

Tim Oates and Paul R. Cohen
Experimental Knowledge Systems Laboratory
Department of Computer Science
Box 34610 LGRC
University of Massachusetts
Amherst, MA 01003-4610
Voice: (413) 577-0669
Fax: (413) 545-1249
oates@cs.umass.edu
cohen@cs.umass.edu

Format: paper or poster, with strong preference for paper

The Role of Dynamics in the Mental Lexicon

The dynamics of our interactions with the physical world are of fundamental importance both epistemologically and linguistically. Neonates are not usually ascribed conceptual thought, and yet somehow sensorimotor experience provides the building blocks for the rich conceptual structures present in adulthood. Simple dynamical features such as the manner of an object's motion provide much of the semantic content of verbs (Talmy 1985), and it has been suggested that those features play an important role in determining the syntactic constructions in which verbs can participate (Pinker 1989).

In this paper we present a dynamical representation, called dynamic maps, sufficiently general and expressive to capture both the meaning of individual verbs and the more abstract commonality of sets of semantically related verbs (Cohen 1998). Our purpose is to develop a complete theory of language acquisition that can be implemented on a physical platform, such as a mobile robot. Dynamic maps are multidimensional spaces, where each dimension represents some measurable physical quantity (Rosenstein 1998). The values of those quantities at any give time define a point in a map, and as they change over time a trajectory is traced through the map. The distance between two objects and the rate of change of that distance are examples of observable quantities that can serve as dimensions of dynamic maps.

We present the results of experiments with a system comprising a physically realistic simulation of interactions between objects and a grammar for generating sentences that describe those interactions as they unfold. Each word in the lexicon (terminal in the grammar) has an associated dynamic map that is updated with the trajectory that occurs when the word is heard. We show that repeated exposure to words and the contexts in which they occur results in dynamic maps with clusters of trajectories that capture the semantics of those words. Because the contents of maps are grounded in observable features of physical interactions, they support language generation. As new interactions unfold, the trajectories they trace can be matched against clusters of trajectories in existing maps, activating those lexical entries with "good" matches. Maps also support language understanding by providing a representation of the physical dynamics that are usually present (trajectory clusters) when a given word is uttered.

Finally, we explored the utility of dynamic maps as a representation for the semantic commonality of syntactically similar words. A distributional clustering technique based on mutual information was applied to a corpus of sentences generated by the aforementioned grammar, yielding a hierarchical clustering of the words (Brown et al. 1992). The leaves of the hierarchy contain individual words, and internal nodes correspond to sets of words that occur in similar contexts, with the size of the sets increasing as you move up the hierarchy. For example, PUSH and SHOVE may be leaves, and PUSH SHOVE may be an internal node. Rather than associating a dynamic map with just the leaves of the hierarchy (the words in the lexicon), a dynamic map is associated with every node in the hierarchy. We demonstrate that the resulting maps capture more abstract dynamical properties that define set membership, such as whether or not the members of the set involve contact between two objects.

Brown, P. F., Della Pietra, V. J., deSouza, P. V., Lai, J. C., and Mercer, R. L. 1992. Class-based n-gram models of natural language. *Computational Linguistics*, 18(4), pages 467 - 479.

Cohen, P. R. 1998. Maps for verbs. Under review.

Pinker, S. 1989. Learnability and cognition: the acquisition of argument structure. MIT Press.

Rosenstein, M. T. and Cohen, P. R. 1998. Concepts from time series. To appear, Fifteenth National Conference on Artificial Intelligence.

Talmy, L. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Shopen, T., ed., *Language Typology and Syntactic Descriptions III*. Cambridge University Press. 57-149.