

---

## Roleo: distributional space visualisation for thematic fit modeling

---

Asad Sayeed  
Universität des  
Saarlandes

Xudong Hong  
Universität des  
Saarlandes

Vera Demberg  
Universität des  
Saarlandes

[asayeed@coli.uni-saarland.de](mailto:asayeed@coli.uni-saarland.de)

[xhong@coli.uni-saarland.de](mailto:xhong@coli.uni-saarland.de)

[vera@coli.uni-saarland.de](mailto:vera@coli.uni-saarland.de)

Donnerstag  
09.03.2017  
10:30 – 11:15  
B4 1, Foyer

We demonstrate Roleo, a web tool for visualizing selectional preferences. Different structured vector space models trained on large amounts of data (e.g., Distributional Memory, word embeddings) are in the back end of the tool and allow us to query different thematic role fillers of a given verb. For example, one can ask for agent, patient, or instrument of “eat” and can compare prototypical role fillers to a given role filler (e.g., “sausage” as a patient or “knife” as an instrument). Our tool provides a variety of options for visualizing such queries and comparing the query results of different models. These vector space models can be used for language applications involving the prediction of predicates and event participants.

Vector-space models differ in dimensionality and many other parameters; visualization helps us compare and hypothesize about fine-grained differences between models on this task. Roleo is a Django-based web tool that allows users to qualitatively compare different approaches to constructing distributional spaces for the selectional preferences task. Roleo’s basic paradigm is to construct a representative role-filler for a given verb-role from the vector space, and then surround this centroid with candidate noun fillers projected down from higher dimensions (via, e.g., singular value decomposition) to a two-dimensional canvas.

The source code for Roleo is provided at <http://github.com/tony-hong/roleo> and can be downloaded and easily customized to include new kinds of distributional spaces. Roleo can be used directly at <http://roleo.coli.uni-saarland.de/>.

**References:** • Baroni, M., Lenci, A. (2010): Distributional memory: a general framework for corpus-based semantics. *Computational Linguistics*: 36(4):673–721. • Sayeed, A., Demberg, V., Shkadzko, P. (2015): An exploration of semantic features in an unsupervised thematic fit evaluation framework. *IJCoL*: 1(1), 25–40. • Sayeed, A., Greenberg, C., Demberg, V. (2016): Thematic fit evaluation: an aspect of selectional preferences. *ACL RepEval*. • Tilk, O., Demberg, V., Sayeed, A., Klakow, D., Thater, S. (2016): Event participant modelling with neural networks. *EMNLP*.