2.4 Äärellisten automaattien minimointi

Voidaan osoittaa, että jokaisella äärellisellä automaatilla on yksikäsitteinen ekvivalentti (so. saman kielen tunnistava) tilamäärältään minimaalinen automaatti.

Annetun äärellisen automaatin kanssa minimointi (ekvivalentin minimiautomaatin määrittäminen) on sekä käytännössä että teoreettiselta kannalta tärkeä tehtävä: siten voidaan esimerkiksi selvittää, tunnistavatko kaksi annettua automaattia saman kielen.

Tehtävä voidaan ratkaista seuraavassa esitettävällä tehokkaalla menetelmällä. Menetelmän perusideana on pyrkiä samaistamaan keskenään sellaiset syötteenä annetun automaatin tilat, joista lähtien automaatti toimii täsmälleen samoin kaikilla merkkijonoilla.

Olkoon

$$M = (Q, \Sigma, \delta, q_0, F)$$

äärellinen automaatti.

Laajennetaan M:n siirtymäfunktio yksittäisistä syötemerkeistä merkkijonoihin: jos $q \in Q$, $x \in \Sigma^*$, merkitään

$$\delta^*(q, x) = \text{ se } q' \in Q, \text{ jolla } (q, x) \underset{M}{\vdash}^* (q', \varepsilon).$$

M:n tilat q ja q' ovat *ekvivalentit*, merkitään

$$q \equiv q',$$

jos kaikilla $x \in \Sigma^*$ on

$$\delta^*(q, x) \in F \quad \text{jos ja vain jos} \quad \delta^*(q', x) \in F;$$

toisin sanoen, jos automaatti q:sta ja q':sta lähtien hyväksyy täsmälleen samat merkkijonot.

Lievempi ekvivalenssiehto: tilat q ja q' ovat k-*ekvivalentit*, merkitään

$$q \overset{k}{\equiv} q',$$

jos kaikilla $x \in \Sigma^*$, $|x| \leq k$, on

$$\delta^*(q, x) \in F \quad \text{jos ja vain jos} \quad \delta^*(q', x) \in F;$$

toisin sanoen, jos mikään enintään k:n pituinen merkkijono ei pysty erottamaan tiloja toisistaan.

Ilmeisesti on:

$$\begin{array}{lll} \text{(i)} \quad q \overset{0}{\equiv} q' & \text{joss} & \text{sekä } q \text{ että } q' \text{ ovat lopputiloja} \\ & & \text{tai kumpikaan ei ole; ja} \\ \text{(ii)} \quad q \equiv q' & \text{joss} & q \overset{k}{\equiv} q' \text{ kaikilla } k = 0, 1, 2, \ldots \end{array} \qquad (1)$$

Esitettävä minimointialgoritmi perustuu syötteenä annetun automaatin tilojen k-ekvivalenssiluokkien hienontamiseen $(k+1)$-ekvivalenssiluokiksi kunnes saavutetaan täysi ekvivalenssi.

Algoritmi MIN-FA [Äärellisen automaatin minimointi]

Syöte: Äärellinen automaatti $M = (Q, \Sigma, \delta, q_0, F)$.

Tulos: M:n kanssa ekvivalentti äärellinen automaatti $\widehat{M}$, jossa on minimimäärä tiloja.

Menetelmä:

1. [Turhien tilojen poisto.] Poista M:stä kaikki tilat, joita ei voida saavuttaa tilasta q_0 millään syötemerkkijonolla.

2. [0-ekvivalenssi.] Osita M:n jäljelle jääneet tilat kahteen luokkaan: ei-lopputiloihin ja lopputiloihin.

3. [k-ekvivalenssi $\to$ $(k+1)$-ekvivalenssi.] Tarkasta, siirrytäänkö samaan ekvivalenssiluokkaan kuuluvista tiloista samoilla merkeillä aina samanluokkaisiin tiloihin. Jos kyllä, algoritmi päättyy ja minimiautomaatin $\widehat{M}$ tilat vastaavat M:n tilojen *luokkia*.

