Proceedings of the 2012 International Conference on Machine Learning and Cybernetics, Xian, 15-17 July, 2012

COMPUTATION OF BILINGUAL WORD SIMILARITIES BY METAPHORICAL PROPERTIES

HAI-BO KUANG[1], BIN LI[1,2], CHEN CHEN[3], PENG JIN[4], XIAO-HE CHEN[1]

[1]Research Center of Language and Informatics, Nanjing Normal University, Nanjing 210000, China
[2] State Key Lab for Novel Software Technology, Nanjing University, Nanjing 210000, China
[3] Department of English Language and Literature, Graduate School of Shanghai International Studies University, Shanghai 200000, China
[4] School of Computer Science, Leshan Normal University, Leshan 614000, China
E-MAIL: lib@nlp.nju.edu.cn

Abstract:

Measuring word similarities is a fundamental issue in NLP, while the measuring procedure is always aided by dictionaries or corpus. However, when figurative or metaphor usages are considered, the situation becomes more complicated. Therefore, based on the Chinese-English bilingual lexical cognitive property knowledgebase we constructed, we design several algorithms to compute word similarities by metaphorical properties. Through experiments based on four different algorithms and over 200 word pairs in the knowledgebase, we find the algorithm DSOMP yields better results despite of the differences of word semantic classes and languages.

Keywords:

Word similarity; Linguistic knowledgebase; Metaphor computation

1. Introduction

Measuring word similarities is a fundamental issue in NLP, and has broad range of application areas[1][2]. Concerning how to measure word similarities automatically, many scholars utilize dictionaries or corpus to achieve this goal[3][4][5]. Besides, the computation of bilingual word similarities is more difficult but practicable by the present language resources[6][7].

However, when figurative or metaphor usages are considered, the situation becomes more complicated. Firstly, the dictionaries and corpus are both lack of metaphor properties themselves. Secondly, the algorithms based on semantic classes may cause errors when we want to compute word similarities by metamorphic properties, as the metamorphic properties are always independent with sematic classes. For instance, by the similarities of semantic classes, two English words *ant* and *elephant* maybe quite similar, while they share opposite metaphoric properties, as *strong* and *weak*. Another example is measuring word similarities of two Chinese words *火箭(rocket)* and *豹(leopard)*. Although they have few similar semantic properties, they share similar metaphoric properties of *快(fast)*.

As far as computation of bilingual word similarities, what we concern about is the metamorphic differences across languages, such as whether *pig* and *猪(pig)*, or *snowflake* and *纸 (paper)* have similar metamorphic properties despite of translations, can we find bilingual words share similar metaphoric properties, can we judge whether bilingual words share similar or different metaphoric properties automatically. In fact, to answer these questions, we also lack of the metaphoric properties of bilingual words, and the algorithms always focus on the semantic classes, while ignoring the metaphoric properties too.

Based on these considerations, we decide to employ the Chinese-English bilingual lexical cognitive property knowledgebase to resolve resource problems, not only in one language, but also across languages. What's more, we design several algorithms to compute word similarities by metaphorical properties, even across semantic classes and languages.

2. Data Resources

To resolve resource problems, we firstly construct the Chinese-English bilingual lexical cognitive property knowledgebase, which contains noun vehicles and their adjective metaphoric properties, as Veale(2007)[8]. Veale and Hao(2007) collected a large scale of English similes to construct the English lexical metaphorical property knowledgebase, which contained word pairs as "noun vehicle-adjective metaphorical property", using search

978-1-4673-1487-9/12/$31.00 ©2012 IEEE

engine and WordNet[1]. The knowledgebase built by Veale mapped 3,769 adjective metaphorical properties to 9,286 noun vehicles. Considering the practical use, they classified the word pairs into two categories: factual (like *horse-strong*) and ironic (like *ant-big*). Based on this, we manually filter out the simple comparisons, and link words to HowNet[2] instead of WordNet to give both Chinese and English expressions and do the comparison across languages.

