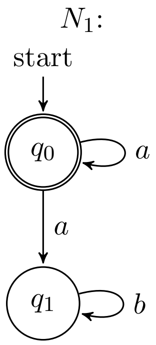
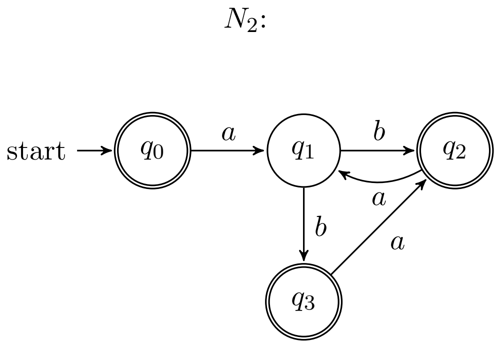
582206 Laskennan mallit, syksy 2012
3. harjoitusten malliratkaisut
Juhana Laurinharju ja Jani Rahkola

1. Olkoon N_1 ja N_2 epädeterministiset automaatit jotka on kuvattu alla. Tunnistavatko automaatit seuraavat sanat?

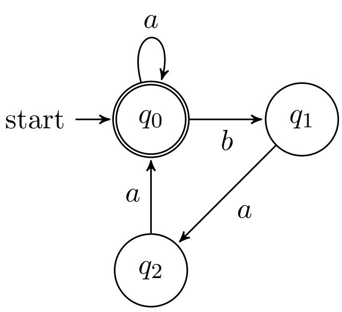
	N_1	N_2
a	kyllä	ei
aa	kyllä	ei
aab	ei	ei
ε	kyllä	kyllä
ab	ei	kyllä
$abab$	ei	kyllä
aba	ei	kyllä
$abaa$	ei	ei





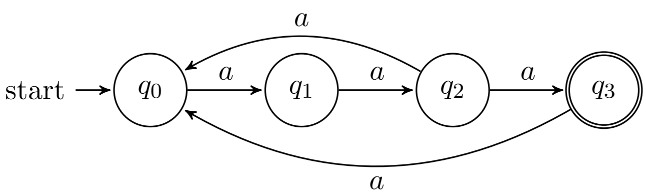
2. Minkälaisia sanoja seuraavat äärelliset epädeterministiset automaatit hyväksyvät?

(a)

