Neurocomputing 219 (2017) 300–311

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Automatic detection and interpretation of nominal metaphor based on the theory of meaning

Chang Su[a,*], Shuman Huang[a], Yijiang Chen[b]

[a] Department of Cognitive Science, Xiamen University, Xiamen 361005, PR China
[b] Department of Computer Science, Xiamen University, Xiamen 361005, PR China

ARTICLE INFO

Communicated by Dr. T. Mu

Keywords:
Nominal metaphor recognition
Metaphor interpretation
Relatedness
Semantic

ABSTRACT

Automatic processing of metaphors can be explicitly divided into two subtasks: recognition and interpretation. This paper presents an approach to recognize nominal metaphorical references and to interpret metaphors by exploiting distributional semantics word embedding techniques and calculating semantic relatedness. In terms of detection, our idea is that nominal metaphors consist of source and target domains and that domains present in metaphors will be less related than domains present in non-metaphors. We represent the meaning of the concept as a vector in high-dimensional conceptual space derived from the corpus and compute the relatedness between the vectors to complete the task of detection. Relatedness here is based on the semantics of concepts. Thus, the model we present deals with metaphors where target and source have the same direct ancestors, such as "*A surgeon is a butcher*".

Then, using the relatedness between target and source domain, based on the properties of source domain and dynamic transfer of properties, we present an approach to interpret metaphors with dynamic transfer. Based on the view that metaphor interpretation is the cooperation of source and target domains, we divide metaphor interpretation into two subtasks: properties extraction and properties transfer. Creatively, we use annotations to express a non-binary evaluation, and we take the degree of the annotators' acceptability to evaluate our interpretation of metaphors.

1. Introduction

A metaphor is a kind of figurative language or trope. For instance, the uses of the noun "*butcher*" in the sentences "*He is a butcher*." and "*A surgeon is a butcher*." are different. In the first sentence, according to WordNet, the noun "*butcher*" means "*a person who slaughters or dresses meat for market*". In the second sentence, "*butcher*" exhibits its metaphorical use of "*someone who makes mistakes because of incompetence*". According to *Metaphor Theory* [1], a metaphor is defined as an analogy between two distinct domains – source and target domains. The target domain is what is actually being talked about; the source domain is the domain used as a basis for understanding the target (e.g., in the metaphor "*Time is money*", "*time*" is a concept in the target domain, and "*money*" is a concept in the source domain.). In the 1980s, Lakoff and Johnson proposed a Conceptual View and emphasized that, rather than being a rare form of creative language, metaphors are primarily a cognitive phenomenon, and metaphorical language serves as evidence for cognitive phenomena.

Metaphor research plays an important role in Natural Language Processing (NLP) [2]. Many sentences convey emotional tendency through underlying meaning [3]. Metaphor research has been applied to many NLP problems, such as machine translating, information retrieving, question answering [4], discourse understanding, and text summarizing. As a widespread phenomenon in natural language and a basic method of human thinking, the way we identify and interpret metaphors attracts the attention of not only linguists, but also cognitive scientists. The method for automatically processing metaphors is a simulation of the way humans identify, interpret and generate metaphors. It is believed that conceptual metaphors are not a barrier to, but a resource for cognition. Metaphors are integral to the human understanding of a myriad of abstract or complex concepts [1].

Following Krishnakumaran and Zhu [5], we divide metaphors into three types: Type I, II and III metaphors. In Type I metaphors(*nominal metaphors*), a noun is associated with another noun through the verb "be", such as in the case of "*Love is a journey*." In Type II metaphors(*verbal metaphors*), a verb acts on a noun such as in the instance, "*He kills a process*." For Type III metaphors(*adjective metaphors*), an adjective acts on a noun. Differing from other authors, who focus on Type II and III metaphors, in this paper, we focus only on Type I metaphors which are subject-verb pairs.

* Corresponding author.
E-mail address: suchang@xmu.edu.cn (C. Su).

http://dx.doi.org/10.1016/j.neucom.2016.09.030
Received 23 November 2015; Received in revised form 20 June 2016; Accepted 14 September 2016
Available online 20 September 2016
0925-2312/

Our method represents the concept by exploiting distributional semantics word embedding techniques and calculates semantic relatedness to determine whether or not the sentence is metaphorical. The core of metaphor interpretation is to extract similarities from target and source domains. Thus, properties of source domain play an important role in the dynamic metaphor process. Our method uses databases to extract a source's properties and then calculates the semantic relatedness between target and source concepts based on those properties. Finally, our method selects the property with the highest relatedness as interpretation output.

Compared with other works in metaphor detection, the main contributions of this paper are as follows:

1. We exploit distributional semantics word embedding techniques and semantic relatedness in the metaphor detection and interpretation fields.
2. Our method is based on the theory of meaning. We consider that the difference between source and target domains is in the semantic level, rather than that the domains belong to two different categories.
3. Our method can be flexibly applied to Chinese and English languages. In the Chinese language, we achieve detection accuracy of 85% and interpretation accuracy of 87%. In the English language, we achieve detection accuracy of 85.2% and interpretation accuracy of 85%.

2. Related work

2.1. Automatic metaphor detection

According to Wilks [6,7], metaphors represent an anomalous breaking of selectional preference in a given context. He believes that an occurrence of a metaphor necessarily leads to a semantic preference violation. Wilks' system divides metaphor understanding into two stages: recognition and interpretation.

One of the first attempts to identify and interpret metaphorical expressions automatically is the work of Fass [8]. The approach of Fass has its origins in the work of Wilks and uses a selectional preference violation technique to detect metaphors. For NLP, Fass introduces Collative Semantics, which extends many of the main ideas of preference semantics. Fass proposes a system ("met*") that discriminates among literalness, metonymy, metaphor and anomaly. However, this system relies on hand-coded declarative knowledge bases and leads to a number of limitations.

The CorMet system, developed by Mason, is the first attempt to discover source-target domain mappings automatically [9], which is accomplished by "finding, in a domain-specific selectional preference, systematic variations that are inferred from large, dynamically mined Internet corpora". Mason built the CorMet system with a statistical approach; the system is a corpus-based system for discovering metaphorical mapping between concepts. The CorMet system dynamically mines domain specific corpora to locate less frequent usages and identifies conceptual metaphors. Verbs selected for a concept in a source domain tend to be selected for their metaphorical equivalent in the target domain.

The method of Gedigian et al. [10] discriminates between literal and metaphorical use. For this purpose, they trained a Maximum Entropy (ME) classifier. They obtained their data by extracting the lexical items, whose frames are related to MOTION and CURE, from FrameNet [11], whereby highly conventionalized metaphors ("dead metaphors") are taken to be negative examples.

Kintsch [12,13] developed a computational system ($CI-LSA$ framework) of "X is Y" metaphoric references using semantic vector space. This system first uses of Latent Semantic Analysis (LSA) [14] and, by computing semantic distances through the bag-of-words representation, attempts to obtain the relevant or similar meaning to X and Y. Then, a Construction-Integration (CI) model [15] is added to select words that have a semantic distance close to the target domain, Y. As a result, words having a high semantic association with X are selected to represent the meaning of the metaphor "X is Y".