Although we have given words in English knowledgebase both Chinese and English expressions, they actually reflect the metaphoric usages of English. Thus, we decide to construct the Chinese knowledgebase[3]. As Jia(2009), we collect "noun vehicle-adjective metaphorical property" word pairs in Chinese, using specific simile sentence, *X 像 Y 一样 P*(which means "*X is as P as Y*") and Chinese search engine Baidu[9]. Then 18,205 word pairs are left after trimming the simple comparisons. We also manually classify word pairs to two categories: factual and ironic, and give both Chinese and English expressions on the basis of HowNet. The difference is that, we reserve the frequencies of word pairs to take a further observation[10].

Now, we can compare two knowledgebase to find out metaphoric differences between Chinese and English. The comparison is very interesting, as we know that the most frequent used noun vehicles in two languages are different[11]. What's more, we can make integration through reserving the word pairs, which have the same definitions in HowNet, to construct a new bilingual knowledgebase. Now, we have 1,065 bilingual word pairs. Thus, it is not difficult to find that the metaphorical properties of some vehicles are bilingually the same. Table 1 shows the top 10 bilingual vehicles which have the largest number of adjective metaphorical properties.

TABLE 1. TOP 10 MOST SIMILAR VEHICLES

ID	Chs Noun	Eng Noun	Properties
1	水晶	crystal	清,清澈,纯,纯净-pure;脆-clear
2	花	flower	新-fresh;甜-sweet;纯真-pure
3	妈妈	mother	好-good; 温柔,柔和-gentle
4	蚂蚁	ant	慢-slow; 渺小,小-tiny
5	蛋糕	cake	甜美,甜蜜-sweet,luscious
6	糕点	cake	甜美,甜蜜-sweet,luscious
7	糖	sugar	甜蜜,好吃-sweet,nice
8	婴儿	baby	裸-bare,naked
9	海洋	ocean	宽广,大-broad
10	针	needle	锋利-sharp,incisive

1 *http://wordnet.princeton.edu*
2 *http://www.keenage.com*
3 *The Chinese knowledgebase is available at http://nlp.nju.edu.cn/lib/cog/ccb_nju.php*

3. Computation of word similarities by metaphorical properties

In section 2, we have constructed the Chinese-English bilingual lexical metaphor property knowledgebase. In this section, we try to compute word similarities by metaphorical properties using several algorithms based on the knowledgebase. Liu(2002) have studied how to compute word similarities with HowNet[5]. Xia(2011) has measured Chinese-English word similarity with HowNet and parallel corpus[6]. So we will use the core algorithm of *sim*, described in Liu's paper, to participate in our computation procedure, using HowNet as well.

Our core idea is to transform the computation of word similarities by metaphorical properties to the computation of word similarities of metaphorical properties, which are mapped by noun vehicles in the knowledgebase. Although we believe that the two computations are not the same, and it is need to measure the adverse effects of the transforming. As far as the computation itself, the transforming has already reduced noises. Here are three examples generally described in Section 1, to show the noise reducing.

3.1. Examples

Example 1:

{Vehicle: ant- property: weak (Factual)}

{Vehicle: elephant- property: strong (Factual)}

Here are two English word pairs. And we want to compute the word similarity of *ant* and *elephant* by metaphorical properties.

Before computing, we have to make rules to control the computation and evaluate the result. We preliminarily suppose the threshold of word similarities by metaphorical properties is between -1 to 1, while the threshold of word similarities by semantic properties is often ruled as 0 to 1.

