3.8 Yhteydettömien kielten rajoituksista

Yhteydettömille kielille on voimassa säännöllisten kielten pumppauslemman vastine. Nyt kuitenkin merkkijonoa on pumpattava samanaikaisesti kahdesta paikasta.

Lemma 3.9 ("uvwxy-lemma") Olkoon L yhteydetön kieli. Tällöin on olemassa sellainen $n \geq 1$, että mikä tahansa $z \in L$, $|z| \geq n$, voidaan jakaa osiin $z = uvwxy$ siten, että

(i) $|vx| \geq 1$,

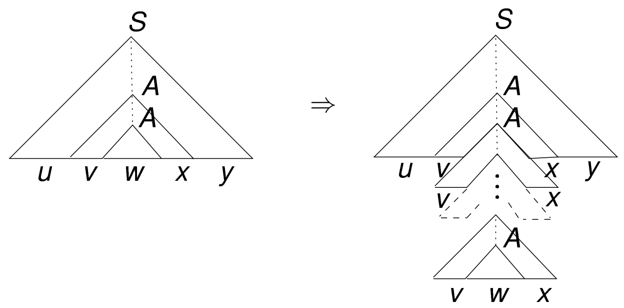
(ii) $|vwx| \leq n$,

(iii) $uv^iwx^iy \in L$ kaikilla $i = 0, 1, 2, \ldots$.

Todistus. Olkoon $G = (V, \Sigma, P, S)$ Chomskyn normaalimuotoinen kielioppi L:lle. Tällöin missä tahansa G:n jäsennyspuussa, jonka korkeus on h, on enintään 2^h lehteä. Toisin sanoen, minkä tahansa $z \in L$ jokaisessa jäsennyspuussa on polku, jonka pituus on vähintään $\log_2 |z|$.

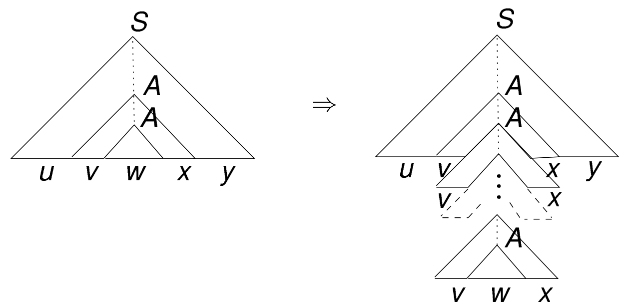
Olkoon $k = |V - \Sigma|$ kieliopin G välikkeiden määrä. Asetetaan $n = 2^{k+1}$. Tarkastellaan jotakin $z \in L$, $|z| \geq n$, ja sen jotakin jäsennyspuuta.

Edellisen nojalla puussa on polku, jonka pituus on $\geq k + 1$; tällä polulla on siis jonkin välikkeen toistuttava — itse asiassa jo polun $k + 2$ alimman solmun joukossa. Olkoon A jokin tällainen välike.



Merkkijono z voidaan nyt osittaa $z = uvwxy$, missä w on A:n alimmasta ilmentymästä tuotettu osajono ja vwx seuraavaksi ylemmästä A:n ilmentymästä tuotettu osajono; osajonot saadaan johdosta

$$S \Rightarrow^* uAy \Rightarrow^* uvAxy \Rightarrow^* uvwxy.$$



Koska siis $S \Rightarrow^* uAy$, $A \Rightarrow^* vAx$ ja $A \Rightarrow^* w$, osajonoja v ja x voidaan "pumpata" w:n ympärillä:

$$S \Rightarrow^* uAy \Rightarrow^* uvAxy \Rightarrow^* uv^2Ax^2y \Rightarrow^* \ldots \Rightarrow^* uv^iAx^iy \Rightarrow^* uv^iwx^iy.$$

Siten $uv^iwx^iy \in L$ kaikilla $i = 0, 1, 2, \ldots$.

Koska kielioppi G on Chomskyn normaalimuodossa ja $A \Rightarrow^* vAx$, on oltava $|vx| \geq 1$.

Koska edelleen välikkeen A valinnan perusteella sen toiseksi ylin ilmentymä on enintään korkeudella $k+1$ jäsennyspuun lehdistä, on tähän ilmentymään juurtuvan alipuun tuotokselle voimassa pituusraja $|vwx| \leq 2^{k+1} = n$. $\square$

Esimerkki. Tarkastellaan kieltä

$$L = \{a^k b^k c^k \mid k \geq 0\}.$$

Oletetaan, että L olisi yhteydetön; valitaan parametri n lemman mukaisesti ja tarkastellaan merkkijonoa $z = a^n b^n c^n \in L$.

