UNIVERSITY OF OSLO

Faculty of Mathematics and Natural Sciences

Exam in INF2820 Computational linguistics
Day of exam: 10 June 2013
Exam hours: 0900-1300
This examination paper consists of 4 page(s).
Appendices: 0
Permitted materials: none

Make sure that your copy of this examination paper is complete before answering.

- **You may answer in English, Norwegian, Danish or Swedish.**
- **You should answer all the questions. The weights of the various questions are indicated.**
- **You should read through the whole set to see whether anything is unclear so that you can ask your questions to the teachers when they arrive.**
- **If you think some assumptions are missing, make your own and explain them!**

Exercise 1

Let G1 be the grammar with main symbol S and the following rules

```
1.  S  -> NP   VP
2.  VP -> IV
3.  VP -> TV   NP
4.  NP -> DET  N
5.  N  -> N    R
6.  R  -> RP   S_NP
7.  S_NP  -> NP_NP  VP
8.  S_NP  -> NP     VP_NP
9.  VP_NP -> TV     NP_NP
10. NP_NP ->
11. IV  -> 'smiled'|'slept'|'danced'
12. TV  -> 'chased'|'saw'|'owned'
13. N   -> 'cat'|'dog'|'child'
14. DET -> 'a'|'the'|'some'
15. RP  -> 'that'
```

S_NP is a regular non-terminal with the same status as for example TV, and so are NP_NP and VP_NP.

Questions 1.1 (10%)

Draw the tree(s) the grammar assigns to

a) the cat that the child owned smiled

Questions 1.2 (5%)

What does it mean for a language to be regular?

Questions 1.3 (10%)

Is the language L(G1), the language generated by G1, a regular language? Give reasons for your answer!

Exercise 2 (10%)

Some of the rules of grammar G1 are similar to other rules, e.g. rule (9) is parallel to rule (3) and rules (7) and (8) are parallel to rule (1). One could therefore reduce the number of rules in the grammar by using features. Construct a feature-based grammar (unification grammar) for the language L(G1) using fewer rules than in G1.

Exercise 3

Questions 3.1 (5%)
Some parsers have problems with the grammar G1. Which parsers? What is the property of G1 that makes a problem for such parsers, and why is this a problem?

Questions 3.2 (10%)
Show how a chart-parser will recognize sentence (b) according to grammar G1.

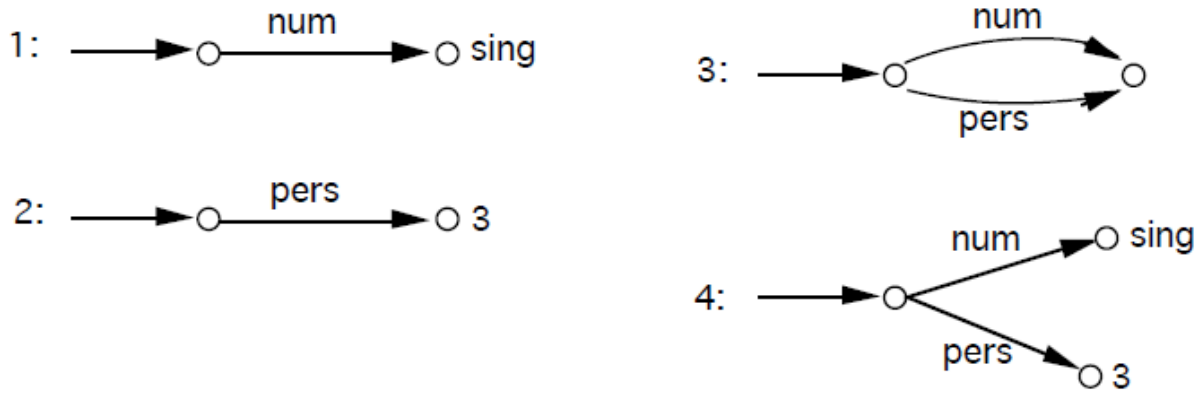
b) the cat smiled

Questions 3.3 (10%)
Find a grammar G2 in Chomsky Normal Form (CNF) for the language L(G1). Show the steps in the transformation of G1 into CNF.

Questions 3.4 (10%)
Show how a CKY-parser will recognize sentence (a) from exercise 1 according to grammar G2.

Exercise 4 (10%)

Consider the following four feature structures. Which pairs of them (A,B) are such that A subsumes B? Which pairs of structures are unifiable? Draw the results of the unification for the pairs that are unifiable. Also show the attribute-value matrix representations of the resulting unifications.



Exercise 5

We will consider semantics. We will simplify a little. We will not consider quantifiers. All NPs will be proper names. We want to represent (c) as (d):

c) Fido walks
d) *walk(f)*

The following rules will produce the desired result:

```
S[SEM = <?vp(?subj)>]   -> NP[SEM=?subj] VP[SEM=?vp]
NP[SEM=?np]             -> PropN[SEM=?np]
VP[SEM=?v]              -> IV[SEM=?v]
PropN[SEM=<f>]          -> 'Fido'
IV[SEM=<\x.walk(x)>]    -> 'walks'
```

We also want sentence (e) to get the representation (f) and sentence (g) to get the representation (h).

e) Fido chases Socks
f) *chase(f, s)*

g) Mary gives John Fido
h) *give(m, j, f)*

Questions 5.1 (10%)
Equip the following rules with features to achieve this result, i.e., fill in for the dots.

VP[SEM=...] -> TV[SEM=...] NP[SEM=...]
VP[SEM=...] -> DTV[SEM=...] NP[SEM=...] NP[SEM=...]
TV[SEM=...] -> 'chases'
DTV[SEM=...] -> 'gives'

Questions 5.2 (10%)
We will also include negation. For example, sentence (i) should get (j) as representation.

i) Socks does not chase Fido
j) *– chase(s, f)*

Extend and modify the grammar to get this representation. Make sure you get the correct grammatical restrictions. Thus the grammar should not generate

k) Socks does not chases Fido
l) Socks chase Fido

THE END