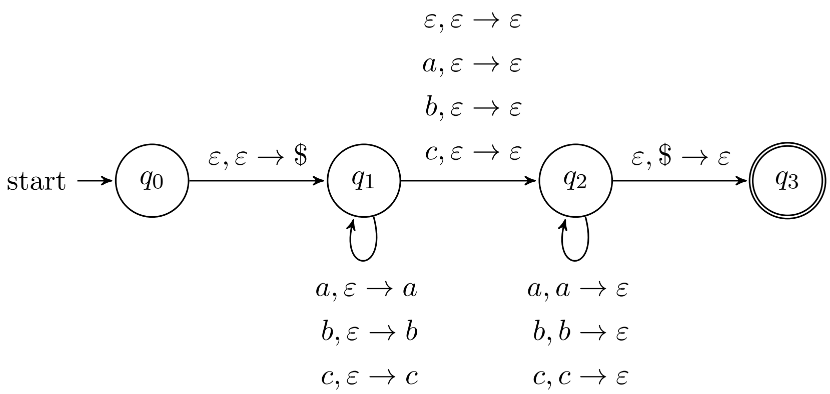
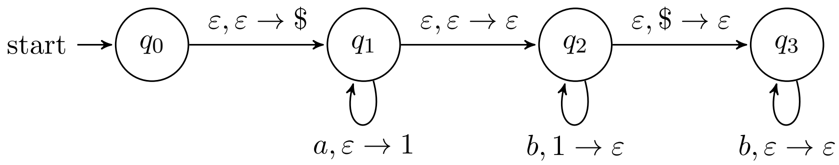
582206 Laskennan mallit, syksy 2012
7. harjoitusten malliratkaisut
Juhana Laurinharju ja Jani Rahkola

1. Esitä pinoautomaatti seuraaville kielille.

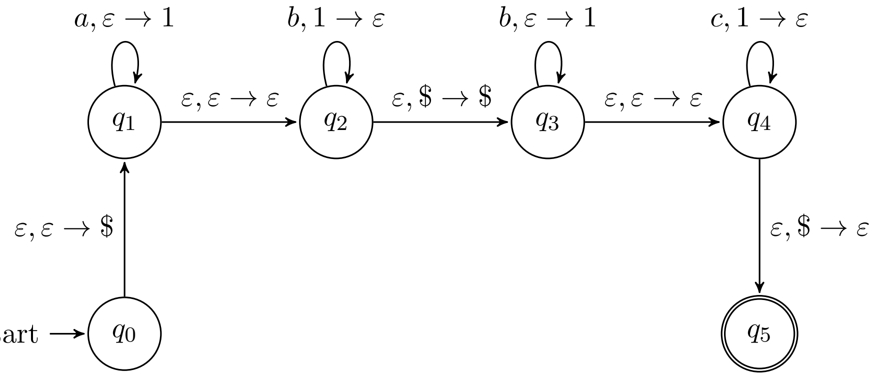
   (a) Kaikki palindromit aakkostosta $\Sigma = \{a, b, c\}$.



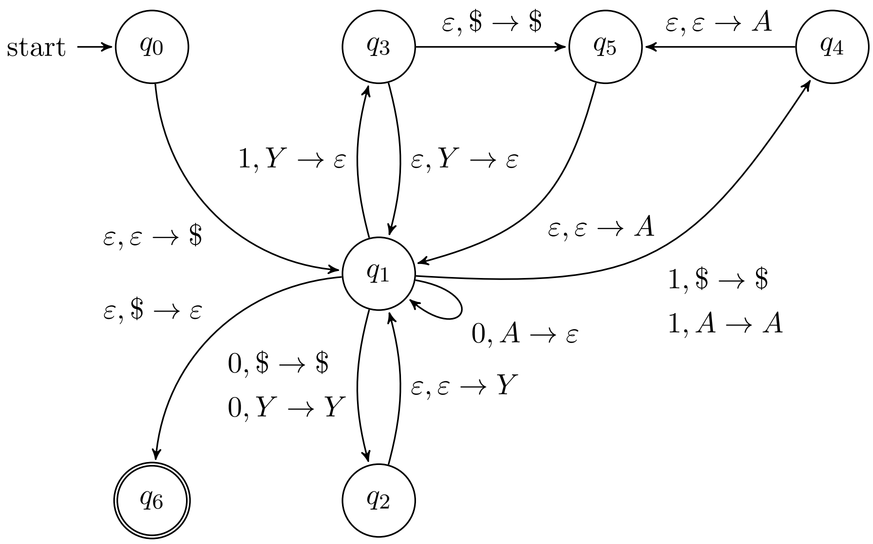
   (b) $\left\{a^i b^j \mid 0 \le i \le j\right\}$ missä $\Sigma = \{a, b, c\}$



