for ordens skyld et sett $(s_i \dots s_{i+n}) = M$.

1. Se på bokstav nummer n i strengen. Dersom denne er inneholdt i M , er d_1 avstanden fra n til siste forekomst av s_n . Dersom den ikke er inneholdt, settes $d_1 = n$.
2. Dersom $d_1 = 0$, se på s_{n-1} , og gjør den samme vurderingen som i forrige instruks.

2 Bad character shift

Vi bestemmer 3 ting:

1. Scan mønster fra venstre til høyre
2. $R(x)$ er posisjonen til den mest høyre forekomst av x i P .
3. i er posisjonen vi står på i mønsteret, målfra høyre til venstre

Dersom første mismatch er funnet i posisjon i , og $T(k)$ er bokstaven som mismatcher $P(i)$, så skift mønsteret med $\max(1, i - R(T(k)))$.