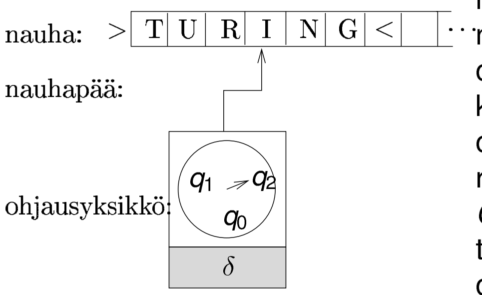
Lemman 3.9 mukaan z voidaan jakaa pumpattavaksi osiin

$$z = uvwxy, \qquad |vx| \geq 1, \qquad |vwx| \leq n.$$

Viimeisen ehdon takia merkkijono vx ei voi sisältää sekä a:ta, b:tä että c:tä. Merkkijonossa $uv^0wx^0y = uwy$ on siten ylijäämä jotakin merkkiä muihin merkkeihin nähden, eikä se voi olla kielen L määritelmässä vaadittua muotoa, vaikka lemman mukaan pitäisi olla $uwy \in L$.

4. TURINGIN KONEET

Alan Turing 1935–36.



Turingin kone on kuin äärellinen automaatti, jolla on käytössään nauha. Kone voi siirtää nauhapäätä vasemmalle tai oikealle; se voi myös lukea tai kirjoittaa nauhapään kohdalla olevan merkin. Nauha on oikealle rajaton.

Churchin–Turingin teesi: Mikä tahansa mekaanisesti ratkeava ongelma voidaan ratkaista Turingin koneella.

Turingin koneen kanssa ekvivalentteja laskentamalleja:

- Gödelin–Kleenen rekursiivisesti määritellyt funktiot (1936),
- Churchin λ-kalkyyli (1936),
- Postin (1936) ja Markovin (1951) merkkijonomuunnossysteemit, kaikki nykyiset ohjelmointikielet.

Turingin koneet $\equiv$ ohjelmointikieli.

Määritelmä 4.1 *Turingin kone* on seitsikko

$$M = (Q, \Sigma, \Gamma, \delta, q_0, q_{acc}, q_{rej}),$$

missä:

- Q on koneen *tilojen* äärellinen joukko;
- Σ on koneen *syöteaakkosto*;
- $\Gamma \supseteq \Sigma$ on koneen *nauha-aakkosto* (ol. että $>, < \notin \Gamma$);
- $\delta : (Q - \{q_{acc}, q_{rej}\}) \times (\Gamma \cup \{>, <\}) \to Q \times (\Gamma \cup \{>, <\}) \times \{L, R\}$ on koneen *siirtymäfunktio*;
- $q_0 \in Q$ on koneen *alkutila*;
- $q_{acc} \in Q$ on koneen *hyväksyvä* ja
- $q_{rej} \in Q$ sen *hylkäävä lopputila*.

Siirtymäfunktion arvon

$$\delta(q, a) = (q', b, \Delta)$$

tulkinta:

Ollessaan tilassa q ja lukiessaan nauhamerkin (tai alku- tai loppumerkin) a, kone siirtyy tilaan q', kirjoittaa lukemaansa paikkaan merkin b, ja siirtää nauhapäätä yhden merkkipaikan verran suuntaan Δ ($L \sim$ “left”, $R \sim$ “right”).

Sallittuja kirjoitettavia merkkejä ja siirtosuuntia on rajoitettu, mikäli a = ‘$>$’ tai ‘$<$’, ja siirtymäfunktion arvo on aina määrittelemätön, kun $q = q_{acc}$ tai $q = q_{rej}$. Joutuessaan jompaan kumpaan näistä tiloista kone pysähtyy heti.

Siis: siirtymäfunktion arvoilta

$$\delta(q, a) = (q', b, \Delta)$$

vaaditaan:

(i) jos $b = >$, niin $a = >$;

(ii) jos $a = >$, niin $b = >$ ja $\Delta = R$;

(iii) jos $b = <$, niin $a = <$ ja $\Delta = L$.

Koneen *tilanne* on nelikko

$$(q, u, a, v) \in Q \times \Gamma^* \times (\Gamma \cup \{\varepsilon\}) \times \Gamma^*,$$

missä voi olla $a = \varepsilon$, mikäli myös $u = \varepsilon$ tai $v = \varepsilon$.