Krishnakumaran and Zhu [5] used hyponymy relation in WordNet [16], and selection preference violation based on knowledge learned from bigram frequencies on the web, to automatically classify sentences into metaphoric or normal usages. They dealt with verbs, nouns and adjectives as parts of speech.

Veale and Hao [17] proffered the argument that the same concepts and properties are described in either case. They automatically acquired a large simile case-base from the web and used the examples to both understand property-attribution metaphors and generate apt metaphors for a given target on demand.

Shutova et al. [18] presented a novel approach to identify metaphors by using verb and noun clustering. Starting from a small seed set of manually annotated metaphorical expressions, the system they presented can distinguish a large number of metaphors of similar syntactic structure from a corpus. This approach is different from former work in that it does not employ any hand-crafted knowledge. In contrast, this approach captures a metaphor by means of verb and noun clustering. The first to employ unsupervised methods for metaphor identification, their system operates with a precision of 0.79.

Based on the view that a metaphor usually involves mapping of a relatively concrete concept to a relatively abstract one, Turney et al. [19] proposed a new approach to identify metaphor expressions from literal usages. Thus, they presented the hypothesis that a metaphorical word is related to the abstractness of the context. Based on this hypothesis, they introduced a method to (1) differentiate metaphorical or literal expressions within a given context and (2) evaluate the algorithm with Type 3 (adjective-nouns) metaphors (as in dark thought) and with the TroFi Example Base for verbs. Therefore, to deal with the identification of metaphorical phrases, they used only one element: the abstractness of nouns in the phrase. Their algorithm has an average accuracy of 0.79 in adjective-nouns metaphors and an accuracy of 0.734 in verbal metaphors.

Neuman et al. [20] described three algorithms for three types of metaphor identification. According to them, the traditional selectional preference applied to metaphor identification has the main problem that using the common sense of phrases as an indication does not work very well in types such as subjective-nouns metaphors. They identified metaphors from two approaches: selectional preference and abstractness-based identification. They emphasize the method of measuring abstractness level, which is viewed as an indirect approximation of a noun's embodied nature. They combine measuring abstractness and selectional preference(Concrete Category Overlap) and first check a noun's selectional preference to obtain its literal sense. Their algorithms achieve an average 0.71 precision for the three types of metaphors. For type I metaphor identification, their approach is to compare the semantic categories of the nouns. They achieve 0.839 precision for Reuters corpus and 0.841 precision for NYT corpus. However, the category-based method is weak in dealing with metaphors of which the source and target concepts come from the same category. Our method avoids such mistakes by calculating the relatedness of nouns.

Tsvetkov et al. [21] used a method of combining abstractness degree and extracted CSF-common semantic features from cross-lingual metaphor distinguishing, especially in Russian and English. The classifier they presented is trained on English expressions first and then applied to another language; therefore, other than English, it does not require any hand-crafted lexical resource, such as TroFi, MRC and WordNet. They distinguish between metaphorical and literal usages by extracting syntactic relations (e.g. subject-verb-object (SVO) and adjective-nouns (AN)) as their features. Regarding SVO relations, they extract three types of features– semantic categories, abstractness degree, and types of named entities – and apply a logistic regression classifier to metaphor detection. If at least one of the relations relates to the target, then the sentence is identified as metaphorical and the sentence is tagged metaphorical. Further, Tsvetkov et al. [22] developed an English metaphor detection system, which added semantic supersenses as conceptual semantic features. They used the model transfer approach to identify metaphoric expressions in other languages, such as Spanish and Farsi.

Wilks [23] presented an algorithm to detect conventionalized metaphors implicit in the lexical data of a resource such as WordNet(WN) or VerbNet(VN). As popularly used, some lexical resources already contain conventionalized metaphors. Consequently, the original selectional preference will not work. Wilks proposed an advanced algorithm of selectional preference by using the senses in WN or VN. If an expression's subjects and objects satisfy selectional preference, the expression is deemed a conventionalized metaphor if the sense of the satisfied word is not the primary sense in WN or VN.

Li et al. [24] proposed a Gaussian Mixture model to detect figurative language (especially idioms) in context. They thought that figurative language exhibits fewer semantic cohesive ties with the context than literal language does. They used Normalized Google Distance to model semantic relatedness and adapt a Bayes decision rule to choose the category by maximizing the probability of fitting the data into different Gaussian models.

Bracewell and Mohler [25] proposed a tiered approach to recognize metaphors. In the first tier, they constructed lexico-syntatic patterns manually, which were automatically expanded and transformed to some dependency patterns with given rules in the second tier. In the last tier, they identified the target and source in the dependency pattern and, by using SVM with some re-defined features, determined whether or not the target-source matches were metaphorical. Applying the method of pattern matching to metaphor recognition, they achieved high precision; application to multi-languages is also an advantage. Their proposed approach achieved 0.738 precision in English. In [26,27], they recognized metaphors by constructing a conceptual space for the target and source. By using Wiki to cluster some words, which are highly related to the target domain, they constructed semantic features of the target domain. These semantic features were used to detect the target concepts of a given metaphor. Through some rules, they obtained a set of possible source concepts. By expanding and clustering the sense with Wordnet and Wiki, they constructed a conceptual space for the target and source of a given metaphor.

2.2. Automatic metaphor interpretation

Martin [28] described a metaphor comprehension system (MIDAS) and applied it to teaching software based on Unix. Martin used KODLAK, the extended sematic system of KL2ONE, as its knowledge interpretation language, which connects elements through an inheritance mechanism and concept hierarchy. When dealing with novel metaphors, MIDAS extends the existing ones to interpret them by the metaphor extend system (MES). Firstly, the metaphor algorithm searches metaphors related to the given novel ones and then selects the most related ones by calculating the concept distance of the two. The most related ones is the interpretation. MIDAS relies on inference and deals with novel metaphors without any corpus.

Veal and Hao [29] described a “fluid knowledge presentation for metaphor interpretation and generation” (Talking Points). Talking Points extracts conceptual properties from WordNet and the web. The properties extracted by Talking Points are then organized in Slipnet, which contains rules of insertions, deletions and substitutions and constructs the connection between concepts, thus completing the interpretation of the metaphors. However, Veale and Hao have not declared the useable range of Talking Points.

According to Shutova [30], the results of interpretation should be directly embedded in other systems. They thus define metaphor interpretation “as a paraphrasing task” and describe a system that

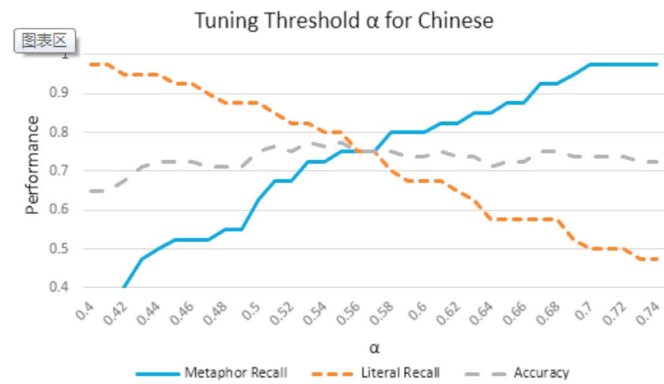


Fig. 1. The tuning process of α for Chinese.

automatically derives a literal interpretation in unrestricted text. The method is divided into two subtasks: generate literal paraphrasing and disambiguate from literal and metaphorical ones. Differing from the normal word disambiguate, paraphrasing must distinguish literal and metaphorical ones from the generated interpretations. The system paraphrases metaphorical expressions with a high accuracy of 0.81.

