Merkkijonomenetelmät (kevät 2005)

Harjoitus 1

Pasi Rastas   ma 7.2 klo 10.15–12   C221

1. Analysoi triviaalialgoritmin 1.0.1 keskimääräistä merkkivertailujen lukumäärää. Oletetaan, että aakkoston Σ kukin merkki esiintyy samalla todennäköisyydellä $1/\sigma$ jokaisessa P:n ja S:n positiossa.

2. (a) Laske KMP-algoritmin $fail$-taulukko hahmolle `IKONITIKITIKO`.
   (b) Esitä tulos myös jononsovitusautomaattina.
   (c) Simuloi hahmontunnistuksen etenemistä itse valitsemassasi tekstissä.

3. Luennolla esitetty KMP-algoritmi on itseasiassa Morris-Pratt-algoritmi. Varsinainen Knuth-Morris-Pratt-algoritmi eroaa siitä vain korjaussiirtymien osalta. Se käyttää seuraavaa $fail'$-taulukkoa:

   - $fail'[i] = k$, kun k on pienin luku, jolla $P[0 \ldots k) = P[i-k \ldots i)$ **ja** $P_k \neq P_i$.

   (a) Kuvaa, miten ja miksi $fail'$-taulukko eroaa $fail$-taulukosta.
   (b) Toista tehtävä 2(a) $fail'$-taulukolle ja totea erot.

4. Modifioi luennolla esitettyä KMP-algoritmin esiprosessointia laskemaan tehtävän 3 $fail'$-taulukko.

5. 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 *lineaarisessa* ajassa toimiva algoritmi, joka tutkii, ovatko annetut jonot x ja y toistensa rotaatioita. (Vihje: käytä lineaarisessa ajassa toimivaa hahmonsovitusalgoritmia (esim. KMP) alirutiinina.)