

The Construction and Validation of Chinese Semantic Space by Using Latent Semantic Analysis

Ming-Lei Chen¹, Hsueh-Cheng Wang², and Hwa-Wei Ko¹

¹Institute of Learning and Instruction, National Central University

²Department of Computer Science, University of Massachusetts at Boston, MA.

Recently, using corpus to create a word net is a new approach in psycholinguistic study. The present study used Latent semantic analysis (LSA) to create a Chinese semantic space. The semantic relationship between words in the Chinese semantic space was estimated by taking the dot product (cosine) between two vectors. The semantic relationship between two sentences or two documents could be estimated in the same way. The results of human data indicate that the Chinese semantic space is a valid way to represent Chinese reader's world knowledge.

Keywords: *latent semantic analysis, semantic relation, semantic space*