

Harjoituksia Pumppauslemmasta ja kontekstittomista kieliopista

- Ovatko seuraavat kielet säännöllisiä? Perustele!
 - $\{w \mid w \text{ on } a\text{:sta ja } b\text{:stä koostuva merkkijono, jonka pituus on } 3\}$
 - $\{ww \mid w \in \{a, b\}^*\}$
 - $\{w^* \mid w\text{:ssä yhtä monta } 1\text{:ä ja } 0\text{:aa}\}$
 - $\{w \mid w\text{:ssä kaksi kertaa niin monta } 0\text{:aa kuin } 1\text{:ä}\}$
- Tiedetään, että yleisessä tapauksessa ongelma $REG(L)$ (ts. onko annettu kieli L säännöllinen vai ei) on ratkeamaton. Mistä tämä johtuu? Voisitko silti laatia ohjelman, joka auttaisi ihmistä osoittamaan Pumppauslemmalla, että kieli on epäsäännöllinen?
- Vahvempi Pumppauslemman variaatio antaa sekä riittävät että välttämättömät ehdot kielen säännöllisyydelle:

Pumppauslemma 2: Kieli $A \in \Sigma^*$ on säännöllinen, jos ja vain jos on olemassa vakio $n \geq 1$ s.e. kaikilla $x \in \Sigma^*$, jos $|x| \geq n$ on olemassa u, v, w s.e. $x = uvw$ ja $|v| \geq 1$ ja kaikilla $i \geq 0$ ja kaikilla $y \in \Sigma^*$ $xy \in A$ jos ja vain jos $uv^iwy \in A$.

Voiko tämän pohjalta laatia ohjelman, joka ratkaisee minkä tahansa kielen A osalta, onko A säännöllinen vai ei?
- Osoita pumppauslemmalla, että seuraavat kielet eivät ole säännöllisiä, ja muodosta kielen kuvaava kontekstiton kielioppi:
 - $\{a^n b^n c^k \mid n, k = 0, 1, \dots\}$
 - $\{a^n b^k c^k \mid n, k = 0, 1, \dots\}$
 - $\{a^n b^n a^m b^m \mid n, k = 0, 1, \dots\}$
- Seuraavien kontekstittomien kielioppien tuottamat kielet ovat säännöllisiä. Kuvaa kielet säännöllisinä lausekkeina:

- $S \rightarrow AS \mid \epsilon$
 $A \rightarrow a \mid b$

- (b) $S \rightarrow SSS \mid a \mid b$
 (c) $S \rightarrow AB$
 $A \rightarrow aAa \mid bAb \mid \epsilon$
 $B \rightarrow aB \mid bB \mid \epsilon$

6. Ulkoavaruuden olioiden kielessä kaikki sanat noudattavat Blurbsin muotoa. Blurb on Whoozit, jota seuraa yksi tai useampi Whatzit. Whoozit on kirjain 'x', jota voi seurata haluttu määrä 'y'-kirjaimia (myös nolla (0) kappaletta). Whatzit on kirjain 'q', jota seuraa joko kirjain 'z' tai 'd', jota seuraa Whoozit. Anna Blurbsin kielen kuvaava kontekstiton kielioppi. Osaatko nyt laatia vastaavan säännöllisen lausekkeen? (Vihje: piirrä ensin automaatti kieliopista!)
7. Säännöllisille lausekkeille pätee seuraava päättelysääntö:
 Jos $r = rs \cup t$, niin $r = ts^*$, kun $\epsilon \notin L(s)$.
 Osoita tämä sääntö todeksi kontekstittomien kielioppien avulla!
8. Laadi oikealle lineaarinen kielioppi kielen $\{w \in \{a, b\}^* \mid w \text{ ei sisällä merkkijonoa } abaa \text{ kuvaamiseen}\}$
9. Laadi kontekstiton kielioppi, joka tuottaa seuraavan esimerkin tapaiset, rajattoman monista sisäkkäisistä for-silmukoista ja alkeisoperaatioista a koostuvat ohjelmointikielen lauseet:

```
for (i = N; i < N; i++) {
  for (j=N; j<N; j++) {
    a;
  }
}
```

missä N on kokonaisluku.

