3D Semantic Knowledge Retrieval

Eduardo San Miguel Martín.
Complutense University of Madrid

A work on Latent Semantic Analysis is presented. Upon the basis of semantic similarities a 3d framework for representing knowledge is built. This framework allows representation of words, sentences and whole texts semantic relationship.

LSA is a computational model of human knowledge representation that approximates semantic relatedness judgements. LSA has proven useful in a variety of comprehension and text processing situations. Although, several things are unsolved yet: representation alternatives, similarity measures, semantic bases choices, etc…

This work is an attempt to represent semantic knowledge in a tri - dimensional fashion using R. Due to the multidimensional nature of the semantic space our work has tried two alternatives: first, representing same knowledge from several perspectives at the same time; second, scaling dimension when possible.

LSA algorithm was fully implemented in R, Porter´s stemming algorithm included. Semantic space was built from a collection of Spanish texts. Measures used were cosine, Euclidean distance inverse and vector length.

Results showed that 3d plots are a useful tool in representing Latent Semantic Similarities. This kind of visualization offers a wide range of exploration capabilities.

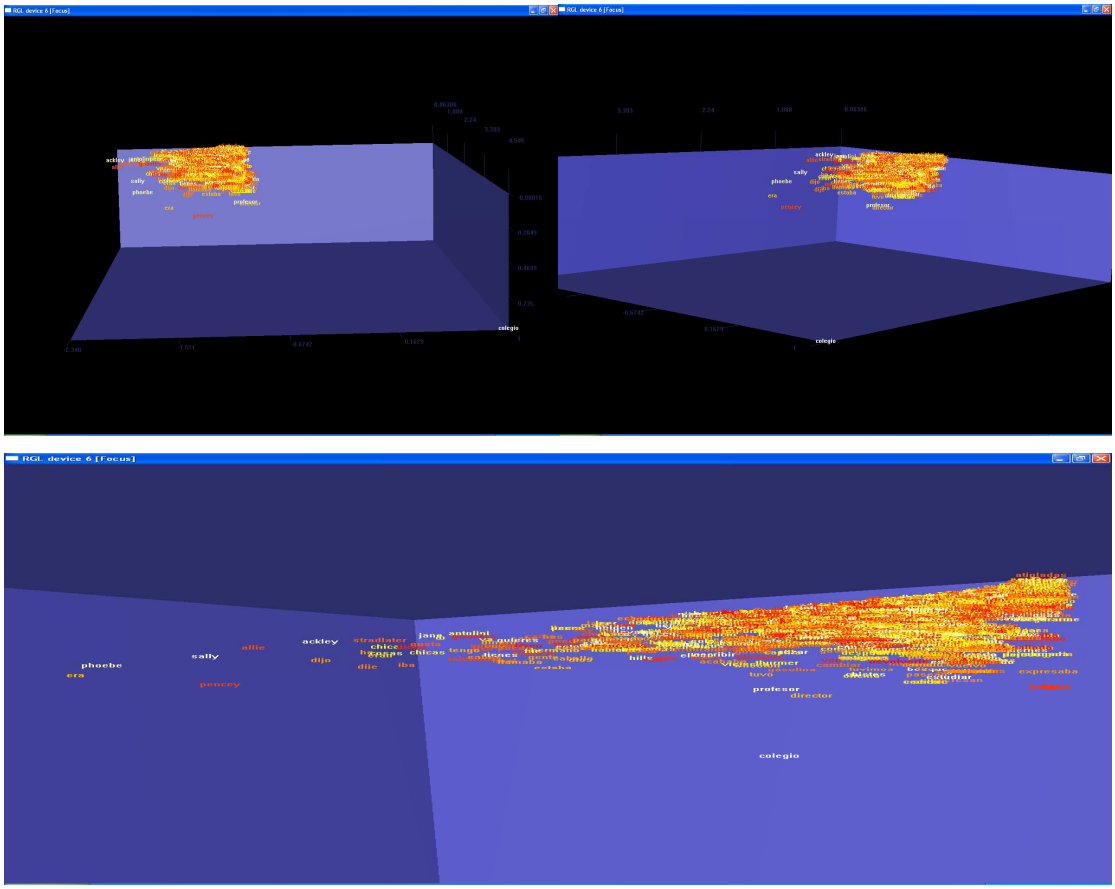

Figure 1.- 3d plot for word “Colegio” (“school”) from a semantic base built from “The catcher in the Rye”