   (c) $\left\{a^i b^j c^k \mid j = i + k\right\}$ missä $\Sigma = \{a, b, c\}$



   (d) Kaikki aakkoston $\Sigma = \{0, 1\}$ merkkijonot joissa nollia on kaksi kertaa niin paljon kuin ykkösiä.



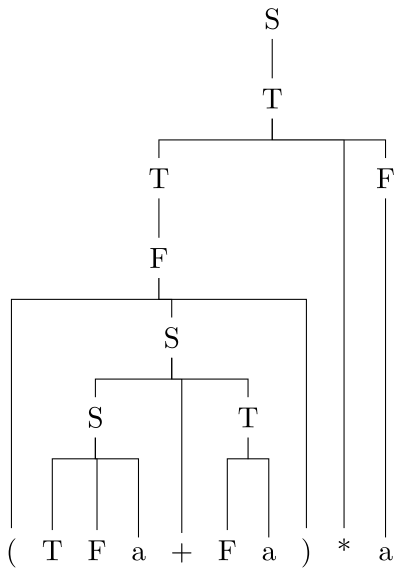
Automaatin ideana on pitää pinossa kirjaa siitä kuinka paljon nolla-merkkien määrässä on yli- tai alijäämää. Pinoon laitetaan A jos siirtymä kerryttää alijäämää ja Y jos se kerryttää ylijäämää. Vastaavasti pinosta poistetaan merkkejä aina tilaisuuden tullen, kun syötemerkin lukeminen tasoittaa nollien ja ykkösten suhdetta.

2. Tarkastellaan kielioppia

$$\begin{aligned} S &\to S + T \mid T \\ T &\to T * F \mid F \\ F &\to (S) \mid a \end{aligned}$$

Muodosta merkkijonon $s = (a + a) * a$ jäsennyspuu tämän kieliopin mukaisesti.

Etsi jäsennyspuusta jokin juuresta lehteen johtava polku, jolla sama muuttuja esiintyy kahdessa solmussa. Muodosta tämän perusteella toistuvuusominaisuuden todistuksen ideaa mukaillen jokin merkkijonon s jako osiin $s = uvxyz$, joilla merkkijono uv^ixy^iz kuuluu tarkasteltavaan kieleen kaikilla $i \in N$.



3. Olkoon A aakkoston $\{0, 1\}$ kieli, joka koostuu niistä merkkijonoista, joissa on sama määrä nollia ja ykkösiä. Tällä kielellä on kontekstiton kielioppi

$$S \to SS \mid 0S1 \mid 1S0 \mid \varepsilon$$

(a) Kielen A eräs toistuvuuspituus on 4. Esitä kieleen A kuuluvalle merkkijonolle $s = 001101$ kaikki eri tavat jakaa se osiin $s = uvxyz$ toistuvuusominaisuuden ehdot toteuttavalla tavalla (lause 2.30; Sipser Theorem 2.34; tässä siis $p = 4$).

u	v	x	y	z
			0011	01
		0	01	101
	0		011	01
	0	0	1	101
	0	01	1	01
	00		11	01
	001		1	01
	0011			01
0			01	101
0			0110	1
0		01	10	1
0	0		1	101
0	0		110	1
0	0	1	1	01
0	01			101
0	01		10	1
0	01	1		01
0	01	10		1
0	011		0	1
0	0110			1
00		1	10	1
00		11	01	
00	1	1	0	1
001			10	1
001		1	01	
001	1		0	1
001	10			1
001	10	1		
0011			01	
0011	0		1	
0011	01			

Yhteensä 31 ehdot täyttävää jakoa.