Shutova and Van de Cruys [31] presented a novel approach to metaphor interpretation with a vector space model, which focuses on verb metaphorical usages. Using a non-negative matrix factorization to compute a meaning list of target verbs, paraphrase candidates are extracted. After annotating the text with UKwac Corpus [32] and Stanford Part-of-Speech Tagger [33], the similar word (candidate paraphrase) is followed by adapting a probability distribution added to some dependency features. Because they assume that target verbs also restrict interpretation, they score the obtained paraphrases with the supplementary target verb itself. As a fully unsupervised approach, the system achieves a precision of 0.52.

Bollegala and Shutova [34] presented a fully unsupervised model of metaphor interpretation using paraphrases extracted from the web. According to them, given a metaphorical verb and its arguments, metaphor interpretation is extracting a paraphrase and replacing it in a literal way. They confirm that the main difference between metaphor interpretation and common paraphrase extraction is how to find paraphrases with literal usage, especially in a given context with given arguments.

Su et al. [35] described an algorithm to interpret the Chinese nominal and verbal metaphors based on latent semantic similarity. Their method relies on the properties of source and target domains from databases. If the target property connects to the source property using the synonyms extended in WordNet, they discover the latent semantic similarity between them, and this property of source forms an interpretation of metaphor. And their system achieves 0.85 accuracy for nominal metaphors and 0.86 accuracy for verbal metaphors.

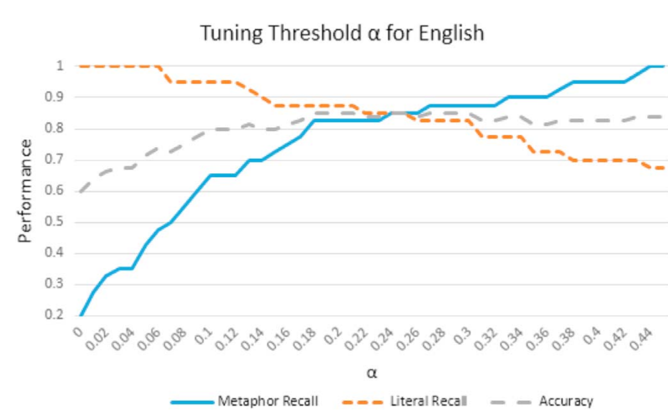


Fig. 2. The tuning process of α for English.

Table 1
The result of identification of Chinese metaphorical references.

Word1	Word2	Relatedness	Machine Judgement	Human's Judgement
儿童 Child	花朵 Flower	0.289613	M	M
流星 Meteor	灯笼 Lantern	0.502267	M	M
石油 Petroleum	能源 Energy	0.69452	L	L
流言 Rumor	瘟疫 Plague	0.698876	L	M
首都 Capital	心脏 Heart	0.155356	M	M
护士 Nurse	天使 Angel	0.407289	M	M
法律 Law	绳子 Line	0.334285	M	M
背 Back	弓 Bow	0.643889	L	M

[1] The degree of *Relatedness* is from 0 (not related) to 1 (related).

[2] M - Metaphorical Usage；L - Literal Usage

Table 2
The result of identification of English metaphorical references.

Word[a]	Word[b]	Relatedness	Machine judgement	Human's judgement
Marriage	Game	0.080653	M	M
Surgeon	Butcher	0.019284	M	M
Life	Journey	0.425209	L	M
Framework	Body	0.265605	L	M
Tourist	Prey	0.041291	M	M
Husband	Brute	0.050127	M	M
Breakfast	Buffet	0.61188	L	L

[a] The degree of *Relatedness* is from 0 (not related) to 1 (related).
[b] M – Metaphorical usage; L – Literal usage.

Table 3
Performance of computation of Chinese referential pairs.

	Accuracy	Precision	Recall	F
Metaphorical Reference	0.85	0.803	0.928	0.861
Normal Reference		0.915	0.772	0.837
Average	0.85	0.859	0.85	0.849

Table 4
Performance of computation of English referential pairs.

	Accuracy	Precision	Recall	F
Metaphorical Reference	0.852	0.839	0.872	0.855
Normal Reference		0.867	0.832	0.849
Average	0.852	0.853	0.852	0.852

Table 5
The comparison with the baselines.

	Hypernym	Relatedness	Hypernym+Relatedness
Chinese			
Accuracy	0.72	0.768	0.85
Precision_M	0.641	0.703	0.803
English			
Accuracy	0.778	0.764	0.852
Precision_M	0.693	0.717	0.839

Table 6
The extracted properties of some nouns and the average acceptability.

Nouns	Properties Set
画 Painting	美丽的 beautiful(5), 生动的 vivid(4.5), 艺术的 artistic(5), 不真实的 unreal(3), 静止的 motionless(3.5), 多彩的 colorful(4.7), 华丽的 gorgeous(3.5)
星星 Star	遥远的 distant(4), 闪烁的 sparkling(5), 明亮的 bright(5), 伟大的 great(1.2), 难以接近的 unapproachable(3.2)
潮水 Tide	激烈的 furious(4.7), 强劲的 strong(4.2),下降的 falling(3.5), 不间断的 relentless(3.7), 迅速的 rapid(4.2), 上涨的 rising(4), 无法停止的 unstoppable(3.2), 惊人的 incredible(3.7)
镜子 Mirror	脆弱的 insubstantial(4), 反射的 reflective(4), 光滑的 smooth(4.7), 有光泽的 shiny(3.2), 清澈的 clear(3.2), 平静的 clam(3.2), 透明的 limpid(3)
噩梦 Nightmare	可怕的 terrible(5), 恐怖 horrible(4.7),不真实的 unreal(3.5), 黑暗的 black(4), 悲伤的 sad(3.5),难以置信的 incredible(3.2)

The average acceptability in brackets is from 1 (low acceptability) to 5 (high acceptability).

However, our method does not require target domain's properties. Our method can deal with a metaphor where the database lacks the target domain's property information.

3. Theoretical foundation

Ogden and Richards [36] gave 22 definitions of “meaning” in *The Meaning of Meaning*. They pointed out that “meaning” is “*an intrinsic property*” and “*a unique unanalysable Relation to other things*”. “An Intrinsic property” points out that the meaning of a word is its basic definitions; “A unique unanalysable Relation to other things” points out that the relatedness is also a kind of word meaning. Actually, according to our assumption, when it is sufficient, relatedness also reflects basic definitions.

