

Chinese Noun Phrase Metaphor Recognition with Maximum Entropy Approach¹

Zhimin Wang, Houfeng Wang, Huiming Duan, Shuang Han, and Shiwen Yu

Department of Computer Science and Technology
Institute of Computational Linguistics, Peking University, 100871, Beijing, China
{wangzm, wanghf, duenhm, yusw}@pku.edu.cn

Abstract. This paper presents a maximum entropy (ME)-based model for Chinese noun phrase metaphor recognition. The metaphor recognizing process will be viewed as a classification task between metaphor and literal meaning. Our experiments show that the metaphor recognizer based on the ME method is significantly better than the Example-based methods within the same context windows. In addition, performance is further improved by introducing additional features into the ME model and achieves good results in window (-2,+2).

1 Introduction

The task of identifying metaphors for a large-scale corpus has received an increasing amount of attention in the computational linguistics literature. Metaphors, one of figurative languages or tropes, can lead to inaccurate translation in Machine Translation systems and irrelevant document retrieval in Information Retrieval systems. For example, the Chinese word for “翅膀” means literally “wing of an animal”. However, when this word appears in a particular context, it has metaphorical expressions. For example,

张开 理想 的 翅膀 (meaning “explore fantasies”)
Spread fantasies of wings

where “翅膀” was not denoted the former literal meaning of “wing”, but has a metaphorical expression of “explore fantasies”. Information Retrieval systems should exclude this metaphorical expression while searching for “翅膀”.

Much research has gone into the processing of metaphors and provides some metaphor understanding systems such as the Met5, which is the first system to recognize examples of metaphors and metonymy under the guidance of preference constraint view⁴, the Structure-Mapping Engine (SME), a program for studying

¹ Supported by the National Grand Fundamental Research 973 Program of China under Grant No. 2004CB318102; the National High-Tech Research and Development Plan of China under Grant Nos. 2001AA114210, 2002AA117010 (863); the National Natural Science Foundation of China Under Grant No.60473138.