We use *SimMeta(v1, v2)* to represent the word similarity of vehicle *v1* and *v2* by metaphorical properties. Thus, we want to get the value of *SimMeta(ant,elephant)*. If we estimate the value through computing word similarities of vehicles by semantic properties, we infer that *SimMeta(ant, elephant)* $\approx$ *Sim(ant,elephant)= 0.094*. In another way, we can compute word similarities by metaphorical properties to estimate the value, that is to say we can infer *SimMeta(ant, elephant)* $\approx$ *Sim(weak, strong)= 0.029*. Obviously, the latter measure reduces some noises and is closer to -1, the right value considered by people.

Example 2:

{Vehicle: 火箭(rocket)- property: 快(fast) (Factual)}

{Vehicle: 豹(leopard)- property: 快(fast) (Factual)}

Here are two Chinese word pairs, and we want to compute the word similarity of *火箭* and *豹* by metaphorical properties. If we estimate the value through computing word similarities of vehicles by semantic properties, we infer that *SimMeta(火箭,豹)≈Sim(火箭,豹)= 0.015*. In another way, we can compute word similarities by metaphorical properties to estimate the value, that is to say we can infer *SimMeta(火箭,豹)≈Sim(快,快)=1.000*. Obviously, the latter measure reduces some noises and is closer to the right value.

Example3:

{Vehicle:纸(paper)-property: 白(white) (Factual)}

{Vehicle: snowflake-property: white (Factual)}

Here are two bilingual word pairs across languages. Suppose we want to judge whether *纸* and *snowflake* share similar word similarity by metaphorical properties. If we estimate the value through computing word similarities by semantic properties across languages, which is possible using the definitions in HowNet, we infer that *SimMeta(纸,snowflake)≈Sim(纸,snowflake)= 0.021*. In our way, we compute word similarities by metaphorical properties to estimate the value. Thus, we infer *SimMeta(纸,snowflake)≈Sim(白,white)=1*. Obviously, the latter measure hints word similarity by metaphorical properties more accurately.

3.2. Computing Formulas

The above examples are quite simple, while the vehicles only each have one factual metaphorical property. Thus we decided to design more complicated algorithms to adapt to the vehicles in the knowledgebase. Before given the core algorithm, we design two less simple algorithms.

a. The Formula of Intersections of Metaphorical Properties (IOMP)

The core idea of IOMP is to reserve the differences about the proportion of the same word types of the metaphorical properties mapped by vehicles as the same relationship and the proportion of the same word types of the metaphorical properties mapped by vehicles as the opposite relationship. The formula is as follows.

$$\mathrm{SimMeta}(\mathrm{v1},\mathrm{v2}) = \left(\frac{P(v1.F)\cap P(v2.F)}{P(v1.F)\cup P(v2.F)} + \frac{P(v1.I)\cap P(v2.I)}{P(v1.I)\cup P(v2.I)}\right) - \left(\frac{P(v1.F)\cap P(v2.I)}{P(v1.F)\cup P(v2.I)} + \frac{P(v1.I)\cap P(v2.F)}{P(v1.I)\cup P(v2.F)}\right) \quad (1)$$

SimMeta(v1,v2) represents the word similarity of vehicle *v1* and *v2* by metaphorical properties, *P(v1.F)* represents the set of the factual metaphorical properties of *v1*, *P(v1.I)* represents the set of the ironic metaphorical properties of *v1*, other elements of the formula share the similar implication.

b. The Formula of Related Field of Metaphorical Properties (RFOMP)

The Core idea of RFOMP is computing the degree of cross-correlation of the metaphorical properties of two vehicles through calculating the differences about the proportion of the same word definitions of the metaphorical properties mapped by vehicles as the same relationship and the proportion of the same word definitions of the metaphorical properties mapped by vehicles as the opposite relationship, which is similar to IOMP to same extent. The formula is as follows.

$$\mathrm{SimMeta}(\mathrm{v1},\mathrm{v2}) = \left(\frac{S(v1.F)\cap S(v2.F)}{S(v1.F)\cup S(v2.F)} + \frac{S(v1.I)\cap S(v2.I)}{S(v1.I)\cup S(v2.I)}\right) - \left(\frac{S(v1.F)\cap S(v2.I)}{S(v1.F)\cup S(v2.I)} + \frac{S(v1.I)\cap S(v2.F)}{S(v1.I)\cup S(v2.F)}\right) \quad (2)$$