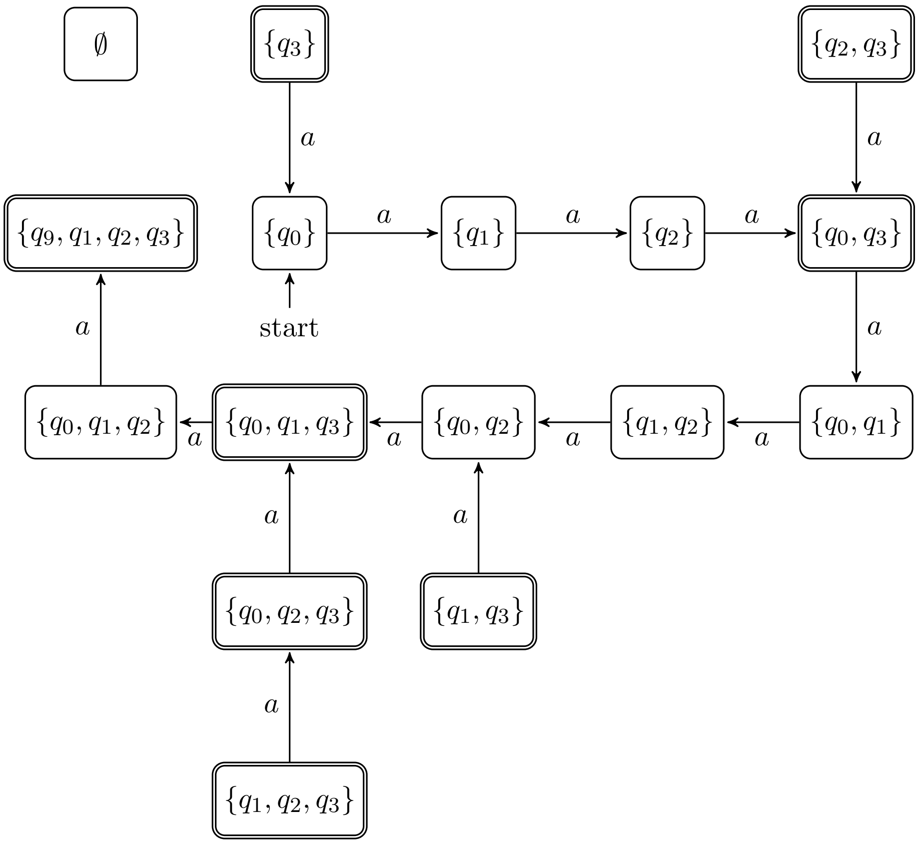


Jokaisen b:n jälkeen on ainakin kaksi a:ta. Automaatin määrittelemä säännöllinen kieli on siis $\{a, baa\}^*$ joka voidaan kuvata myös säännöllisellä lausekkeella $(a \cup baa)^*$.

(b)



Rakennetaan automaattia vastaava deterministinen automaatti.



Deterministisestä automaatista voidaan huomata, että automaatti hyväksyy kaikki merkkijonot, joissa on kolme, kuusi, seitsemän tai vähintään yhdeksän a:ta. Säännöllisenä lausekkeena kielen voi kirjoittaa esimerkiksi $aaa(aaa \cup aaaa)^*$ (muodostettu vastaamaan epädeterminististä automaattia) tai $(a^3 \cup a^6 \cup a^7 \cup a^9 a^*)$ (muodostettu deterministisen automaatin antamasta lukumääräintuitiosta)

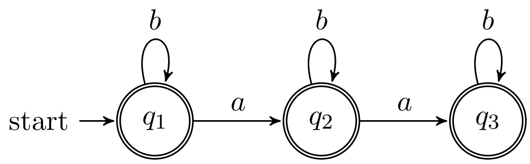
3. Piirrä epädeterministiset automaatit tiloineen ja siirtymänuolineen seuraaville kielille.

   (a) $L = \{w \in \{a,b\}^* \mid w$ sisältää korkeintaan kaksi a:ta$\}$

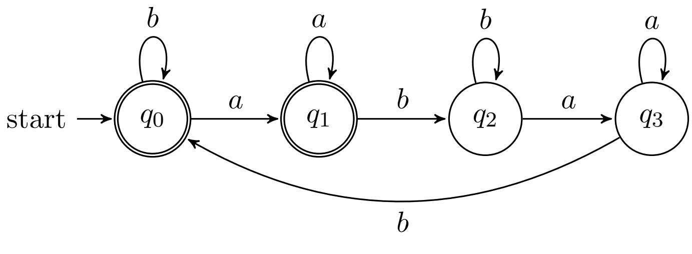
   (b) $L = \{w \in \{a,b\}^* \mid w$ sisältää parillisen määrän alimerkkijonoa $ab\}$

   (c) $L = \{w \in \{a,b\}^* \mid w$:n ensimmäinen ja viimeinen kirjain ovat samat$\}$

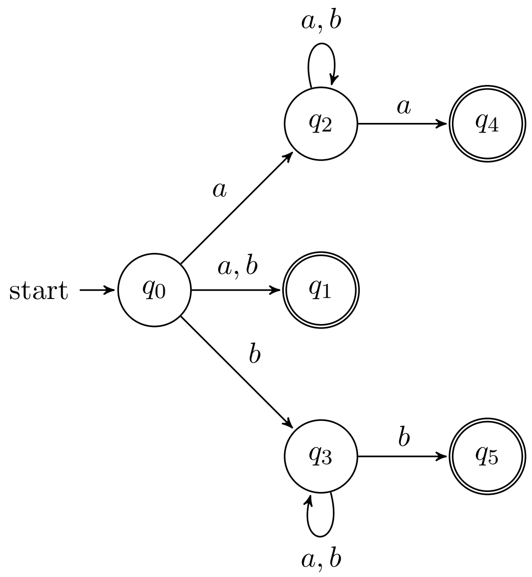
(a)



