

# Oppgaver INF 2820 V2011

## Obligatorisk innleveringsoppgave 2, deloppgave A

*Frist 14.3 – strikt – se beskrivelse på semestersiden*

### Oppgave 1:

La G1 være følgende grammatikk

1. S → NP VP
2. NP → NP og NP
3. NP → NP som VP
4. VP → IV
5. VP → TV NP
6. NP → studenten | læreren | vennen | boka | eplet
7. IV → sov | smilte | spiste | leste
8. TV → spiste | leste | så | skrev | kjente

Og la (A) være følgende ordsekvens

A) Studenten og vennen som kjente læreren som smilte sov.

a. En av setningsstrukturene som G1 tilskriver til (A) kan skrives med parentesnotasjon som følgende: (kall denne A1):

```
(S (NP (NP Studenten)
      og
      (NP (NP vennen)
          som
          (VP (TV kjente)
              (NP (NP læreren)
                  som
                  (VP (IV smilte)))))))
  (VP (IV sov)))
```

Hvor mange flere strukturer tilskriver G1 til strengen (A)? Vi tar her ikke hensyn til at ikke alle analysene er god norsk. Tegn opp tilsvarende parentesnotasjon for alle de øvrige analysene av (A)

b. Vis hvordan en SR-parser kan anerkjenne (A) med strukturen (A1). Bruk følgende format:

```
| studenten og vennen som kjente læreren som smilte sov
studenten |og vennen som kjente læreren som smilte sov      (FLYTT)
NP |og vennen som kjente læreren som smilte sov            (REDUSER)
NP og | vennen som kjente læreren som smilte sov          (FLYTT)
Osv.
```

Du behøver bare vise et vellykket parse. Du trenger ikke å backtracke eller å finne mer enn den ene analysen.

c. Vi skal nå bruke en tabellparser i stedet. Vi vil først bruke CKY. Da må grammatikken være på CNF (Chomsky-normalform). Finn en grammatikk G2 på CNF som er svakt ekvivalent til G1.

d. Gjør et fullstendig CKY-parse av (A) med G2. Dette kan du gjøre ved å lage en tabell og fylle ut rutene med symboler. Nummerer symbolene etter hvert som de legges inn i tabellen slik at det er mulig å se hvordan du har gått frem.

e. Vi skal nå bruke en chart-parser og grammatikk G1 til å analysere streng (A). Du kan selv velge hvilken variant av chartparse du bruker, men skriv opp hvilken strategi du har valgt. Du skal levere inn fullstendig chart og agendaen som er brukt. Chartet vil bestå av en mengde kanter. De kan enten legges inn i en tabell a la en CKY-tabell (og være på form som  $(S \rightarrow NP . VP)$ ) eller de kan skrives etter hverandre (og da være på form  $(S \rightarrow NP . VP, [0,1])$ ). Kantene skal nummereres i den rekkefølgen de legges inn i chartet. Agendaen skal bestå av alle kanter som har vært innom agendaen, på formen  $(S \rightarrow NP . VP, [0,1])$ . De skal nummereres med samme nummer som de får når de legges over i chartet. Dette nummeret markerer altså når de tas ut av agendaen. I tillegg skal det på hver av dem markeres hvilke kanter det er som har konstruert dem. For eksempel slik

9.  $(S \rightarrow NP VP, [0,0])$  fra 3

...

47.  $(S \rightarrow NP . VP, [0,1])$  fra 9 og 15

48.  $(S \rightarrow NP VP ., [0,9])$  fra 28 og 47

## Oppgave 2:

(= Oppgave 13.3, s. 458 i brun bok:) Modifiser CKY-algoritmen på figur 13.10 slik at den kan takle unære produksjonsregler, altså regler på formen  $A \rightarrow B$ , der B er en ikke-terminal.

Hint:

- I konstruksjonen av en celle, la unære regler virke etter andre regler.
- Vær obs på at grammatikken kan inneholde cykliske unære produksjoner, for eksempel  $A \rightarrow B, B \rightarrow C, C \rightarrow A$ . Dette må algoritmen kunne takle. Samtidig må den terminere.

- SLUTT -