10. Kuvaa seuraavanlaiset C-kielen ohjelmat kontekstittomana kielioppina.

- funktio on muotoa

```
fkttyyppi fuktnimi(partyyppi par (,partyyppi par)*)
{
  fktrunko
}
```

missä fkttyyppi voi olla void tai int ja partyyppi on int

- fktrunko on muotoa iflause tai sijoitus
- sijoitus on muotoa muuttuja=arvo;
- par ja muuttuja ovat 1-kirjaimisia muuttujan nimiä
- arvo on kokonaisluku
- iflause on muotoa
if (ehto) iflause sijoitus; tai
if (ehto) iflause sijoitus; else iflause sijoitus; tai
tyhjä merkkijono
- ehto on muotoa (muuttuja==arvo)

Exercises about Pumping Lemma and context-free grammars

1. Are the following languages regular? State the reasons for your answer.
 - (a) $\{w|w \text{ is a character string of length 3 consisting of } a\text{'s and } b\text{'s}\}$
 - (b) $\{ww|w \in \{a, b\}^*\}$
 - (c) $\{w^*|\text{there are as many 1's as 0's in } w\}$
 - (d) $\{w|\text{there are twice as many 0's as 1's in } w\}$
2. We know that in general the problem $REG(L)$ (i.e. is the given language L regular or not) is unsolvable. What is the reason for that? Could you still invent a program, which would help people to prove by the Pumping Lemma that the given language is nonregular?
3. The stronger version of the Pumping Lemma gives both sufficient and necessary conditions for the regularity of a language:

Pumping Lemma 2: The language $A \in \Sigma^*$ is regular if and only if there is a constant $n \geq 1$ such that for all $x \in \Sigma^*$, if $|x| \geq n$ then there exists u, v, w such that $x = uvw$ and $|v| \geq 1$, and for all $i \geq 0$ and all $y \in \Sigma^*$ $xy \in A$ if and only if $uv^iwy \in A$.

Based on this, could you make a program, which decides for any given language A , if it is regular or not?

4. Prove with the help of the Pumping lemma that the following languages are not regular, and create a context-free grammar describing the language.

(a) $\{a^n b^n c^k \mid n, k = 0, 1, \dots\}$

(b) $\{a^n b^k c^k \mid n, k = 0, 1, \dots\}$

(c) $\{a^n b^n a^m b^m \mid n, k = 0, 1, \dots\}$

5. The languages produced by the following context-free grammars are regular. Describe the languages with regular expressions.

(a) $S \rightarrow AS \mid \epsilon$
 $A \rightarrow a \mid b$

(b) $S \rightarrow SSS \mid a \mid b$

(c) $S \rightarrow AB$
 $A \rightarrow aAa \mid bAb \mid \epsilon$
 $B \rightarrow aB \mid bB \mid \epsilon$

6. All words of the language of outerspace aliens follows the Blurbs normalform. Blurb is a Whoozit, followed by one or more Whatzit. Whoozit is letter 'x', which can be followed by any number of letters 'y' (also zero). Whatzit is letter 'q', which is followed by either letter 'z' or 'd', followed by Whoozit.

Give the context-free grammar, which describes the Blurbs language. Can you now give the corresponding regular expression? (Hint: construct first the automaton from the grammar!)

7. For regular expression holds the following rule::

If $r = rs \cup t$, then $r = ts^*$, when $\epsilon \notin L(s)$.

Show, why this rule holds with the help of context-free grammars!

8. Create a grammar that is linear to the right and describes the language

$$\{w \in \{a, b\}^* \mid w \text{ does not contain character string } abaa\}$$

9. Create a context-free grammar that produces programming language expressions consisting of an infinite number of 'for' loops inside each other and elementary operations a , in the same way as in the example.

```

for (i = N; i < N; i++) {
    for (j=N; j<N; j++) {
        a;
    }
}

```

in which N is an integer.

10. Describe the following C-programs as context-free grammar!

```

for (i = N; i < N; i++) {
    for (j=N; j<N; j++) {
        a;
    }
}

```

in which N is integer.

- the function takes the form

```

fctype fctname(partype par (,partype par)*)
{
    fcttrunk
}

```

in which fctype may be void or int and partype is int.

- fcttrunk is an ifexpression or assignment
- the assignment takes the form variable=value;
- par and the variable are variable names consisting of one letter
- the value is an integer
- the ifexpression takes the form
if (condition) ifexpression assignment; or
if (condition) ifexpression assignment; else ifexpression assignment; or
an empty character string
- the condition takes the form (variable=value)