

REPORT ON WORKSHOP AND PANEL ON SEMANTIC ANNOTATION

The workshop and panel discussion on semantic annotation has been organized as a joint event of the FLaReNet project (European) and the SILT project (U.S.A.) as part of the 5th *International Conference on Generative Approaches to the Lexicon*, in Pisa at the Institute of Computational Linguistics “A. Zampolli” - CNR. The panel discussion has been a moment in which the participants of the conference had the opportunity to discuss openly with the panelists some of the most compelling issues in the field of semantic annotation, exchange ideas and identify blueprints of actions for the development of the field in the very next years. The presence of some of the most important experts in this field of study from the international scientific community such as James Allen (panelist), Patrick Hanks (panelist), Nancy Ide (panelist), Aravid Joshi (panelist), Bonnie Webber (panelist), Martha Palmer (discussant), Pierette Bouillon (discussant), Bolette Pedersen (discussant), Emanuele Pianta (discussant), Chu-Ren Huang (discussant), Joan Bos (discussant), Nuria Bel (discussant), Malvina Nissim (discussant), Piek Vossen (discussant), Luca Dini (discussant), Nicoletta Calzolari (chair) and James Pustejovsky (chair), have provided a unique *milieu* for the discussion and exchange of ideas. Recently, works in semantic annotation have taken into account increasingly more complex linguistic phenomena. Nowadays, semantic annotation involves among others word sense disambiguation, discourse structure, eventualities and their internal structures, semantic roles labelling, temporal relations, detection of textual entailment, spatial relations.

The development of international standards under the ISO TC37/ SC4 initiative suggests that the field has reached a good level of maturity and a substantial convergence of opinion and practice in the research community. The availability of standards is considered a viable solution to deal with current issues in the production of semantic language resources and to improve their efficiency for the performance of NLP systems and for their interoperability.

However, some compelling issues must be tackled by the community in order to provide reliable solutions which overcome current limitations and can guarantee interoperability and good quality. Some of these issues have been discussed in this panel and are reported below.

REPRESENTATION FORMAT FOR MULTILAYERED LANGUAGE RESOURCES

Semantic annotation can be better described as a high level annotation. Semantic information is usually not expressed on the surface, as is instead part-of-speech or syntax. Thus the goal of semantic annotation is to make explicit linguistic information at a deeper level.

Semantic annotation has so far adopted a methodology of low level annotations, in particular it is often still very close to the text level, referring to objects in linear order and where links signify constituency and not relevant/specific named/semantic relations.

To boost semantic annotation a shift in the annotation paradigm is necessary. In particular, the way annotation is represented should be at a more abstract level and it should benefit from other annotation layers which are more surface based. The relevant use of semantic information requires that all other more surface-based annotation levels are available and that this information can be combined and extracted on demand.

Different, stand-off, semantic annotation layers are therefore required to be linked one to another in order to extend the informative power of the language resources created. The links between the different levels should be available in a representation format which allows for different and interlinked queries in order to extract only the relevant information for a given purpose. It appears that a good solution to overcome these limits is represented by annotations in a graph-based format.

BEST PRACTICES AND GOOD QUALITY

The development of annotation schemes and language resources needs to be accompanied by best practices to ensure good quality annotations. The annotation process and, consequently, the creation of language resources is a compromise based on constraints which, on the one hand, allow the annotators to perform the required task, and, on the other hand, do not bias the knowledge of a certain linguistic phenomenon both for the development of linguistic theories or for the performance of applications.. The ideal language resource is the one which presents the perfect balance of the constraints. The development of best practices is thus a compelling issue for any kind of annotation and especially for ensuring the good quality of semantic language resources. Standardisation and interoperability are part of these best practices but are not sufficient. It becomes more and more important that the release annotation schemes and language resources be accompanied by the guidelines that have been followed and, moreover, by a document explaining “the compromises” that have been made in order to face practicals problems encountered at the annotation time. This will introduce a new level of best practice which, together with standardisation and interoperability, could represent a key point for the development of good and appropriate semantically annotated language resources and applications.

SHIFTING PERSPECTIVES ON EVALUATION

Evaluation is a necessary corollary for the advancement of language technology. Classical evaluation measures for language resources are not always applicable as such for semantic annotation. The results which are by means of these measures are hardly interpretable as measures of good quality semantic annotations, or to evaluate the goodness of applications in the semantic domain. To boost semantically annotated language resources and applications new and more specific measures based on innovative methodologies for evaluating the reliability of semantic annotations are required. So far, the research community has not reached a consensus on which new different evaluations should be considered as providing the *gold standards* over which language resources and applications can be compared. This *facet* of semantic annotation requires a collaborative effort of the scientific community.