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

Säännölliset kielet

1. Osoita seuraavat kielet epäsäännöllisiksi käyttäen pumppauslemmaa (tai jollain muulla haluamallasi tavalla):

(a)

$$A = \{a^m b^n c^n \mid n, m \geq 1\}$$

Ensinnäkin huomataan, että A on säännöllinen täsmälleen silloin, kun $A^{\mathcal{R}}$ on säännöllinen, sillä aikaisemmin ollaan laskuharjoituksissa näytetty, että säännölliset kielet ovat suljettu kääntämisen suhteen. Riittää siis näyttää kieli $A^{\mathcal{R}}$ epäsäännölliseksi, missä

$$A^{\mathcal{R}} = \{c^n b^n a^m \mid n, m \geq 1\}.$$

Tehdään vastaoletus. Oletetaan, että kieli $A^{\mathcal{R}}$ on säännöllinen. Tällöin sillä on jokin pumppauspituus p. Nyt voidaan valita merkkijono

$$s = c^p b^p a \in A^{\mathcal{R}},$$

jolla pätee $|s| \geq p$. Säännöllisten kielten pumppauslemman nojalla s voidaan nyt jakaa kolmeen osaan siten, että

$$\begin{aligned} s &= xyz, \\ |xy| &\leq p \text{ ja} \\ |y| &> 0 \end{aligned}$$

joten

$$\begin{aligned} xy &= c^k && \text{jollain } k \leq p, \\ z &= c^{p-k} b^p a && \text{ja erityisesti} \\ y &= c^n && \text{jollain } n > 0. \end{aligned}$$

Nyt pumppauslemman nojalla myös merkkijonon

$$xz = c^{k-n} c^{p-k} b^p a = c^{p-n} b^p a$$

tulisi kuulua kieleen $A^{\mathcal{R}}$. Nyt kuitenkin $n > 0$, joten

$$xz = c^{p-n} b^p a \notin A.$$

Tämä on ristiriidassa säännöllisten kielten pumppauslemman kanssa, joten kielellä $A^{\mathcal{R}}$ ei ole pumppausominaisuutta, eikä se siten voi olla säännöllinen, joten myöskään kieli A ei ole säännöllinen.

(b)

Väite. *Kieli* $A = \{w \in \Sigma^* \mid w \textit{ on palindromi}\}$ *on epäsäännöllinen, missä* $\Sigma = \{a, b, c\}$

Todistus. Jos kieli A on säännöllinen, niin sillä on jokin pumppauspituus p. Valitaan kielestä A merkkijono $s = a^p b a^p$, jolla $|s| \geq p$. Nyt pumppauslemman nojalla s voidaan jakaa kolmeen osaan $s = xyz$, joilla pätee ehdot $|xy| \leq p$ ja $|y| > 0$. Nyt siis $x = a^n$, $y = a^k$ ja $z = a^{p-(n+k)} b a^p$, missä $k > 0$. Pumppauslemman ensimmäisen ehdon nojalla myös merkkijonon $xz = a^n a^{p-(n+k)} b a^p = a^{p-k} b a^p$ pitäisi kuulua kieleen A, mutta koska $k > 0$, niin tämä ei pidä paikkansa. Siis kieli A ei voi olla säännöllinen. $\square$

(c)

Väite. *Kieli* $A = \{0^n 1 0^n \mid n \in N\}$ *on epäsäännöllinen.*

Todistus. Oletetaan vastoin, että A on säännöllinen. Nyt on olemassa pumppauspituus $p \in \mathbb{N}$. Valitaan A:n merkkijono

$$s = 0^p 1 0^p$$

jolla selvästi $|s| \geq p$. Nyt jos

$$s = xyz,\ |xy| \leq p \text{ ja } |y| > 0,$$

niin $xy = 0^k$ jollain $0 < k \leq p$ ja erityisesti $y = 0^n$ jollain $n \geq 1$. Kuitenkin

$$xz = 0^{k-n} 0^{p-k} 1 0^p = 0^{p-n} 1 0^p \notin A$$

sillä $p - n \neq p$, mikä on ristiriidassa pumppauslemman kanssa. Siis A ei voi olla säännöllinen. $\square$

2. Mitkä seuraavista kielistä ovat säännöllisiä, mitkä eivät (kielillä A_1 ja A_2 aakkostona $\{0, 1\}$, muilla $\{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$):

$$A_1 = \{0^n 1^m 0^n \mid n, m \in N\} \qquad A_2 = \{0^n 0^n \mid n \in N\}$$
$$A_3 = \left\{ww^{\mathcal{R}} \mid w \in \Sigma^*\right\} \qquad A_4 = \left\{wuw^{\mathcal{R}} \mid w, u \in \Sigma^+\right\}.$$
$$A_5 = \left\{wxw^{\mathcal{R}} \mid w \in \Sigma^*, x \in \Sigma\right\} \qquad A_6 = \{\mathrm{abca}^n \mathrm{b}^n \mathrm{c}^n \mid n \in N\}$$

Perustele. Voit käyttää hyväksi kaikkia tunnettuja säännöllisiä kieliä koskevia ominaisuuksia, etenkin edellisen tehtävän tuloksia.

A_1 ei ole säännöllinen, sillä merkkijono $s = 0^p 1 0^p$ ei pumppaudu.

A_2 on säännöllinen, sillä $L((00)^*) = A_2$.

A_3 ei ole säännöllinen, sillä merkkijono $s = a^p bba^p$ ei pumppaudu.

A_4 on säännöllinen, sillä $L(a\Sigma^+ a \cup b\Sigma^+ b \cup c\Sigma^+ c) = A_4$

A_5 ei ole säännöllinen, sillä merkkijono $s = a^p ba^p$ ei pumppaudu.

A_6 ei ole säännöllinen, sillä $A_6^{\mathcal{R}}$ ei ole säännöllinen. Tämä voidaan nähdä sillä, että $s = c^p b^p a^p cba \in A_6^{\mathcal{R}}$ ei pumppaudu.

Kontekstittomat kielet

3. Esitä kontekstittomat kieliopit, jotka tuottavat seuraavat aakkoston $\Sigma = \{0, 1\}$ kielet:

(a) parittoman mittaiset merkkijonot

$$\begin{aligned} S &\to MT \\ T &\to MMT \mid \varepsilon \\ M &\to 0 \mid 1 \end{aligned}$$

(b) merkkijonot, joilla on osamerkkijono 111

$$\begin{aligned} S &\to H111H \\ H &\to MH \mid \varepsilon \\ M &\to 0 \mid 1 \end{aligned}$$

(c) merkkijonot, joissa on ainakin kaksi merkkiä ja joiden ensimmäinen ja viimeinen merkki ovat samat

$$\begin{aligned} S &\to 1H1 \mid 0H0 \\ H &\to MH \mid \varepsilon \\ M &\to 0 \mid 1 \end{aligned}$$

(d) parittoman mittaiset merkkijonot, joiden ensimmäinen ja keskimmäinen merkki ovat samat.

$$\begin{aligned} S &\to 1T_1M \mid 0T_0M \\ T_1 &\to MT_1M \mid 1 \\ T_0 &\to MT_0M \mid 0 \\ M &\to 0 \mid 1 \end{aligned}$$

4. Esitä kontekstittomat kieliopit seuraaville kielille:

(a) $01^* \cup 10^*$

$$\begin{aligned} S &\to 0T_1 \mid 1T_0 \\ T_1 &\to 1T_1 \mid \varepsilon \\ T_0 &\to 0T_0 \mid \varepsilon \end{aligned}$$

(b) $\{0^n1^m \mid m, n \in N \text{ ja } m \geq n\}$

$$\begin{aligned} S &\to T_{01}T_1 \\ T_{01} &\to 0T_{01}1 \mid \varepsilon \\ T_1 &\to 1T_1 \mid \varepsilon \end{aligned}$$

(c) $\left\{0^n1^k0^m \mid m, n, k \in N \text{ ja } k = n + m\right\}$

$$\begin{aligned} S &\to T_{01}T_{10} \\ T_{01} &\to 0T_{01}1 \mid \varepsilon \\ T_{10} &\to 1T_{10}0 \mid \varepsilon \end{aligned}$$

(d) $\{a^nb^mc^m \mid m, n \in N\}$

$$\begin{aligned} S &\to AT_{bc} \\ A &\to aA \mid \varepsilon \\ T_{bc} &\to bT_{bc}c \mid \varepsilon \end{aligned}$$

(e) aakkoston $\{0, 1\}$ merkkijonot, joissa on yhtä paljon nollia ja ykkösiä.

$$S \to 0S1S \mid 1S0S \mid \varepsilon$$

5. Täydennä Jyrkin luentojen lauseen 2.3 todistus (s. 140) osoittamalla, että kontekstiton kielten luokka on suljettu myös konkatenaation ja tähtioperaation suhteen. Esitä todistus samalla tarkkuustasolla kuin luentomuistiinpanoissa esitetty yhdisteen tapaus.

Olkoon A ja B aakkoston Σ yhteydettömiä kieliä ja $A = L(G_A)$ ja $B = L(G_B)$ kieliopeilla $G_A = (V_A, \Sigma, R_A, S_A)$ ja $G_B = (V_B, \Sigma, R_B, S_B)$. Oletetaan $V_A \cap V_B = \emptyset$.

Väite. *Kieli $A \circ B$ on yhteydetön.*

Todistus. Luodaan uusi kielioppi, jossa on uusi lähtösymboli $S \notin V_A \cup V_B$, ja sääntö $S \to S_A S_B$ jolla tästä uudesta lähtösymbolista voi tuottaa alkuperäisten kielten lähtösymbolien katenaation.

$$\begin{aligned} G_{A\circ B} &= (V_{A\circ B}, \Sigma, R_{A\circ B}, S) \\ V_{A\circ B} &= V_A \cup V_B \cup \{S\} \\ R_{A\circ B} &= R_A \cup R_B \cup \{S \to S_A S_B\} \end{aligned}$$

□

Väite. *Kieli A^* on yhteydetön.*

Todistus. Luodaan uusi kielioppi, jossa on uusi lähtösymboli $S \notin V_A \cup V_B$, ja sääntö $S \to S_A S \mid \varepsilon$ joka mahdollistaa A:n merkkijonojen toistamisen.

$$\begin{aligned} G_{A^*} &= (V_{A^*}, \Sigma, R_{A^*}, S) \\ V_{A^*} &= V_A \cup \{S\} \\ R_{A^*} &= R_A \cup \{S \to S_A S \mid \varepsilon\} \end{aligned}$$

□

6. Voidaan osoittaa, että kieli $A = \{a^n b^n c^n \mid n \in N\}$ ei ole kontekstiton. (Tähän palataan myöhemmin kurssilla.) Käyttäen tätä tietoa hyväksi osoita, että kontekstiton kielten luokka ei ole suljettu leikkauksen suhteen. (*Vihje:* esitä A kahden kontekstittoman kielen leikkauksena.) Päättele edelleen, että kontekstittomien kielten luokka ei ole suljettu komplementoinnin suhteen.

