Evaluation of Similarity Measure employing Point-of-View Reinforcement

Kenji Nagamatsu and **Hidehiko Tanaka**
University of Tokyo
Hongo 7-3-1, Bunkyo-ku, Tokyo, Japan, 113

1 Introduction

Two different words may not be similar in general, rather they are similar under some aspects or point-of-views. This paper proposes a new similarity measure between words based on point-of-views. The method utilizes co-occurrence probability-based similarity as a basis and extends it by weighting the values according to the relevance between input words and point-of-view words(called *point-of-view reinforcement*).

2 Similarity with Point-of-View

Based on both corpus- and feature-based measures the formulation of our similarity $Sim(w_1, w_2; w_p)$ is defined.

$$Sim(w_1, w_2; w_p) = \sum_{\forall w \in Co(w_1) \cap Co(w_2)} \frac{Pr(w|w_1; w_p) + Pr(w|w_2; w_p)}{2}$$

$Pr(w|w'; w_p)$ denotes the co-occurrence probability of w conditioned by w' and reinforced by a point-of-view w_p, $Co(w)$ the set of words co-occurring with w.

The *point-of-view reinforcement* is responsible for *modulating* this basic similarity by point-of-view words.

$$Pr(w|w'; w_p) = \frac{\alpha^{\mu(w_p, w)} f(w|w')}{(\alpha^{\mu(w_p, w)} - 1) f(w|w') + \sum_{\forall x \in Co(w')} f(x|w')}$$

$f(w|w')$ denotes the normal co-occurrence frequency and α is a parameter controlling how the relatedness between two point-of-views(w_p, w) affects the similarity.

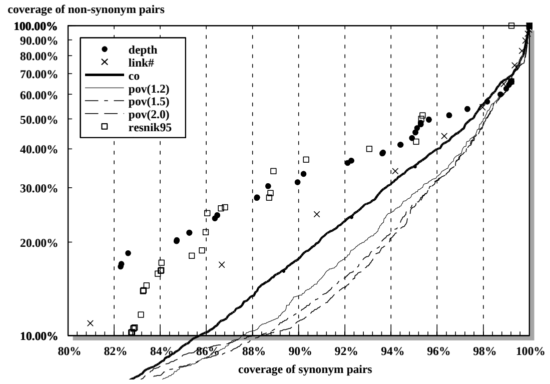
$\mu(w_p, w)$ is the factor indicating the relatedness between input words and a point-of-view word. It is defined as the mutual information content between w_p and w and approximated with another type of co-occurrence data extracted from a tagged corpus.

3 Experiments

One experiment is a selectivity test([Nagamatsu and Tanaka, 1996]) with large word-pair sets of synonyms and non-synonyms. This evaluates the whole attitude of similarity measures(see the figure).

The result shows clearly that the corpus-based measures(co, pov*) are superior to the thesaurus-based ones(link#, depth). Moreover, among these corpus-based measures, employing the point-of-view reinforcement(pov) makes the selectivity higher than its original co(the lower a data sequence is located, the higher the selectivity of the measure becomes).

The other is a experiment employing human subjects. This shows the correlation between similarity values and rating scores by human subjects(see the table).



Sim measure	Whole	IPAL	Bunrui-Goi-Hyo
resnik95	$\underline{0.426}$	$\underline{0.235}$	0.420
pov(2.0)	0.424	0.232	$\underline{0.495}$
pov(1.2)	0.390	0.210	0.415
depth	0.380	0.164	0.449
link#	0.365	0.104	0.442
co	0.344	0.211	0.306

This experiment shows that the thesaurus-based measures(depth, link#, resnik95) have higher correlation with human judgment than the corpus-based ones(co). By employing point-of-view reinforcement, however, the derived measures(pov*) have become even better than the thesaurus-based measures and when the parameter α is adequately selected, the highest correlation has been achieved.

4 Conclusion

From the experiments it is concluded that the proposed similarity measure can distinguish synonym pairs from non-synonym pairs better than other similarity measures(selectivity test) and that the measure has high correlation with the rating scores by human subjects.

References

[Nagamatsu and Tanaka, 1996] Kenji Nagamatsu and Hidehiko Tanaka. Estimating point-of-view-based similarity using pov reinforcement & similarity propagation. In *Proceedings of PACLIC 11 Language, Information and Computation*, pages 373–382, December 1996.

[Resnik, 1995] Philip Resnik. Using information content to evaluate semantic similarity in a taxonomy. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, volume 1, pages 448–453, 1995.