Kövocses [37] divided metaphors into three levels; the one of interest to us in this paper is the sub-individual level. In the sub-individual level, Kövocses pointed out that metaphors are perceptual, cultural and category-based correlations between elements that involve human physiology.

This paper proposes the idea that nominal metaphors consist of source and target domains, and that domains present in metaphors are less related than domains present in non-metaphors. In this paper, we measure the relatedness of the two concepts by using distributional semantic similarity measures. The idea is based on the Distributional Hypothesis [38] that linguistic items with similar distributions have similar meanings. A concept has a certain relatedness to another if their distributional information is similar.

4. Automatic metaphor detection

4.1. The method

4.1.1. Strategy

Because we focus on nominal metaphors, our identification task is viewed as classifying according to whether the subject-object relationship is used metaphorically or literally. The target for nominal metaphor identification is to detect an anomaly from common sense in the sentences. In this paper, we exploit corpus and quantify the extent to which concepts share information.

Most nominal metaphor computational models depend on hand-coded knowledge bases and work on a few examples. Our approach is inspired by the desire to strengthen concept representation with context of concepts. To capture more context, we use word embedding to obtain vector representations of the concepts. The word embedding approach has been applied to many tasks [39,40]; we apply it to metaphor tasks in this paper. Important semantic information of concepts is implied in word representations, such as the relations and properties of the concepts. After we obtain the relatedness between two concepts, we judge whether or not the reference is metaphorical.

4.1.2. Relatedness algorithm based on corpus

Given a sentence, after dependency parse, we first transform the reference into a concepts-pair, $\langle w_u, w_v \rangle$, where w_u is the subject and w_v is the object.

Table 7
The interpretation of some nominal metaphors with our method.

Target Domain	Source Domain	Interpretation
风景 Landscape	画 Painting	风景是美丽的 The landscape is beautiful.
眼睛 Eye	星星 Star	眼睛是明亮的 The eye is bright.
爱 Love	潮水 Tide	爱是无法停止的 The love is unstoppable.
湖水 Lake	镜子 Mirror	湖水是平静的 The Lake is calm.
灾难 Disaster	噩梦 Nightmare	灾难是可怕的 The disaster is terrible.

Table 8
Comparison with the three baselines.

	Random	R_{source}	$R_{target+source}$	Our method
Accuracy of Chinese	0.42	0.71	0.82	0.87
Accuracy of English	0.28	0.37	0.69	0.85

Table 9
Compare our method with the method of Su et al. [35].

	Method of Su et al. [35]	Our method
Accuracy	0.85	**0.86**

In order to measure semantic relatedness between words, a word, w, is represented by a vector, $\vec{w}$, as follows:

$$\vec{w} = \langle c_1, c_2, \ldots, c_q \rangle \tag{1}$$

where, q is the dimension of vector; $c_i (1 \le i \le q)$ is the value of dimension i.

Mikolov et al. [41] proposed the following two model architectures for computing continuous vector representations of words from large data sets: Continuous Bag-of-Words (CBOW) and distributed Skip-gram models. They measured the quality of these representations in a word similarity task and compared their methods with different types of neural networks. The results revealed their methods had a large improvement in accuracy at a much lower computational cost. In this paper, we apply the CBOW model[1] to obtain vector representations of words.

To compute the semantic relatedness of the concept-pairs, we compare the concept vectors using the cosine metric [42].

$$Relatedness(w, w') = \frac{\sum_{j=1}^{q} c_j c'_j}{\sqrt{\sum_{j=1}^{q} c_j'^2}\sqrt{\sum_{j=1}^{q} c_j'^2}} \tag{2}$$

where, w is represented as $\vec{w} = \langle c_1, c_2, \ldots, c_q \rangle$, w' is represented as $\vec{w'} = \langle c_{\prime 1}, c_{\prime 2}, \ldots, c_{\prime q} \rangle$.

Cosine similarity is an improvement of the vector inner product. In practice, as a method of calculation, cosine similarity is the most widely used .

After comparing the relatedness of the concept-pair, our system queries whether w_u is a hyponym or hypernym of w_v in WordNet. If w_u is a hyponym or hypernym of w_v, the sentence is determined to be of literal usage.

4.1.3. Judgment for nominal metaphor

In this paper, a reference is determined as a nominal metaphor, only if the subject and object share little information between them. When $Relatedness < \alpha$ and concept-pair is not a hyponym/hypernym relation, the reference is determined to be metaphorical; otherwise, it is literal. Here α is the threshold of *relatedness*.

4.2. Experimental data

4.2.1. Corpus

In this paper, we choose *Reader Corpus*[2] as the Chinese corpus, and use *Segtag*[3] to support the CBOW model in computing the vector representations of Chinese words. We choose *BNC Corpus*[4] as the English corpus.

4.2.2. Threshold estimation

Taking a set of tuning data containing forty metaphorical and forty literal sentences in Chinese, we tune the threshold, α. In no instances do the tuning and test sets overlap. The *Accuracy*, *Metaphor's Recall and Literal's Recall* are shown in Fig. 1. To enable the system to achieve its best performance, we change the values of α. The

[1] https://code.google.com/p/word2vec/.

[2] A Chinese corpus. www.duzhe.com.
[3] A word segmentation tool of NLP Lab of Xiamen University.
[4] The British National Corpus (BNC) is a 100 million word collection of samples of written and spoken language. http://www.natcorp.ox.ac.uk/.

performance of the system depends upon the values of threshold, α. According to the results, when the threshold, α, is 0.575, the divergence between the Chinese metaphor and literal recalls is minimum and the accuracy nearly maximized..

Taking a set of tuning data containing forty metaphorical and forty literal sentences in English, we tune the threshold for English, α, as shown in Fig. 2. According to the results, we choose 0.235 as the threshold, α, for English..

4.2.3. Test data

To test the performance of our system, we conducted our experiments on the data of nominal metaphorical sentences, most of which are novel or active metaphors accepted by most people. Our Chinese test data is annotated from the web, newspaper, blog, and books. The English test data is from the *Collins Cobuild Corpus* and *Master Metaphor List*.[5] The Chinese test data contains 250 metaphorical usages and 250 literal usages; the English test data contains 250 metaphorical usages and 250 literal usages, within the following genres: news/journal articles, politics, finance, essays, fiction, and speech. Appendices A and B show a subset of our test data.

We tested the reliability of this annotated data using multiple annotators on the corpus.

Annotators: Five independent annotators, all native speakers of Chinese with some linguistic knowledge, were given the same text from our test data and asked to determine whether the sentences were metaphorical or literal.

Inter-annotator agreement: We evaluated the reliability of our annotation scheme by assessing the inter-annotator agreement in terms of κ [43]. The classification was performed with an agreement of 0.66(κ), which is considered reliable.

4.3. Experiments and results

4.3.1. Experiments

Based on the above method, we propose the following steps in metaphor recognition processing:

Step 1: Pre-process Test Data: Based on the corpus, the corresponding vector of each concept is acquired by the CBOW model.

