Comparison of two weights of relation between words; recollection weight and relevancy one

Kaname Kasahara **Inago Nozomu**
Nippon Telephone and Telegraph Corporation
Tomoko Kanasugi **Chiharu Nagamori**
NTT Advanced Technology Corporation
Tsuneaki Kato
Language and Information Sciences, The University of Tokyo

Major theme of the paper: Analysis of language sense in the society
Address: NTT Communication Science Laboratories, 2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237 Japan
phone numbers: +81 774 93 5353(tel), +81 774 93 5353(fax)
Email: kaname@cslab.kecl.ntt.co.jp

Abstract

In this paper, two sorts of weights of relation between words, recollection weight and relevance one, are compared to each other and discussed based on the results of two experiments employing human subjects. We selected synonymy, antinomy, and associativity as the relations to be tested and found that there were basically middle degrees of correlation between the two sorts of weight in synonym and antinomy. However, there was low degree of correlation in associativity which suggested there will be much more words which can be judged to be relevant associative to a stimulus word than ones which are simply recollected.

1 Introduction

Simulations of semantic judgments about words, such as retrieving associative words and clustering words based on similarity (for example, (Schutze, 1992; Kasahara et al., 1997)), are useful for realizing intelligent text processing technologies like information retrieval(Deerwester et al., 1990; Schutze and Pedersen, 1995), machine translation(Ikehara et al., 1991), and text segmentation(Bessho, 2001). When studying a method for the simulation, it is required to evaluate how the result of the method is close to human judgment.

Usually, a large-scaled database of human judgments, which contains word pairs and names of relations that connect each pair are used for that purpose. In English, WordNet(Fellbaum, 1998) is one of the famous databases and there are Nihongo-GoiTaikei(Ikehara et al., 1997) and EDR(Japan Electronic Dictionary Research Institue, 1993) in Japanese. These databases were constructed based on the knowledge of a few linguists or psychologists. Therefore, they contain words and relations which even native speakers of a language may hardly use in their life. Moreover, information about weight of relation between words, which represents how strongly the words are connected in a given kind of relation such as synonym, is lacked in the databases. This weight information is important when a database of human judgments about words is used to get the most related word to a given word or get a ranked word list based on the weights of relations between a stimulus word and target words.

In the studies of psychology, there are several norms of word association (Moss and Older, 1996; Nelson et al., 1998; Umemoto, 1969), which were build through human subjects experiments where the free association method was used. They contain weights of associated words, which are estimated based on how many subjects recollected the same words. We call the weight 'recollection weight.' Meanwhile, the other sort of weight of a relation, which can be estimated from how many subject judge the target word has a relevant relationship with a stimulus word, may exist. We call it 'relevancy weight.' If there is always high degree of correlation between the recollection weight and the relevancy one in any kind of relation, experiment for acquiring only one of the two weight is enough to build the database of human judgments about a relation between words. However

it has not been clear whether there is correlation or not.

In this paper, we report the result of measuring correlation between the two sorts of weight in the relations of synonymy, antinomy, and associativity. We also show the database made through them is effective by applying it to evaluate the method of simulating judgment of synonymy between words employing machine readable dictionaries(Kasahara et al., 1997).

2 Experiments

Stimulus words
We used the database of word familiarity(Amano and Kondo, 1999) to select 200 Japanese daily-used words as stimuli. The database includes 60,000 words each of which has degree of its familiarity, from one to seven, decided from psychological experiments. The words whose degree of familiarity are higher than five were randomly selected from the database; odoru('dance'), uma('hourse'), dainamikku('dynamic'), and so on.

Kinds of relation
It has been thought that there are several kinds of relation between two words. For example, in WordNet(Fellbaum, 1998), seventeen relations such as HYPERNYM and SYNSET are listed. As the first step of our study, the following relations, synonymy, antinomy, and associativity, were tested.

Experiment 1
100 subjects of university students were asked to write down recollected words for every stimulus word of the 200 words to questionnaire sheets in a given relation of the three ones. They were required to do as many words as they can in ten seconds for one stimulus word. For each stimulus word and each one of the relations, frequency of how many subjects wrote the same associated word was counted.

Experiment 2
For each word recollected at the Experiment 1, every pair of the stimulus word and the recollected target word was shown to other seventy six subjects and they judged whether the target word has a given relation with the stimulus word or not. Frequency of how many subjects commonly judge the pair to be relevant was counted.

3 Results

Table 1 shows the correlation coefficients between two sorts of frequency. These figures may suggest that the recollection weight and relevancy one have middle degree of correlation in synonymy and antinomy. In associativity, however, the degree of correlation has low degree of correlation.

Table 1: Correlation between recollection frequency and relevancy frequency

relation	correlation coefficient
synonymy	0.465
antinomy	0.418
associativity	0.294

Figure 2 shows correlation between the recollection frequency and averaged relevant frequencies of word pairs, which have the same recollection frequency, in the relation of association. Even when several words were recollected by only one of the 100 subjects, these words were judged to be relevant to have a associative relations by more than half of other 76 subjects on average. This result suggests that there is many possibly relevant words except for recollected words in the relation of free associativity. Detailed analysis of the results of two experiments will be required.

