

Enriching FrameNet with Scalar Information

Josef Ruppenhofer & Jasper Brandes
Universität Hildesheim

The questions that sentiment analysis tries to answer about a stretch of text are the following, ordered by increasing complexity and finer granularity (Pang & Lee 2008): (a) Is the text mainly descriptive and factual or does it contain speculations, uncertainty, or evaluations? (b) If a stretch of text is subjective and in particular evaluative, does it convey positive or negative sentiment? (c) Who [opinion source] expresses how strong [intensity] a positive or negative [polarity] opinion about what [opinion target] ?

Unsurprisingly, FrameNet, has been considered as a resource to use in this task. Ruppenhofer et al. 2008 argued that enriching FrameNet with mappings between frame roles and the "opinion" roles Source and Target would allow the use of automatic semantic role labeling systems such as SEMAFOR (Das et al. 2010) for mostly solving the task of identifying OPINION HOLDERS and TARGETS (cf. (c) above). Ruppenhofer & Rehbein 2012 extend the earlier proposal, arguing for frame semantics as an anchor representation for sentiment analysis.

Here we take up one particular aspect of representation that needs to be elaborated in FrameNet, namely scale structures. For instance, FrameNet has no information on the difference in strength between excellent and good in the Expertise frame (cf. (1) vs. (2)).

- (1) Kim is a good programmer.
- (2) Kim is an excellent programmer.

While data-driven statistical methods have been developed for assigning intensity scores to adjectives (Liu & Seneff 2009, de Marneffe et al. 2010), these methods are not linguistically well-grounded. Therefore, a key question to investigate is how to represent the scales on which predicates are ordered. We report on our representational decisions and on some first experiments to semi-automatically add scale-related information to frames and lexical units in FrameNet.

Our main desiderata include the following. Among gradable adjectives, dimensional adjectives should be distinguished from evaluative adjectives: the former usually have clear antonyms (*tall-short*), while for the latter there tend to exist bundles without clear antonyms (*bold, brave, cowardly, timid, fearful*) (Bierwisch 1989:90). The type of scale, in the sense of Kennedy & McNally's (2005) distinction between closed and open scales also needs to be recorded as it goes along with different types of modification as well as different entailments. Finally, the orderings of adjectives should closely track human understanding. In many cases, relative orderings of adjectives can be derived from particular features such as their scale types: for instance, absolute adjectives such as *awful* should be stronger than relative ones such as *bad*. However, there exist some questions as to how well the scale types proposed in the literature fit actual language use. For instance, in particular institutional frames, orderings may be imposed where none exist in the general language. An example of this are the verbalizations of school grades. E.g. on the traditional German 6-point scale, shown in (3), the adjective *ausreichend* is defined as a lesser grade than *befriedigend*, although both evoke the idea of meeting a minimum standard in the general language.

- (3) *sehr gut* 'very good' > *gut* 'good' > *befriedigend* 'satisfactory' > *ausreichend* 'sufficient' > *mangelhaft* 'insufficient/deficient' > *ungenügend* 'insufficient/unsatisfactory'

Another issue is that contextually evaluative use may have an effect on scale structure (Croft & Cruse 2004:189-190).