T-79.148 **Syksy 2002**

Tietojenkäsittelyteorian perusteet

Harjoitus 5

Demonstraatiotehtävien ratkaisut

Säännölliset lausekkeet määritellään induktiivisesti:

- $\emptyset$ ja kaikki $a \in \Sigma$ ovat säännöllisiä lausekkeita.
- Mikäli α ja β ovat säännöllisiä lausekkeita, niin myös $(\alpha\beta)$, $(\alpha \cup \beta)$ ja α^* ovat säännöllisiä lausekkeita.
- Mitkään muut lausekkeet eivät ole säännöllisiä.

Säännöllisestä lausekkeesta $\emptyset^*$ käytetään lyhennysmerkintää ε (myös e ja λ esiintyvät kirjallisuudessa). Yleensä jätetään ylimääräiset sulut pois: $((a(bc)) \cup (cd^*)) = abc \cup cd^*$.
Säännöllinen lauseke α määrittelee kielen $L(\alpha)$, joka on joukon Σ^* osajoukko. Esimerkiksi:

$$L(ab \cup ac^*d) = \{ab, ad, acd, accd, accd, acccd, accccd, \dots\} \ .$$

Mikäli sekaantumisen vaaraa ei ole[1], jätetään $L()$ melkein aina pois säännöllisen lausekkeen ympäriltä, ja käytetään lauseketta itseään tarkoittamaan niiden sanojen joukkoa, jotka kuuluvat sen kuvaamaan kieleen.
Säännölliset lausekkeet on ehkä helpointa ajatella muottina. Sana kuuluu kieleen, jos sen saa sovitettua lausekkeen antamaan muottiin:

- $ab \Rightarrow$ otetaan ensin a, sitten b
- $a \cup b \Rightarrow$ otetaan joko a tai b
- $a^* \Rightarrow$ otetaan kuinka monta a:ta tahansa (myös 0).

4. **Tehtävä**: Sievennä seuraavia säännöllisiä lausekkeita (so. konstruoi yksinkertaisemmat lausekkeet samojen kielten kuvaamiseen):

   1. $(\emptyset^* \cup a)(a^*)^*(b \cup a)b^*$
   2. $(a \cup b)^* \cup \emptyset \cup (a \cup b)b^*a^*$
   3. $a(b^* \cup a^*)(a^*b^*)^*$

   Vastaus:

   (a)

$$\begin{aligned}(\emptyset^* \cup a)(a^*)^*(b \cup a)b^* =& (\varepsilon \cup a)(a^*)^*(b \cup a)b^* \\ =& (\varepsilon \cup a)a^*(b \cup a)b^* \\ =& a^*(b \cup a)b^*\end{aligned}$$

   (b)

$$(a \cup b)^* \cup \emptyset \cup (a \cup b)b^*a^* = (a \cup b)^*$$

   Sievennyksen tuloksen näkee suoraan siitä, että $L((a \cup b)^*) = \Sigma^*$, joten jo ensimmäinen alilauseke generoi kaikki aakkoston merkkijonot.

   (c)

$$a(b^* \cup a^*)(a^*b^*)^* = a(a \cup b)^*$$

   Tässä huomataan, että alilauseke $R_2 = (a^*b^*)^*$ generoi kaikki merkkijonot, jotka $R_1 = (b^* \cup a^*)$ generoi, joten R_1 voidaan poistaa.

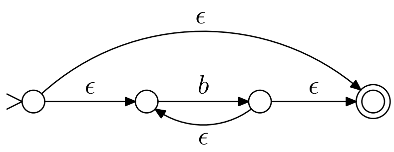
[1] Ja valitettavan usein myös silloin, kun sekaantumisen vaara on olemassa.

5. **Tehtävä**: Ratkaise, kuvaavatko säännölliset lausekkeet $r_1 = b^*a(a^*b^*)^*$ ja $r_2 = (a \cup b)^*a(a \cup b)^*$ saman kielen muodostamalla lausekkeita vastaavat (minimaaliset) deterministiset tilakoneet.

