

# Chapter 12

## A Lexical Knowledge Representation Model for Natural Language Understanding

**Ping Chen**

*University of Houston-Downtown, USA*

**Wei Ding**

*University of Massachusetts-Boston, USA*

**Chengmin Ding**

*IBM Business Consulting, USA*

### ABSTRACT

*Knowledge representation is essential for semantics modeling and intelligent information processing. For decades researchers have proposed many knowledge representation techniques. However, it is a daunting problem how to capture deep semantic information effectively and support the construction of a large-scale knowledge base efficiently. This paper describes a new knowledge representation model, SenseNet, which provides semantic support for commonsense reasoning and natural language processing. SenseNet is formalized with a Hidden Markov Model. An inference algorithm is proposed to simulate human-like natural language understanding procedure. A new measurement, confidence, is introduced to facilitate the natural language understanding. The authors present a detailed case study of applying SenseNet to retrieving compensation information from company proxy filings.*

### INTRODUCTION

A natural language represents and models information of real world entities and relations. There exist a large number of entities in the world, and the number of relations among entities is even

higher. Entities and relations together make a highly complex multiple dimensional lattices. It is not a surprise that it usually takes a lot of training for a human being to speak, write and understand a natural language even with the fact that the computation power packed in a small human brain surpasses the most powerful super-computer in many aspects.

DOI: 10.4018/978-1-4666-0261-8.ch012

Human beings receive information through vision, hearing, smelling and touching, and send information through facial and body expressions, talking and writing. Of these communication channels, reading (from human vision), hearing, talking and writing are related to natural languages. All of them are temporally one-dimensional, and only one signal is sent out or received at a certain time point, so a natural language is communicated one dimensionally. With one-dimensional natural languages used by human being, in order to understand and describe a highly dimensional environment a series of filtering and transformations are necessary as illustrated in Figure 1. These transformations can be N-dimensional to N-dimensional or one-dimensional to N-dimensional in input process, and N-dimensional to one-dimensional or N-dimensional to N-dimensional in an output process. After these transformations information should be ready to be used by the central processing unit directly. Effectiveness and efficiency of these transformations are very important to knowledge representation and management.

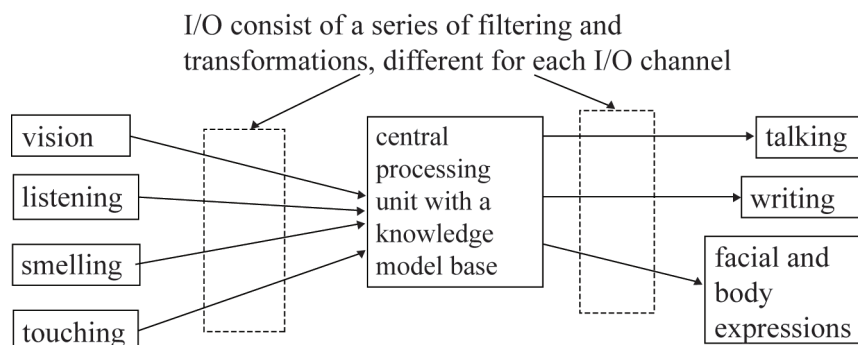
A knowledge model describes structure and other properties of a knowledge base which is part of a central processing system. A knowledge representation model is simply a mirror of our world, since one important requirement for a model is its accuracy. In this sense there is hardly any intelligence in a knowledge model or a knowledge base. Instead it is the communication

process consisting of filtering and transformations that shows more intelligent behaviors. As expressed by Robert C. Berwick, et al., in a white paper of MIT Genesis project (Berwick, et. al., 2004), “The intelligence is in the I/O”. As shown in Figure 1, a knowledge model may be the easiest component to start since its input has been filtered and transformed tremendously from the original format, and is ready to be stored in the knowledge base directly. On the other hand, a knowledge representation (KR) model plays a central role to any knowledge-based systems, and it eventually decides how far such a system can go. Furthermore, knowledge and experience can make the process of filtering and transformations more efficient and effective.

A KR model captures the properties of real world entities and their relationships. Enormous amounts of intervened entities constitute a highly complex multi-dimensional structure. Thus a KR method needs powerful expressiveness to model such information.

Many cognitive models of knowledge representation have been proposed in cognitive informatics. Several cognitive models are discussed in (Wang & Wang, 2006). Object-Attribute-Relation model is proposed to represent the formal information and knowledge structures acquired and learned in the brain (Wang, 2007). This model explores several interesting physical and physiological aspects of brain learning and gives a plausible estimation of

*Figure 1. Communication process for a knowledge-based system*



16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/lexical-knowledge-representation-model-natural/65130](http://www.igi-global.com/chapter/lexical-knowledge-representation-model-natural/65130)

## Related Content

---

### Recombination Operators in Permutation-Based Evolutionary Algorithms for the Travelling Salesman Problem

Camelia Chira and Anca Gog (2012). *Logistics Management and Optimization through Hybrid Artificial Intelligence Systems* (pp. 268-285).

[www.irma-international.org/chapter/recombination-operators-permutation-based-evolutionary/64925](http://www.irma-international.org/chapter/recombination-operators-permutation-based-evolutionary/64925)

### Comparative Study Between the MySQL Relational Database and the MongoDB NoSQL Database

Houcine Matallah, Ghalem Belalemand Karim Bouamrane (2021). *International Journal of Software Science and Computational Intelligence* (pp. 38-63).

[www.irma-international.org/article/comparative-study-between-the-mysql-relational-database-and-the-mongodb-nosql-database/280516](http://www.irma-international.org/article/comparative-study-between-the-mysql-relational-database-and-the-mongodb-nosql-database/280516)

### Performance Analysis of Multi-Hop Routing Protocol With Optimized Grid-Based Clustering for Wireless Sensor Network

Saloni Dhiman, Deepti Kakkar and Gurjot Kaur (2020). *Nature-Inspired Computing Applications in Advanced Communication Networks* (pp. 254-282).

[www.irma-international.org/chapter/performance-analysis-of-multi-hop-routing-protocol-with-optimized-grid-based-clustering-for-wireless-sensor-network/240958](http://www.irma-international.org/chapter/performance-analysis-of-multi-hop-routing-protocol-with-optimized-grid-based-clustering-for-wireless-sensor-network/240958)

### A High Level Model of a Conscious Embodied Agent

Jirí Wiedermann (2010). *International Journal of Software Science and Computational Intelligence* (pp. 62-78).

[www.irma-international.org/article/high-level-model-conscious-embodied/46147](http://www.irma-international.org/article/high-level-model-conscious-embodied/46147)

### Factors Determining the Success of eHealth Innovation Projects

Antonio Hidalgo, Nerea Pérez and Isaac Lemus-Aguilar (2022). *International Journal of Software Science and Computational Intelligence* (pp. 1-22).

[www.irma-international.org/article/factors-determining-the-success-of-ehealth-innovation-projects/309709](http://www.irma-international.org/article/factors-determining-the-success-of-ehealth-innovation-projects/309709)