Muussa tapauksessa hienonna ositusta jakamalla kunkin äskeistä ehtoa rikkovan ekvivalenssiluokan tilat uusiin, pienempiin ekvivalenssiluokkiin sen mukaan, mihin luokkaan kustakin tilasta siirrytään milläkin aakkosella, ja toista kohta 3 uudella osituksella.

On helppo osoittaa, että askelen 3 $(k+1)$:nnen suorituskerran $(k = 0, 1, \dots)$ alussa kaksi tilaa kuuluu samaan muodostetun osituksen luokkaan, jos ja vain jos ne ovat k-ekvivalentteja.

Tästä seuraa, että algoritmin suorituksen päättyessä, kun ositus ei enää hienone, muodostuneet tilaluokat ovat täsmälleen M:n tilojen $\equiv$-ekvivalenssiluokat (vrt. ominaisuus (1.ii)).

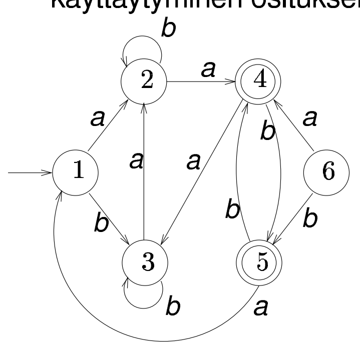
Algoritmin suoritus päättyy välttämättä aina, sillä kullakin askelen 3 suorituskerralla, viimeistä lukuunottamatta, vähintään yksi tilaluokka ositetaan pienemmäksi.

Lause 2.1 Algoritmi MIN-FA muodostaa annetun äärellisen automaatin M kanssa ekvivalentin äärellisen automaatin $\widehat{M}$, jossa on minimimäärä tiloja. Tämä automaatti on tilojen nimeämistä paitsi yksikäsitteinen. □

Esimerkki. Tarkastellaan seuraavan automaatin minimointia:

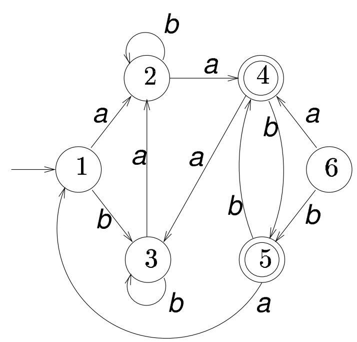
Vaiheessa 1 automaatista poistetaan tila 6, johon ei päästä millään merkkijonolla.

Vaiheessa 2 ositetaan automaatin tilat 1–5 ei-lopputiloihin (luokka I) ja lopputiloihin (luokka II), ja tarkastetaan siirtymien käyttäytyminen osituksen suhteen:



			a	b
I :	$\to$	1	2, I	3, I
		2	4, II	2, I
		3	2, I	3, I
II :	$\leftarrow$	4	3, I	5, II
	$\leftarrow$	5	1, I	4, II

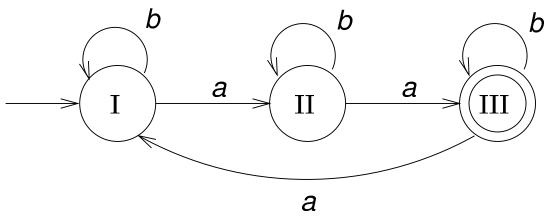
Luokassa I on nyt kahdentyyppisiä tiloja ({1,3} ja {2}), joten ositusta täytyy hienontaa ja tarkastaa siirtymät uuden osituksen suhteen:



			a	b
I :	$\to$	1	2, II	3, I
		3	2, II	3, I
II :		2	4, III	2, II
III :	$\leftarrow$	4	3, I	5, III
	$\leftarrow$	5	1, I	4, III

Nyt kunkin luokan sisältämät tilat ovat keskenään samanlaisia, joten minimointialgoritmi päättyy.

