

Computational Syntax

Exam, 31 May 2011 at 13.30–17.30 in Viktoriagatan 30

Language Technology Master, University of Gothenburg

Teacher: Aarne Ranta (tel. 772 10 82, email aarne@chalmers.se)

Grading scale: Max = 60p, VG = 48p, G = 24p.

Aids: the book *Grammatical Framework* (Aarne Ranta, CSLI 2011); markings in the book allowed.

Question 1. Write a BNF (= context-free) grammar that covers the following fragment of English:

- **utterances:** sentences and questions
- **questions:** sentential questions with a noun phrase and a verb phrase, e.g. *do I sleep*
- **sentences:** with a noun phrase and a verb phrase, e.g. *I sleep*
- **verb phrases:** consisting of a one-place verb, e.g. *sleep*
- **noun phrases:** *I, you*
- **one-place verbs:** *sleep*

Notice that the lexicon is very small, and the only utterances recognized are *do I sleep, do you sleep, I sleep, you sleep*. But you must use separate categories for each of the six items listed above (utterances, etc), so that the grammar scales up to a bigger lexicon. (7p)

Show the parse trees of each of the four utterances recognized by your grammar. (3p)

Question 2. Divide the context-free grammar of Question 1 into a GF grammar with abstract and concrete syntax. (4p)

Add a German concrete syntax, where the English utterances get the translations *schlafe ich, schläfst du, ich schlafe, du schläfst* (using the familiar singular *you*). Thus

$$I = \textit{ich}, \textit{we} = \textit{wir}, \textit{sleep} = \textit{schlafe/schläfst}$$

The form of the verb depends on the subject by agreement. Sentences have the order SV, and the corresponding questions have the order VS. Notice that you have to introduce a parameter in German. (6p)

Question 3. Write a bigger version of the grammar of Question 2 by adding

- **two-place verbs:** *love* with German translation *liebe/liebst*
- verb phrases formed from a two-place verb and its noun-phrase complement, e.g. *love you = liebe dich*

Thus the following new utterances can be formed, in English and German:

<i>I love me</i>	<i>ich liebe mich</i>
<i>I love you</i>	<i>ich liebe dich</i>
<i>you love me</i>	<i>du liebst mich</i>
<i>you love you</i>	<i>du liebst dich</i>
<i>do I love me</i>	<i>liebe ich mich</i>
<i>do I love you</i>	<i>liebe ich dich</i>
<i>do you love me</i>	<i>liebst du mich</i>
<i>do you love you</i>	<i>liebst du dich</i>

Notice that both English and German now have an accusative case of noun phrases (*me* of *I*, *you* of *you*, *mich* of *ich*, *dich* of *du*). Also notice that the German word order is SVO for sentences and VSO for questions. (10p)

Question 4. Show the parse trees, abstract syntax tree, and word alignment for the sentences *do you love me* and *liebst du mich* in the grammar of Question 3. (10p)

Question 5. Write a GF grammar for the language $a^n b^n c^n$, that is, for the set of strings with some number of *a*'s followed by the same number of *b*'s and *c*'s, in this order. Examples are

- the empty string
- *a b c*
- *a a b b c c*
- *a a a b b b c c c*

But the strings *a*, *a a b c c*, *b c* are *not* in this language. (7p)

Show the abstract syntax tree and the parse tree of the string *a a a b b b c c c*. (3p)

Question 6. The following context-free grammar defines English noun phrases with various determiners.

```
DetCN. NP ::= Det CN
ModCN. CN ::= AP CN
a.     Det ::= "a"
the.   Det ::= "the"
this.  Det ::= "this"
my.    Det ::= "my"
house. CN ::= "house"
big.   AP ::= "big"
```

All such noun phrases with at most one adjective are shown in the following table, with their Swedish translations.

<i>a house</i>	<i>ett hus</i>
<i>a big house</i>	<i>ett stort hus</i>
<i>the house</i>	<i>huset</i>
<i>the big house</i>	<i>det stora huset</i>
<i>this house</i>	<i>det här huset</i>
<i>this big house</i>	<i>det här stora huset</i>
<i>my house</i>	<i>mitt hus</i>
<i>my big house</i>	<i>mitt stora hus</i>

This table shows considerable parametric variation in Swedish:

- the noun *hus* ("house") has the **indefinite** form *hus* and the **definite** form *huset*
- so does the adjective *stort* ("big"): indefinite *stort* and definite *stora*
- the form of the noun and the adjective depends on the determiner, which are translated as follows: *a* = *ett*, *the* = *det*, *this* = *det här*, *my* = *mitt*
- the determiner *the* disappears if the CN has no adjective

In addition to this, Swedish has variation in gender, but we don't care about it in this question. The task is now simply:

1. divide the context-free grammar into an abstract syntax and English concrete syntax (2p)
2. write a Swedish concrete syntax for the same abstract syntax, so that the examples in the above table get translated correctly (8p)

Hint: you have to define suitable parameters for definiteness and for the complexity of CN, and to use them in a clever way.