Automaatit

Automaatit ovat teoreettisia koneita, jotka käsittelevät muodollisia sanoja. Automaatti lukee muodollisen sanan kirjain kerrallaan, vasemmalta oikealle, ja joko hyväksyy tai hylkää sanan. Täten jokaista automaattia vastaa muodollinen kieli, joka on täsmälleen niiden sanojen joukko, jotka automaatti hyväksyy.

Automaatit

Esimerkki. (Vertaa kirjan[1] esimerkki 153.) Olkoon aakkosto Σ suomen kielen aakkosten joukko. Olkoon $L \subseteq \Sigma^*$ kieli, johon kuuluu täsmälleen kirjaimella a alkavat ja kirjaimeen $\ddot{o}$ päättyvät Σ-sanat. Muodostetaan abstrakti kone, joka hyväksyy kielen L. Alussa kone on *alkutilassa* q_0, jossa se lukee sanan ensimmäisen kirjaimen. Jos tämä kirjain ei ole a, kone M siirtyy *hylkäystilaan* q_1. Jos taas ensimmäinen kirjain on a, kone siirtyy tilaan q_2, jossa se lukee seuraavan kirjaimen. Aina ollessaan tilassa q_2, kone joko siiryy tilaan q_3 tai pysyy tilassa q_2. Jos luettu kirjain on $\ddot{o}$, siirtyy kone *hyväksymistilaan* q_3. Muulloin kone pysyy tilassa q_2. Ollessaan tilassa q_3, kone joko pysyy tilassa q_3 tai siirtyy takaisin tilaan q_2 sen mukaan, onko seuraava luettu kirjain $\ddot{o}$ vai ei. Havainnollistamme konetta M tiladiagrammilla (ks. kirja).

[1] Merikoski et. al. "Johdatus diskreettiin matematiikkaan"

Automaatin formaali määritelmä

Muodollisesti, *äärellinen automaatti* on yhdelmä

$$(Q, \Sigma, \delta, q_0, F).$$

1. Q on epätyhjä ja äärellinen joukko *tiloja*.
2. Σ on epätyhjä ja äärellinen aakkosto.
3. $\delta : Q \times \Sigma \longrightarrow Q$ on *siirtymäfunktio*, joka määrittelee automaatin toiminnan.
4. $q_0 \in Q$ on automaatin *aloitustila*.
5. $F \subseteq Q$ on joukko *hyväksyviä lopputiloja*.

Jos $a \in \Sigma$ ja $q \in Q$, niin $\delta(q, a)$ on tila, johon automaatti siirtyy luettuaan tilassa q kirjaimen a.

Automaatin formaali määritelmä

Automaatti siis aloittaa lukemaan sanaa kirjain kerrallaa vasemmalta oikealle. Alussa automaatti on alkutilassa q_0. Automaatti muuttaa tilaa lukiessaan sanaa siirtymäfunktion δ kuvaamalla tavalla. Jos $a \in \Sigma$ ja $q \in Q$, niin $\delta(q, a)$ on tila, johon automaatti siirtyy luettuaan tilassa q kirjaimen a. Lopuksi, luettuaan kaikki sanan kirjaimet, automaatti hyväksyy jos ja vain jos se on jossakin hyväksyvässä lopputilassa $q \in F$.

Automaatteja on kätevää havainnollistaa tiladiagrammilla. Tiladiagrammissa alkutila merkitään vinoviivoittamalla se. Hyväksyvät lopputilat ympyröidään kahdesti. (vrt. kirja)

Automaatit

Määritellään yllä kuvailtu kone $M = (Q, \Sigma, \delta, q_0, F)$ formaalisti.

1. $Q = \{q_0, q_1, q_2, q_3\}$
2. Σ on suomen kielen aakkosto.
3. $\delta(q_0, a) = q_2$; $\delta(q_0, \varphi) = q_1$ kun $\varphi \neq a$;
   $\delta(q_1, \varphi) = q_1$ kaikilla $\varphi \in \Sigma$;
   $\delta(q_2, ö) = q_3$; $\delta(q_2, \varphi) = q_2$ kun $\varphi \neq ö$;
   $\delta(q_3, ö) = q_3$; $\delta(q_3, \varphi) = q_2$ kun $\varphi \neq ö$;
4. $F = \{q_3\}$.

Harjoitustehtävä.(vrt. kirja, esimerkki 154.) Piirrä sellaisen automaatin tiladiagrammi, joka hyväksyy täsmälleen ne bittijonot (eli sanat $\alpha \in \{0,1\}^*$), jotka päättyvät vähintään kahteen peräkkäiseen nollaan tai kahteen peräkkäiseen ykköseen. Kirjoita säännöllinen lauseke, joka määrittelee kyseisen automaatin hyväksymän kielen.

Tiladiagrammeja piirrettäessä ei välttämättä tarvitse piirtää roskatilaan johtavia nuolia. Sovimme, että ne nuolet joita ei ole piirretty johtavat roskatilaan.

Automaatit.

Harjoitustehtävä. Piirrä sellaisen automaatin tiladiagrammi, joka hyväksyy kielen $11 \cup 00$.

Harjoitustehtävä. Piirrä sellaisen automaatin tiladiagrammi, joka hyväksyy kielen $(1 \cdot 1) \cdot (0 \cdot 0)^*$.

Yllä olevista kielistä ensimmäinen on äärellinen ja toinen ääretön.