(b)

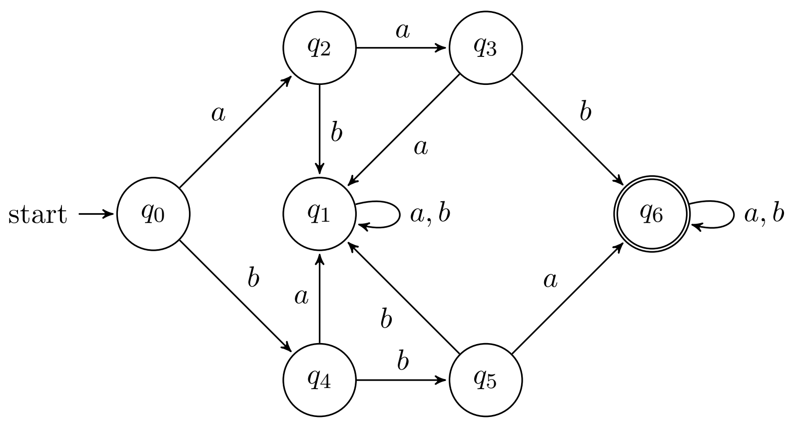


(c)

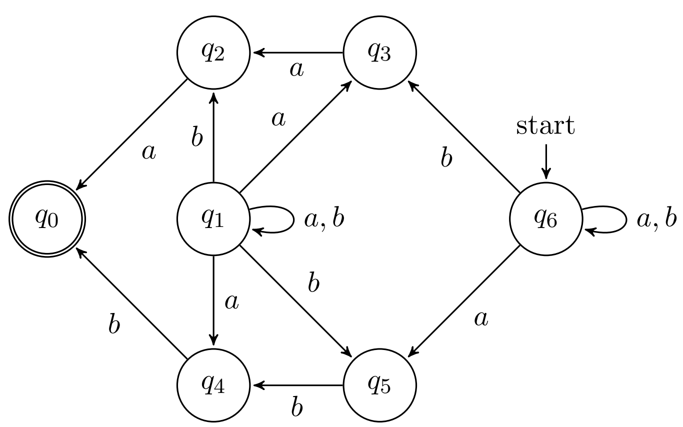


4. Merkkinon $w = w_1 w_2 \dots w_n$ *käänteismerkkijono* on $w^{\mathcal{R}} = w_n w_{n-1} \dots w_1$. Olkoon kielen A *käänteiskieli* $A^{\mathcal{R}} = \{w^{\mathcal{R}} \mid w \in A\}$. Näytä että jos A on säännöllinen niin myös $A^{\mathcal{R}}$ on säännöllinen (vihje: käytä epädeterministisyyttä apuna). Tee myös pienet esimerkit.

   Muodostetaan ensin pari esimerkkiä. Tässä automaatti, joka tunnistaa kaikki merkkijonot, jotka alkavat joko aab:llä tai bba:lla.

