

CG / Syntax course

Constraint Grammar exercises

For the CG exercise section, we will use VISL's CG lab, at <http://beta.visl.sdu.dk/cglab.htm> , linked from the general CG page at http://beta.visl.sdu.dk/constraint_grammar.html .

Suggested reading: For doing the actual exercises, you don't have to be a Constraint Grammar expert, The Summer School syntax module should have provided you with some basics about the CG formalism as such, and the Visl_{cg}/CG-3 grammar formalism in particular. Hopefully you'll also have a rough idea by now, from rule examples or Corpus examples, about the categories used in the VISL parsers. If you need help while doing the exercises, you can always check the details in the following references:

Visl_{cg} formalism: http://beta.visl.sdu.dk/cg3_howto.html

Grammatical categories: <http://beta.visl.sdu.dk/da/info/>

For the exercises, we suggest that you copy the example grammar file (http://beta.visl.sdu.dk/cg_template) to your computer, and use it as a point of departure to add your own rules. You can then make changes and upload the changed versions for use with CG lab. Alternatively, you can copy & paste the grammar into CG lab's grammar field, and change it there.

1. Inspection

For a start, run the example grammar *as-is*, with the following example sentence, and try to understand which rule does what. In the parser menu, use first 'tagger' and then 'morphological analyzer'. The latter will give your grammar ambiguous morphological input, while 'tagger' will give it disambiguated input.

(a) Dette er en test. (This is a test.)

(b) .

2. Noun phrases

Next, try to make a little grammar for noun phrases, with the examples from the ambiguity exercise as a point of departure, using 'morphological analyzer'. The idea is to disambiguate part of speech by looking at the immediate np-context. E.g.:

SELECT (N) (-1 (ART)) ;

LIST PRE-N = ART DET ADJ STA ;

SELECT (N) (*-1 (ART) BARRIER (*) - PRE-N) ;

vi må finde en køber.
den nye køber har aldrig set skibet.

3. Place & Time semantic roles

Finally, using the full syntactic parser ('Standard Parser') as input, try to add semantic roles for place and time to prepositions (in effect, marking the whole pp). Since at this level of analysis, parses already contain @tags (for syntax), you have to use a different prefix, the VISL convention being '%' for semantic roles. Try some of the following semantic roles:

%TOP-PL = place/topological ... i Danmark

%TOP-GL = goal/direction til Danmark

%TOP-SRC = source ... fra Danmark

%TEMP-PL = point/period in time ... i 1999, før 1999, mellem tirsdag og torsdag

%TEMP-GL = time/temporal goal ... til 1999!

%TEMP-SRC = time/temporal source ... siden 1999

%TEMP-DUR = durative ... under krigen, i 4 timer, på 10 minutter

For a start, simply write MAP rules into the grammar window - CG lab will automatically add the DELIMITER, SETS and MAPPING lines. You can exploit the fact, that time/place preposition should already be marked as @<ADVL, @ADVL>, @<AS, @<AO etc. by the syntactic parser. The following rule, for instance, will mark the preposition 'i' for place-hood, given a few set definitions:

LIST N/PROP-LOC = <top> <civ> <inst> <L> <Lh> <Lciv> <Lwater> <Lpath> <build> <BB> ;

LIST @ADVL = @<ADVL @ADVL> @<ADVS @ADVS> @<ADVO @ADVO> ;

MAP (%TOP-PL) TARGET N/PROP-LOC (0 @P<) (*-1 PRP LINK 0 ("i") + @ADVL) ;

An example of a time rule:

MAP (%TIME-PL) TARGET (NUM @P<) (*-1 PRP LINK 0 ("i")) ;

The following sentences are from the Corpus2000. All contain at least the preposition 'i', but you are of course welcome to mark other prepositions, too.

Han understreger, at danske og grønlandske embedsmand har holdt to møder, siden forslaget dukkede op i februar.

Billedet spænder dermed over et bredt spektrum af betydninger, som i de forløbne år har knyttet sig til ideen om 'befrielse'.

Uddannelsen på de forskellige niveauer er i for høj grad ude af trit med den virkelighed, samfundet efterspørger.

Vi valgte f.eks. hårde hvidevarer i stål, selv om vi kunne have fået dem til den halve

pris i hvidt, men det ville fuldstændig havde ødelagt det hele".

Da den skrappe omstilling blev indledt i 1997, var der ialt 6599 store og mellemstore statsselskaber, som gav underskud.

Den knap to år gamle han-odder Thor, der snart er kønsmoden, blev fjernet fra sin mor og forenet med en særlig udvalgt hun-odder, som Aqua Ferskvands Akvarium har været i Holland for at hente.

Mellem disse to poler ligger det store område, som faktisk ikke har det så godt i disse år, og som ikke prioriteres videre højt i kulturpolitikken.

Trafikmængden på vejene er jo blevet mangedoblet gennem de seneste tyve år i det nordlige Århus og langs banestrækningen i midtbyen, så de mange baneoverkørsler udgør et alvorligt sikkerhedsproblem," understreger Hans Schiøtt.