

INF2820 Datalingvistik (vår 2011)

— Prøveeksamen —

1 Endelige tilstandamaskiner og regulære uttrykk

- (a) Vi skal lage en FSA for talluttrykk i norsk som *to hundre og femtitre*. Vi skal lage en FSA som skriver tallene med orddeling, *to hundre og femti tre*, og går ut i fra at en annen modul står for sammenskrivning. Alfabetet vil altså bestå av norske ord

$$A = \{en, to, \dots, ti, \dots, nitten, tjue, tretti, \dots, nitti, hundre, tusen, \dots, og\}$$

Lag en FSA for alle tall mellom en og ni hundre og nittini. Du kan bruke forkortelser som $EN = \{en, to, \dots, ni\}$.

- (b) Lag en FSA for alle uttrykk opp til ni hundre og nittini tusen ni hundre og nittini.
- (c) Er FSA-en du lagde deterministisk? Hvis ikke, lag en deterministisk FSA (en DFA) for det samme språket.
- (d) Lag et regulært uttrykk for språket. Du kan bruke de samme forkortelsen som for FSA-ene.

2 Kontekstfri Grammatikk

- (a) Et uttrykk som *faren til prinsen sin kone* er flertydig. Hvis prinsen er William, kan uttrykket enten betegne Michael Middleton eller Camilla Parker Bowles, hertuginnen av Cornwall. Vi vil prøve å lage en grammatikk som fanger dette:

$$\begin{aligned} S &\rightarrow NP VP \\ NP &\rightarrow DET NI \\ NP &\rightarrow NP PP \\ NP &\rightarrow ND \\ PP &\rightarrow P NP \\ DET &\rightarrow NP G \\ VP &\rightarrow smilte|lo|danset \\ ND &\rightarrow prinsen|faren|mora|søstera|broren \\ NI &\rightarrow kone|mann|mor|søster|bror \\ P &\rightarrow til \\ G &\rightarrow sin \end{aligned}$$

Tegn opp de to trærne grammatikken tilskriver setningen

$$Faren\ til\ prinsen\ sin\ kone\ smilte. \quad (1)$$

- (b) Noen parsere vil ha problemer med denn grammatikken. Hvilke parsere? Forklar kort hva problemet består i.
- (c) Vis hvordan en chart-parser for denne grammatikken vil akseptere setning (1).
- (d) Er språket beskrevet av grammatikken regulært? Begrunn svaret kort.
- (e) (En liten utfordring:) Se på setningen

$$Faren\ til\ faren\ til\ prinsen\ sin\ kone\ danset. \quad (2)$$

Hvor mange forskjellige betydninger mener du denne har? Hvor mange forskjellige syntaktiske analyser får den i grammatikken.

(Hint: Hvor mange betydninger har *faren til faren til prinsen* og hvor mange grammatiske analysser får den?)

3 Unifikasjonsgrammatikk (i LKB)

- (a) Vis trekkstruktur-representasjonen av følgende regel slik den brukes i våre LKB-grammatikker. Angi også resultatet av type-ekspansjon på denne regelen (type-ekspansjon er prosessen LKB anvender ved kompilering av grammatikker), det vil si inkluder relevant informasjon som arves fra type-hierarkiet. Bruk gjerne forenklende liste-notasjon (hjørneparenteser) i din beskrivelse av den ekspanderte trekkstrukturen.

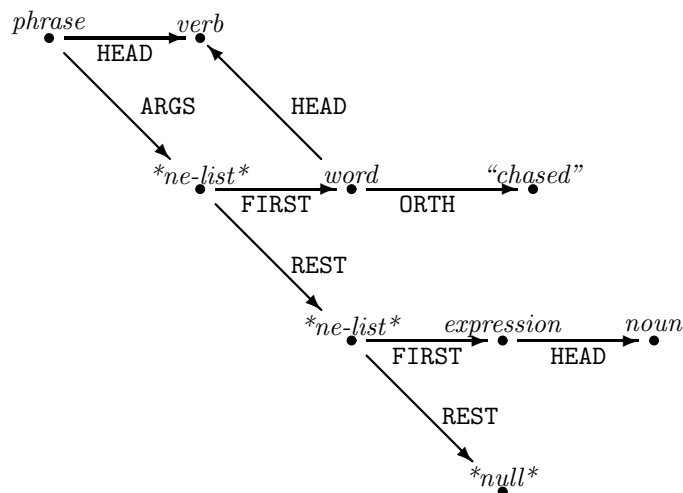
```
mystery := head-final &
[ SPR <>,
  COMPS #comps,

  ARGS < #1,
        phrase & [ SPR < #1 >, COMPS #comps & <> ] > ].
```

- (b) Hvilket navn har regelen over i vår grammatikk? Beskriv i noen få setninger hvilken funksjon den har, og gi to eksempler på typer av fraser som er bygget opp ved hjelp av denne regelen.

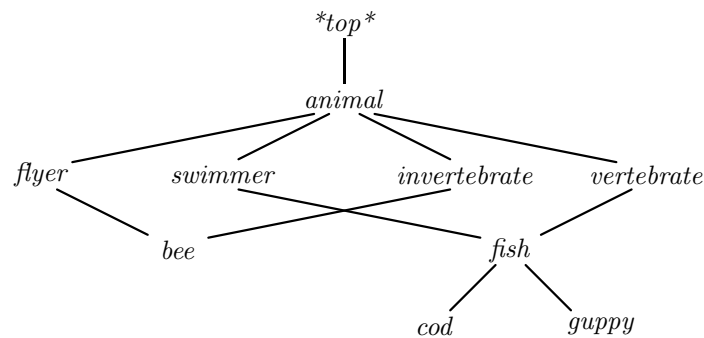
4 Typete trekkstrukturer

- (a) Tegn opp følgende DAG (*directed acyclic graph*) som en vanlig trekkstruktur, det vil si som en attributt-verdi-matrise (AVM).



- (b) Gi en kort beskrivelse (i høyst to setninger) av forholdet mellom elementene i trekkstrukturen og elementene i DAGen.

(c) Anta følgende type-hierarki.



Hvilke av de følgende par av typer kan unifiseres? For hvert unifiserbart par, angi deres største (felles) nedre grense (*greatest lower bound*, i.e. glb).

- (i) *animal* and *bee*
- (ii) *flyer* and *invertebrate*
- (iii) *flyer* and *vertebrate*
- (iv) *swimmer* and *whale*

5 Noe grunnleggende syntaks

- (a) Beskriv kort noen forskjeller mellom på den ene siden spesifikatorer og komplementer (*specifiers* and *complements*) og på den andre siden modifikatorer (*modifiers*).
- (b) Forklar tvetydigheten i følgende engelske setninger. Beskriv i dine egne ord hva de forskjellige lesningene betyr, og forklar uformelt hvordan disse tvetydighetene oppstår.
 - (i) *the dog chased that cat near those aardvarks*
 - (ii) *my report on the table arrived*
- (c) Skisser konstituenttrær (*constituent trees*) for hver mulig lesning av setningene (i) og (ii) over. Navngi noder ved hjelp av forkortelser som 'Det', 'N', 'NP', 'VP', etc., og angi på hver forgrening hvorvidt konstituenten under fungerer som hode (*head*), spesifikator (*specifier*), komplement (*complement*) eller modifikator (*modifier*).