Tulkinta: kone on tilassa q, nauhan sisältö sen alusta nauhapään vasemmalle puolelle on u, nauhapään kohdalla on merkki a ja nauhan sisältö nauhapään oikealta puolelta käytetyn osan loppuun on v.

Mahdollisesti on $a = \varepsilon$, jos nauhapää sijaitsee aivan nauhan alussa tai sen käytetyn osan lopussa. Ensimmäisessä tapauksessa ajatellaan, että kone “havaitsee” merkin ‘$>$’ ja toisessa tapauksessa merkin ‘$<$’.

Alkutilanne syötteellä $x = a_1 a_2 \ldots a_n$ on nelikko

$$(q_0, \varepsilon, a_1, a_2 \ldots a_n).$$

Tilannetta (q, u, a, v) merkitään yleensä yksinkertaisemmin $(q, u\underline{a}v)$, ja alkutilannetta syötteellä x yksinkertaisesti $(q_0, \underline{x})$.

Tilanne (q, w) *johtaa suoraan* tilanteeseen (q', w'), merkitään

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

jos jokin seuraavista ehdoista täyttyy: kaikilla $q, q' \in Q$, $u, v \in \Gamma^*$, $a, b \in \Gamma$ ja $c \in \Gamma \cup \{\varepsilon\}$:

jos $\delta(q, a) = (q', b, R)$, niin $(q, u\underline{a}cv) \underset{M}{\vdash} (q', ub\underline{c}v)$;

jos $\delta(q, a) = (q', b, L)$, niin $(q, uc\underline{a}v) \underset{M}{\vdash} (q', u\underline{c}bv)$;

jos $\delta(q, >) = (q', >, R)$, niin $(q, \underline{\varepsilon}cv) \underset{M}{\vdash} (q', \underline{c}v)$;

jos $\delta(q, <) = (q', b, R)$, niin $(q, u\underline{\varepsilon}) \underset{M}{\vdash} (q', ub\underline{\varepsilon})$;

jos $\delta(q, <) = (q', b, L)$, niin $(q, uc\underline{\varepsilon}) \underset{M}{\vdash} (q', u\underline{c}b)$;

jos $\delta(q, <) = (q', <, L)$, niin $(q, uc\underline{\varepsilon}) \underset{M}{\vdash} (q', u\underline{c})$.

Tilanteet, jotka ovat muotoa (q_{acc}, w) tai (q_{rej}, w) eivät johda mihinkään muuhun tilanteeseen. Näissä tilanteissa kone *pysähtyy*.

Tilanne (q, w) *johtaa tilanteeseen* (q', w'), merkitään

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

jos on olemassa tilannejono $(q_0, w_0), (q_1, w_1), \ldots, (q_n, w_n)$, $n \geq 0$, siten että

$$(q, w) = (q_0, w_0) \underset{M}{\vdash} (q_1, w_1) \underset{M}{\vdash} \cdots \underset{M}{\vdash} (q_n, w_n) = (q', w').$$

Turingin kone M *hyväksyy* merkkijonon $x \in \Sigma^*$, jos

$$(q_0, \underline{x}) \underset{M}{\vdash}^* (q_{\text{acc}}, w) \qquad \text{jollakin } w \in \Gamma^*;$$

muuten M *hylkää* x:n.

Koneen M *tunnistama kieli* on:

$$L(M) = \{x \in \Sigma^* \mid (q_0, \underline{x}) \underset{M}{\vdash}^* (q_{\text{acc}}, w) \text{ jollakin } w \in \Gamma^*\}.$$

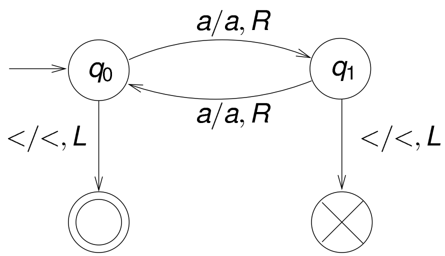
Esimerkki 1. Kieli $\{a^{2k} \mid k \geq 0\}$ voidaan tunnistaa Turingin koneella

$$M = (\{q_0, q_1, q_{\text{acc}}, q_{\text{rej}}\}, \{a\}, \{a\}, \delta, q_0, q_{\text{acc}}, q_{\text{rej}}),$$

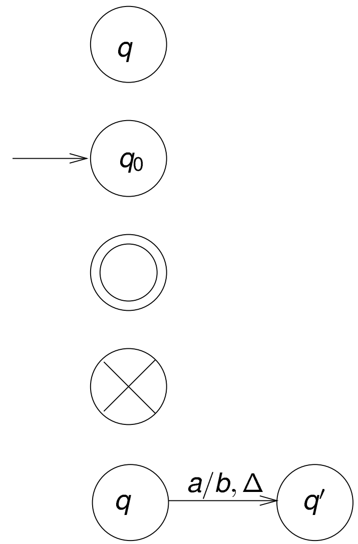
missä

$$\begin{aligned} \delta(q_0, a) &= (q_1, a, R), \\ \delta(q_1, a) &= (q_0, a, R), \\ \delta(q_0, <) &= (q_{\text{acc}}, <, L), \\ \delta(q_1, <) &= (q_{\text{rej}}, <, L). \end{aligned}$$

Kaavioesitys:



Kaavioesityksessä käytetyt merkinnät:



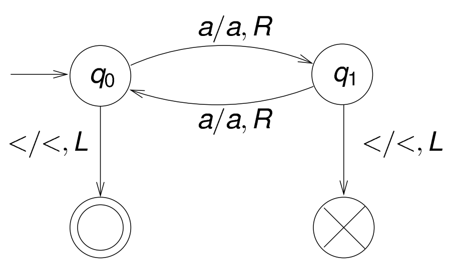
Tila q

Alkutila

Hyväksyvä lopputila (q_{acc})

Hylkäävä lopputila (q_{rej})

Tilasiirtymä $\delta(q, a) = (q', b, \Delta)$

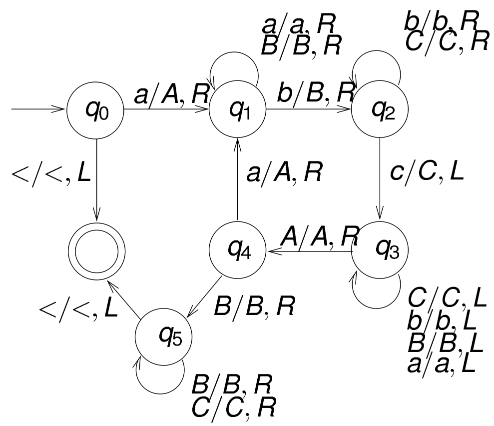


Koneen M laskenta esimerkiksi syötteellä *aaa* etenee seuraavasti:

$$\begin{aligned}(q_0, \underline{a}aa) \;&\underset{M}{\vdash}\; (q_1, a\underline{a}a) \underset{M}{\vdash} (q_0, aa\underline{a}) \\ &\underset{M}{\vdash} (q_1, aaa\underline{\varepsilon}) \underset{M}{\vdash} (q_{\mathrm{rej}}, aa\underline{a}).\end{aligned}$$

Kone pysähtyy tilassa q_{rej}, joten $aaa \notin L(M)$.

Esimerkki 2. Kielen $\{a^k b^k c^k \mid k \geq 0\}$ tunnistava Turingin kone:



Selkeyden vuoksi ei koneen hylkäävää lopputilaa q_{rej} ole tässä esitetty eksplisiittisesti. Tulkinta on tällöin, että kaikki kaaviosta "puuttuvat" kaaret johtavat tähän tilaan.

Koneen laskenta syötteellä *aabbcc*:

$(q_0, \underline{a}abbcc) \vdash (q_1, A\underline{a}bbcc) \vdash$
$(q_1, Aa\underline{b}bcc) \vdash (q_2, AaB\underline{b}cc) \vdash$
$(q_2, AaBb\underline{c}c) \vdash (q_3, AaB\underline{b}Cc) \vdash$
$(q_3, Aa\underline{B}bCc) \vdash (q_3, A\underline{a}BbCc) \vdash$
$(q_3, \underline{A}aBbCc) \vdash (q_4, A\underline{a}BbCc) \vdash$
$(q_1, AA\underline{B}bCc) \vdash (q_1, AAB\underline{b}Cc) \vdash$
$(q_2, AABB\underline{C}c) \vdash (q_2, AABBC\underline{c}) \vdash$
$(q_3, AABB\underline{C}C) \vdash (q_3, AAB\underline{B}CC) \vdash$
$(q_3, AA\underline{B}BCC) \vdash (q_3, A\underline{A}BBCC) \vdash$
$(q_4, AA\underline{B}BCC) \vdash (q_5, AAB\underline{B}CC) \vdash$
$(q_5, AABB\underline{C}C) \vdash (q_5, AABBC\underline{C}) \vdash$
$(q_5, AABBCC\underline{\varepsilon}) \vdash (q_{\mathrm{acc}}, AABBC\underline{C})$.