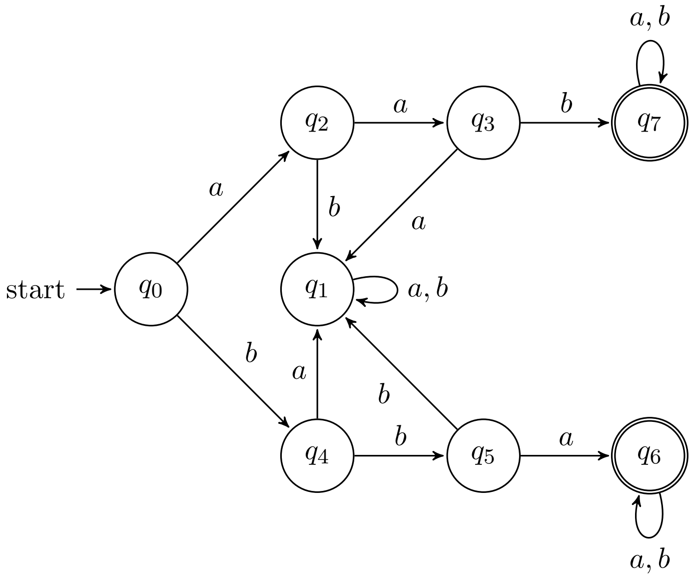


   Tälle voidaan nyt helposti muodostaa käänteiskielen automaatti asettamalla lopputila alkutilaksi, alkutila lopputilaksi ja kääntämällä kaikki nuolet ympäri.



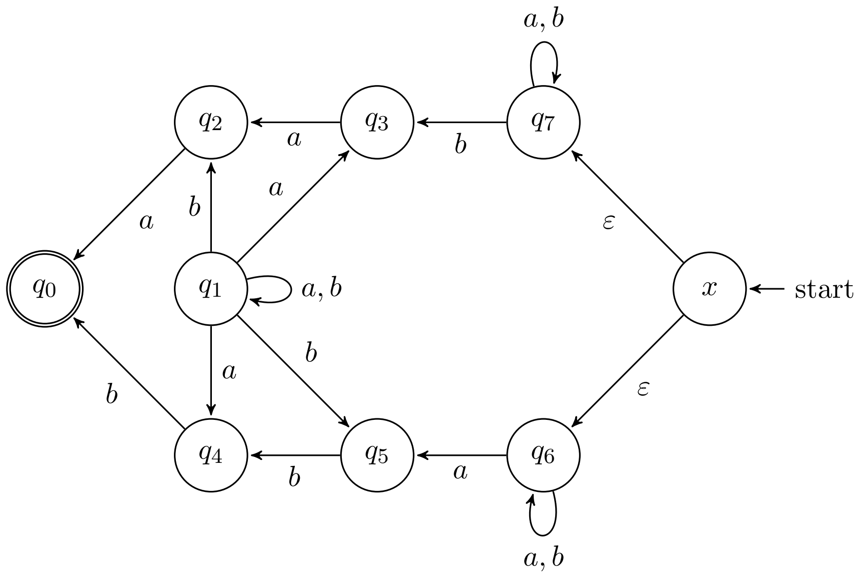
Ja näin saatiin epädeterministinen automaatti, joka tunnistaa kaikki merkkijonot jotka loppuvat joko *baa* tai *abb*.

Alkuperäinen automaatti oltaisiin kuitenkin voitu muodostaa myös seuraavasti:



Nyt voidaan huomata, että automaattia kääntäessä tarvittaisiin kaksi aloitustilaa, sillä alkuperäisessä automaatissa on useampia hyväksyviä tiloja.

Ongelma voidaan korjata lisäämällä käänteiskielen automaattiin uusi aloitustilan, josta tulee ε-siirtymät jokaiseen haluttuun aloitustilaan.



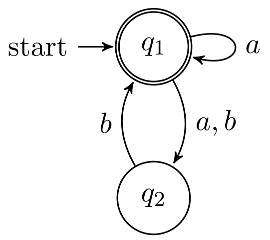
Kun nyt ollaan saatu intuitio siitä miten kääntäminen pitäisi tehdä, voidaan tämä vielä kirjoittaa formaalisti.

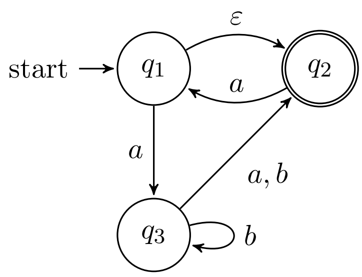
Olkoon $M = (Q, \Sigma, \delta, s, F)$ kielen A tunnistava deterministinen automaatti. Muodostetaan nyt yllä olevien esimerkkien motivoimana epädeterministinen automaatti $M^{\mathcal{R}}$ käänteiskielelle $A^{\mathcal{R}}$.

