Tik-79.148 **Kevät 1998**
Tietojenkäsittelyteorian perusteet Laskuharjoitus 5
Ratkaisut

1. (2.3.7a / 2.5.1a)

   Konstruktion perusajatuksena on käydä systemaattisesti läpi kaikki mahdolliset reitit alkutilasta kaikkiin mahdollisiin lopputiloihin. Aluksi kaikki tilat numeroidaan yksikäsitteisesti, ja sen jälkeen muodostetaan reitti käyttäen rekursiivisia kaavoja.

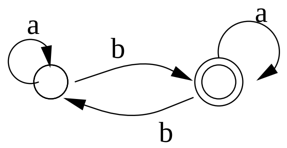
   Rekursiiviset säännöt ovat:

$$R(i,j,1) = \begin{cases} \{\sigma \in \Sigma | \delta(q_i,\sigma) = q_j\} & i \neq j \\ \{e\} \cup \{\sigma \in \Sigma | \delta(q_i,\sigma) = q_j\} & i = j \end{cases}$$
$$R(i,j,k+1) = R(i,j,k)R(k,k,k)^*R(k,j,k)$$

   Merkintä $R(i,j,k)$ tarkoittaa kaikkia niitä reittejä tilasta q_i tilaan q_j, jotka eivät käy välillä tilassa q_k tai sitä suuremmassa tilassa.

   Ensimmäinen säännöistä sanoo, että mikäli välillä ei saa käydä missään tilassa, ainoat reitit tilojen q_i ja q_j välillä ovat suorat transitiot.

   Säännöistä jälkimmäinen tarkoittaa, että reitit tilasta q_i tilaan q_j käymättä välillä tilassa q_{k+1}, tai sitä suuremmassa tilassa, voidaan jakaan kahteen ryhmään: niihin, joissa ei käydä edes tilassa q_k, ja niihin, joissa käydään tilassa q_k.



$$\begin{aligned} L(M) &= R(1,2,3) \\ R(1,2,3) &= R(1,2,2) \cup R(1,2,2)R(2,2,2)^*R(2,2,2) \\ R(1,2,2) &= R(1,2,1) \cup R(1,1,1)R(1,1,1)^*R(1,2,1) \\ R(2,2,2) &= R(2,2,1) \cup R(2,1,1)R(1,1,1)^*R(1,2,1) \end{aligned}$$

$$\begin{aligned} R(1,2,1) &= R(2,1,1) = b \\ R(1,1,1) &= R(2,2,1) = e \cup a \end{aligned}$$

Lopuksi sijoitetetaan saadut lausekkeet aikaisempiin yhtälöihin, ja sievennetään tulokset:

$$\begin{aligned}
R(1,2,2) &= b \cup (e \cup a)(e \cup a)^* b = b \\
R(2,2,2) &= (e \cup a) \cup b(e \cup a)^* b = e \cup a \cup ba^* b \\
R(1,2,3) &= a^* b + a^* b(e \cup a \cup ba^* b)^* (e \cup a \cup ba^* b) \\
&= a^* b(a \cup ba^* b)^*
\end{aligned}$$

2. (2.4.5 / 2.6.5)

   Tilakoneiden pumppauslemman mukaan äärettömälle säännölliselle kielelle L löytyy aina merkkijonot x, y ja z ($x, z \in \Sigma^*, y \in \Sigma\Sigma^*$) siten, että $xy^n z \in L, n \geq 0$.

   Toisin sanoen, mikäli säännöllinen kieli on ääretön, täytyy olla olemassa jokin merkkijono, jonka voi lisätä sanan keskelle mielivaltaisen monta kertaa.

   Tehtävässä piti todistaa, että $L = \{ww^R | w \in \{a,b\}^*\}$ ei ole säännöllinen. Määritellään kieli $L' = L \cap (ab)^*(ba)^*$. Koska säännöllisten kielten joukko on suljettu leikkauksen suhteen, on L säännöllinen vain jos L' on säännöllinen.

   Tarkastelemalla kieleen L kuuluvia sanoja huomataan, että $L' = (ab)^n(ba)^n$. Jotta L' voisi olla säännöllinen, pitäisi siis löytää y, joka voidaan lisätä sanaan mielivaltaisen monta kertaa.

   Nyt y voi olla muotoa:

   1° $y = (ab)^r$ tai $y = (ba)^r, r > 0$ Nyt kuitenkin toinen puolisko kasvaisi pidemmäksi kuin toinen, joten syntyneet sanat eivät kuulu enää kieleen.

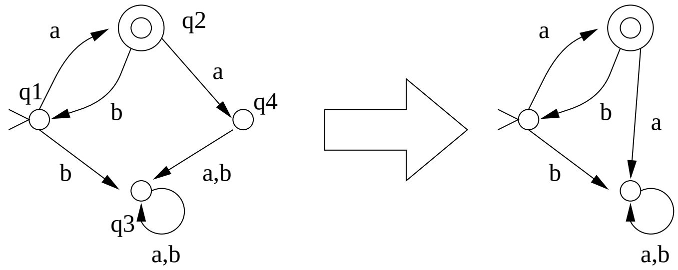
   2° $y = bb, y = aa, y = a(ba)^r$ tai $(ba)^r b, r \geq 0$ eivät myöskään kelpaa, sillä kielen sanoissa on aina yhtä monta a- kuin b-kirjainta.

