

A Multilingual Functional Repository for Cognitive Linguists

Srini Narayanan & Jisup Hong
International Computer Science Institute

We describe a functional repository and knowledge base of conceptual metaphors, schemas, and example linguistic manifestations. The MetaNet repository supports content-based search, programmatic access through JAVA and Python APIs, and human editing, browsing, and visualization through a variety of tools and mechanisms. It is currently populated by a variety of automatic metaphor extraction tools, by linguists using the Embodied Construction Grammar (ECG) semantic analysis framework (Feldman et al. 2009), and through structured entry by cognitive linguists using a semantically enhanced wiki for collaborative development (Krötzsch et al. 2007). In addition, the repository integrates a variety of knowledge-based and inferential extensions that have proven crucial to the kind of deep, context-sensitive, metaphor-based analysis required to model scenarios and narratives and to infer the intentions and world views of participants in them (Lakoff and Narayanan 2010).

The design and specification of the repository meets several functional requirements. First, given that metaphors represent mappings between source and target schemas, the repository is able to represent schemas (including frames), roles, fillers, constraints, and role-to-role mappings. Metaphors are structured hierarchical objects arranged in networks, which link multiple conceptual levels (primary, basic, compositional, generic) with their linguistic manifestations (constructions and specific instances in text). Second, populating the repository requires APIs to multiple sources, such as metaphor extraction algorithms (which automatically extract instances of conceptual metaphors from text), the semantic wiki (for collaborative development by linguists), the FrameNet relational database (in MySQL/XML) (Ruppenhofer et al. 2006), and to other ontological and linguistic knowledge expressed in ECG. Third, use of the repository is facilitated by the ability to write to relational databases (in MySQL), to provide linked data access (in RDF/XML) to its contents, and by programmatic access through JAVA and Python APIs. Fourth, the repository supports methods for browsing, navigation, visualization, and content-based semantic search. Furthermore, the repository provides standard database functions including relational joins, version control, access control, scalable stores, and combining linked and structured data. Finally, the repository is flexible and extensible to support inferential functions, including the encoding of the probability of map access and use as well as the ability to support context-sensitive dynamic simulations, which are essential for the deep analysis of metaphor in building systems capable of Natural Language Understanding. Figure 1 displays an overview of the architecture of our current repository.

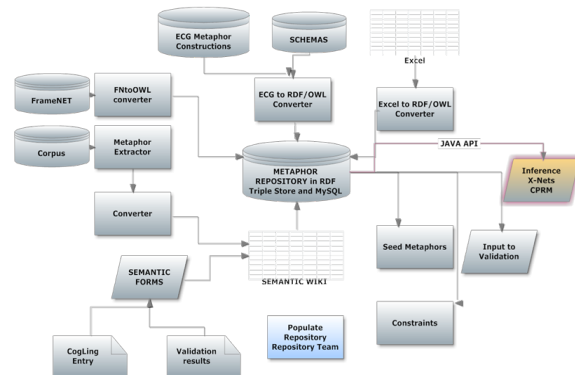


Figure 1

Presently, we have completed the first version of the repository, including knowledge bases, semantic wikis, and automatic extraction, in four languages—English, Persian, Spanish, and Russian—with each encoding a hierarchical and relational network of metaphors. The relations are typed into different categories, and semantic content search functions allow for database queries by relation type, metaphor family, source/target schemas, etc. Our semantic wiki serves as a major front end to the repository. We hope to demo and release the first version of the repository and the semantic tools for use and cooperative development by cognitive linguists by the time of the conference.

References

- Feldman, Jerome, Ellen Dodge and John Bryant. 2009. Embodied Construction Grammar. In Heine B., Narrog H., (eds). *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford U Press, p. 111–38.
- Krötzsch, Markus, Denny Vrandečić, Max Völkel, Heiko Haller, and Rudi Studer. 2007. Semantic Wikipedia. In *Journal of Web Semantics* 5 (4), pp. 251–261. Elsevier.
- Lakoff, George and Srini Narayanan. 2010. Toward a computational model of narrative. *Computational Models of Narrative: Papers from the AAAI Fall Symposium (FS-10-04)*
- Ruppenhofer, Josef, Michael Ellsworth, Miriam R. L. Petrucci, Christopher R. Johnson, and Jan Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice*. Berkeley, CA: FrameNet.