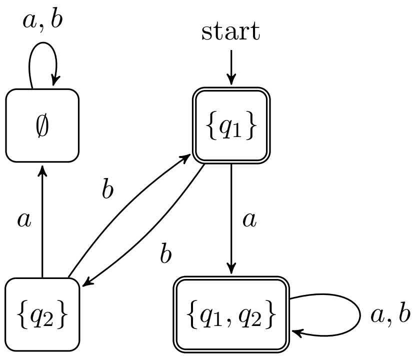
$Q^{\mathcal{R}} = Q \cup \{x\}$, missä $\notin Q$		uusi alkutila
$\Sigma^{\mathcal{R}} = \Sigma \cup \{\epsilon\}$		$M^{\mathcal{R}}$ epädeterministinen
$\delta^{\mathcal{R}}(q_i, \sigma) = \{q_j \in Q \mid \delta(q_j, \sigma) = q_i\}$		kaikki M:n tilat, joista pääsee σ:lla nykyiseen tilaan
$\delta^{\mathcal{R}}(x, \epsilon) = F$		uudesta alkutilasta pääsee suoraan M:n kaikkiin lopputiloihin
$\delta^{\mathcal{R}}(x, \sigma) = \emptyset$		uudesta alkutilasta ei ole siirtymiä Σ:n merkeillä
$F^{\mathcal{R}} = \{s\}$		$M^{\mathcal{R}}$:n ainoa hyväksyvä tila on M:n alkutila

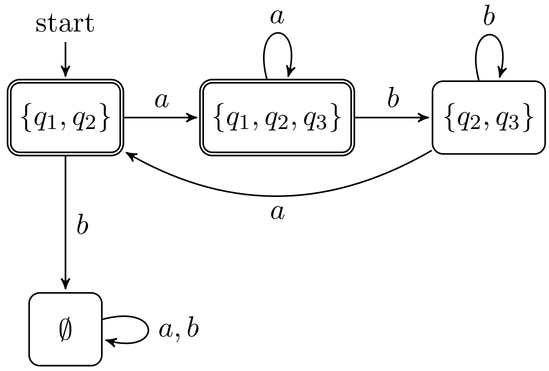
Nyt ollaan siis muodostettu automaatti, joka aloittaa alkuperäisen automaatin hyväksyvistä tiloista, lopettaa alkuperäisen automaatin alkutilaan ja kääntää kaikki nuolet ympäri.

5. Muunna seuraavat epädeterministiset automaatit deterministisiksi käyttämällä lauseen 1.39 todistusta apuna.









6. Olkoon M **deterministinen** automaatti, missä on n tilaan, ja $L(M) = A$. (vihje ajattele syklejä)

   (a)

   Väite. *$A \neq \emptyset \Leftrightarrow \exists w \in A$ mille $|w| < n$.*

   Todistus.

   "$\Leftarrow$"

   Koska on olemassa $w \in A$, ei A ole tyhjä.

   "$\Rightarrow$"

   Lähdetään osoittamaan väitettä seuraavasti:

   - Otetaan merkkijono $w \in A, w = w_1 w_2 \dots w_k$
   - Tarkastellaan tilajonoa $\bar{q} = q_1 q_2 \dots q_{k+1}$ jonka läpi automaatti M kulkee merkkijonolla w.
   - Poistetaan tilajonosta $\bar{q}$ kaikki mahdollisesti löytyvät silmukat.
   - Edellisen kohdan tuloksena saatua tilajonoa vastaa jokin merkkijono $u \in A$ jolla on haluttu ominaisuus.