Saadun minimiautomaatin tilakaavio on seuraava:

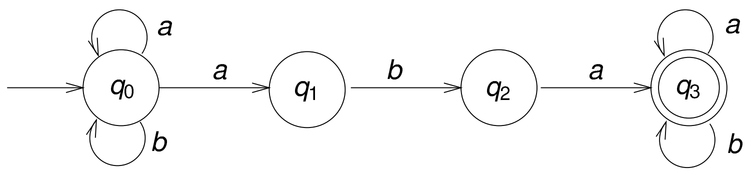


2.5 Epädeterministiset äärelliset automaatit

Epädeterministiset automaatit ovat muuten samanlaisia kuin deterministiset, mutta niiden siirtymäfunktio δ ei liitä automaatin vanhan tilan ja syötemerkin muodostamiin pareihin yksikäsitteistä uutta tilaa, vaan *joukon* mahdollisia seuraavia tiloja. Epädeterministinen automaatti hyväksyy syötteensä, jos *jokin* mahdollisten tilojen jono johtaa hyväksyvään lopputilaan.

Vaikka epädeterministisiä automaatteja ei voi sellaisinaan toteuttaa tietokoneohjelmina, ne ovat tärkeä päätösongelmien *kuvausformalismi*.

Esimerkki. Epädeterministinen automaatti, joka tutkii sisältääkö syötejono osajonoa *aba*:



Automaatti hyväksyy esim. syötejonon *aaba*, koska sen on mahdollista edetä seuraavasti:

$$(q_0, aaba) \vdash (q_0, aba) \vdash (q_1, ba) \vdash (q_2, a) \vdash (q_3, \varepsilon).$$

Automaatti voisi päätyä myös hylkäävään tilaan:

$$(q_0, aaba) \vdash (q_0, aba) \vdash (q_0, ba) \vdash (q_0, a) \vdash (q_0, \varepsilon),$$

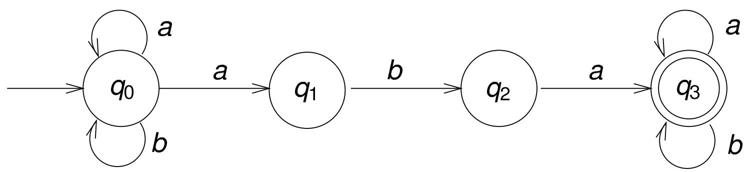
mutta tällä ei ole merkitystä— voidaan ajatella, että automaatti osaa “ennustaa” ja valita aina parhaan mahdollisen vaihtoehdon.

Määritelmä 2.2 Epädeterministinen äärellinen automaatti on viisikko

$$M = (Q, \Sigma, \delta, q_0, F),$$

missä

- Q on äärellinen *tilojen* joukko;
- Σ on *syöteaakkosto*;
- $\delta : Q \times \Sigma \to \boxed{\mathcal{P}(Q)}$ on automaatin joukkoarvoinen *siirtymäfunktio*;
- $q_0 \in Q$ on *alkutila*;
- $F \subseteq Q$ on *(hyväksyvien) lopputilojen* joukko.



Esimerkiksi *aba*-automaatin siirtymäfunktio:

		a	b
$\rightarrow$	q_0	$\{q_0, q_1\}$	$\{q_0\}$
	q_1	$\emptyset$	$\{q_2\}$
	q_2	$\{q_3\}$	$\emptyset$
$\leftarrow$	q_3	$\{q_3\}$	$\{q_3\}$

Taulukosta nähdään, että esimerkiksi $\delta(q_0, a) = \{q_0, q_1\}$ ja $\delta(q_1, a) = \emptyset$.

Epädeterministisen automaatin tilanne (q, w) voi *johtaa suoraan* tilanteeseen (q', w'), merkitään

$$(q, w) \underset{M}{\vdash} (q', w'),$$

jos on $w = aw'$ ($a \in \Sigma$) ja $q' \in \delta(q, a)$. Sanotaan myös, että tilanne (q', w') on tilanteen (q, w) mahdollinen *välitön seuraaja*.