Automaatit

Harjoitustehtävä. Piirrä sellaisen automaatin tiladiagrammi, joka hyväksyy kielen $(1 \cdot 0 \cdot 1)^*$.

Automaattien ja säännöllisten kielten yhteys

Olkoon $\Sigma = \{a, b\}$ aakkosto. Piirrä jokaista seuraavaa yksinkertaista säännöllistä lauseketta vastaava automaatti (vrt. kirja, sivu 170).

1. $\emptyset$
2. ε
3. a
4. $a \cup b$
5. $a \cdot b$
6. a^*

Automaattien ja säännöllisten kielten yhteys

Lause. Kieli on säännöllinen jos ja vain jos se voidaan tunnistaa automaatilla.

Sivuutamme lauseen todistuksen.

Epsilon-siirrot ja epädeterministiset automaatit

Epädeterministinen epsilon-automaatti on yhdelmä $(Q, \Sigma, \delta, q_0, F)$. Tässä Q, δ, q_0 ja F ovat kuten tavallisen (deterministisen) automaatin tapauksessa. Olio δ on nyt funktio

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

Intuitio on, että ollessaan tilassa q, automaatti voi luettuaan symbolin a siirtyä mihin hyvänsä tilaan $q' \in \delta(q, a)$. Nyt siis parilla (q, a) on (mahdollisesti) useampi kuin yksi seuraajatila. *Lisäksi*, automaatti voi siirtyä *lukematta symbolia a* mihin tahansa tilaan $q'' \in \delta(q, \varepsilon)$. Tällaista tilanvaihtoa, jossa luettavana oleva symboli ei lainkaan vaihdu, kutsutaan epsilon-siirroksi.

Deterministinen epsilon-automaatti

Myös deterministisen automaatin siirtymäfunktiota δ voi laajentaa epsilon-siirroilla. Tällöin saadaan *deterministinen epsilon-automaatti*.

Epsilon-siirrot ja epädeterministiset automaatit

Epädeterministinen epsilon-automaatti hyväksyy sanan α jos ja vain jos on ainakin yksi tapa lukea sana α siten että kun sana on luettu, automaatti on hyväksyvässä lopputilassa.

Periaate epäderminististä epsilon-automaattia vastaavan tiladiagrammin piirtämisessä on se, että tila q yhdistetään nuolella $\varphi \in \Sigma \cup \{\varepsilon\}$ jokaiseen tilaan $q' \in \delta(q, \varphi)$. (Katso myös kirja, sivu 171.)

Epsilon-siirrot ja epädeterministiset automaatit

Esimerkki. (Vrt. kirjan Esimerkki 155.) Määritellään epädeterministinen epsilon-automaatti $M = (Q, \Sigma, \delta, q_0, F)$, missä

1. $Q = \{q_0, q_1, q_2, q_3\}$, $\Sigma = \{a, b, c\}$, $F = \{q_3\}$.
2. $\delta(q_0, a) = \{q_1\}$; $\quad \delta(q_0, \varepsilon) = \{q_0, q_2\}$;
   $\delta(q_1, b) = \{q_2, q_3\}$; $\quad \delta(q_2, c) = \{q_3\}$;

Määritellään tavallinen deterministinen automaatti $M' = (P, \Sigma, \delta', p_0, F')$, missä

1. $Q = \{p_0, p_1, p_2, p_3\}$, $\Sigma = \{a, b, c\}$, $F = \{p_2, p_3\}$.
2. $\delta'(p_0, a) = p_1$; $\quad \delta(p_1, b) = p_2$; $\quad \delta'(p_0, c) = \delta'(p_2, c) = p_3$;

Automaatit M ja M' tunnistavat saman kielen $\{ab, abc, c\}$.

Determinististen automaattien ja epädeterministiten epsilon-automaattien suhde

Lause. Kieli L voidaan tunnistaa epädeterministisellä epsilon-automaatilla jos ja vain jos se voidaan tunnistaa tavallisella deterministisellä automaatilla.
Sivuutamme lauseen todistuksen.
Yllä olevan lauseen avulla on helppo todistaa, että jos kieli on säännöllinen, se voidaan tunnistaa tavallisella deterministisellä automaatilla. On nimittäin helpohko nähdä, että jokaista säännöllisen kielen muotoilussa käytettävää operaattoria $(\emptyset, \varepsilon, \cup, \cdot, ^*, a \in \Sigma)$ kohden on olemassa vastaava automaattikonstruktio. Esim. konkatenaatiota vastaa kahden automaatin kytkeminen sarjaan epsilon siirtojen avulla. Yhdistettä $\cup$ vastaa kahden automaatin kytkeminen rinnan epsilon siirtojen avulla. Emme kuitenkaan käy näitä konstruktioita läpi tässä esitettyä täsmällisemmin.

Harjoitustehtävä. Olkoon Σ aakkosto. Olkoon M (tavallinen deterministinen) automaatti, joka tunnistaa kielen $L \subseteq \Sigma^*$. Miten muodostetaan automaatti, joka tunnistaa kielen L komplementin $\overline{L} = \Sigma^* \setminus L$.

Vastaus. Olkoon $M = (Q, \Sigma, \delta, q_0, F)$. Automaatti $\overline{M} = (Q, \Sigma, \delta, q_0, \overline{F})$, missä $\overline{F} = Q \setminus F$, tunnistaa kielen $\overline{L}$.