Koska $A \neq \emptyset$, niin löytyy $w \in A, w = w_1 w_2 \dots w_k$, ja olkoon $M = (Q, \Sigma, \delta, s, F)$. Määritellään tilajono $\bar{q} = q_1 q_2 \dots q_{k+1}$ missä $q_1 = s$ ja $q_{i+1} = \delta(q_i, w_i)$.

$$\begin{array}{ccccccc} q_1 & \xrightarrow{w_1} & q_2 & \xrightarrow{w_2} & \dots & \xrightarrow{w_n} & q_{n+1} \\ s & & \delta(s, u_1) & & & & \end{array}$$

i. Jos $|w| < n$, niin u on etsitty merkkijono.

ii. Jos $|w| \geq n$, niin lokeroperiaatteen nojalla löytyy $q_i = q_j$ jollain $i, j \in \{1 \dots n\}$, $i < j$. Siis tilajonosta $\bar{q}$ löytyy silmukka.

$$\dots \xrightarrow{w_{i-2}} q_{i-1} \xrightarrow{w_{i-1}} \overset{w_i \dots w_{j-1}}{\overset{\circlearrowleft}{q_i = q_j}} \xrightarrow{w_j} q_{j+1} \xrightarrow{w_{j+1}} \dots$$

Erityisesti

$$\begin{aligned} & \delta(q_{i-1}, w_{i-1}) = q_i && = q_j \\ \text{ja} \quad & \delta(q_i, u_j) && = \delta(q_j, u_j) = q_{j+1}. \end{aligned}$$

Tällöin merkkijono

$$v = w_1 \dots w_{i-1} w_j \dots w_k$$

kulkee tilajonon

$$\bar{p} = q_1 \dots q_i q_{j+1} \dots q_{k+1}$$

läpi. Koska tilajonoissa $\bar{q}$ ja $\bar{p}$ on sama viimeinen tila $q_{k+1} \in F$, hyväksyy automaatti M myös merkkijonon v. Lisäksi $|v| \leq |u| - 1$, eli merkkinono lyhenee ainakin yhden merkin verran. Toistamalla tätä menetelmää korkeintaan $(k-n)+1$ kertaa, löydetään automaatin M hyväksymä merkkijono u, jolla merkkijono lyhenee ainakin $(k-n)+1$ merkin verran.

$$\begin{aligned} |u| &\leq |u| - ((k-n)+1) \\ &= k - k + n - 1 \\ &= n - 1 \\ &< n \end{aligned}$$

Siis $|u| < n$ ja täten merkkijono u toteuttaa halutun ehdon.

□

(b) Määritellään seuraavaa todistusta varten δ-funktion yleistys $\delta^* : Q \times \Sigma^* \to Q$

$$\delta^*(q, w_1 \dots w_n) = \begin{cases} \delta(q, \varepsilon) & \text{kun } n = 0 \\ \delta(q, w_1) & \text{kun } n = 1 \\ \delta(\delta^*(q, w_1 \dots w_{n-1}), w_n) & \text{kun } n > 1 \end{cases}$$

Funktio δ^* ottaa siis tilan ja merkkijonon, ja palauttaa sen tilan johon automaatti päätyy kun se käy merkkijonon läpi merkki kerrallaan.

Väite. *Kieli A on ääretön* $\Leftrightarrow \exists w \in A$ *mille* $n \leq |w| < 2n$

Todistus.

"$\Leftarrow$"

Idea:

- Olkoon $w = w_1 \dots w_k \in A$ jokin ehdon täyttävä merkkijono.
- Nyt erityisesti $|w| \geq n$.
- Muodostetaan jälleen merkkijonolla w automaatin läpikäymä tilajono

$$\bar{q} = q_1 \dots q_{k+1}$$

missä $k + 1 \geq n$.

- Nyt erityisesti $k > n$, joten tilajonossa on silmukka.
- Näin löydettyä silmukkaa toistamalla saadaan aina toinen toistaan pidempiä merkkijonoja jotka kuuluvat kieleen.