Useamman askelen mittaiset tilannejohdot, merkkijonojen hyväksyminen ja hylkääminen ym. käsitteet määritellään samoin sanoin kuin deterministisillä automaateilla. Koska yhden askelen johdon määritelmä kuitenkin nyt on toinen, niiden sisältö muovautuu hieman erilaiseksi.

Lause 2.2 Olkoon $A = L(M)$ jonkin epädeterministisen äärellisen automaatin M tunnistama kieli. Tällöin on olemassa myös deterministinen äärellinen automaatti $\widehat{M}$, jolla $A = L(\widehat{M})$.

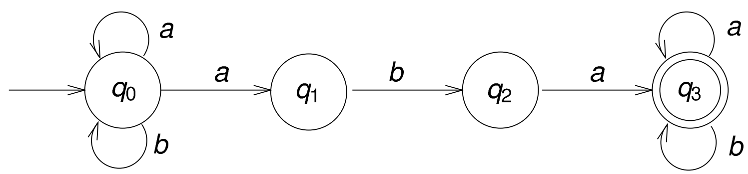
Todistus. Olkoon $A = L(M)$, $M = (Q, \Sigma, \delta, q_0, F)$. Todistuksen ideana on laatia deterministinen äärellinen automaatti $\widehat{M}$, joka simuloi M:n toimintaa kaikissa sen kullakin hetkellä mahdollisissa tiloissa *rinnakkain*.

Formaalisti: automaatin $\widehat{M}$ tilat vastaavat M:n tilojen *joukkoja*:

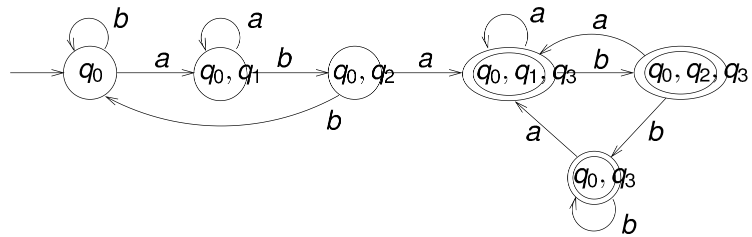
$$\widehat{M} = (\hat{Q}, \Sigma, \hat{\delta}, \hat{q}_0, \widehat{F}),$$

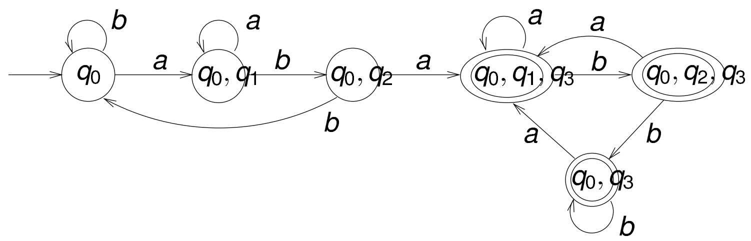
missä

$$\begin{aligned}
\hat{Q} &= \mathcal{P}(Q) = \{S \mid S \subseteq Q\}, \\
\hat{q}_0 &= \{q_0\}, \\
\widehat{F} &= \{S \subseteq Q \mid S \text{ sisältää jonkin } q_f \in F\}, \\
\hat{\delta}(S, a) &= \bigcup_{q \in S} \delta(q, a).
\end{aligned}$$

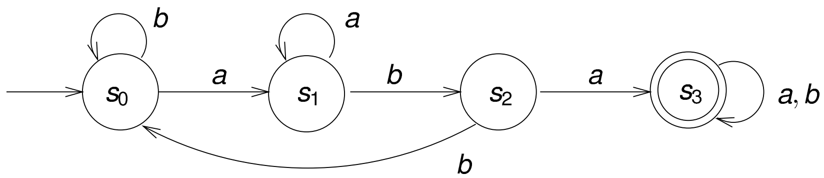


Esimerkiksi *aba*-automaattiin sovellettuna em. konstruktio tuottaa seuraavan deterministisen automaatin (vain alkutilasta saavutettavat tilat esitetty):