   Väite. *Yhteydettömien kielten luokka ei ole suljettu leikkauksen suhteen.*

   Todistus. Tehtävässä 3 osoitimme antamalla kieliopin, että kieli $A = \{a^n b^m c^m \mid n, m \in \mathbb{N}\}$ on yhteydetön. Vastaavasti voidaan osoittaa yhteydettömäksi kieli $B = \{a^n b^n c^m \mid n, m \in \mathbb{N}\}$. Kieli A on siis merkkijonot joissa b ja c merkkejä on yhtämonta. Vastaavasti kieli B on merkkijonot joissa a ja b merkkejä on yhtämonta. Nyt leikkauskieli $A \cap B = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ josta tiedämmä ettei se ole yhteydetön. Siispä yhteydettömien kielten luokka ei ole suljettu leikkauksen suhteen. □

   Väite. *Yhteydettömien kielten luokka ei ole suljettu komplementin suhteen.*

   Todistus. Oletetaan vastoin, että yhteydettömät kielet ovat suljettu komplementin suhteen. Olkoon nyt A ja B yhteydettömiä kieliä. Tällöin

$$\begin{aligned} A \cup B \text{ yhteydetön } &\Rightarrow \overline{(A \cup B)} \text{ yhteydetön} \\ &\Rightarrow \overline{A} \cap \overline{B} \text{ yhteydetön} \\ &\Rightarrow A \cap B \text{ yhteydetön} \end{aligned}$$

   mikä on ristiriita edellä osoitetun kanssa. Siispä yhteydettömien kielten luokka ei voi olla suljettu komplementin suhteen. □

7. Osoita, että seuraavien aakkoston $\{a, b, c\}$ kielten komplementit ovat kontekstittomia:

   (a) $A_1 = \{a^n b^n \mid n \in N\}$

   Kielen A_1 komplementti $\overline{A_1}$ voidaan ilmaista kolmen kielen yhdisteenä:

$$\begin{aligned} \overline{A_1} =& \{a^n b^m \mid n, m \in \mathbb{N}, n \neq m\} \cup \overline{L(a^* b^*)} \\ =& \{a^n b^m \mid n, m \in \mathbb{N}, n < m\} \\ \cup& \{a^n b^m \mid n, m \in \mathbb{N}, n > m\} \\ \cup& \overline{L(a^* b^*)} \end{aligned}$$

Tässä kieli $\overline{L(a^*b^*)}$ on säännöllisen kielen komplementtina säännöllinen ja siten sille on jokin kontekstiton kielioppi. Lisäksi kontekstittomat kielet ovat suljettu yhdisteen suhteen, joten riittää keksiä kielioppi lopulle osalle yllä olevaa yhdistettä. Tälle kielelle voidaan antaa seuraavanlainen kielioppi

$$\begin{aligned} S &\to aT_aT_{ab} \mid T_{ab}T_bb \\ T_a &\to aT_a \mid \varepsilon \\ T_{ab} &\to aT_{ab}b \mid \varepsilon \end{aligned}$$

(b) $A_2 = \{a^nb^nc^n \mid n \in N\}$

Kieli A_2 komplementti voidaan ilmaista seuraavanlaisena yhdisteenä:

$$\begin{aligned} \overline{A_2} = &\left\{ a^nb^mc^k \mid n, m, k \in \mathbb{N} : n \neq m \text{ tai } m \neq k \right\} \\ &\cup \overline{L(a^*b^*c^*)}. \end{aligned}$$

Kieli $\overline{L(a^*b^*c^*)}$ on säännöllisen kielen komplementtina säännöllinen ja täten sille löytyy jokin kontekstiton kielioppi G. Koska konstekstittomat kielet ovat suljettuja yhdisteen suhteen, riittää enää kirjoittaa kielioppi unionin ensimmäiselle kielelle.

$$\begin{aligned} S &\to aT_aT_{ab}T_c \mid T_{ab}T_bbT_c \\ &\quad \mid T_abT_bT_{bc} \mid T_aT_{bc}T_cc \\ T_a &\to aT_a \mid \varepsilon \\ T_b &\to bT_b \mid \varepsilon \\ T_c &\to cT_c \mid \varepsilon \\ T_{ab} &\to aT_{ab}b \mid \varepsilon \\ T_{bc} &\to bT_{bc}c \mid \varepsilon \end{aligned}$$