Step 2: Calculate the Relatedness of Concept-Pairs: The relatedness between concept-pairs is attained by the method mentioned in Section 4.1.2.

Step 3: Identify the Metaphorical References: Only if $Relatedness < \alpha$, and a concept-pair is not a hyponym/hypernym relation, then the reference is judged to be a metaphor.

Tables 1 and 2 list partial results obtained by applying our methodology to classify Chinese and English test data into a metaphorical or normal reference. The values in Tables 1 and 2 exhibit a special phenomenon: if the metaphors are used by people just like common phrases, the values of *Relatedness* are greater. This phenomenon suggests that the application of metaphors to linguistics changes the cognition of human beings, enforces the source and target that share the same properties, and transforms the source into the target's synonym. This phenomenon also demonstrates the life of a metaphor (i.e. a metaphor from birth to death [44]).

4.3.2. Accuracy and error analysis

The performance of the presented method is evaluated based on precision, recall and F-measure. Precision is the number of correct recognized references divided by the total number of recognized references. Recall is the number of correct recognized references divided by the total number of actual references in the manual evaluation data. The F-measure is calculated by using precision and recall

$$F-measure = \frac{2 \times precision \times recall}{precision + recall} \tag{3}$$

Tables 3, 4 demonstrate the efficiency of our method for recognizing metaphors.

The model incorrectly detects some metaphors because of the following reasons:

1. Given a *dead metaphor* [1,45,46], some of its usages are already literal. This phenomenon also leads to incorrectness.
2. As a limitation of corpus, if a concept does not appear in the corpus, the system cannot evaluate the relatedness of the concept-pair.
3. If a literal sentence has a low relatedness in corpus, the sentence is determined as metaphorical, e.g., *My father is a policeman*.

4.3.3. Baseline and comparison

In this paper, we compare our approach against two baselines: (1) using the WordNet hypernym information; (2) using the Relatedness score with a tuned threshold, α. We compare the quality of our system in the following aspects: *Accuracy*, *Metaphor's Precision* (See Table 5).

In terms of *Accuracy*, our method combining WordNet hypernym with relatedness is better than two baselines. Because the source concept is usually not a hypernym/hyponym to the target concept in a metaphor, the method combining relatedness with WordNet hypernym improves the precision of the literal identification.

5. Automatic metaphor interpretation

5.1. The core of metaphor interpretation

Different from metaphor detection, metaphor interpretation is a kind of translation, which translates an abstract expression into a more concrete paraphrase. A metaphor's source and target domains come from two different domains and, in some way, contain similarities. Thus, how to extract the similarities is the key point of metaphor interpretation. We propose that metaphor interpretation is the cooperation of the source and target domains in the following three ways: (1) The source and target share common properties; (2) The properties of the source and those of the target have some similarities; (3) The target is matched to one of the source domain's properties. The properties of source and target play an important role in metaphor interpretation.

In this paper, we propose a property transfer process to describe this cooperation theory. In this process, the target domain is active, and the source domain is passive. The source domain provides some properties, and the target domain decides how to use these properties to describe itself. We divide the property transfer process into the following two steps: Firstly, we extract the properties, C, of the source concept, S; secondly, we choose the source's properties that cooperate with the target domain, T. Thus, we take (T, C) as the interpretation of the metaphor.

Compared with previous work, we divide the metaphor interpretation into two subtasks: properties extraction and properties transfer. The first subtask is to extract the significant properties of the source domain; the second subtask is to decide if the target domain and source domain will, and how they will, cooperate, which means interpretation of the metaphor.

5.2. Extraction of properties of source

The properties of the source concept in our system are extracted from the *Property Database*[6] and *Sardonicus*.[7] The *Property*

[5] http://www.lang.osaka-u.ac.jp/~sugimoto/MasterMetaphorList/metaphors/index.html.

[6] A database developed by NLP Lab of Xiamen University.

[7] http://afflatus.ucd.ie/sardonicus/tree.jsp.

Table A1

Concept$_A$	Concept$_B$	Relatedness	Machine's Judgement	Human's Judgement
希望 Hope	种子 Seed	0.47107	M	M
文明 Culture	风景 Scenery	0.412248	M	M
未来 Future	太阳 Sun	0.366289	M	M
晚年 Old Age	黄昏 Dusk	0.442744	M	M
头发 Hair	雪 Snow	0.374517	M	M
市场 Market	奶酪 Cheese	0.304053	M	M
世界 World	花园 Garden	0.253757	M	M
时间 Time	金钱 Money	0.420658	M	M
时光 Time	医生 Doctor	0.222693	M	M
诗歌 Poetry	武器 Weapon	0.296259	M	M
关系 Relationship	网络 Net	0.422295	M	M
生命 Life	花 Flower	0.317591	M	M
建筑 Building	音乐 Music	0.413861	M	M
婚姻 Marriage	战争 War	0.528083	M	M
历史 History	小说 Fiction	0.491628	M	M
国家 Country	机器 Machine	0.290591	M	M
婚姻 Marriage	坟墓 Tomb	0.504028	M	M
文字 Word	花 Flower	0.294981	M	M
儿童 Child	花朵 Flower	0.289613	M	M
护士 Nurse	天使 Angel	0.407289	M	M
法律 Law	绳子 Line	0.334285	M	M
流言 Rumor	瘟疫 Plague	0.698876	L	M
家庭 Family	围城 City	0.379533	M	M
钢笔 Pen	武器 Weapon	0.524251	M	M
记忆 Memory	烙印 Brand	0.627262	L	M
夫妻 Couple	鸳鸯 Mandarin duck	0.393798	M	M
生命 Life	小草 Grass	0.325493	M	M
民族 Nation	花 Flower	0.170053	M	M
经济 Economy	泡沫 Bubble	0.155673	M	M
理想 Dream	灯塔 Lighthouse	0.443574	M	M
考试 Exam	尺子 Ruler	0.419441	M	M
情感 Emotion	酒 Wine	0.243577	M	M

Database was constructed by extracting properties of entity concepts from *Reader Corpus* based on phrase patterns, such as "*Adjective +Noun*". *Sardonicus* is an adjective taxonomy that knows the exemplary properties of different objects in the real world. *Sardonicus* has acquired this knowledge by sifting the contents of the web in search of meaningful comparisons.

To test the reliability of the perceptual features in *Property Database* and *Sardonicus*, we invited fifteen independent volunteers, who had sparse or no linguistic knowledge. All received the same part of our extracted perceptual features from *Property Database* and *Sardonicus* and were asked to rate the acceptability of the features using a scale from one to five (One being very low acceptability; Five being very high acceptability). Table 6 shows the results.

5.3. The property transfer: properties of source domain transferred to target domain

After extracting the source's properties, we transfer those properties to the target concept. Thus, we select an property, c_i of *C*, to match the target domain, *T*.

