

Merkkijonomenetelmät (syksy 2008)

Harjoitus 1

Jarkko Toivonen ti 09.09. klo 10.15–12 CK111

- Luennolla esitetty KMP-algoritmi on itseasiassa Morris-Pratt-algoritmi. Varsinainen Knuth-Morris-Pratt-algoritmi eroaa siitä vain korjaussiihtymien osalta. Se käyttää seuraavaa *fail'*-taulukkoa:
 - $fail'[i] = k$, kun k on suurin luku, jolla $P[0 \dots k] = P[i - k \dots i]$ ja $P_k \neq P_i$.
 - Kuvaa, miten ja miksi *fail'*-taulukko eroaa *fail*-taulukosta.
 - Laske *fail*- ja *fail'*-taulukot hahmolle IKONITIKITIKO ja totea erot.
- Modifioi luennolla esitettyä KMP-algoritmin esiprosessointia (Algoritmi 1.1.3) laskemaan tehtävän 1 *fail'*-taulukko.
- BM-algoritmin molemmat heuristiikat antavat pienimmän siirtymän, joka täyttää tietyn ehdon. On myös mahdollista määritellä yhdistelmäheuristiikka, joka antaa pienimmän siirtymän, joka toteuttaa molemmat ehdot.
 - Anna esimerkki tilanteesta, jossa yhdistelmäheuristiikka antaa paljon pidemmän siirtymän kuin luentojen algoritmi (eli maksimi kahden peruseuristikan siirtymästä).
 - Onko yhdistelmäheuristiikassa mitään haittapuolia?
- Olkoon $\Sigma = \{A, B\}$, ja numeroidaan aakkoston Σ merkit seuraavasti:

A . . . 0

B . . . 1

Olkoon $q = 7$. Simuloi Karp-Rabin-algoritmin (Algoritmi 1.4.1) toimintaa hahmolle $P = BAAA$ ja tekstillä $S = ABAABAAABAABA$. Montako oikeaa/väärää täsmäystä löytyy? Montako merkkivertailua tehdään?

- Merkkijonot x ja y ovat toistensa *rotaatioita*, jos on olemassa jonot u ja v , joilla $x = uv$ ja $y = vu$. Esimerkiksi $abcde$ ja $deabc$ ovat toistensa rotaatioita. Hahmottele *linearisessa* ajassa toimiva algoritmi, joka tutkii, ovatko annetut jonot x ja y toistensa rotaatioita. (Vihje: käytä lineaarisessa ajassa toimivaa hahmonsovitusalgoritmia (esim. KMP) alirutiinina.)