T-79.1001 **Kevät 2006**
Tietojenkäsittelyteorian perusteet T
Harjoitus 8
Demonstraatiotehtävien ratkaisut

Lemma (Säännöllisten kielten pumppauslemma). Olkoon A säännöllinen kieli. Tällöin on olemassa $n \geq 1$ siten, että kaikki A:n merkkijonot x , joiden pituus $|x| \geq n$ ovat ilmaistavissa muodossa $x = uvw$, missä $|uv| \leq n$, $|v| \geq 1$ ja merkkijonot muotoa uv^iw kuuluvat kieleen A kaikilla $i \geq 0$.
Kompaktimmin, painottaen pumppauslemman asettamia vaatimuksia, voimme kirjoittaa seuraavasti:

$\forall$ säännöllisille kielille A

$\exists n \geq 1$ s.e.

$\forall x \in A : |x| \geq n$

$\exists$ osiinjako $x = uvw$, missä $|uv| \leq n$ ja $|v| \geq 1$

$\forall i \geq 0 \; uv^iw \in A.$

Pumppauslemmaa voidaan käyttää hyväksi, kun halutaan osoittaa kieli L *ei-säännölliseksi.* Tehdään ensin vastaoletus, eli oletetaan L säännölliseksi kieleksi. Tavoitteena on päästä ristiriitaan tämän oletuksen kanssa seuraten pumppauslemman asettamia vaatimuksia säännöllisille kielille.
Pumppauslemmaa käytettäessä täytyy aina muistaa, että sillä voi osoittaa vain kielen epäsäännöllisyyden, ja sitä *ei* voi käyttää toiseen suuntaan. Esimerkiksi kieli

$$I = \{c^i a^n b^n \mid i > 0 \wedge n \geq 0\} \cup L(a^*b^*)$$

ei ole säännöllinen, mutta kaikki siihen kuuluvat sanat (tyhjää sanaa lukuunottamatta) voidaan osittaa pumppauslemman ehtojen mukaisesti. Näin ollen kieltä I ei voida suoraan todistaa epäsäännölliseksi, vaan todistuksessa täytyy käyttää apuna säännöllisten kielten sulkeumaominaisuuksia.
Jos halutaan osoittaa kieli säännölliseksi, voidaan muodostaa sen hyväksyvä äärellinen automaatti, sillä pätee: Kieli L on säännöllinen $\Leftrightarrow$ on olemassa äärellinen automaatti M, joka hyväksyy kielen L (merkitään $L(M) = L$).

4. **Tehtävä**: Osoita, että yhteydettömien kielten luokka on suljettu yhdiste-, katenaatio- ja sulkeumaoperaatioiden suhteen, so. jos kielet $L_1, L_2 \subseteq \Sigma^*$ ovat yhteydettömiä, niin samoin ovat myös kielet $L_1 \cup L_2$, L_1L_2 ja L_1^*.

   Vastaus: Olkoon L_1 ja L_2 yhteydettömiä kieliä. Tällöin on olemassa kieliopit $G_1 = (V_1, \Sigma_1, R_1, S_1)$ ja $G_2 = (V_2, \Sigma_2, R_2, S_2)$, siten, että $L(G_1) = L_1$ ja $L(G_2) = L_2$. Vaaditaan lisäksi, että $(V_1 - \Sigma_1) \cap (V_2 - \Sigma_2) = \emptyset$, eli kieliopeissa ei esiinny samoja välikkeitä. Koska kieliopin välikkeet voidaan tarvittaessa nimetä uudelleen, ei tämä aseta oleellista rajoitusta.

   Unioni: Olkoon S uusi välike ja $G = (V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, R_1 \cup R_2 \cup \{S \to S_1 \mid S_2\}, S\}$. Nyt $L(G) = L(G_1) \cup L(G_2) = L_1 \cup L_2$. Näin on, koska S:stä voidaan johtaa vain S_1 tai S_2, joista voidaan edelleen johtaa vain sanoja jotka kuuluvat jompaan kumpaan aiemmista kielistä (sääntöjen sekaannukselta vältytään, koska välikejoukot ovat pistevieraita).

   Katenaatio: Tällä kertaa uusi kielioppi $G = (V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, R_1 \cup R_2 \cup \{S \to S_1S_2\}, S\}$. Nyt $L(G) = L_1L_2$.

   Kleenen tähti: Tällä kertaa uusi kielioppi $G = (V_1 \cup \{S\}, \Sigma_1, R_1 \cup \{S \to \epsilon \mid SS_1\}, S\}$. Nyt $L(G) = L_1^*$

5. **Tehtävä**: Osoita, että yhteydettömien kielten luokka ei ole suljettu leikkausten eikä komplementtien suhteen. (*Vihje:* Esitä kieli $\{a^k b^k c^k \mid k \geq 0\}$ kahden yhteydettömän kielen leikkauksena.)

   Vastaus: Olkoon kieli $L = \{a^k b^k c^k \mid k \geq 0\}$. Opetusmonisteessa (s.75) on todistettu, että tämä kieli ei ole yhteydetön. Osoitetaan, että yhteydettömät kielet eivät ole suljettuja leikkauksen suhteen esittämällä L:n kahden yhteydettömän kielen leikkauksena.

   Olkoon $L_1 = \{a^i b^k c^k \mid i, k \geq 0\}$ ja $L_2 = \{a^k b^k c^i \mid i, k \geq 0\}$. Nyt sekä L_1 että L_2 ovat yhteydettömiä, mutta $L = L_1 \cap L_2$, joten yhteydettömät kielet eivät ole suljettuja leikkauksen suhteen.

   Tuloksesta seuraa suoraan se, että yhteydettömät kielet eivät voi olla suljettuja komplementin suhteen, sillä ne ovat suljettuja unionin suhteen ja DeMorganin sääntöjen perusteella $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$.