SimMeta(v1,v2) represents the word similarity of vehicle *v1* and *v2* by metaphorical properties, *S(v1.F)* represents the set of the definitions of the factual metaphorical properties of *v1* in HowNet, *S(v1.I)* represents the set of the definitions of the ironic metaphorical properties of *v1* in HowNet, other elements of the formula share the similar implication.

c. The Formula of Degree of the Similarity of Metaphor Properties (DSOMP)

Our idea is illustrated by the formula as follows. This formula transforms the computation of word similarities of vehicles by metaphorical properties to the computation of word similarities of metaphorical properties, which are mapped by vehicles.

$$\mathrm{SimMeta}(\mathrm{v1},\mathrm{v2}) = \left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} sim(v1.Fi, v2.Fj)}{m\times n} + \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} sim(v1.Ii, v2.Ij)}{m\times n}\right) - \left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} sim(v1.Fi, v2.Ij)}{m\times n} + \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} sim(v1.Ii, v2.Fj)}{m\times n}\right) \quad (3)$$

SimMeta(v1,v2) represents the word similarity of vehicle *v1* and *v2* by metaphorical properties, *Sim(v1.Fi,v2.Fj)* represents the word similarity of factual metaphorical property *i* of *v1* and factual metaphorical property *j* of *v2*, other elements of the formula share the similar implication.

We should also note that, in the whole computation processing, we abandon the value of *sim=-1*, which hits the appearance of the out-of-vocabulary words in HowNet, described in Liu's Paper[5].

4. Experiments and Analysis

The experiments actually consist of three steps, experiment 1 is choosing 10 groups of English vehicles and computing to judge if each group shares similar metaphorical properties. Experiment 2 is to choose 16 Chinese vehicles, in a semantic class or across semantic classes, then compute word similarities by metaphorical properties to arrange them by word similarities of metaphoric properties. At last, we select 10 bilingual vehicles in table 1 and another 10 bilingual vehicles randomly. Then we will compute word similarities of these vehicles by metaphorical properties across languages. All these experiments will utilize our three algorithms.

Experiment 1 : The computation of English word similarities

We choose 10 groups of English vehicles shown in table 2. It is easy to find out that 10 groups can be classified to 5 larger groups. Specifically, the vehicles in first two groups share similar metaphorical properties and belong to one semantic class. The vehicles in group 3, 4 share opposite metaphorical properties in the same semantic class. While the vehicles in group 5, 6 share similar metaphorical properties in different semantic classes, and the vehicles in group 7, 8 share opposite metaphorical properties in different semantic classes. In the last two groups, not the same as the former 8 groups, we find out-of-vocabulary words of HowNet in it.

TABLE 2. 10 GROUPS OF ENGLISH VEHICLES IN THE ENGLISH KNOWLEDGEBASE

Group	V1	V2
1	leopard	tiger
2	candy	fruit
3	leopard	tortoise
4	knife	pillow
5	athlete	leopard
6	lotus	pearl
7	professor	monkey
8	comet	snail
9	feather	artillery_shell
10	ice_water	jalapeno_pepper

Instead of displaying the calculation process, we show the results in table 3. From Analyzing the result, we find that *sim(v1,v2)* is sensitive to the semantic classes of vehicles and will be helpless to find out word similarities by metaphorical properties. On the other hand, IOMP, RFOMP and DSOMP are all helpful to judge if each group shares similar metaphorical properties. The accuracy rates are even 100%. Furthermore, the distribution of DSOMP is more evenly to make the further observation.

