Minutes of the bilateral Meeting of DFKI and University of Tilburg, 14th-16th December 2005, in Saarbrücken

Attendees: Amanda Schiffrin (Tilburg),
                Thierry Declerck, Mirjam Kessler and Anette Frank (DFKI)

The meeting was set up in order to discuss issues related to WP4 (Meta-model and data categories for semantic annotation) in LIRICS, as well as issues closely related to the ad hoc Task Domain Group TDG 3 devoted to this area, within Technical Committee 37, Subcommittee 4 (Language Resources Management), where a subgroup is also dealing with the topic "predicate argument structures and semantic roles". The meeting was also set up in order to facilitate collaboration and knowledge transfer between DFKI and University of Tilburg.

The main topic of the meeting was the investigation of how to extract/derive candidate semantic roles from the FrameNet corpus (the English version as well as the German version, which has been developed in the Salsa Project at the University of Saarland). We have now obtained a research licence for the English corpus and will soon also obtain one for the German corpus.

Wednesday 14th December 2005

In the morning, work by Mirjam Kessler describing the syntactic part of the FrameNet corpus was presented. A paper version of this work has been finalised and the suggestion is to include some or all of this in a new version of the Deliverable 4.1.

Following this, Thierry Declerck presented a tool that extracts special information and statistics from the FrameNet corpus in an automated fashion and presents it in a formatted way. This tool provides empirical verification of hypotheses about candidate semantic roles that can be extracted from the FrameNet corpus. The idea is to abstract over labels, using the frequency scores as a guide. It is also proposed to take a look at PropBank, with a view to seeing whether a similar process can be used on this corpus and to make a comparison with the results from FrameNet. A manual comparative study will also be carried out by Amanda Schiffrin and Olga Petukhova at Tilburg University. After discussions at the meeting in Nancy (8th-9th December 2005), the following questions arise: (a) Is this approach good enough? The FrameNet corpus does not by any means provide a complete list of all frames and frame elements. How then can one be sure that the extracted/abstracted list is comprehensive? (b) Is frequency of use/occurrence a good enough criterion for the detection of generic semantic role labels? What other criteria might there be?

In the afternoon, Anette Frank presented the state of the art of the German FrameNet Corpus as developed in the Salsa project at the University of Saarland. She also demonstrated the use of manual annotation tools that are used to build onto the syntactic tree bank "Tiger". This tool is freely available from the Salsa website (the only restriction on use is that the corpus for annotation has to be in "Tiger" XML format). Some of the work in the Salsa Project is also dedicated to providing tools for automatically annotating large documents with FrameNet semantics. Anette further explained how the work of the Salsa project is looking at cross-lingual frame

alignment. For these reasons, Salsa is of considerable interest for our approach to data categories for semantic roles, which would support the interoperability of FrameNet annotation with other formats and languages. Other useful pointers: Dan Gildea's work on the simplification of FrameNet frame elements (although it is unclear as yet what the take-up is on this work) and a list of relevant publications on the Salsa web-pages.

Thursday 15th December 2005

In the morning, Amanda Schiffrin and Thierry Declerck discussed a possible strategy for semi-automatically extracting/deriving semantic roles from the FrameNet corpus. This is outlined below:

- A first step consists in calculating the frequency and description of so-called frame elements (FE) in the whole corpus. Frame elements that occur more frequently than a certain threshold (to be fixed) might be considered candidates. In this first phase we will also try to take into account certain typographic and also labelling variants.

- In a second step we will check (with the assistance of the analysis tool developed) which frame elements occur in a very large number of frames (and lexical units). FEs that occur in a higher number of frames than the threshold (to be fixed) are considered candidate semantic roles.

- Merge the candidates from step 1 and 2

- In a fourth step tune the process and look at FEs that did not meet the threshold. See if they can be in fact considered as "subtypes" of already existing FEs (there are also indications about this in various FrameNet documents), or at least as specialisation.

The proposed candidates for semantic roles will be compared with the roles associated with syntactic arguments in the PropBank corpus, in order to test the possible interoperability of our candidates.

In a second study, we will extract from the corpus all pairs of FE and Grammatical Functions in order to study regularities for the syntax-semantics interface. This is particularly relevant for the TDG 3 work. There was also discussion about the possibility of extracting a set of use constraints for semantic roles to see if there are some mandatory roles and also to see if the presence of certain roles will determine the presence of others.

In the afternoon, Amanda Schiffrin began to extract frames from FrameNet for a preliminary cursory comparison with PropBank. In the process of this work, she made a couple of potentially interesting observations.

Firstly, there seems to be fairly strong evidence for the inclusion of a data category for the generic role of *Agent* at least; this is borne out by the statistics extracted by the analysis tool, and also by the mistaken use of *Agent* instead of the more specific *Avenger* in the *Revenge* frame in an example taken from the FrameNet book. The frame elements in the Revenge frame are given as *Avenger*, *Punishment*, *Offender*,

Injury, and *Injured_party*: these are used correctly in examples (1) and (3) below, but not in (2):

(1) [His brothers *Avenger*] **avenged** [him *Injured_party*].

(2) With this, [El Cid *Agent*] at once **avenged** [the death of his son *Injury*].

(3) [Hook *Avenger*] tries to **avenge** [himself *Injured_party*] [on Peter Pan *Offender*] [by becoming a second and better father *Punishment*].

It is not clear from the documentation whether this is simply an error, or whether it is caused by the existence of a more generic FE, *Agent*, from which more action-specific FEs are derived. Either way, we conclude that *Agent* is a likely candidate for a data category.

The second observation made from studying the FrameNet annotation is that negation (as well as other types of modification and quantification) is ignored. For example:

(46) The commander `ordered` [the troops *Avenger*] not to **retaliate** [against the rebels *Offender*].

It seems ultimately very strange to talk about an *Avenger* in a non-event of retaliation. This would presumably be marked up in exactly the same way given the positive formulation.

The omission of negation, etc., seems to reflect the fact that FrameNet is a purely relational approach to semantic annotation, and so will not be an exhaustive source for all the goals of WP4.

Friday 16th December 2005

Mainly a day used to finalise and concretise discussions and work over the previous two days, and to agree on the minutes of the meeting. In the next period, we agreed that we will carry out the comparison with PropBank, and identify sources for the semantic annotation of negation, quantification, etc.

We will also produce a small document on the LIRICS work on Semantic Annotation to be included in a reading for a tutorial on semantic annotation to be held by Sigfried Handschuh (Karlsruhe) and Thierry Declerck (Saarbruecken at the 3rd Eurpean Conference for Semantic Web).

Next meetings on these issues are foreseen to take place at the ISO TC37/SC4 meeting in Jeju, January 2006. The TDG3 meeting will be fixed in March/April in the US in order to discuss results with American colleagues.