(b) Onko kielellä A pienempiä toistuvuuspituuksia kuin 4? Perustele.

4. (a) Koostukoon aakkoston $\{a, b, c\}$ kieli A merkkijonoista, joissa on yhtä monta a-, b- ja c-merkkiä.

Väite. *Kieli A ei ole yhteydetön.*

Todistus. Oletetaan vastoin väitettä, että kieli A on yhteydetön. Nyt tehtävässä 7 todistetun nojalla leikkaus $A \cap L(a^*b^*c^*) = \{a^nb^nc^n \mid n \in \mathbb{N}\}$ on yhteydetön. Tämä on kuitenkin tunnetusti ristiriita. Siis kieli A ei ole yhteydetön. □

(b) Osoita, että kieli $\{0^n1^n0^n1^n \mid n \in \mathbb{N}\}$ ei ole yhteydetön.

Väite. *Kieli $A = \{0^n1^n0^n1^n \mid n \in \mathbb{N}\}$ ei ole yhteydetön.*

Todistus. Oletetaan vastoin väitettä, että kieli A on yhteydetön. Nyt pumppauslemman nojalla sillä on pumppauspituus p. Valitaan merkkijono $s = 0^p1^p0^p1^p$. Nyt merkkijonon jakojen $s = uvxyz$ tulisi täyttää pumppauslemman ehdot.

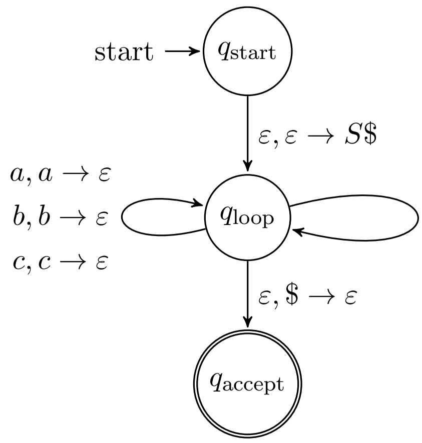
Merkkijonon s pituus on $4p$ ja koska osaan vxy ei saa tulla yli p:tä merkkiä, täytyy osiin u ja z jäädä yhteensä vähintään $3p$ merkkiä. Nyt siis joko osaan u jäävät ainakin kaikki merkkijonon p esimmäistä nollaa, tai osaan z jäävät ainakin kaikki merkkijonon p viimeistä ykköstä.

Jos kieleen A kuuluvassa merkkijonossa on p kappaletta peräkkäisiä nollia tai p kappaletta peräkkäisiä ykkösiä, täytyy merkkijonon pituuden olla $4p$. Pumppauslemman mukaan merkkijonon $uv^0xy^0z = uxz$ tulisi kuulua kieleen A. Koska osat u ja z pitävät sisällään joko p nollaa tai p ykköstä, tulisi merkkijonon uxz pituuden olla edelleen $4p$ jotta se voisi kuulua kieleen A. Koska puuttuvat osat v ja y eivät kuitenkaan saaneet molemmat olla tyhjiä, täytyy merkkijonon uxz pituuden olla aidosti vähemmän kuin $4p$. Siis uxz ei kuulu kieleen A, mikä on ristiriita. Täten kieli A ei ole yhdeydetön. $\square$

5. Anna yhteydetön kielioppi, joka tuottaa kielen $\{a^i b^j c^k \mid i = 2j \text{ tai } j = 2k\}$. Muodosta apulauseen 2.21 mukaisesti kieliopistasi pinoautomaatti, joka tunnistaa saman kielen.

$$
\begin{aligned}
S &\to T_{aab}T_c \mid T_aT_{bbc} \\
T_{aab} &\to aaT_{aab}b \mid \varepsilon \\
T_c &\to cT_c \mid \varepsilon \\
T_a &\to aT_a \mid \varepsilon \\
T_{bbc} &\to bbT_{bbc}c \mid \varepsilon
\end{aligned}
$$