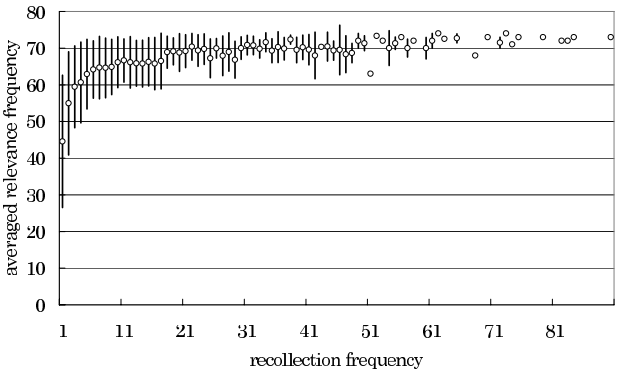

Figure 1: Correlation in associativity

Figure 2 is the result of applying the database made from the two experiments to evaluation of the method for simulating judgment of semantic similarity between words(Kasahara et al., 1997). This method represents a word concept in a vector form whose weights of attributes are decided from word definitions in machine read-

able dictionaries. The degree of similarity between two word become a cosine of an angle between their word vectors. Basically, the closer definitions of two word are to each other, the higher the degree of similarity become. When this method is applied to simulate similar word retrieval, the degrees of similarity between a given word and all the words which have word vectors are calculated and some words which have high degrees of similarity are output as the result.

For a stimulus word, we assumed that words, whose recollection frequency are more than two and whose relevant weights are more than half of the number of the subject, can be regarded as standard related words. For each stimulus of 200 words, the result of above mentioned simulation and the standard related words were compared by using a well-known method of evaluating a result of retrieval (Buckley, 1992) and values of precision and recall were acquired in eleven points. Figure 2 is the precision–recall curves averaged on 200 retrievals in the relation, synonymy, antinomy, or associativity. It shows that this method simulates judgment of synonymy better than antinomy and associativity. The result also suggest that the other method of simulation or the other text data to build the word vectors is necessary for simulating antinomy or associativity.

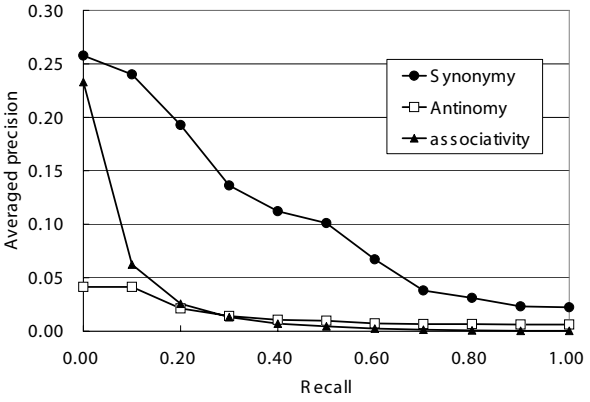

Figure 2: Evaluation of a method of the simulation

References

Shigenari Amano and Kimihisa Kondo. 1999. *Goi-Tokusei (Lexical properties of Japanese) Vol. 1.* Sanseido, Tokyo.

Katsuji Bessho. 2001. Text segmentation using word conceptual vectors (in japanese). *Trans. of IPSJ*, 42(11):2650 – 2662.

Chris Buckley. 1992. Smart version 11.0. ftp://ftp.cs.cornell.edu/pub/smart.

S. Deerwester, S. Dumais, G. Furnas, T. Landauer, and R. Harshman. 1990. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41:391 – 407.

Christiane Fellbaum, editor. 1998. *WordNet : an electronic lexical database.* The MIT Press.

Satoru Ikehara, Satoshi Shirai, Akio Yokoo, and Nakaiwa Hiromi. 1991. Toward an mt system without pre-editing -effects of new methods in alt-j/e-. In *MT Summit '91*, pages 101–106.

Satoru Ikehara, Masahiro Miyazaki, Satoshi Shirai, Akio Yokoo, Hiromi Nakaiwa, Kentaro Ogura, Yoshifumi Oyama, and Yoshihiko Hayashi, editors. 1997. *The Semantic System, volume 1 of Goi-Taikei – A Japanese Lexicon.* Iwanami Shoten.

Ltd. Japan Electronic Dictionary Research Institue, 1993. *EDR Electronic Dictionary Technical Guide*, tr-042 edition.

Kaname Kasahara, Kazumitsu Matsuzawa, and Tsutomu Ishikawa. 1997. A method for judgment of semantic similarity between daily-used words by using machine readable dictionaries (in japanese). *Trans. of IPSJ*, 38(7):1272 – 1283.

Helen Moss and Liann Older. 1996. *Birkbeck Word Association Norms.* Psychology Press.

D. L. Nelson, C. L. McEvoy, and T. A Schreiber. 1998. The university of south florida word association, rhyme, and word fragment norms. In *http://www.usf.edu/FreeAssociation/.*

H. Schutze and J.O. Pedersen. 1995. Information retrieval based on word senses. In *Fourth Annual Sympo. on Document Analysis and Information Retrieval*, pages 161–175.

H. Schutze. 1992. Dimensions of meaning. In *Proceedings of Supercomputing 92*, pages 787–796.

Takeo Umemoto. 1969. *Renso Kijun Hyou.* University of Tokyo Press.