

E-HowNet- a Lexical Knowledge Representation System

Keh-Jiann CHEN

Institute of Information Science, Academia Sinica

kchen@iis.sinica.edu.tw

E-HowNet is a frame-based entity-relation model extended from HowNet to define lexical senses and achieving compositional semantics. The followings are major extension features of E-HowNet to achieve the goal. a) Word senses (concepts) are defined by either primitives or any well-defined concepts and conceptual relations; b) A uniform sense representation model for content words, function words and phrases; c) Semantic relations are explicitly expressed; and d) Near-canonical representations for lexical senses and phrasal senses. We demonstrate the above features and show how coarse-grained semantic composition can be carried out under the framework of E-HowNet. Possible applications of E-HowNet are also suggested. We hope that the ultimate goal of natural language understanding will be accomplished after future improvement and evolution of the current E-HowNet.

We also design an ontology to express the taxonomic relations between concepts and the attributes of concepts. To establish the taxonomic relations between word senses, we introduce a strategy that constructs the E-HowNet ontology automatically. We then implement the lexical ontology as a Web application to demonstrate the taxonomy and the search functions for querying key-terms and E-HowNet expressions in the lexicon, which contains more than 88,000 lexical senses.