

# A New Life for Semantic Annotations?

**Harry Bunt**

**Tilburg University (The Netherlands)**

email: `harry.bunt@uvt.nl`

---

Semantic annotation has so far been approached in essentially the same way as annotation at other levels of linguistic information, namely as the business of labeling text with certain tags which add certain information to the text, in this case, semantic information. Semantic role labeling is a case in point. This may be very useful, for instance for determining the variety of ways in which certain types of semantic information tend to be expressed, but it seems to me that semantic annotations can and should have a deeper significance and a more important role to play.

Since semantic annotations are intended to capture some of the meaning of the annotated text, it ought to be possible to use such annotations in reasoning, and hence to apply that information in language processing tasks. However, reasoning with semantic annotations presupposes that the annotation language has a formal semantics. (In fact, one may wonder how much sense it makes to use a semantic annotation language *without* a semantics, since there is *a priori* little reason to assume that semantically undefined annotations would capture the meanings of natural language expressions any better than the expressions themselves.)

Still, existing work in this area, for instance on semantic role annotation (as in the FrameNet and PropBank initiatives) or on the annotation of temporal information (as in the TimeML effort) makes use of uninterpreted annotation languages.

In this talk, I will discuss some of the possibilities, perspectives, and problems in defining semantic annotation languages with a well-defined semantics. I will do this by starting from an attempt to integrate intermediate results from the design of a standard for temporal annotation in the International Organisation for Standards (ISO), and from the definition of annotation schemas for coreference annotation and semantic role labeling in the European project LIRICS. I will indicate the requirements for integrated and multilayered semantic annotation approaches, and certain general principles for semantic annotation. Moreover, I hope to address issues concerning the relations between using semantic annotations with a formal semantics on the one hand, and using underspecified semantic representations on the other. Finally, I will consider some of the potential applications of the use of interpreted semantic annotations in areas such as information extraction, paraphrase generation, and textual entailment.