Olkoon $M = (Q, \Sigma, \delta, s, F)$ ja $w \in A$ jolla $n \leq |w| < 2n$. Nyt

$$w = w_1 w_2 \dots w_k \text{ missä } n \leq k < 2n.$$

Määritellään automaatin M merkkijonolla w läpikäymä tilajono $\bar{q}$,

$$\bar{q} = q_1 q_2 \dots q_{k+1}, \text{ missä } q_1 = s \text{ ja } q_{i+1} = \delta(q_i, w_i).$$

Lokeroperiaatteen nojalla nyt löytyy sellaiset $i, j \in \{1 \dots n\}, i < j$, joilla $q_i = q_j$. Tilajonossa $\bar{q}$ on siis silmukka. Määritellään nyt merkkijonon w osamerkkijonot

$$\begin{aligned} a &= w_1 \dots w_{i-1} \\ b &= w_i \dots w_{j-1} \\ c &= w_j \dots w_k. \end{aligned}$$

Nyt $w = abc$ ja

$$\begin{aligned} \delta^*(q_1, a) &= q_i \\ \delta^*(q_j, b) &= q_j = q_i \\ \delta^*(q_j, c) &= q_{k+1}. \end{aligned}$$

Nyt näyttäisi siltä, että voimme toistaa merkkijonoa b mielivaltaisen monta kertaa ja saamme aina jonkin automaatin M hyväksymän merkkijonon joka siis kuuluu kieleen A.

Automaatti M hyväksyy merkkijonon $ab^m c$ jos $\delta^*(s, ab^m c) \in F$. Näytetään tämä todeksi. Ensinnäkin a:n, b:n ja c:n määritelmistä seuraa heti, että

$$\begin{aligned} \delta^*(s, ab^m c) &= \delta^*(\delta^*(s, a), b^m c) \\ &= \delta^*(q_i, b^m c) \\ &= \delta^*(\delta^*(q_i, b^m), c) \end{aligned}$$

ja

$$\delta^*(q_j, c) = q_{k+1}.$$

Ainoa puuttuva palanen on näyttää, että $\delta^*(q_i, b^m) = q_j$. Jos tämä pitää paikkansa niin $\delta^*(q_1, ab^m c) = q_{k+1} \in F$. Todistetaan tämä induktiolla.

Väite. *Kaikilla $m \in \mathbb{N}$ pätee $\delta^*(q_i, b^m) = q_j$.*

Todistus.

Alkuaskel: $m = 1$ Tilajonon määritelmän nojalla $\delta^*(q_i, b^1) = q_j$.

Induktioaskel

Induktio-oletus $\delta^*(q_i, b^m) = q_j$

Induktioaskeleen väite $\delta^*(q_i, b^{m+1}) = q_j$

Induktioaskeleen todistus

$$\begin{aligned}
\delta^*(q_i, b^{m+1}) &= \delta^*(q_i, bb^m) \\
&= \delta^*(\delta^*(q_i, b), b^m) && \delta^*\text{:n määritelmä} \\
&= \delta^*(q_j, b^m) && \text{alkuaskel} \\
&= \delta^*(q_i, b^m) && q_i = q_j \\
&= q_j && \text{induktio-oletus}
\end{aligned}$$

□

Nyt kaikilla $m \in \mathbb{N}$ pätee

$$\begin{aligned}
\delta^*(q_1, ab^m c) &= \delta^*(\delta^*(q_1, a), b^m c) && \delta^*\text{:n määritelmä} \\
&= \delta^*(q_i, b^m c) && a\text{:n määritelmä} \\
&= \delta^*(\delta^*(q_i, b^m), c) && \delta^*\text{:n määritelmä} \\
&= \delta^*(q_j, c) && \text{äskeinen induktiotoditus} \\
&= q_{k+1} \in F && c\text{:n ja } \bar{q}\text{:n määritelmät}
\end{aligned}$$

Siis voidaan muodostaa erilaiset merkkijonot

$$ac, abc, ab^2c, \ldots$$

joista jokainen kuuluu kieleen A, joten A on ääretön.

