

Miriam Butt und Maribel Romero
G220/5109, G212/2728
{miriam.butt|maribel.romero}@uni-konstanz.de

Advanced Computational Semantics
Sommersemester 2008, Hauptseminar

Übung 8, 19.06.2008

- Nutcracker installieren (Anleitung für MacOS X)
- zum Testen: Einige Textual Entailments ausprobieren

Aufgabe 1)

a) Nutcracker ist im Subversion-Repository von candc / Boxer bereits enthalten. Du musst es aber noch kompilieren. Das machst du mit

```
make bin/nc
```

Das Kommando sollte zunächst WordNet herunterladen und installieren und dann Nutcracker selbst kompilieren. Sollte das Kommando nicht funktionieren, fehlt dir wahrscheinlich das Unix-Programm wget. In diesem Fall solltest du mit

```
sudo fink install wget
```

wget installieren (das wiederum setzt voraus, dass du Fink installiert hast¹). Hast du Nutcracker kompiliert, fehlen dir noch die Programme Bliksem (Theorem-Prover), Mace und Paradox (beides Model Builder). Die Binaries dieser Tools müssen sich am Ende im Verzeichnis /candc/ext/bin befinden. Die Makefile von Nutcracker kann benutzt werden, um die Programme herunterzuladen und zu installieren; dazu musst du zunächst das Verzeichnis /candc/ext/bin erstellen und dann die folgenden Kommandos ausführen:

```
make ext/bin/bliksem  
make ext/bin/mace
```

¹Siehe <http://finkproject.org/>.

```
make ext/bin/paradox
```

Das Problem ist, dass diese Kommandos momentan nur auf Linux problemlos funktionieren, jedoch auf MacOS X Fehler auftreten können (bei mir war das der Fall). Wenn du vorher das Verzeichnis erstellt hast und trotzdem Fehler auftreten, musst du die Tools Bliksem, Mace und Paradox von den jeweiligen Seiten im Internet laden. Gehe hierzu auf diese Seite:

<http://svn.ask.it.usyd.edu.au/trac/candc/wiki/nutcracker>

Dort sind alle Links zu den Programmen angegeben.

Hier noch die Versionsnummern, die ich getestet habe. Von Bliksem gibt es nur eine Version im Netz. Für Mace das Paket mit Otter (Dateiname: otter-3.3f.tar.gz), und für Paradox paradox2-3.gz (diese Datei musst du nur entpacken, und du erhältst eine Binary).

Wenn die Sachen alle kompiliert sind, ist nur noch wichtig, dass sie im richtigen Verzeichnis liegen, nämlich candc/ext/bin.

b) Um Nutcracker zu benutzen, sollte man zuerst ein eigenes Arbeitsverzeichnis anlegen (/candc/working/rte wird empfohlen²). In diesem Verzeichnis speichert man 2 einfache Textdateien ab; eine sollte den Namen 't' tragen, die andere 'h' (ohne Endung, tokenisiert). 't' steht für Text, 'h' für Hypothese. In 't' sollte also der Basistext, in 'h' die zu überprüfende Hypothese stehen. Sind die Dateien vorhanden, kann man Nutcracker starten:

```
bin/nc --dir working/rte
```

Man gibt dem Programm also den Ordner als Input, in dem sich die Textdateien befinden. Nutcracker wird dann versuchen, eine Beziehung zwischen den beiden Aussagen in 't' und 'h' herzustellen: Entailment ('t' hat 'h' zur Folge), informativ ('t' hat 'h' nicht zur Folge) oder inkonsistent ('t' und 'h' sind unvereinbar).

Nutcracker schreibt die relevanten Dateien in das angegebene Verzeichnis: die DRS-Strukturen, die Modelle, das Hintergrundwissen und die Aussage zur Beziehung von 'h' zu 't'.

²RTE steht für Recognizing Textual Entailment.

Aufgabe 2)

a) Finde je 3 Beispielsätze zu den Relationen Entailment, informativ und inkonsistent. Du solltest vielleicht Unterordner für die einzelnen Relationen und Beispielsätze erstellen:

```
/candc/working/rte/{ent|incons|inf}/{ent1,...|incons1,...|inf1,...}
```

Schicke am besten deinen kompletten Ordner /candc/working/rte als Zip-Datei ein.