SIAM-LSA: An interactive activation model of sentence similarity

Peter Wiemer-Hastings

peterwh@cti.depaul.edu

DePaul University

School of Computer Science,

Telecommunications, and Information Systems

243 S. Wabash

Chicago IL 60604

Introduction

The ability to assess the similarity of objects in the world is fundamentally important to our survival. A variety of theories have been proposed for modeling human similarity judgments. Most of these theories involve comparing the sets of features of different objects to determine the overlap between them. But most of the theories also ignore the structure of the objects and relationships between the parts. Goldstone's SIAM system uses a connectionist architecture to create correspondences between objects and their features in different scenes (Goldstone, 1994). Excitatory connections reinforce coherent mappings between objects (e.g. ObjectA to ObjectC and ObjectB to ObjectD). Inhibitory connections fight against redundant or contradictory mappings (e.g. ObjectA to ObjectC and Object A to ObjectD). Likewise, connections between the features of objects either support or inhibit each other and the corresponding object-object connections. Siam's connectionist architecture allows it to take into account the structure of the scenes and the objects as well as the similarity of the features.

Recent language processing research has used similarity judgments between sentences as a method for "understanding" student utterances in an intelligent tutoring system (Wiemer-Hastings, Wiemer-Hastings, & Graesser, 1999). The system matches what the student says to its set of expected answers for the current question. A significantly high match indicates that the student has addressed that particular answer. Latent semantic analysis (LSA) uses a corpus-derived vector representation to compare texts. LSA takes no account for word order or sentence structure. By combining a simple surface parse with comparison of components, structured LSA (SLSA) has produced improved similarity judgments (Wiemer-Hastings & Zipitria, 2001). The current research merges the SIAM approach with SLSA to compare texts based on a connectionist representation of the correspondences between their components.

Mapping from objects to sentences.

In our first experiment, we took Goldstone's simple approach which calculates the similarity of scenes which contain simple objects (Goldstone, 1994). This approach represents a scene as a list of objects. Each object has a set of parts each of which has some value. For example, one of Goldstone's butterflies could be represented as: `(object1 (head square) (tail zig-zag) (body-shading white) (wing-shading checkered))`. To map this approach to sentences, we broke the inputs into subject, verb, and object parts. For example, the sentence "The dog bit a man" would be represented as `(object1 (verb ''bit'') (subject ''The dog'') (object ''a man''))`.

In the second experiment, we included the relationships between objects as Goldstone did in his later experiments. This allows Siam to account for human's tendency to favor maintaining the spatial relations between items. This led us to "promote" the status of a sentence to that of an entire scene. Here the previous sentence would be represented as: `[objects: (subject1 (head ''The dog'')), (object1 (head ''a man'')), relations: ''bit'' (subject1, object1)]`. With this approach the system can use the similarity between a subject and an object as part of the overall similarity judgment.

References

Goldstone, R. (1994). Similarity, Interactive Activation, and Mapping. *Journal of Experimental Psychology, 20*(1), 3–28.

Wiemer-Hastings, P., Wiemer-Hastings, K., & Graesser, A. (1999). Improving an intelligent tutor's comprehension of students with Latent Semantic Analysis. In Lajoie, S., & Vivet, M. (Eds.), *Artificial Intelligence in Education*, pp. 535–542 Amsterdam. IOS Press.

Wiemer-Hastings, P., & Zipitria, I. (2001). Rules for Syntax, Vectors for Semantics. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* Mahwah, NJ. Erlbaum. In press.