"$\Rightarrow$"

Olkoon $M = (Q, \Sigma, \delta, s, F)$ deterministinen äärellinen automaatti ja A automaatin M tunnistama ääretön säännöllinen kieli. Koska korkeintaan n:n mittaisia merkkijonoja on vain äärellinen määrä, löytyy merkkijono

$$w = w_1 w_2 \ldots w_k \in A \text{ missä } k \geq n.$$

Jos $k < 2n$, niin w on haettu merkkijono. Tarkastellaan vielä tapausta, missä $k \geq 2n$. Olkoon $q_1 \ldots q_{k+1}$ tilajono, jonka automaatti M käy läpi syötteellä w. Etsitään tilajonon lyhin sykli. Lokeroperiaatteen nojalla löytyy ainakin yksi (i, j) pari, jolla

$$i, j \in 1 \ldots n \text{ ja } i < j.$$

Valitaan näistä (i, j) pareista sellainen, jolla erotus $j - i$ on pienin. Erityisesti tilajonosta löytyy nyt sykli $q_i \ldots q_j$ missä $q_i = q_j$ ja joka on pituudeltaan mahdollisimman lyhyt. Merkkijono w voidaan nyt jakaa kolmeen osaan. Sykliä edeltävään, sykliin ja sykliä seuraavaan osaan seuraavasti:

$$w = \underbrace{(w_1 \ldots w_{i-1})}_{\text{alkuosa}} \underbrace{(w_i \ldots w_{j-1})}_{\text{silmukka}} \underbrace{(w_j \ldots w_k)}_{\text{loppuosa}}$$

Kohdan a) nojalla alkuosasta $w_1 \ldots w_{i-1}$ voidaan muodostaa merkkijono a, jolla

$$|a| < n \text{ ja } \delta^*(q_1, a) = q_i.$$

Vastaavasti voidaan muodostaa merkkijono c, jolla

$$|c| < n \text{ ja } \delta^*(q_j, c) = q_{k+1}.$$

Lisäksi voidaan määritellä valittua lyhintä silmukkaa vastaava merkkijono

$$b = w_i \dots w_{j-1}.$$

Nyt $|b| \leq n$, sillä jos $|b| > n$, niin

$$\begin{aligned} |q_i \dots q_{j-1}| &= j - (i-1) \\ &= j - i + 1 \\ &= |b| > n \end{aligned}$$

joten lokeroperiaatteen nojalla on olemassa $i', j' \in \{1 \dots n\}$ joilla $i' < j'$ ja $q_{i'} = q_{j'}$. Nyt $j' - i' < j - i$ mikä on ristiriidassa parin (i, j) valinnan kanssa.
Nyt merkkijono $ab^kc \in A$ kaikilla $k \in \mathbb{N}$. Haluaisimme lisäksi, että $n \leq |ab^kc| < 2n$ jollain $k \in \mathbb{N}$. Haluttaisiin siis sellainen $k \in \mathbb{N}$, jolla pätisi

$$\begin{array}{rccl|l} & n \leq & |ab^kc| & < 2n & |b^k| = k|b| \\ \Leftrightarrow & n \leq & k|b| + |ac| & < 2n & -|ac| \\ \Leftrightarrow & n - |ac| \leq & k|b| & < 2n - |ac| & \div|b| \\ \Leftrightarrow & \frac{n-|ac|}{|b|} \leq & k & < \frac{2n-|c|}{|b|} & \end{array}$$

Valinta

$$k = \left\lceil \frac{n - |ac|}{|b|} \right\rceil$$

vaikuttaisi turvalliselta valinnalta. Tarkistetaan vielä että se toteuttaa molemmat epäyhtälöt.

$$\frac{n - |ac|}{|b|} \leq \left\lceil \frac{n - |ac|}{|b|} \right\rceil$$

ja

$$\begin{aligned} \left\lceil \frac{n - |ac|}{|b|} \right\rceil &< \frac{n - |ac|}{|b|} + 1 \\ &= \frac{n - |ac| + |b|}{|b|} \\ &\leq \frac{n - |ac| + n}{|b|} \\ &= \frac{2n - |ac|}{|b|}. \end{aligned}$$

Voidaan siis valita $k = \left\lceil \frac{n-|ac|}{|b|} \right\rceil$. Nyt siis $ab^kc \in A$ ja $n \leq ab^kc < 2n$ kuten haluttiin.

$\square$