Assuming source domain *Source* has the property, c_i of *C*, then the semantic relatedness between target domain, *T*, and source domain, *S*, based on the property c_i, $Rel(T, S, c_i)$, is computed as follows:

$$Rel(T, S, c_i) = Relatedness(T, c_i) \quad (4)$$

Considering the data sparse problem in corpus, we extend the synonyms of the property, c_i, using the *Tongyi Cilin* (*Extended*)[8] or WordNet. The synonyms set of c_i is represented as $Syn = \{v_1, v_2, \ldots, v_{|Syn|}\}$, where $|Syn|$ is the number of synonyms. Then, the semantic relatedness, $Rel(T, S, c_i)$, is computed as follows:

$$Rel(T, S, c_i) = \frac{1}{|Syn|}\sum_{i=1}^{|Syn|}(Relatedness(T, v_i)) \tag{5}$$

We obtain the semantic relatedness between target and source domains based on all the properties of the source domain. Then we choose the property, **c**, which makes the relatedness between target and source domains maximum as the comprehension result, as shown in Eq. (6).

$$\begin{aligned}\mathbf{c} &= \underset{c_i}{\text{argmax}}\{Rel(T, S, c_i)\}\\ &= \underset{c_i}{\text{argmax}}\{Relatedness(T, c_i)\}\\ &= \underset{c_i}{\text{argmax}}\left\{\frac{1}{|Syn|}\sum_{i=1}^{|Syn|}Relatedness(T, v_i)\right\}\end{aligned} \tag{6}$$

The most related property, $\mathbf{c} = c_i$, of source domain forms the interpretation, (T, c_i), which is expressed briefly as "T is c_i.

5.4. The output of interpretation

Different from metaphor identification, the result of the metaphor interpretation algorithm chooses the most related source's properties. We formalize the output "T is c_i". Thus, (*rumor*, *fast*) is formalized to "*The rumor is fast*". Table 7 shows the interpretation of some nominal metaphors with our method.

5.5. Evaluation and result analysis

5.5.1. Test data

We focus on Type I metaphors. Our test data is annotated from the web, newspapers, blogs, and books. The test data contains one hundred Chinese metaphorical usages and one hundred English metaphorical usages (nominal metaphors).

5.5.2. Accuracy

To determine accuracy, we invited five human volunteers to evaluate the interpretation results. The volunteers rated the acceptability of the interpretation results (Results and ratings are shown in Appendices C and D; based on these ratings, we calculated accuracy. The rating scale, used by the volunteers, ranges from one to five as follows: One: Highly unacceptable; Two: Unacceptable; Three: Neutral; Four: Acceptable; and Five: Highly Acceptable.

We evaluated the reliability of our annotation scheme for metaphor interpretation task by assessing the inter-annotator agreement in terms of κ [43]. The results was performed with an agreement of 0.39(κ), which is considered reliable.

We divide the acceptability into five levels instead of a simple binary decision (accept/decline), because the five-level method makes the evaluation finer-grained compared with a binary decision. We then evaluate the system's performance against their judgments in terms of accuracy. We consider results with an average acceptability below three as incorrect results and those over three as correct results. The final system achieved 87% accuracy for Chinese and 85% accuracy for English.

5.5.3. Baseline and comparison

To test the quality of our approach, we compare it against the following three baselines: (1) Random: Using a random property. We randomly select one feature of the source and output the interpretation with the "*target is feature*"; (2) R_{source}: Using the most related property with the source; (3) $R_{target+source}$: Using the most related property with the target and the source. In our method we use the most related property with the target. After human annotators evaluate the results, we compare the accuracies with the algorithm proposed in this paper (See Table 8).

5.5.4. Comparison of our work with pervious work

In this section, we compare our method with previously related method of Su et al. [35], which seeks the latent semantic similarity between a metaphor's target and source domains. If the target's property connects to the source's property using the synonyms extension in WordNet, they discover the latent semantic similarity between them, and this property of source forms an interpretation of metaphor. We test our method on their test data which includes 85 instances of nominal metaphors. Most of their test data are simple metaphors without contextual information, such as "A is B". In Table 9.

The results show the accuracy of two methods is close. And our method has a strong advantage over their method: their method relied on both target and source domains' properties which limited its applicability. It is a common phenomenon in the metaphor that people are not familiar with the target domain. The database may not provide available property information of target domain. Their method can not get a result in the absence of the target domain's properties. Our method only needs source's properties which are always familiar, and our method utilizes the relatedness between target and source's properties to guide the interpretation.

5.5.5. Analysis result

Some of the interpretation is not reliable; the reasons are listed as follows:

1. Some of the nominal metaphors are not easily interpreted by language.
2. The extraction of a source's properties is not accurate, which leads to

[8] A Chinese Thesaurus, http://ir.hit.edu.cn/.

Table B1

$Concept_A$	$Concept_B$	Relatedness	Machine's Judgement	Human's Judgement
Wheel	Business	0.111275	M	M
Building	Idea	0.052411	M	M
Credit	Tool	0.195582	M	M
Test	Tool	0.221593	M	M
Problem	Container	0.077293	M	M
Anger	Heat	0.188719	M	M
Music	Building	0.030961	M	M
Belief	Partner	0.097632	M	M
Competition	War	0.188785	M	M
Tower	Strength	0.164615	M	M
Belief	Disease	0.046816	M	M
Feeling	Touch	0.176307	M	M
Surgeon	Butcher	0.019284	M	M
Emotion	Force	0.11791	M	M
Fear	Cold	0.171942	M	M
Harm	Journey	0.050816	M	M
Hope	Light	0.027891	M	M
Idea	Food	0.108783	M	M
Mind	Eye	0.476182	L	M
Mind	Travel	0.152437	M	M
Customer	God	0.0626003	M	M
Love	Journey	0.184947	M	M
Cleanliness	Morality	0.156285	M	M
Emotion	Battery	0.028818	M	M
People	Plant	0.1351	M	M
Marriage	Game	0.080653	M	M
Husband	Brute	0.0501265	M	M
Chain	Relationship	0.043864	M	M
Wall	Silence	0.070401	M	M

some mistakes. In future work, we will concentrate on improving this work.
3. Some properties cannot cooperate with the target concept directly. Future work will focus on how to map properties to a target domain.

6. Conclusions

Our system, the first of its kind, is capable of identifying metaphorical expressions with high precision (85% in Chinese, 85.2% in English). In contrast to other work, our system exploits word embedding to obtain concepts' vectors and evaluate their semantic relatedness. Our system presents a new approach for identifying metaphorical expressions. We evaluated this measure relative to human judgment and found it to be reasonably correlated. Then, using the relatedness between target and source domain, based on the properties of source domain and dynamic transfer of properties, we presented an approach to interpret metaphors with dynamic transfer. Our results show that the relatedness of semantic information is efficient to metaphor recognition and interpretation.

Acknowledgments

We would like to thank all the referees for their constructive and helpful suggestions. Funding was provided by the National Natural Science Foundation of China under Grant no. 61075058.

Appendix A. The result of Chinese metaphor recognition (a subset)

See Table A1.