Jos automaateissa saisi laittaa monta aakkosta pinoon kerralla, näyttäisi automaatti seuraavalta:

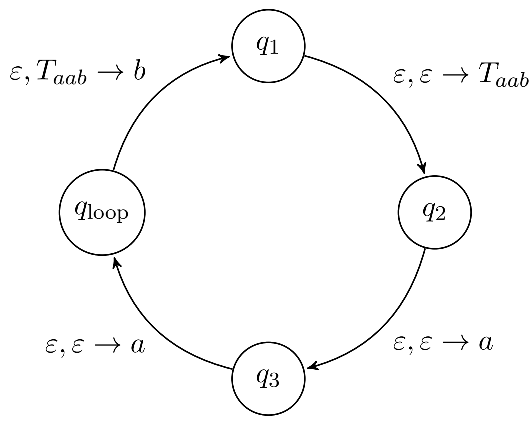


ε, S	$\to T_{aab}T_c$
ε, S	$\to T_aT_{bbc}$
ε, T_{aab}	$\to aaT_{aab}b$
ε, T_{aab}	$\to \varepsilon$
ε, T_c	$\to cT_c$
ε, T_c	$\to \varepsilon$
ε, T_a	$\to aT_a$
ε, T_a	$\to \varepsilon$
ε, T_{bbc}	$\to bbT_{bbc}c$
ε, T_{bbc}	$\to \varepsilon$

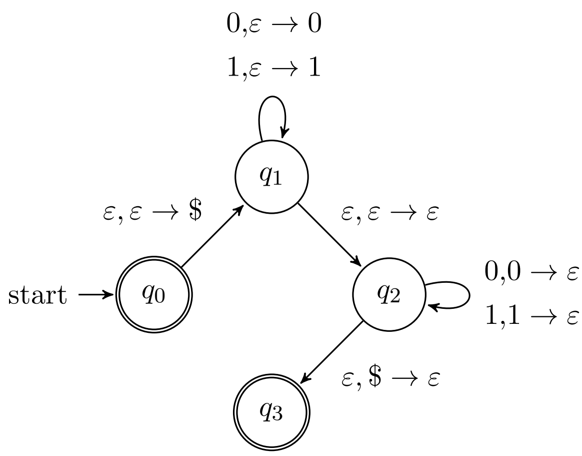
Jokainen sääntö, jossa pinoon lisätään monta symbolia kerralla voidaan avata silmukaksi. Esimerkiksi sääntö

$$\varepsilon, T_{aab} \to aaT_{aab}b$$

voidaan toteuttaa tavallisella pinokoneella seuraavasti:



6. Tee alla olevasta pinoautomaatista Apulauseen 2.27 mukaisesti kielioppi.



7. (a) Osoita, että jos A on yhteydetön ja B säännöllinen kieli, niin $A \cap B$ on yhteydetön.

Vihje: muodosta pinoautomaatin ja äärellisen automaatin leikkausautomaatti samaan tapaan kuin Jyrkin luentojen lauseessa 1.1 (luentomateriaalin sivut 48–50).

Olkoon A yhteydetön kieli ja $M_A = (Q_A, \Sigma, \Gamma, \delta_A, q_{A0}, F_A)$ automaatti joka tunnistaa kielen A. Olkoon B säännöllinen kieli ja $M_B = (Q_B, \Sigma, \delta_B, q_{B0}, F_B)$ deterministinen automaatti joka tunnistaa kielen B.

Väite. *Kieli $A \cap B$ on säännöllinen.*

Todistus. Leikkauksen tunnistava automaatti luodaan samankaltaisella menetelmällä kuin säännöllisten kielten tapauksessa. Ero säännöllisten kielten tapaukseen on siirtymäfunktion $\delta_{A\cap B}$ määrrittelyssä.