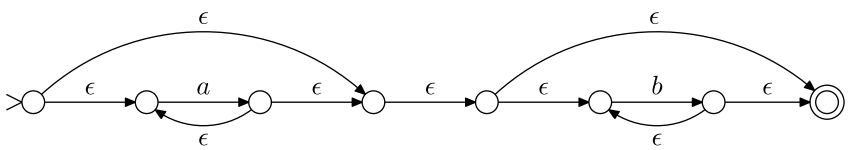
Vastaus: Kutakin säännöllistä kieltä vastaa tilojen nimeämistä lukuunottamatta yksikäsitteinen minimaalinen deterministinen automaatti. Mikäli kahta säännöllistä lauseketta vastaavat minimiautomaatit ovat samat, ovat ne ekvivalentteja.

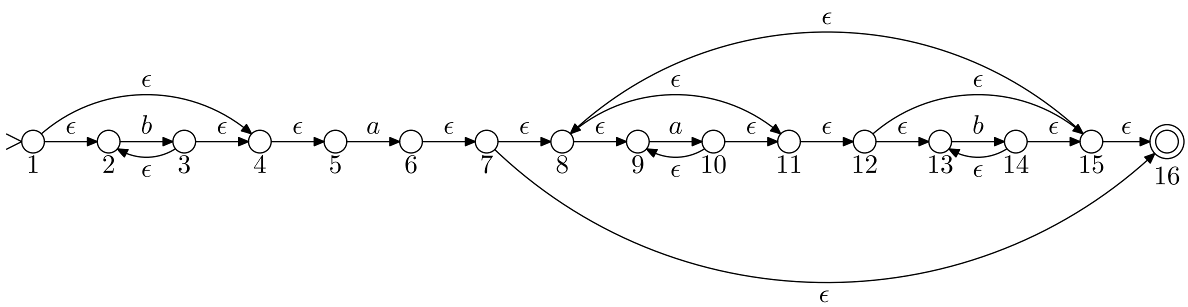
Rakennetaan vaiheittain säännöllistä lauseketta r_1 vastaava tilakone.

b^*:

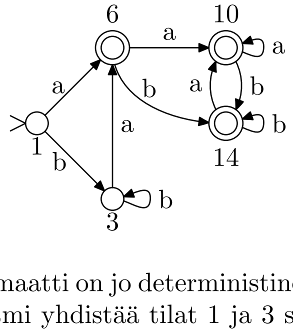


a^*b^*:

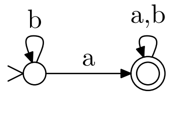




Poistetaan automaatista tyhjät siirtymät:

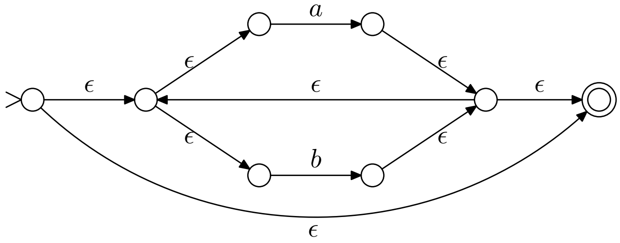


Huomataan, että syntynyt automaatti on jo deterministinen, joten voidaan siirtyä suoraan minimointiin. Minimointialgoritmi yhdistää tilat 1 ja 3 sekä tilat 6, 10 ja 14. Näin ollen pienimmäksi lausekkeen r_1 tunnistavaksi tilakoneeksi saadaan:

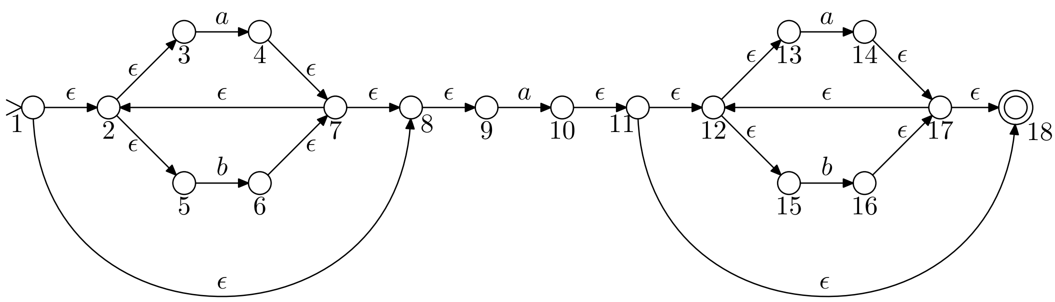


Rakennetaan vaiheittain r_2:a vastaava automaatti:

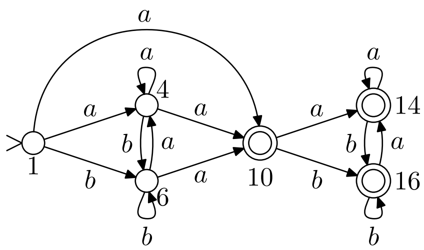
$(a \cup b)^*$



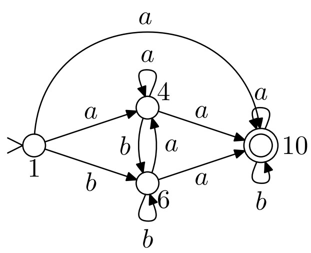
$(a \cup b)^* a (a \cup b)^*$



Poistetaan tyhjät siirtymät:

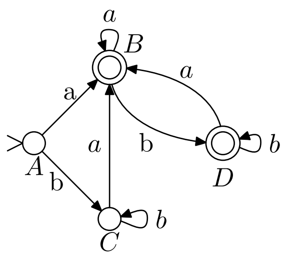


Huomataan, että konetta voidaan sieventää yhdistämällä kaikki lopputilat yhdeksi tilaksi:

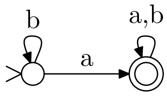


Determinisoidaan tämä tilakone:

det. tila	vast. tilat	a	b	
A	$\{1\}$	$\{4, 10\}$	$\{6\}$	
B	$\{4, 10\}$	$\{4, 10\}$	$\{6, 10\}$	$\times$
C	$\{6\}$	$\{4, 10\}$	$\{6\}$	
D	$\{6, 10\}$	$\{4, 10\}$	$\{6, 10\}$	$\times$



Minimointi muuttaa automaatin muotoon:



Koska tulokseksi saatiin sama tilakone kummallekin lausekkeelle, r_1 ja r_2 ovat ekvivalentit.

6. **Tehtävä**: Olkoon L säännöllinen kieli. Osoita, että kieli Ł$' = \{xy \mid x \in L, y \notin L\}$ on säännöllinen.

   Vastaus: Kielen todistaminen säännölliseksi onnistuu yleensä helpoiten käyttämällä kieliperheen sulkeumaominaisuuksia: säännöllisten kielten joukko on suljettu unionin, katenaation, Kleenen tähden, komplementoinnin ja leikkauksen suhteen.

   Tehtävässä on annettuna jokin säännöllinen kieli L, jonka avulla muodostetaan kieli

   $$L' = \{xy \mid x \in L \text{ ja } y \notin L\}$$

   Kieli L' on muodostettu katenaation avulla kielestä L ja sen komplementista $\overline{L}$ ($y \notin L \Rightarrow y \in \overline{L}$). Koska säännöllisten kielten joukko on suljettu katenaation ja komplementin suhteen ja L on säännöllinen, on myös L' säännöllinen.

   Toinen tapa esittää todistus olisi rakentaa kielen L' hyväksyvä tilakone. Koska L on säännöllinen, on olemassa deterministinen tilakone M, joka hyväksyy sen. Muodostetaan nyt deterministinen tilakone $\overline{M}$ siten, että se on muuten samanlainen kuin M, mutta hyväksyvien ja hylkäävien lopputilojen joukot vaihdetaan keskenään, eli $\overline{M}$ hyväksyy kaikki ne sanat, jotka M hylkää ja päinvastoin. Yhdistetään nyt nämä tilakoneet epädeterministiseksi koneeksi M', siten, että kaikista M:n lopputiloihin lisätään tyhjä siirtymä $\overline{M}$:n alkutilaan. Nyt M' hyväksyy kielen L', joten L' on säännöllinen.