Table C1

Metaphors	Interpretations	Acceptability					
		1.0	2.0	3.0	4.0	5.0	Average
Eye is star. 眼睛是星星	Eye is bright. 眼睛是明亮的	5.0	5.0	5.0	5.0	4.0	4.8
Full moon is dish. 满月是盘	Full moon is round. 满月是圆的	5.0	5.0	5.0	5.0	4.0	4.8
Lawyer is fox. 律师是狐狸	Lawyer is crafty. 律师是狡猾的	5.0	5.0	5.0	5.0	4.0	4.8
the flower of life 生命之花	Live is brilliant. 生命是灿烂的	5.0	3.0	5.0	3.0	5.0	4.2
Swan is cloud. 天鹅是白云	Swan is white. 天鹅是洁白的	5.0	5.0	5.0	5.0	4.0	4.8
Time is gold. 光阴是金	Time is precious. 光阴是宝贵的	5.0	5.0	5.0	5.0	4.0	4.8
Disaster is nightmare. 灾难是噩梦	Disaster is terrible. 灾难是可怕的	5.0	5.0	5.0	5.0	4.0	4.8
Love is sugar. 爱情是糖	Love is sweet. 爱情是甜蜜的	5.0	5.0	5.0	5.0	4.0	4.8
Marriage is war. 婚姻是战争	Marriage is serious. 婚姻是严重的	3.0	5.0	3.0	3.0	5.0	4.0
New moon is sickle. 新月是镰刀	New moon is bent. 新月是弯的	5.0	5.0	5.0	5.0	4.0	4.8
New York is the heart of USA. 纽约是美国的心脏	New York is important. 纽约是重要的	5.0	5.0	5.0	5.0	4.0	4.8
Swan is ship. 天鹅是帆船	Swan is floating. 天鹅是漂浮的	5.0	4.0	5.0	5.0	4.0	4.6
pass is foreign country. 过去是异国他乡	pass is distant. 过去是遥远的	5.0	5.0	5.0	5.0	3.0	4.6
Road is the snake. 公路是蛇	Road is winding. 公路是蜿蜒的	5.0	5.0	5.0	5.0	5.0	5.0
Market is the prairie. 市场是草原	The market is vast. 市场是广阔的	5.0	5.0	5.0	5.0	5.0	5.0

Table D1

Metaphors	Interpretations	Acceptability					
		1	2	3	4	5	Average
Atom is solar system.	Atom is complex.	5.0	5.0	3.0	5.0	3.0	4.2
Landscape is poetry.	Landscape is beautiful.	5.0	5.0	5.0	4.0	5.0	4.8
Essay is flower.	Essay is graceful.	5.0	5.0	4.0	4.0	5.0	4.6
Courtyard is box.	Courtyard is square.	5.0	5.0	5.0	5.0	5.0	5.0
Friendship is cloud.	Friendship is mutable.	5.0	4.0	3.0	4.0	3.0	3.8
Friendship is flame.	Friendship is passionate.	5.0	5.0	5.0	4.0	5.0	4.8
Inspiration is spark.	Inspiration is instantaneous.	5.0	5.0	5.0	5.0	5.0	5.0
Rumor is plague.	Rumor is epidemic.	3.0	5.0	5.0	4.0	5.0	4.4
Economy is bubble.	Economy is frail.	5.0	5.0	4.0	4.0	5.0	4.6
Young is mountain.	Young is mighty.	5.0	5.0	3.0	4.0	4.0	4.2
Elder is children.	Elder is powerless.	5.0	5.0	4.0	4.0	3.0	4.2
Hope is bubble.	Hope is frail.	5.0	5.0	5.0	3.0	4.0	4.4
Brain is machine.	Brain is intellectual.	5.0	5.0	5.0	5.0	3.0	4.6
Working is machine.	Working is intricate.	5.0	5.0	5.0	5.0	4.0	4.8
Star is diamond.	Star is brilliant.	5.0	4.0	5.0	5.0	4.0	4.6
Life is travel.	Life is attractive.	5.0	4.0	3.0	4.0	4.0	4.0
Dew is pearl.	Dew is luminescent.	5.0	5.0	4.0	4.0	4.0	4.4
Moon is diamond.	Moon is shiny.	5.0	5.0	5.0	5.0	5.0	5.0
Branch is arm.	Branch is thick.	5.0	4.0	5.0	4.0	5.0	4.6
Flower is torch.	Flower is red.	5.0	5.0	5.0	5.0	5.0	5.0
Motherhood is flame.	Motherhood is passionate.	5.0	5.0	4.0	5.0	5.0	4.8
Feather is porcelain.	Feather is white.	5.0	5.0	5.0	5.0	5.0	5.0
Swan is fairy.	Swan is beautiful.	5.0	5.0	5.0	5.0	5.0	5.0
Snow is blanket.	Snow is soft.	5.0	5.0	5.0	5.0	5.0	5.0
Maple is banner.	Maple is scarlet.	5.0	5.0	1.0	5.0	3.0	3.8
History is river.	History is abundant.	4.0	5.0	5.0	5.0	4.0	4.6
Childhood is song.	Childhood is joyful.	5.0	5.0	5.0	5.0	5.0	5.0
City is grave.	City is lonely.	5.0	5.0	5.0	2.0	4.0	4.2
Life is dew.	Life is lovely.	5.0	4.0	5.0	1.0	3.0	3.6
Girl is rose.	Girl is beautiful.	5.0	5.0	5.0	5.0	5.0	5.0

Appendix B. The result of english metaphor recognition (a subset)

See Table B1.

Appendix C. The acceptability of the Chinese interpretation results (a subset)

The degree of acceptability is from 5 (easy to be accepted) to 1 (hard to be accepted) (Table C1).

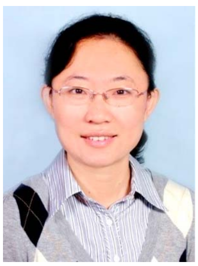
Appendix D. The acceptability of the English interpretation results (a subset)

The degree of acceptability is from 5 (easy to be accepted) to 1 (hard to be accepted) (Table D1).

