

IA161 Final homework
Task : due on January 6th 2012
Deadline: January 13th 2012, 05:00 p.m

1. Installation and evaluation of the parsers

Download all the materials (parsers, instruction, data) from the following web page:

<https://is.muni.cz/el/1433/podzim2011/IA161/30264269/>

For evaluating the RASP, CCG Parser and Stanford Parser, the same input sentences will be used. Download **data.tar.gz** which include the input sentences to be parsed and unpack it. The input sentences are extracted from the NYT.

1.1 RASP

Download the RASP **rasp3os.tar.gz** and unpack it to get *rasp3os* directory.

```
$ tar xzvf rasp3os.tar.gz
$ cd rasp3os/scripts
```

Run the RASP with the following command:

```
$/rasp.sh -m <test.txt> result.txt
```

1.2 CCG Parser

Download the CCG Parser **candc-linux-1.00.tgz** and **models-1.02.tbz2** in two steps.

First, unpack **candc-linux-1.00.tgz** to get *candc-1.00* directory which include the software tool, then unpack **models-1.02.tbz2** to get *models* directory which include the trained models from the Penn Treebanks under the *candc-1.00* directory.

```
$ tar xzvf candc-linux-1.00.tgz
$ cd candc-1.00
$ tar xzvf ../models-1.02.tbz2
```

Run the CCG Parser with the following command:

```
$ candc-1.00/bin/candc --models=candc-1.00/models \
--input data/test --output data/test.candc
```

1.3 Stanford Parser

Download the Stanford Parser **stanford-parser-2011-09-14.tgz** and unpack it to get *stanford-parser-2011-09-14* directory.

Run the Stanford Parser with the following command:

```
$ ./lexparser.sh data/test.txt
```

2. Analysis and comment on the obtained results

- a) Analyze the quality of parsing outputs. Compare and comment at least 5 phenomena contained in the input sentences, for example, parentheses, punctuation, reflexive verb, relative clause, modal verb followed by a verb in the infinitive form, prepositional clausal modifier, temporal modifier, multi-word expression, etc. You can evaluate the parsing results manually by using your linguistic intuition.
- b) Report the errors that you found during the analysis of the outputs.
- c) Compare the parsing method (techniques, theories) used in the three parsers. Remind you the technical and theoretical backgrounds presented in the lectures of this course.

Attention!

- a) Please attach your outputs (screen-shots) on the final file.
- b) Please send one compressed file which include all the documents: copies of outputs, analysis of the results, etc to qkang@fi.muni.cz.