

Oppgaver INF 2820 V2011

Gruppe 10.3

Vi har sett at en RD-parser, som virker ovenifra og ned (TD) og fra venstre mot høyre, har problemer med venstrerekursjon. Forrige uke så vi at vi til en gitt grammatikk G kan finne en svakt ekvivalent grammatikk som unngår venstrerekursjon. På forelesning har vi sagt at en chartparser unngår problemene med venstrerekursjon selv om den innfører aktive kanter ovenifra og ned. Vi skal se mer på dette i dag og i oblig. 2a. Å løse denne oppgaven vil være til stor hjelp for å løse oblig. 2a.

Selv om en chartparser er forholdsvis effektiv, blir det en del arbeid når en skal parse for hånd. Vi skal derfor bruke et forenklet eksempel. Vi bruker adverb som eksempel. Vær klar over at plasseringen av adverb i norske setninger kan være litt vanskelig å beskrive enkelt med en kontekstfri grammatikk. Følgende grammatikk er derfor forenklet og ikke riktig lingvistisk motivert.

La G være grammatikken:

1. S → NP VP
2. NP → DET NOM
3. VP → IV
4. IV → IV ADV
5. NOM → N
6. DET → ϵ
7. N → studenten | læreren
8. DET → den | denne
9. IV → smilte | lo | sov
10. ADV → ikke | ofte | raskt | lenge | sjelden

Oppgave 1

Se på setningen

Studenten smilte ikke ofte lenge

Parse setningen med en chart-parser. Bruk samme oppsett som på innleveringssett 2a, oppgave 1e. Bruk en TD-strategi for innføring av aktive (inkomplette) kanter. Legg inn alle leksikalske oppslag i agenda først som initialisering.

Oppgave 2

I oppgave 1 bruker vi den venstrerekursive regelen $IV \rightarrow IV\ ADV$ gjentatte ganger. Gir dette problemer med venstrerekursjon? Observer hvorfor det ikke oppstår problem.

Oppgave 3

Anta nå at du skal gjøre oppgave 1 med en annen strategi der du skal innføre de aktive kantene BU. Hvorfor og hvordan er det problematisk? (Du behøver ikke gjennomføre parset.)

Oppgave 4

Samme oppgave som oppgave 1, men med noen endringer. Du bruker fremdeles en TD-strategi for innføring av aktive kanter, men de leksikalske oppslagene innføres ikke ved initialisering. I stedet går

du fra venstre mot høyre. Når agenda er tom, flytter du et ord mot høyre og legger de tilhørende leksikalske kantene til agenda.

Oppgave 5

Sammenlikn det du gjør i oppgave 3 med Jurafsky og Martins beskrivning av Earleys algoritme. Kan du gjenkjenne subprosedyrene deres i det du gjør?

Blir det tid til overs, arbeid med oblig 2A.

- SLUTT -