
In support of self-assessment – exploiting available information from tools

Donnerstag
09.03.2017
12:45 – 13:45
B4 1, Foyer

Markus Gärtner
Universität Stuttgart

markus.gaertner@ims.uni-stuttgart.de

Kerstin Eckart
Universität Stuttgart

kerstin.eckart@ims.uni-stuttgart.de

Many automatic tools for natural language processing tend to produce less reliable results if presented with data which differs from an expected standard. Users in search of the most suitable tool for their out-of-domain data have to either perform some additional error detection after processing (cf. Dickinson, 2015, for an automatic approach) or invest significant manual effort into the detailed evaluation of several tools.

Interestingly enough, automatic tools are often internally aware of a relative reliability of their output, since they make use of probabilities and forced guessing to decide on a single analysis. Such information can be understood as the internal confidence for the complete analysis, e.g. n-best lists of outputs correspond to a ranking according to confidence.

We argue for an approach where tool output is transparent with respect to its internal confidence estimation (e.g. BitPar; Schmid 2004). A single confidence value does not have to cover the whole analysis, but can refer to subparts, such as a dependency relation or a specific label.

We suggest that tool output includes confidence values as an additional annotation layer which will be an advantage when handling large data sets that rely on automatic annotation: Users can restrict their queries to more reliable parts or find interesting cases by inspecting parts with low confidence. That is, transparent confidence estimation will not increase the quality of the output as such, but its usability in that it helps users to assess if an analysis, or a specific part of it, is sufficiently reliable.

Moreover, providing confidence values as annotation raises the awareness with regard to reliability, which we think can foster the application of state-of-the-art tools on out-of-domain data also in related fields such as the Digital Humanities.

References: • Dickinson, M. (2015): Detection of annotation errors in corpora. *Language and Linguistics Compass* 9(3), 119–138. • Schmid, H. (2004): Efficient parsing of highly ambiguous context-free grammars with bit vectors. In: *Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004)*. Geneva, Switzerland.