TABLE 3. THE RESULT OF STEP 1

Group	word similarity	sim(v1,v2)	IOMP	RFOMP	DSOMP
1	similar	0.950	0.095	0.098	0.097
2	similar	0.019	0.125	0.250	0.268
3	opposite	0.094	-0.034	-0.096	-0.104
4	opposite	0.498	-0.048	-0.167	-0.199
5	similar	-INF[4]	0.052	0.085	0.093
6	similar	0.019	0.039	0.088	0.063
7	opposite	0.035	-0.019	-0.157	-0.105
8	opposite	0.019	-0.037	-0.136	-0.145
9	opposite	-INF	-0.091	-0.125	-0.121
10	opposite	-INF	-0.167	-1	-0.485

Experiment 2 : The computation of Chinese word similarities

We choose 8 Chinese vehicles belong to the same semantic class and another 8 Chinese vehicles across different semantic classes in the Chinese knowledgebase, shown in table 4, to conduct our experiment. We should point out that to make the experiment better understanding, metaphorical properties of these chosen vehicles are clear, as the vehicles convey the metaphoric properties of *fierce* or *not fierce* in group 1 and the metaphoric properties of *sweet* or *bitter* in group 2.

TABLE 4. CHINESE VEHICLES IN TWO GROUP

Group 1	**The same semantic domain**	虎(tiger),山羊(goat),猎豹(leopard),狮(lion),野猪(boar),熊猫(panda),乌龟(tortoise),蜗牛(snail)
Group 2	**The different semantic domains**	蜂蜜(honey),林依晨(Ariel Lin),中药(traditional Chinese medicine),香烟(cigarette),蛋糕(cake),黄连(coptis),甜点(dessert),苦瓜(momordica)

Technically, we make the value of word similarities of *虎* and the value of *蜂蜜* equal to 1.000, to set an initial point for the two groups. Then, we compute word similarities of all the vehicle pairs like *SimMeta(虎,兔)* and *SimMeta(蜂蜜,甜点)* by metaphorical properties. All these values can be checked in table 5, 6 with four different algorithms.

TABLE 5. CHINESE VEHICLES IN THE SAME SEMANTIC DOMAIN

vehicles	sim(v1,v2)	IOMP	RFOMP	DSOMP
虎	1.000	1.000	1.000	1.000
山羊	0.950	-0.056	-0.111	-0.313
猎豹	-INF	0.056	0.111	0.210
狮	0.950	0.000	0.333	0.479
野猪	0.685	0.000	0.000	0.204
熊猫	0.950	-0.050	-0.083	-0.081
乌龟	0.094	-0.125	-0.333	-0.324
蜗牛	0.094	-0.071	-0.333	-0.324

TABLE 6. CHINESE VEHICLES IN DIFFERETN SEMANTIC DOMAINS

vehicles	sim(v1,v2)	IOMP	RFOMP	DSOMP
蜂蜜	1.000	1.000	1.000	1.000
林依晨	-INF	0.000	0.000	0.950

[4] *-INF represents infinite negative as the appearance of out-of-vocabulary word in the computation*

中药	-INF	0.000	-1.000	-0.971
香烟	0.044	0.000	0.000	0.016
蛋糕	-INF	0.125	0.333	0.658
黄连	-INF	-0.250	-1.000	-1.000
甜点	-INF	0.000	0.167	0.329
苦瓜	0.021	0.000	-1.000	-0.971

Now, we can arrange all the vehicles in two groups using the values in table 5, 6. The result based on DSOMP is in table 7, while most people agree with this arrangement. Based on this arrangement, we can evaluate other algorithms. The accuracy of *sim(v1,v2)* is 18.75%, the accuracy of IOMP is 37.50%, and the accuracy of RFOMP is 81.25%. Thereby, our algorithm DSOMP displays its efficiency in this experiment.