References

[1] M. Johnson, G. Lakoff, Metaphors We Live by, University of Chicago.
[2] C.-L. Zhou, Y. Yang, X.-X. Huang, Computational mechanisms for metaphor in languages: a survey, J. Comput. Sci. Technol. 22 (2) (2007) 308–319.
[3] E. Cambria, P. Gastaldo, F. Bisio, R. Zunino, An elm-based model for affective analogical reasoning, Neurocomputing 149 (2015) 443–455.
[4] T. Sagara, M. Hagiwara, Natural language neural network and its application to question-answering system, Neurocomputing 142 (2014) 201–208.
[5] S. Krishnakumaran, X. Zhu, Hunting elusive metaphors using lexical resources, in: Proceedings of the Workshop on Computational Approaches to Figurative Language, Association for Computational Linguistics, 2007, pp. 13–20.
[6] Y. Wilks, A preferential, pattern-seeking, semantics for natural language inference, Artif. Intell. 6 (1) (1975) 53–74.
[7] S. Wilks, Making preferences more active, Artif. Intell. 11 (3) (1978) 197–223.
[8] D. Fass, met*: a method for discriminating metonymy and metaphor by computer, Comput. Linguist. 17 (1) (1991) 49–90.
[9] Z.J. Mason, Cormet: a computational, corpus-based conventional metaphor extraction system, Comput. Linguist. 30 (1) (2004) 23–44.
[10] M. Gedigian, J. Bryant, S. Narayanan, B. Ciric, Catching metaphors, in: Proceedings of the Third Workshop on Scalable Natural Language Understanding, Association for Computational Linguistics, 2006, pp. 41–48.
[11] C.F. Baker, C.J. Fillmore, J.B. Lowe, The berkeley framenet project, in: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1, Association for Computational Linguistics, 1998, pp. 86–90.
[12] W. Kintsch, Metaphor comprehension: a computational theory, Psychon. Bull. Rev. 7 (2) (2000) 257–266.
[13] W. Kintsch, A.R. Bowles, Metaphor comprehension: what makes a metaphor difficult to understand, Metaphor Symb. 17 (4) (2002) 249–262.
[14] P.W. Foltz, W. Kintsch, T.K. Landauer, The measurement of textual coherence with latent semantic analysis, Discourse Process. 25 (2–3) (1998) 285–307.
[15] W. Kintsch, Comprehension: A Paradigm for Cognition, Cambridge University Press, 1998.
[16] C. Fellbaum, Wordnet: An Electronic Database, 1998.
[17] T. Veale, Y. Hao, Comprehending and generating apt metaphors: a web-driven, case-based approach to figurative language, in: AAAI, vol. 2007, 2007, pp. 1471–1476.
[18] E. Shutova, L. Sun, A. Korhonen, Metaphor identification using verb and noun clustering, in: Proceedings of the 23rd International Conference on Computational Linguistics, Association for Computational Linguistics, 2010, pp. 1002–1010.
[19] P.D. Turney, Y. Neuman, D. Assaf, Y. Cohen, Literal and metaphorical sense identification through concrete and abstract context, in: Proceedings of the 2011 Conference on the Empirical Methods in Natural Language Processing, 2011, pp. 680–690.
[20] Y. Neuman, D. Assaf, Y. Cohen, M. Last, S. Argamon, N. Howard, O. Frieder, Metaphor identification in large texts corpora, PLoS One 8 (4) (2013) e62343.
[21] Y. Tsvetkov, E. Mukomel, A. Gershman, Cross-lingual metaphor detection using common semantic features, Meta4NLP (2013).
[22] Y. Tsvetkov, L. Boytsov, A. Gershman, E. Nyberg, C. Dyer, Metaphor detection with cross-lingual model transfer, in: Proceedings of the Annual Meeting of the Association for Computational Linguistics, 2014, pp. 248–258.
[23] Y. Wilks, L. Galescu, J. Allen, A. Dalton, Automatic metaphor detection using large-scale lexical resources and conventional metaphor extraction, Meta4NLP 2013 (2013) 36.
[24] L. Li, C. Sporleder, Using gaussian mixture models to detect figurative language in context, in: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, 2010, pp. 297–300.
[25] D.B. Bracewell, M.T. Tomlinson, M. Mohler, B. Rink, A tiered approach to the recognition of metaphor, in: Computational Linguistics and Intelligent Text Processing, Springer, 2014, pp. 403–414.
[26] M. Mohler, D. Bracewell, D. Hinote, M. Tomlinson, Semantic signatures for example-based linguistic metaphor detection, Meta4NLP 2013 (2013) 27.
[27] D.B. Bracewell, M.T. Tomlinson, M. Mohler, Determining the conceptual space of metaphoric expressions, in: Computational Linguistics and Intelligent Text Processing, Springer, 2013, pp. 487–500.
[28] J.H. Martin, A corpus-based analysis of context effects on metaphor comprehension, Trends Linguist. Stud. Monogr. 171 (2006) 214.
[29] T. Veale, Y. Hao, A fluid knowledge representation for understanding and generating creative metaphors, in: Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1, Association for Computational Linguistics, 2008, pp. 945–952.
[30] E. Shutova, Automatic metaphor interpretation as a paraphrasing task, in: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, 2010, pp. 1029–1037.
[31] E. Shutova, T. Van de Cruys, A. Korhonen, Unsupervised metaphor paraphrasing using a vector space model, in: COLING (Posters), Citeseer, 2012, pp. 1121–1130.
[32] M. Baroni, S. Bernardini, A. Ferraresi, E. Zanchetta, The wacky wide web: a collection of very large linguistically processed web-crawled corpora, Lang. Resour. Eval. 43 (3) (2009) 209–226.
[33] K. Toutanova, D. Klein, C.D. Manning, Y. Singer, Feature-rich part-of-speech tagging with a cyclic dependency network, in: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1, Association for Computational Linguistics, 2003, pp. 173–180.
[34] D. Bollegala, E. Shutova, Metaphor interpretation using paraphrases extracted from the web, PLoS One 8 (9) (2013) e74304.
[35] C. Su, J. Tian, Y. Chen, Latent semantic similarity based interpretation of chinese metaphors, Eng. Appl. Artif. Intell. 48 (2016) 188–203.
[36] C.K. Ogden, I.A. Richards, B. Malinowski, F.G. Crookshank, The Meaning of Meaning, Harcourt, Brace & World, New York, 1946.
[37] Z. Kovecses, Metaphor: A Practical Introduction, Oxford University Press, 2002.
[38] M. Sahlgren, The distributional hypothesis, Ital. J. Linguist. 20 (1) (2008) 33–54.
[39] B. Hu, B. Tang, Q. Chen, L. Kang, A novel word embedding learning model using

the dissociation between nouns and verbs, Neurocomputing 171 (2016) 1108–1117.
[40] P. Wang, B. Xu, J. Xu, G. Tian, C.-L. Liu, H. Hao, Semantic expansion using word embedding clustering and convolutional neural network for improving short text classification, Neurocomputing 174 (2016) 806–814.
[41] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, arXiv preprint arXiv:1301.3781.
[42] J. Zobel, A. Moffat, Exploring the similarity space, in: ACM SIGIR Forum, vol. 32, ACM, 1998, pp. 18–34.
[43] J.L. Fleiss, Measuring nominal scale agreement among many raters, Psychol. Bull. 76 (5) (1971) 378.
[44] D.G. Brian, F. Bowdle, The career of metaphor, Psychol. Rev. 112 (1) (2005) 216.
[45] G. Lakoff, The contemporary theory of metaphor, Metaphor Thought 2 (1993) 202–251.
[46] .G. Lakoff, Master Metaphor List, University of California, 1994.



Chang Su received her Ph.D. degree from Xiamen University, Xiamen, China, in 2008. She is now an associate professor in the Cognitive Science Department of Xiamen University. Her research interests include natural language understanding, metaphor computation and semantic computing.

Shuman Huang is currently a Master candidate at the School of Information Science and Engineering, Xiamen University. Her research interests include natural language processing and metaphor computation.

Yijiang Chen received his Ph.D. degree from Xiamen University, Xiamen, China. He is now an associate professor in the Computer Science Department of Xiamen University. His research interests include natural language processing and machine learning.