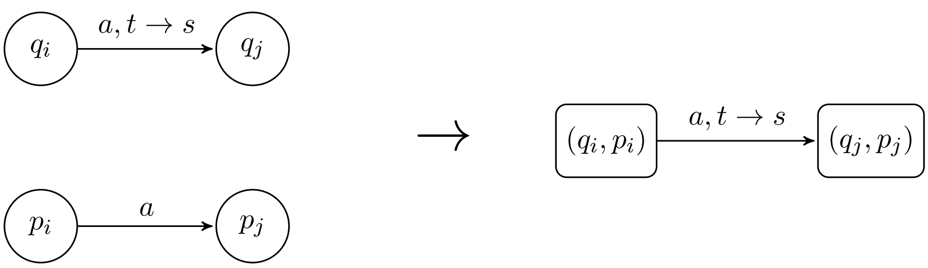
Muodostetaan siis automaatti

$$M_{A\cap B} = (Q_A \times Q_B,\ \Sigma,\ \Gamma,\ \delta_{A\cap B},\ (q_{A0}, q_{B0}),\ F_A \times F_B)$$

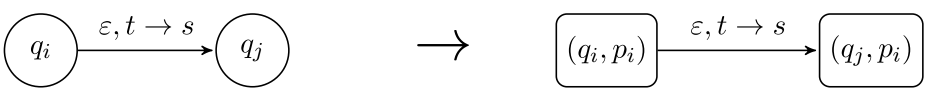
missä siirtymäfunktio $\delta_{A\cap B}$ on määritelty seuraavasti.

$$\delta_{A\cap B}((q_i, p_i), a, t) = \begin{cases} \{((q_j, p_i), s) \mid \delta_A(q_i, \varepsilon, t) = (q_j, s)\} & \text{kun } a = \varepsilon \\ \{((q_j, p_j), s) \mid \delta_B(q_i, a) = q_j \text{ ja } \delta_A(p_i, a, t) = (p_j, s)\} & \text{muulloin} \end{cases}$$

Kaikki uuden automaatin tilat ovat siis muotoa (q, p) missä $q \in Q_A$ ja $p \in Q_B$. Siirtymät noudattavat parin ensimmäisen alkion kohdalla automaatin M_A siirtymäfunktiota ja toisen alkion kohdalla automaatin M_B siirtymäfunktiota. Pinon käsittely noudattaa aina automaatin M_A siirtymäfunktiota, sillä automaatissa M_B ei ole pinoa.



Pinoautomaatti M_A on epädeterministinen, mutta M_B ei. Pinoautomaatin epädeterminististen siirtymien kohdalla uudessa automaatissa tilaparin jälkimmäinen alkio ei muutu. Ensimmäinen alkio noudattaa pinoautomaatin M_A siirtymäfunktiota.



Luotu automaatti $M_{A\cap B}$ hyväksyy merkkijonon w jos ja vain jos M_A ja M_B hyväksyvät merkkijonon w. Siis $M_{A\cap B}$ tunnistaa kielen $A \cap B$. $\square$

(b) Tiedetään, että kieli L on yhteydetön ja R säännöllinen. Voidaanko tästä päätellä, että $L - R$ on yhteydetön? Entä $R - L$? Perustele.

Väite. *Olkoon L yhteydetön ja R säännöllinen kieli. Nyt $L - R$ on yhteydetön.*

Todistus. Joukko-opista tiedämme, että $L - R = L \cap \overline{R}$. Lisäksi tiedämme, että säännölliset kielet ovat suljettuja komplementin suhteen. Nyt siis edellisen kohdan nojalla $L \cap \overline{R}$ on yhteydetön, ja siten myös $L - R$ on yhteydetön. $\square$

Toinen suunta ei päde yleisesti. Koska yhteydettömät kielet eivät ole suljettuja komplementin suhteen, on olemassa yhteydetön kieli jonka komplementti ei ole yhteydetön. Olkoon L jokin tällainen kieli. Olkoon nyt $R = \Sigma^*$ joka tunnetusti säännöllinen. Nyt siis L on yhteydetön ja R säännöllinen, mutta $R - L = \Sigma^* - L = \overline{L}$ joka oletuksen mukaan ei ole yhteydetön.