TABLE 7. CHINESE VEHICLES IN DIFFERETN SEMANTIC DOMAINS

Group 1	**The same semantic domain**	虎(tiger),狮(lion),猎豹(leopard),野猪(boar),熊猫(panda),山羊(goat),乌龟(tortoise),蜗牛(snail)
Group 2	**The different semantic domains**	蜂蜜(honey),林依晨(Ariel Lin),蛋糕(cake),甜点(dessert),香烟(cigarette),中药(traditional Chinese medicine),苦瓜(momordica),黄连(coptis)

Experiment 3：The computation of bilingual word similarities

Base on the knowledgebase, we have found some vehicles in different languages do have similar metaphorical properties, shown in table 1[11]. Now we want to prove this by our core algorithm DSOMP. Meanwhile, we will choose another 10 bilingual words to enrich our experiment with the same algorithm. The result has been show in table 8, 9.

TABLE 8. THE WORD SIMILARITIES OF METAPHORICAL PROPERTIES OF TOP 10 MOST SIMILAR VEHICLES

vehicles	水晶	花	妈妈	蚂蚁	蛋糕	糕点	糖	婴儿	海洋	针
crystal	.618	.211	.179	.026	.267	.267	.026	.024	.025	.029
flower	.283	.246	.129	.026	.221	.221	.191	.024	.027	.027
mother	.089	.030	.514	.033	.025	.025	.024	.024	.029	.029
ant	.027	.026	.039	.562	.025	.025	.024	.024	.034	.026
cake	.143	.350	.179	.025	.675	.675	.505	.024	.025	.029
cake	.143	.350	.179	.025	.675	.675	.505	.024	.025	.029
sugar	.119	.191	.342	.032	.356	.356	.517	.024	.025	.029
baby	.025	.024	.025	.024	.023	.023	.023	1.000	.025	.026
ocean	.030	.027	.027	.027	.026	.026	.024	.024	.514	.024
needle	.027	.026	.029	.025	.025	.025	.024	.023	.512	-INF

From table 8, we find that 80% of the values of word similarities of bilingual vehicles, such as *水晶* and *crystal*, shown in the grey cell, is over 0.500, which is quite larger than the other pairs. This will definitely demonstrate these vehicles share similar metaphoric properties. As far as why the values are not 1.000 while these vehicles express same concepts, the answer is the differences of numbers of metaphoric properties they map to and the translation differences of metaphoric properties.

Besides, some other vehicles share similar metaphoric properties across languages, such as *花* and *cake,* as the values of them are around 0.350. On the other hand, some vehicles share totally different metaphoric properties, such as *妈妈* and *baby*, which possess the value of 0.025. Obviously, these will help us judge whether words share similar or different metaphoric properties across semantic classes and languages automatically.

On the other hand, not surprisingly, some bilingual vehicles in table 9 share similar metaphorical properties across languages, such as *坟场* and *cemetery*, *花瓣* and *petal, etc*. More importantly, we also find some vehicles share similar metaphorical properties, while they do not share same semantic properties. For instances, the word similarity of *图书馆* and *graveyard* by metaphorical properties is 1.000, which hints that they have similar metaphoric properties of *peaceful*. And another two interesting pairs is *蜜蜂(or bee)* and *football_fans(or 足球迷)*, *蜜蜂(or bee)* and *market(or 市场)*. They show similar word similarities by metaphorical properties, which is not so clear if we just compute word similarities by semantic properties.

TABLE 9. THE WORD SIMILARITIES OF METAPHORICAL PROPERTIES OF TOP 10 MOST SIMILAR VEHICLE

vehicles	墓地	坟场	图书馆	婴儿	钻石	砖	花瓣	蜜蜂	足球迷	市场
graveyard	.956	1	1	1	.021	.021	.000	.024	-.971	.033
cemetery	.956	1	1	.514	.028	.028	.000	.029	-.967	.036
library	.956	1	1	1	.021	.022	.000	.024	-.971	.033
baby	.319	.351	.350	.512	.022	.024	.000	.026	-.324	.028
diamond	.000	.028	.026	.025	.757	.595	-.239	.029	.000	.029
brick	.000	.029	.029	.027	.514	.461	-.470	.059	.004	.029
petal	.000	.030	.029	.029	.035	.033	.648	.352	-.002	.028
bee	.000	.032	.029	.029	.030	.033	-.142	.517	.214	.353

football-fans	-.298	.000	.001	.016	.009	.012	-.010	.127	.196	.167
market	-.905	.004	.002	.032	.026	.029	.000	.352	.289	.489

