

Grammar Engineering: Minimal Recursion Semantics

06.2012

1. Download the starting grammar <http://www.coli.uni-saarland.de/~yzhang/ge/grammar/grammar05.tgz>.

This is a small Matrix-based HPSG grammar for English. The basic phrase structure rules, lexical and morphological rules and lexical entries have been provided. Basic types are also available to handle subject-verb agreements. Go through various files in the grammar and understand how it works.

Also have a look at the test suite file “*items*”. With the version of the grammar you downloaded, you should be able to parse all positive items and block negative ones.

2. Parse the positive sentences from the test suite. From the pop-up parse tree window, right click the analysis, and select to display the MRS of the analysis. Carefully examine the displayed structures. Is it a good MRS structure for the given input? For those parts of semantic composition that already work, please refer to “MATRIX.TDL” for detailed implementation. Document the semantic composition process with the example sentence “the dog sleeps”. Point out what information still needs to be added/removed/modified.
3. To make the grammar output better semantics, several changes need to be made to the grammar. The following instructions should be understood as a set of sub-tasks. This is not meant to be a detailed step-by-step tutorial. You should not be limited by what is mentioned here.
 - Introduce relation names in lexical entries. The MRS representation is built upon the semantic units called EPS, which have re-

lation names from lexical entries ¹ The relation name can be introduced via the feature-path `SYNSEM.LKEYS.KEYREL.PRED` in the lexical entry (whose value is usually a string). The naming convention we use for open-category word relations is *"_stem_pos_rel"*. For instance, the noun lexical entry for “cat” should introduce a relation *"_cat_n_rel"*.

- Specify `ARG-ST` for lexical types. Recall the binding theory in HPSG. The `ARG-ST` feature should be properly composed according to the valence list. For noun and determiner lexical types, you can inherit from Matrix types *basic-one-arg* and *basic-zero-arg*. For verbs, further token identity is required from the `XARG` of the verb to the `INDEX` of the subject.
- Constrain the `INDEX` of the `MRS` structure, so that it equals to the `ARG0` of the main verb (the main event of the sentence).

Please document your changes to the grammar, and note down the final `MRS` structures for all positive test items.

¹In principle, EP can also come from phrase rules via `C-CONT`, but this is not touched in this exercise.