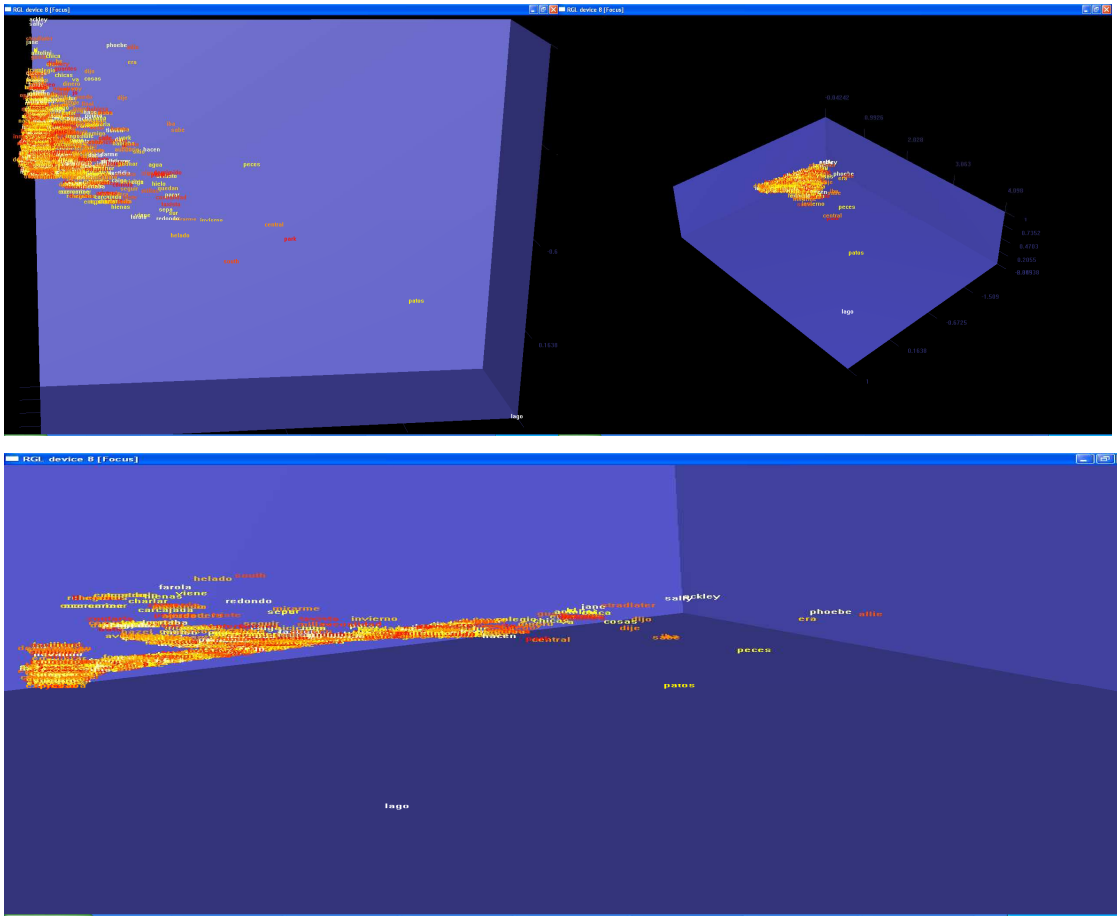

Figure 1.- 3d plot for word “Lago” (“lake”) from a semantic base built from “The catcher in the Rye”.