Towards Word Sense Disambiguation by Reasoning

Javier Álvez[1], Itziar Gonzalez-Dios[2], and German Rigau[3]

[1] LoRea Group, University of the Basque Country (UPV/EHU)
javier.alvez@ehu.eus
[2] Ixa Group, University of the Basque Country (UPV/EHU)
itziar.gonzalezd@ehu.eus
[3] Ixa Group, University of the Basque Country (UPV/EHU)
german.rigau@ehu.eus

In this proposal we describe a practical application of Vampire [5] for Word Sense Disambiguation (WSD) [1]. In particular, we propose a method for the automatic disambiguation of the semantic relations in BLESS [3], which is a dataset designed to evaluate models of distributional semantics. It compiles 200 concrete nouns (100 animate and 100 inanimate nouns) from different classes and each noun is linked to other words via the semantic relations *hyperonymy*, *cohyponymy*, *meronymy*, *attributes*, *events* and *random*.

For the disambiguation of the word pairs and semantic relations of BLESS, we have to map the words to WordNet synsets [4]. WordNet is a large lexical database of English where nouns, verbs, adjectives and adverbs are grouped into sets of synonyms (synsets). Each synset denotes a distinct concept and they are interlinked among them by means of lexical-semantic relations such as *synonymy*, *antonymy*, *hyponymy*, *meronymy* or morphosemantic relations.

Since a word can have different meanings, we need to choose which synset it belongs to, that is, the one that better fits the semantics of the word in the given context. To that end, we plan to use the knowledge in Adimen-SUMO [2], which is obtained by means of a suitable transformation of the knowledge in the core of SUMO[1] [6] into first-order logic (FOL) and enables its use by FOL automated theorem provers such as Vampire. WordNet and SUMO (and therefore Adimen-SUMO) are connected in a semantic mapping [7] by means of three semantic relations: *equivalence*, *subsumption* and *instance*. By exploiting this mapping, we will automatically create a set of conjectures for each word pair by considering the semantic relations provided by BLESS. Then, these conjectures will be evaluated using Vampire. According to the outcomes, each word will be connected to a single synset and, consequently, disambiguated.

Finally, we plan to compare the results provided by our proposal and different disambiguation systems that can be found in the literature.

References

[1] E. Agirre and P. Edmonds. *Word sense disambiguation: Algorithms and applications*, volume 33. Springer Science & Business Media, 2007.

[2] J. Álvez, P. Lucio, and G. Rigau. Adimen-SUMO: Reengineering an ontology for first-order reasoning. *Int. J. Semantic Web Inf. Syst.*, 8(4):80–116, 2012.

[3] M. Baroni and A. Lenci. How we blessed distributional semantic evaluation. In *Proceedings of the GEMS 2011 Workshop on GEometrical Models of Natural Language Semantics*, pages 1–10. Association for Computational Linguistics, 2011.

[4] C. Fellbaum, editor. *WordNet: An Electronic Lexical Database*. MIT Press, 1998.

[5] L. Kovács and A. Voronkov. First-order theorem proving and Vampire. In N. Sharygina and H. Veith, editors, *Computer Aided Verification*, LNCS 8044, pages 1–35. Springer, 2013.

---

[1] http://www.ontologyportal.org

[6] I. Niles and A. Pease. Towards a standard upper ontology. In Guarino N. et al., editor, *Proc. of the 2nd Int. Conf. on Formal Ontology in Information Systems (FOIS 2001)*, pages 2–9. ACM, 2001.

[7] I. Niles and A. Pease. Linking lexicons and ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In H. R. Arabnia, editor, *Proc. of the IEEE Int. Conf. on Inf. and Knowledge Engin. (IKE 2003)*, volume 2, pages 412–416. CSREA Press, 2003.