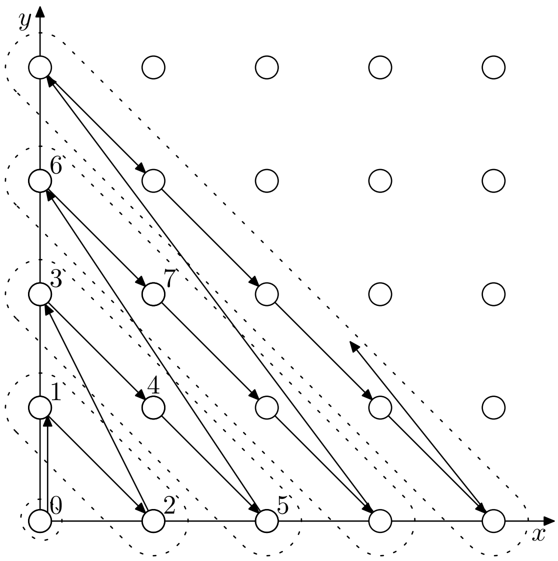
   Osoitetaan vielä lopuksi, että L_1 ja L_2 ovat todellakin yhteydettömiä muodostamalla niitä vastaavat kieliopit. Kielen L_1 generoi yhteydetön kielioppi $G_1 = (\{S, A, B, a, b, c\}, \{a, b, c\}, P_1, S)$, missä $P_1 = \{S \to AB, A \to aA \mid \varepsilon, B \to bBc \mid \varepsilon\}$.

   Kielen L_2 generoiva kielioppi $G_2 = (\{S, A, B, a, b, c\}, \{a, b, c\}, P_2, S)$, $P_2 = \{S \to AB, A \to aAb \mid \varepsilon, B \to cB \mid \varepsilon\}$.

6. **Tehtävä:** Osoita, että karteesinen tulo $\mathbb{N} \times \mathbb{N}$ on numeroituvasti ääretön. (*Vihje:* Ajattele parit $(m, n) \in \mathbb{N} \times \mathbb{N}$ sijoitetuiksi euklidiseen (x, y)-tasoon $\mathbb{R}^2$. Numeroi parit suoran $y = -x$ suuntaisin vinorivein.) Päättele tämän tuloksen perusteella, että myös rationaalilukujen joukko $\mathbb{Q}$ on numeroituvasti ääretön.

   Vastaus: Joukko S on numeroituvasti ääretön, mikäli voidaan muodostaa bijektio $f : \mathbb{N} \to S$. Intuitiivisesti tämä tarkoittaa sitä, että kaikille joukon S alkioille voidaan asettaa yksikäsitteinen järjestysnumero.

   Joukon $\mathbb{N} \times \mathbb{N}$ alkiot $(x, y) \in \mathbb{N} \times \mathbb{N}$ voidaan numeroida seuraavan kuvan mukaisesti:



   Ajatuksena on siis järjestää kaikki lukuparit suoran $y = -x$ suuntaisiin jonoihin, ja numeroida nämä jonot alkioittain yksi kerrallaan lyhyimmästä alkaen. Tässä numerointia ei voida suorittaa x-akselin suuntaisesti, sillä silloin kaikki indeksit kuluisivat jo y-akselin läpikäyntiin eikä yhtään lukuparia $(x, y), y > 0$ saavutettaisi koskaan.

   Ylläoleva numerointi voidaan määritellä seuraavasti:

$$f(x, y) = x + \sum_{k=1}^{x+y} k = x + \frac{(x + y)(x + y + 1)}{2}$$

Esimerkiksi $f(3,1) = 13$, eli lukuparin $(3,1)$ järjestysnumero on 13. Todettakoon vielä, että funktio $f(x,y)$ on todellakin bijektio, eli kutakin järjestysnumeroa vastaa yksikäsitteinen lukupari. Koordinaattiparin laskeminen annetusta indeksistä on kuitenkin hankalampaa, ja se jätetäänkin vastausten lopussa olevaan liitteeseen tiedoksi asiasta kiinnostuneille.

Positiiviset rationaaliluvut $\mathbb{Q}^+$ voidaan esittää $\mathbb{N} \times \mathbb{N}$ -lukuparina s.e.

$$(x,y) \equiv \frac{x}{y}, \qquad y \neq 0$$

Tämä on numeroituvasti äärettömän joukon $\mathbb{N}\times\mathbb{N}$ aito osajoukko, joten $\mathbb{Q}^+$ on joko äärellinen tai numeroituvasti ääretön. Jos $\mathbb{Q}^+$ olisi äärellinen, olisi olemassa joku rationaaliluku $\frac{x}{y}, x \in \mathbb{N}, y \in \mathbb{N}, y \neq 0$, jolla olisi suurin järjestysluku $n < \infty$ ($\mathbb{Q}^+$:n numeroinnissa). Kuitenkin voidaan aina löytää yo. kuvan perusteella rationaaliluku, jolla olisi järjestysluku $n' > n$, joten tämä on ristiriita oletuksen $\mathbb{Q}^+$ on äärellinen kanssa. Näin ollen $\mathbb{Q}^+$ on numeroituvasti ääretön.

Joukko $\mathbb{Q}^+$ voidaan laajentaa kaikkien rationaalilukujen joukoksi määrittelemällä joukko:

$$\mathbb{Q}^- = \{(-x,y) \mid (x,y) \in \mathbb{Q}^+\} \ .$$

Koska joukko $\mathbb{Q} = \mathbb{Q}^+ \cup \mathbb{Q}^-$ on kahden numeroituvasti äärettömän joukon yhdiste, on myös se numeroituvasti ääretön.