5. Summary and Future Work

In this paper, we design several algorithms, especially DSOMP, to compute word similarities of different words by metaphoric properties based on the Chinese-English bilingual lexical cognitive property knowledgebase automatically, even across domains and languages. The result implies that the algorithm we design can hint word similarities by metaphoric properties to some extent and achieve better accuracy. This will be a plus to the computation of word similarities and lay a good foundation of metaphor computation.

Our next step will continue to enrich the knowledgebase to take more experiments. For instance, we try to add the frequency information of word pairs during the computation. Furthermore, we will try to use the word similarities by metaphoric properties in some application boundaries, such as machine translation, cross-language information retrieval and dictionary compilation, etc.

Acknowledgements

This paper is supported in part by National Social Science Fund of China under contract 10CYY021, 11CYY030, State Key Lab. for Novel Software Technology under contract KFKT2011B03, China PostDoc Fund under contract 2012M510178, Jiangsu PostDoc Fund under contract 1101065C, Priority Academic Program Development of Jiangsu Higher Education Institutions, National Natural Science Fund of China under contract 61003206, 61073119.

References

[1] Yael Karov. & Shimon Edelman. "Similarity-based Word Sense Disambiguation". Computational Linguistics, Volume 24, Number 1, 1998.

[2] Katharina Probst. & Ralf Brown. "Using Similarity Scoring To Improve the Bilingual Dictionary for Word Alignment". Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics(ACL), Philadelphia, July 2002, pp. 409-416.

[3] Dagan I., Lee L. and Pereira F., "Similarity-based models of word cooccurrence probabilities", Machine Learning, Special issue on Machine Learning and Natural Language, 34(1-3), 1999, pp. 43-69

[4] Agirre E. and Rigau G., "A proposal for word sense disambiguation using conceptual distance", Proc. of International Conference Recent Advances in Natural Language Processing (RANLP) , 1995, pp. 258-264, Tzigov Chark, Bulgaria.

[5] LIU Qun , LI Sujian. "Word Similarity Computing Based on How-net". Computational Linguistics and Chinese Language Processing. Vol. 7, No. 2, August 2002, pp. 59-76

[6] Yunqing Xia, Taotao Zhao, Jianmin Yao, and Peng Jin. "Measuring Chinese-English cross-lingual word similarity with HowNet and parallel corpus". In Proceedings of the 12th international conference on Computational linguistics and intelligent text processing–Volume Part II (CICLing'11), pp 221-233.

[7] Takaaki TANAKA. "Measuring the Similarity between Compound Nouns in Different Languages Using Non-Parallel Corpora". Proceedings of the 19th international conference, 2002.

[8] Veale,T. & Hao,Y.F. "Learning to Understand Figurative Language: From Similes to Metaphors to Irony". Proceedings of CogSci 2007, 2007.

[9] Jia Y. X. and Yu S. W. "Instance-based Metaphor Comprehension and Generation". Computer Science, vol. 36, no.3, pp.138-41, 2009.

[10] Kuang Haibo & Libin, etc. "The comparative analysis and the technology of mutual-gain on lexical metaphor properties of Chinese-English", CNCCL2011.

[11] Bin Li, Haibo Kuang, Xiaohe Chen, Xuri Tang, Chen Chen. "Construction of a Bilingual Cognitive Property Knowledgebase". CIS, Sanya, China, 2011.12.