Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 60 (2015) 55 – 62

Procedia Computer Science

19th International Conference on Knowledge Based and Intelligent Information and Engineering Systems

A Simile Recognition System using a Commonsense Sensory Association Method

Eriko Yoshimura[a]*, Misako Imono [a], Seiji Tsuchiya [a], Hirokazu Watabe [a]

[a]*Dept. of Intelligent Information Engineering & Sciences, Faculty of Science and Engineering, Doshisha University, Kyo-Tanabe, Kyoto, Japan*

Abstract

We propose a method for comprehending the significance of metaphorical expressions by using an intuitive sensory association method implemented on computer. The metaphors we address in this work are similes. When simile expressions are used in human conversation, it appears that the listener uses intuitive sensory associations, cultivated through experience, to recall the characteristics of the subject and the predicate and comprehend the meaning of the metaphor by replacing the predicate with another appropriate word to describe the subject. A sensory association method has been proposed that is capable of clarifying these sorts of intuitive sensory relationships between nouns and their characteristics. The sensory association method outputs the sensations and impressions that humans naturally feel in response to a given noun. In this work, we construct a simile comprehension system based on the sensory association method and seek to use it to get computers to understand similes. In this paper, we define comprehension of a simile—such as "cheeks like apples"—as to convert the simile into the phrase "red cheeks." The capacity to perform this conversion demonstrates the computer's understanding that the two expressions are synonymous. The results of our tests indicated an accuracy of 65.7%; thus, by introducing a sensory association method we were able to exceed the accuracy achieved in a previous study.

Peer-review under responsibility of KES International

Keywords: Semantic understanding; Commonsense understanding, Natural language processing;

1. Introduction

We propose a method for comprehending the significance of metaphorical expressions by using a commonsense sensory association method implemented on computer. The study of metaphorical expressions originated in the

* Corresponding author. *E-mail address:* eyoshimura@indy.doshisha.ac.jp

Peer-review under responsibility of KES International
doi:10.1016/j.procs.2015.08.104

fields of linguistics [1] and psychology [2], and in recent years has been taken up as a subject of research in a variety of fields under the framework of natural-language processing [3][4]. About the study of metaphorical judgement, a study [5] gives 65.6% of effects by the pattern classification using the meaning classification.

The metaphors we address in this work are similes. These are metaphors that contain within the expression a phrase such as "like" or "as" that clearly indicate that the expression is a metaphor; a property of these metaphors is that the word being compared (the *subject* of the metaphor) and the word to which it is being likened (the *predicate*) are clearly identified. Similes are the primary object of study in research on metaphors. As one approach of study, there is a method using semantic network structure for neural network and the precision is around 30% [6]. The other, they express a set of properties indicating the respective characteristics of the subject and the predicate together with property values codifying the details of those properties, and the main focus of research involves quantitative computation regarding their relevance [3]. Techniques for representing sets of properties and property values include methods based on probability distribution values [7], methods rated based on cognitive psychology experiments [8], and methods using large-scale text data [9]. In addition, some systems [11] attempt to understand the meaning of similes using concept bases [10] that define the notion of words—which express sets of properties and property values—as sets of words (properties) that express the meaning characteristics.

These methods have enjoyed some degree of success. However, to utilize only knowledge expressed in terms of probability distributions of sets using these types of large-scale text data is to fail to take the intuitive senses of human beings into account, thereby limiting the capacity of such methods to understand the meaning of similes. As an example, consider the metaphorical expression "Her cheeks were like apples." Here the use of the word *apple* is intended not to convey the standard meaning of the fruit, apple, but instead is used to express an example of a red object. In a conversation between humans, the participants are able to grasp the meaning of the comparison being made in light of common sense. But it is just this element of *intuition*—which may be regarded as shared, mutually-understood information between humans—that is so difficult to reproduce in computer conversation.

When simile expressions are used in human conversation, it appears that the listener uses commonsense sensory associations, cultivated through experience, to recall the characteristics of the subject and the predicate and comprehend the meaning of the metaphor by replacing the predicate with another appropriate word to describe the subject. A sensory association method [12] has been proposed that is capable of clarifying these sorts of commonsense *sensory* relationships between nouns and their characteristics. The sensory association method outputs the sensations and impressions that humans naturally feel in response to a given noun. In this work, we construct a simile comprehension system based on the sensory association method and seek to use it to get computers to understand similes. In this paper, we define comprehension of a simile—such as "cheeks like apples"—as to convert the simile into the phrase "red cheeks." The capacity to perform this conversion demonstrates the computer's understanding that the two expressions are synonymous.

2. Sensory Association Method [12]

The sensory association method is a technique proposed for the purpose of outputting impressions and sensations naturally felt by human beings. In response to an input noun, the system attempts to commonsensible associate a given word with its characteristic, then outputs appropriate adjectival phrases.

Sensory words are adjectives that describe the sensations and impressions that humans experience in daily life; examples include "bitter," "painful," "cute," and "dirty." We have defined 215 words as sensory words.

2.1. Sensory knowledge base

A sensory knowledge base is a knowledge base that preserves relationships between words and sensations, constructed in reference to the layout of a thesaurus. The thesaurus used in this study [13] incorporates some 130,000 words; it uses a tree structure to represent local/global relationships and upper-level/lower-level relationships among some 2700 significant attributes that express the meaning conventions of general-purpose nouns. Our sensory knowledge base contains 2066 representative words selected from thesaurus leaves and 696 classifying words selected from thesaurus nodes, each with assigned sensory words. By turning the knowledge base into a thesaurus, it is possible to inherit the sensory words assigned to upper-level nodes.