Minimoimalla ja nimeämällä tilat uudelleen tämä yksinkertaistuu muotoon:



[Todistus jatkuu.] Tarkastetaan, että automaatti $\widehat{M}$ todella on ekvivalentti M:n kanssa, so. että $L(\widehat{M}) = L(M)$.

Määritelmien mukaan on:

$$x \in L(M) \ \text{ joss } \ (q_0, x) \vdash_M^* (q_f, \varepsilon) \ \text{ jollakin } q_f \in F$$

ja

$$x \in L(\widehat{M}) \ \text{ joss } \ (\{q_0\}, x) \vdash_{\widehat{M}}^* (S, \varepsilon)$$

ja S sis. jonkin $q_f \in F$.

Osoitetaan siis, että kaikilla $x \in \Sigma^*$ ja $q \in Q$ on:

$$(q_0, x) \vdash_M^* (q, \varepsilon) \ \text{ joss } \ (\{q_0\}, x) \vdash_{\widehat{M}}^* (S, \varepsilon) \text{ ja } q \in S. \qquad (2)$$

Väite (2):

$$(q_0, x) \vdash_M^* (q, \varepsilon) \ \text{ joss } \ (\{q_0\}, x) \vdash_{\widehat{M}}^* (S, \varepsilon) \text{ ja } q \in S.$$

Väitteen (2) todistus tehdään induktiolla merkkijonon x pituuden suhteen.

(i) Tapaus $|x| = 0$:

$(q_0, \varepsilon) \vdash_M^* (q, \varepsilon)$ joss $q = q_0$;

samoin $(\{q_0\}, \varepsilon) \vdash_{\widehat{M}}^* (S, \varepsilon)$ joss $S = \{q_0\}$.

Väite (2): $(q_0, x) \vdash_M^* (q, \varepsilon)$ joss $(\{q_0\}, x) \vdash_{\widehat{M}}^* (S, \varepsilon)$ ja $q \in S$.

(ii) Induktioaskel:

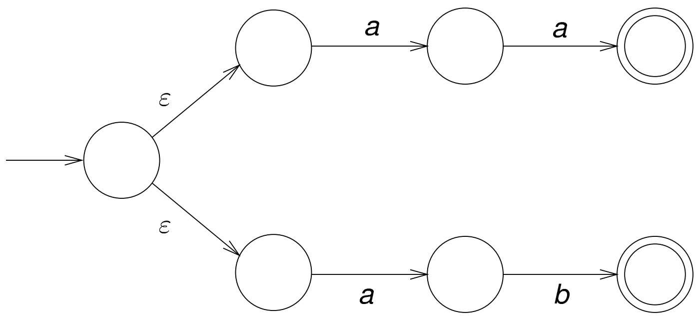
Olkoon $x = ya$; oletetaan, että väite (2) pätee y:lle. Tällöin:

$(q_0, x) = (q_0, ya) \vdash_M^* (q, \varepsilon)$ joss
$\exists q' \in Q$ s.e. $(q_0, ya) \vdash_M^* (q', a)$ ja $(q', a) \vdash_M (q, \varepsilon)$ joss
$\exists q' \in Q$ s.e. $(q_0, y) \vdash_M^* (q', \varepsilon)$ ja $(q', a) \vdash_M (q, \varepsilon)$ joss (ind.ol.)
$\exists q' \in Q$ s.e. $(\{q_0\}, y) \vdash_{\widehat{M}}^* (S', \varepsilon)$ ja $q' \in S'$ ja $q \in \delta(q', a)$ joss
$(\{q_0\}, y) \vdash_{\widehat{M}}^* (S', \varepsilon)$ ja $\exists q' \in S'$ s.e. $q \in \delta(q', a)$ joss
$(\{q_0\}, y) \vdash_{\widehat{M}}^* (S', \varepsilon)$ ja $q \in \bigcup_{q' \in S'} \delta(q', a) = \hat{\delta}(S', a)$ joss
$(\{q_0\}, ya) \vdash_{\widehat{M}}^* (S', a)$ ja $q \in \hat{\delta}(S', a) = S$ joss
$(\{q_0\}, ya) \vdash_{\widehat{M}}^* (S', a)$ ja $(S', a) \vdash_{\widehat{M}} (S, \varepsilon)$ ja $q \in S$ joss
$(\{q_0\}, x) = (\{q_0\}, ya) \vdash_{\widehat{M}}^* (S, \varepsilon)$ ja $q \in S$. $\square$

ε-automaatit

Jatkossa tarvitaan vielä yksi äärellisten automaattien mallin laajennus: epädeterministinen äärellinen automaatti, jossa sallitaan ε-*siirtymät*. Tällaisella siirtymällä automaatti tekee epädeterministisen valinnan eri jatkovaihtoehtojen välillä lukematta yhtään syötemerkkiä.

Esimerkiksi kieli $\{aa, ab\}$ voitaisiin tunnistaa seuraavalla ε-automaatilla:



Formaalisti: ε-automaatti on viisikko

$$M = (Q, \Sigma, \delta, q_0, F),$$

missä siirtymäfunktio δ on kuvaus

$$\delta : Q \times (\Sigma \cup \{\varepsilon\}) \to \mathcal{P}(Q).$$

Muut määritelmät ovat kuten tavallisilla epädeterministisillä äärellisillä automaateilla, paitsi suoran tilannejohdon määritelmä: ε-automaattien tapauksessa relaatio

$$(q, w) \underset{M}{\vdash} (q', w')$$

on voimassa, jos on

(i) $w = aw'$ $(a \in \Sigma)$ ja $q' \in \delta(q, a)$; tai
(ii) $w = w'$ ja $q' \in \delta(q, \varepsilon)$.

Lemma 2.4 Olkoon $A = L(M)$ jollakin ε-automaatilla M. Tällöin on olemassa myös ε-siirtymätön epädeterministinen automaatti $\widehat{M}$, jolla $A = L(\widehat{M})$.

Todistus. Olkoon $M = (Q, \Sigma, \delta, q_0, F)$ jokin ε-automaatti. Automaatti $\widehat{M}$ toimii muuten aivan samoin kuin M, mutta se "harppaa" ε-siirtymien yli suorittamalla kustakin tilasta lähtien vain ne "aidot" siirtymät, jotka ovat siitä käsin jotakin ε-siirtymäjonoa pitkin saavutettavissa.

Formaalisti määritellään annetun tilan $q \in Q$ ε-*sulkeuma* $\varepsilon^*(q)$ automaatissa M kaavalla

$$\varepsilon^*(q) = \{q' \in Q \mid (q, \varepsilon) \underset{M}{\vdash}^* (q', \varepsilon)\},$$

so. joukkoon $\varepsilon^*(q)$ kuuluvat kaikki ne automaatin M tilat, jotka ovat saavutettavissa tilasta q pelkillä ε-siirtymillä.

Automaatin $\widehat{M}$ siirtymäsäännöt voidaan nyt kuvata seuraavasti:

$$\widehat{M} = (Q, \Sigma, \hat{\delta}, q_0, \widehat{F}),$$

missä

$$\hat{\delta}(q, a) = \bigcup_{q' \in \varepsilon^*(q)} \delta(q', a);$$

$$\widehat{F} = \{q \in Q \mid \varepsilon^*(q) \cap F \neq \emptyset\}.$$

Poistamalla edellisen konstruktion mukaisesti ε-siirtymät ε-automaatista saadaan tavallinen epädeterministinen automaatti, esim.: