Pumppauslemmaopas

Juhana Laurinharju Jani Rahkola

12. lokakuuta 2012

Pumppauslemma

Jokaisella säännöllisellä kielellä on seuraava *pumppauslemmana* tunnettu ominaisuus.

Määritelmä (pumppauslemma)**.** Olkoon A säännöllinen kieli. Tällöin A:lla on jokin pumppauspituus $p \in \mathbb{N}$, $p > 0$. Nyt kaikki vähintään p:n mittaiset merkkijonot $s \in A$, $|s| \geq p$ voidaan jakaa kolmeen osaan $s = xyz$ siten, että seuraavat kolme ehtoa ovat voimassa.

1. $xy^iz \in A$ jokaisella $i \in \mathbb{N}$, erityisesti myös kun $i = 0$.
2. $|y| > 0$, eli toistettava osa y ei saa olla tyhjä merkkijono ε.
3. $|xy| \leq p$

Otetaan tästä konkreettinen esimerkki. Kieli $A = L(a^*b^*)$ on säännöllinen ja sillä on täten jokin pumppauspituus p. Tämän kielen kohdalla eräs mahdollinen pumppauspituus on $p = 3$. Tämän pumppauspituuden saa esimerkiksi siitä, että kielen voi tunnistaa deterministisellä äärellisellä automaatilla jossa on kolme tilaa. Tarkastellaan jotain riittävän pitkää kielen A merkkijonoa. Valitaan $s = abbb$. Nyt pumppauslemman nojalla löytyy *jokin* jako $s = xyz$, jolle yllä olevat kolme ehtoa pätevät.

Tarkastellaan merkkijonon s mahdollisia jakoja.

- Kokeillaan ensin seuraavaa jakoa:

$$\begin{aligned} x &= \varepsilon, \\ y &= ab \\ \text{ja } z &= bb \end{aligned}$$

  Nyt ehdot 2 ja 3 ovat voimassa, mutta ensimmäinen ehto ei täyty, sillä esimerkiksi merkkijono $xyyz = ababbb$ ei kuulu kieleen A.

- Ensimmäinen jako ei siis täyttänyt kaikkia pumppauslemman ehtoja. Pumppauslemma ei kuitenkaan takaa, että nämä ehdot täyttyisivät jokaisella jaolla. Ainoa tae on se, että löytyy jokin ehdot täyttävä jako. Tällainen on esimerkiksi seuraava jako:

$$\begin{aligned} x &= a, \\ y &= b \\ \text{ja } z &= bb \end{aligned}$$

Nyt $|xy| = 2 \leq p$, $|y| = 1 > 0$. Entä miltä näyttää merkkijono $xy^i z$? Tarkastellaan tätä ensin i:n arvoilla 0, 1 ja 2:

$$\begin{aligned} xy^0 z &= xz = abb \in A \\ xy^1 z &= xyz = abbb \in A \\ xy^2 z &= xyyz = abbbb \in A \end{aligned}$$

Ja yleisessä tapauksessa, kun $i \in \mathbb{N}$, niin

$$xy^i z = ab^i bb = ab^{i+2} \in L(a^* b^*) = A$$

Tehtävä 1. Olkoon

$$A = L((ab)^*)$$

säännöllinen kieli. Tällä kielellä on pumppauspituus $p = 2$. Valitaan kielestä merkkijono

$$s = ababab$$

joka on pidempi kuin pumppauspituus $p = 2$. Anna kaikki merkkijonon s jaot $s = xyz$ jotka täyttävät pumppauslemman ehdot 2 ja 3, eli

$$\begin{aligned} |y| &> 0 \text{ ja} \\ |xy| &\leq p. \end{aligned}$$

Tehtävä 2. Mitkä edellisen tehtävän jaoista $s = xyz$ toteuttavat pumppauslemman ensimmäisen ehdon? Ensimmäinen ehto on

$$xy^i z \in A \text{ kaikilla } i \in \mathbb{N}$$

eli keskikohtaa y voi toistaa.

Pumppauslemman käyttäminen

Pumppauslemma on hyödyllinen, koska sillä voidaan näyttää monia kieliä epäsäännöllisiksi. Mutta kuinka tämä onnistuu työkalulla, joka ei puhu mitään epäsäännöllisistä kielistä? Pumppauslemman jos-niin rakenteeseen piiloutuu kuitenkin myös väite epäsäännöllisyydestä. Seuraava lemma on nimittäin yhtäpitävä pumppauslemman kanssa.