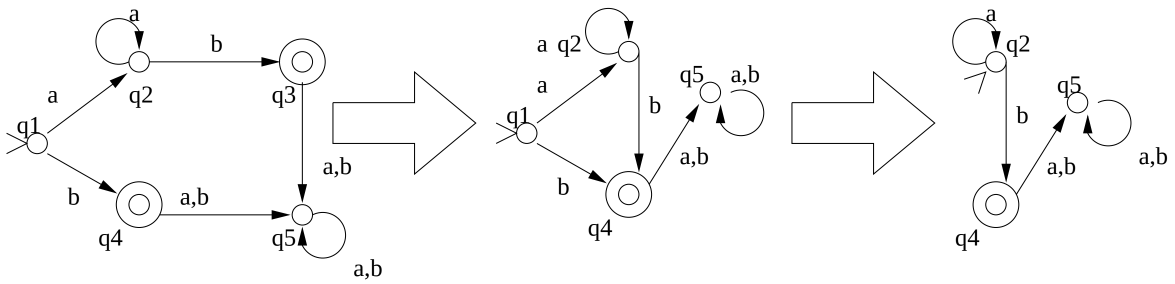
   3° $y = a(ba)^r bb(ab)^r b$ ei käy, koska tällöin sanassa esiintyy ba ennen viimeisiä ab:ta.

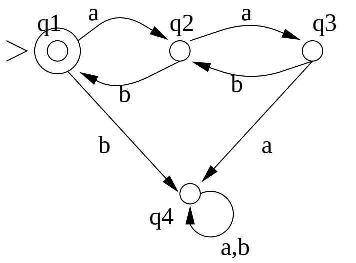
   Näin huomattiin, että mikään y:n arvo ei ole mahdollinen. L' ei siis ole säännöllinen, joten myöskään L ei voi olla säännöllinen.

3. (2.5.3 / -)

   Tilakonetta minimoitaessa etsitään koneesta ekvivalenttejen tilojen joukkoja, ja yhdistetään ne yhdeksi tilaksi. Kaksi tilaa on ekvivalentteja, mikäli niistä lähtevät kaaret vievät samoihin tiloihin ja mikäli joko molemmat kuuluvat lopputiloihin tai kumpikaan ei kuulu.







a) Tilojen q_3 ja q_4 molemmat kaaret menevät tilaan q_3. Lisäksi $q_3, q_4 \notin F$, joten ne voi yhdistää. Kaikki siirtymät, joilla ennen pääsi jompaan kumpaan tiloista siirretään osoittamaan yhdistettyyn tilaan.

b) Tiloista q_4 ja q_3 johtavat molemmat kaaret tilaan q_5, joten ne voi yhdistää. Yhdistämisen tekemisen jälkeen myös q_1 ja q_2 ovat muuttuneet ekvivalenteiksi, joten nekin voi yhdistää.

c) Tilakone C ei pienene lainkaan.

4. (3.1.2)

Yhteydettömässä kieliopissa $G = (V, \Sigma, R, S)$ V on kieleen kuuluvien merkkien joukko, $\Sigma \subset V$ on kielen terminaalisymbolien joukko. Kaikki kieleen kuuluvat sanat koostuvat pelkästään terminaalisymboleista, joukkoon $V - \Sigma$ kuuluvia symboleita (nonterminaaleja) käytetään pelkästään kielen sanojen johtamisessa. $R \in (V - \Sigma) \times V^*$ on joukko muunnossääntöjä. $S \in V - \Sigma$ on kielen aloitussymboli.

Tehtävän kielioppi on

$$
\begin{aligned}
V &= \{a, b, S, A\} \\
\Sigma &= \{a, b\} \\
R &= \{S \to aAa, S \to bAb, S \to e, A \to SS\}
\end{aligned}
$$

Sana *baabbb* voidaan johtaa tällä seuraavasti:

$$
\begin{aligned}
S &\to bAb \to bSSb \to baAaSb \to baSSaSb \\
&\to baSaSb \to baaSb \to baabAbb \\
&\to baabSSbb \to baabSbb \to baabbb
\end{aligned}
$$

Kielen sanoissa esiintyvät a ja b aina pareittain, siten että jokaisen parin sisällä voi olla muita pareja.

5. (3.1.3)

Yksinkertaisen ohjelmointikielen kuvaava yhteydetön kielioppi $G = (V, \Sigma, R, S)$ on:

$$
\begin{aligned}
\Sigma &= \{id, if, then, while, do, goto, label, begin, end :=, <, :, ;\} \\
V &= \Sigma \cup \{Stmt, Stmts, E\} \\
S &= Stmt
\end{aligned}
$$

Tässä kielen aakkosina ovat kaikki mahdolliset tokenit, millä tarkoitetaan yhteen kokonaisuuteen liittyviä merkkijonoja. Siten esim. kaikki mahdolliset muuttujannimet on sisällytetty tokeniin *id*.

Nonterminaaleista *Stmt* tarkoittaa yhtä lausetta, *Stmts* montaa lausetta ja E aritmeettista lauseketta. E:n kuvaava kielioppi on kirjan esimerkissä 3.1.3, ja sitä ei toisteta tässä.

$$
\begin{aligned}
R = \{&Stmt \to id := E \\
&Stmt \to if\ E < E\ then\ Stmt \\
&Stmt \to while\ E < E\ do\ Stmt \\
&Stmt \to goto\ label \\
&Stmt \to label : Stmt \\
&Stmt \to begin\ Stmts\ end \\
&Stmts \to Stmt\ ;\ Stmts \\
&Stmts \to Stmt\}
\end{aligned}
$$