

Harjoituksia LL(1)-kieliopeista, CYK-agoritmista ja Chomskyn normaalimuodosta

1. Täyttääkö seuraava kielioppi LL(1)-ehdon?

$$\begin{aligned} S &\rightarrow (L)|p|q \\ L &\rightarrow LandS|LorS|S \end{aligned}$$

Millaisen kielen kielioppi kuvaa?

2. Laadi (pseudokoodina) rekursiivinen jäsentäjä kieliopin kuvaamalle kielelle (Huom! Muunna ensin tarvittaessa LL(1):een!).
3. Olkoon L kieli, jonka sanat koostuvat mistä tahansa tekstistä sekä laillisista sulkurakenteista. Ts. kuten ongelmassa 8, mutta nyt joukossa voi esiintyä tekstiä. Esimerkiksi lause ”(suuri) kissaeläin (joko leijona tai tiikeri { jotka harvinaisia } tai sapelihammastiikeri { joka sukupuuttoon kuollut [ks. tarkemmin Kurten] – harmi –} tai leopardi) on viehättävä (mutta vaarallinen!) tuttavuus.” kuuluu kieleen. Voit olettaa yksinkertaisuuden vuoksi, että tekstiosuudet sisältävät vain pieniä kirjaimia a..z sekä välilyöntejä.

Anna kielen kuvaava LL(1)-kielioppi ja suunnittele sille rekursiivinen jäsentäjä!

4. Osoita, että seuraavat kielet ovat deterministisiä:

$$\begin{aligned} \text{a)} & \{a^m b^n \mid m \neq n\} \\ \text{b)} & \{wcw^R \mid w \in \{a, b\}^*\} \\ \text{c)} & \{a^m c b^m\} \cup \{a^m d b^{2m}\} \end{aligned}$$

5. Muunna kielioppi

$$\begin{aligned} S &\rightarrow (S)|A \\ A &\rightarrow SS|\epsilon \end{aligned}$$

Chomskyn normaalimuotoon. Esitä myös välivaiheet (ϵ -produktioiden poisto, yksikköproduktioiden poisto)!

6. Muunna kielioppi

$$\begin{aligned} S &\rightarrow ABC|a \\ A &\rightarrow aAaa|\epsilon \\ B &\rightarrow bBbb|\epsilon \\ C &\rightarrow cCa|c \end{aligned}$$

Chomskyn normaalimuotoon. Esitä myös välivaiheet!

7. Simuloi CYK-algoritmin toimintaa sen ratkaistessa, kuuluvatko merkkijonot bbaab, ababab ja aabba kieliopin

$$S \rightarrow AS|b$$

$$A \rightarrow SA|a$$

tuottamaan kieleen. Myönteisissä tapauksissa esitä merkkijonoille kieliopin mukaiset jäsennesspuut.

8. Millä tavalla jäsentäisit seuraavat kielet? Perusstrategia riittää.
- HTML
 - C- tai Pascal-ohjelmakoodi
 - SQL-kysely
9. Olkoon $L_1 = \{a^n b^{2n} c^m | n, m \geq 0\}$ ja $L_2 = \{a^n b^m c^{2m} | n, m \geq 0\}$. Onko $L_1 \cap L_2$ kontekstiton kielioppi? Perustele vastauksesi!
10. Laadi tehokas algoritmi, joka tutkii, onko välikesymboli tyhjentyvä (ts. voiko se tuottaa ϵ :in jossain lausejohdoksessa)!
11. Osoita, että kontekstittomat kielet ovat suljettuja yhdisteen, katenaation ja sulkeuman suhteen. Ts. jos L_1 ja L_2 ovat kontekstittomia, niin
- $L_1 \cup L_2$
 - $L_1 L_2$
 - L_1^*
- ovat kontekstittomia. (Vihje: oletta, että on olemassa pinoautomaatit $M(L_1)$ ja $M(L_2)$ ja laadi sulkeumakielille pinoautomaatit.)
12. Kieltä $\{a^n b^n c^n | n \geq 0\}$ ei voida tunnistaa tavallisella pinoautomaatilla. Voisiko sen tunnistaa, jos käytössä olisi kaksipinoinen automaatti? Jos voi, niin piirrä ko. automaatin siirtymäkaavio ja simuloi sen toimintaa! Jos ei, niin perustele, miksi ei!
13. Laadi kielioppi, joka kuvaa html:n lailliset $\langle ol \rangle$ - ja $\langle ul \rangle$ -listarakenteet, joissa listatut kohdat ovat mitä tahansa tekstiä (ei erikoismerkkejä) tai uusia alilistoja. Ts. esim. $\langle ol \rangle \langle li \rangle$ Kissa $\langle li \rangle$ Koirat $\langle ul \rangle \langle li \rangle$ Susi $\langle li \rangle$ Kettu $\langle li \rangle$ Hyeena $\langle /ul \rangle \langle li \rangle$ Kani $\langle li \rangle$ Rotta $\langle /ol \rangle$ kuuluu kieleen. Millaisella jäsennessalgoritmilla tutkisit, onko annettu tekstipätkä kieliopin mukainen?