ONTOLOGY-BASED EVENT RELATION PREDICTION
A SUMO BASED STUDY OF MANDARIN VV COMPOUNDS[*]

Jia-Fei Hong **Chu-Ren Huang**

National Taiwan Normal University   *The Hong Kong Polytechnic University*

ABSTRACT

This paper explores the interaction between eventive information and morpho-syntax based on Chinese VV compounds. Chinese VV Compounds' identical morpho-syntactic structure represents different event relations between the two component words and the correct interpretation of the meaning of these compounds relies on the prediction on their event relations. Without overt syntactic clues, we propose that ontology-based conceptual classification can be used to predict the event relation between the two component words. Compounding is the most productive way to research multi-word expressions in Mandarin Chinese. A Mandarin VV compound can be classified according to the eventive relation between two simplex verbs, which specifies how the eventive meanings of the two simplex verbs combine to form the meaning of the compound. The way in which two events combine with each other depends upon their event types, and the three types of eventive relations that we deal with in this paper are coordinate, modificational, and resultative. Using an ontology-based prediction approach, we hypothesized that the eventive relations could be predicted by the conceptual classification of the two simplex verbs' event types. First, we utilized SUMO and Sinica BOW to classify each simplex verb. Next, the correlation between the

ontology-based classification of each verb position and each eventive type was scored using a manually tagged lexical database and a training set was established. Finally, we encoded the ontological information of each VV compound in a 3-tuple based on these correlation scores. This 3-tuple was represented as a three-dimensional vector and was used to predict the eventive type of the new VV compounds. The results of our findings show that the classification experiments on event relation of unknown VV compounds can be reliably predicted based on the ontological classification of their component words.

知识本体驱动的中文复合词关系预测
SUMO 为本的研究

洪嘉馡   **黄居仁**

台湾师范大学   香港理工大学

提要

本文探讨概念结构和语法形态之间的相互作用。特别的是，我们将呈现以知识本体为本的概念分类，可以用来预测中文复合动词中两个组合词汇的语义关系。汉语中，复合词汇的组合是最常用来呈现多种词汇的表达。汉语复合动词的分类可以根据两个单一动词的事件关系，指定如何将两个单一动词结合的事件意义组合成一个复合动词，这个方法是取决于他们事件类型的结合。本文中，我们将处理三种不同的事件关系类型：协调一致的、修饰性的、结果性的。使用以知识本体的预测研究方法，我们假设事件关系可以靠着两个单一动词的事件类型的概念分类而被预测。首先，我们使用 SUMO 和 Sinica BOW 来对于每一个单一动词作分类；接着，在每一个动词基于知识本体为本的分类和事件类型的相关联性，取得以人工标记的词汇数据库，并建立训练语料；最后，我们根据这 3 组关联性的分数，编码每一个复合动词的知识本体信息。这 3 组关联性的分数代表一个三维向量并用来预测新复合动词的事件类型。我们的研究成果显示，对于未知复合动词的分类实验，其结果是可以得到确信的召回率和精确率。

主题词
中文复合动词   **SUMO**   **知**识本体   **概**念结构   **结**构—语法

Department of Applied
Chinese Language and Culture
National Taiwan Normal University
162, Heping East Road Section 1
Taipei, Taiwan
[jiafeihong@ntnu.edu.tw]