Lemma. Olkoon A jokin kieli. Oletetaan lisäksi, että kaikilla luonnollisilla luvuilla $p > 0$ jotain kielen A merkkijonoa s, $|s| \geq p$ ei voida jakaa *millään tapaa* kolmeen osaan $s = xyz$ siten, että seuraavat ehdot olisivat voimassa:

1. $xy^iz \in A$ jokaisella $i \in \mathbb{N}$, erityisesti myös kun $i = 0$.
2. $|y| > 0$, eli toistettava osa y ei saa olla tyhjä merkkijono ε.
3. $|xy| \leq p$

Nyt kieli A on epäsäännöllinen.

Todistus. Olkoon A kieli jolla ei ole pumppauspituutta. Oletetaan vastoin väitettä, että A on säännöllinen. Nyt pumppauslemman nojalla kielellä A kuitenkin tulisi olla pumppauspituus. Siis A ei voi olla säännöllinen. □

Pumppauslemman avulla voidaan siis todistaa epäsäännöllisiksi vain sellaiset kielet, joilla ei ole pumppauspituutta. Tulee kuitenkin muistaa, että joillain epäsäännöllisilläkin kielillä on olemassa pumppauspituus.

Käydään läpi todistus erään jo tutun kielen epäsäännöllisyydelle. Olkoon kieli $A = \{0^n1^n \mid n \in \mathbb{N}\}$. Osoitetaan tämän kielen epäsäännöllisyys yllä todistetun lemman avulla.

Väite. *Kieli* $A = \{0^n1^n \mid n \in \mathbb{N}\}$ *on epäsäännöllinen.*

Todistus. Olkoon p positiivinen luonnollinen luku. Nyt haluttaisiin jokin merkkijono $s \in A$, jolla $|s| \geq p$. FIXME: Tää virke on liian pitkä Merkkijonon s tulisi olla sellainen, että jos sen jakaa osiin $s = xyz$ niin, että pumppauslemman ehdot $|xy| \leq p$ ja $y \neq \varepsilon$ ovat voimassa, niin toistettava osa y sisältää kielen A epäsäännöllisen ominaisuuden. Tämän ominaisuuden ei tulisi säilyä kun osaa y toistaa monta kertaa tai sen poistaa.

Valitaan merkkijono
$$s = 0^p1^p \in A$$
joka kuuluu kieleen A ja sisältää vähintään p merkkiä. Haluamme osoittaa, että mikään merkkijonon s jako $s = xyz$, joka noudattaa lemman ehtoja 2 ja 3, ei voi noudattaa ehtoa 1. Merkkijonoon s valittiin p kappaletta nollia ja ykkösiä, jotta kaikki ensimmäiset p merkkiä olisivat samoja, tässä tapauksessa nollia. Koska y on jokin ensimmäisen p:n merkin alimerkkijono, y on muotoa 0^k ja kaikki 1-merkit jäävät loppuosaan z. Nyt jos osaa y toistaa, eli

muodostaa esimerkiksi merkkijonon $xyyz$, niin 0-merkkien lukumäärä muuttuu. Kaikki 1-merkit ovat kuitenkin loppuosassa z, joten niiden lukumäärä säilyy ennallaan eikä merkkijono $xyyz$ enää toteuta kielen A ehtoa.

Merkkijonon s ensimmäiset p merkkiä ovat nollia, joten kaikki mahdolliset ehtojen 2 ja 3 mukaiset jaot ovat olennaisesti samaa muotoa. Ne voidaan kuvata seuraavasti

$$\begin{aligned} x &= 0^n \\ y &= 0^m,\ m > 0 \\ \text{ja } z &= 0^{p-(m+n)}1^p \end{aligned}$$

missä $n, m \in \mathbb{N}$. Jotta y ei olisi tyhjä merkkijono, täytyy m:n olla positiivinen. Jotta merkkijonon xy pituus olisi korkeintaan p, tulee summan $n + m$ olla korkeintaan p. Tarkastellaan merkkijonoa, jossa osa y esiintyy kaksi kertaa

$$\begin{aligned} xy^2z &= xyyz \\ &= 0^n 0^{2m} 0^{p-(m+n)} 1^p \\ &= 0^{n+2m+(p-(m+n))} 1^p \\ &= 0^{n+2m+p-m-n} 1^p \\ &= 0^{p+m} 1^p \end{aligned}$$

Koska m on aidosti positiivinen luku, niin merkkijonossa xy^2z on enemmän nollia kuin ykkösiä. Se ei siis kuulu kieleen A. Ehto 1 ei voi olla voimassa, joten lemman nojalla A on epäsäännöllinen. □

Tehtävä 3. Osoita kieli $A = \{0^n 1^m \mid n, m \in \mathbb{N}, n < m\}$ epäsäännölliseksi lemman avulla.

Tehtävä 4. Osoita kieli $A = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ epäsäännölliseksi lemman avulla.