Engineering Applications of Artificial Intelligence 48 (2016) 188–203

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Latent semantic similarity based interpretation of Chinese metaphors

Chang Su [a,*], Jia Tian [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

Article history:
Received 4 June 2015
Received in revised form
17 September 2015
Accepted 27 October 2015

Keywords:
Metaphor interpretation
Natural language processing
WordNet

ABSTRACT

As a ubiquitous usage in both text and speech, metaphor now attracts more and more attention. Automatic metaphor processing can be divided into two subtasks: metaphor detection and metaphor interpretation. This paper describes an algorithm to interpret the Chinese nominal and verbal metaphors based on latent semantic similarity which we define in this paper. Our method extends the perceptual features of the source and target concepts using the synonyms in WordNet to discover the latent semantic similarity between them and thereby generates the interpretation of nominal metaphors. It is considered that if two words are latent semantic similar, not only is there an extension path in WordNet from one to the other, but also their sentiments should be consistent. So the sentiment of the word is used to constrain the extension. Without a context, we think that the results of interpretation may be multiple because there are several features of the source that can be used to describe the target. Thus, we use Google Distance to rank the interpretation results. This model achieves 85% accuracy in nominal metaphors and 86% accuracy in verbal metaphors.

1. Introduction

Metaphor research, which has been applied to many NLP problems, such as machine translation, information retrieval, question answering, discourse understanding and text summarizing, plays an important role in Natural Language Processing (NLP). Research shows that some errors in machine translation and word segmentation are caused by metaphors. Thus, a good approach, dealing with metaphors, will effectively improve the performance and precision of the existing models in other NLP items, such as machine translation. Metaphor is a widespread phenomenon in natural language, and a basic method of human thinking, and the way we identify and interpret metaphors attracts not only the attention of linguists, but also that of cognitive scientists. The method for automatically processing metaphors simulates 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 (Lakoff and Johnson, 1980).

The automatic processing of metaphors can be divided into two subtasks: metaphor detection and metaphor interpretation. Metaphor detection is to distinguish between metaphorical and literal usages; metaphor interpretation is to identify the intended literal meaning of a metaphorical expression. We consider that interpretation is the following three-way correspondence between source domain and target domain: (1) the source and target share common properties; (2) the properties of source and those of target have some similarities; and (3) the target is matched to one of the source domain's properties. In this paper, we define (1) as shallow semantic similarity and (2) as latent semantic similarity (see Section 3).

Following Krishnakumaran and Zhu (2007), we divide metaphors into three types: Type I (Nominal Metaphors), II (Verbal Metaphors) and III (Adjective Metaphors). For example:

Type I : A noun is associated with another noun through the verb "be", e.g., "Love is a journal."

Type II: A verb acts on a noun such as in the instance "He kills a process."

Type III: An adjective and the noun it describes, e.g., "sweet child," or "The book is dead." (Gandy et al., 2013)

In this paper, we focus on Type I and Type II metaphors. Regarding Type III metaphors, we take the metaphors formed as "⟨target⟩ BE [a/an/the] ⟨source⟩" as our database. Regarding Type II metaphors, we take the ones formed as "Verb–DirectObject" and "Subject–Verb" as our database.

In this paper, we claim that a metaphor is a process that creates similarity between two concepts in cases where such similarity does not already exist. If it were simply a matter of highlighting existing similarity, many relations between concepts which are not

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

http://dx.doi.org/10.1016/j.engappai.2015.10.014

metaphoric would be predicted to be metaphoric. Thus, the similarity between concepts may not be present in a knowledge-base, unless, like WordNet (Miller, 1995), a metaphor is already implicitly present. Therefore, in this paper, WordNet is used to obtain the latent semantic similarity between concepts.

Our method is divided into three steps: (1) the extraction of perceptual features of the source and target concepts; (2) the synonymous extension of extracted features using WordNet; and (3) the ranking of interpretation results based on Google Distance. In the section on extracting the perceptual features, we use the features extracted by the online database, *Sardonicus* (Veale and Hao, 2007). Then we use the synonyms in WordNet to extend the perceptual features to interpret the metaphors. Because the extension is non-directional, we use the sentiment of each feature to constrain the extension's direction, guaranteeing the extension's consistency. At last, when there is no context, we think that there should be more than one interpretation result. There should be different interpretation results related to the different features of the source concept. Thus, we use Google Distance to rank the results.

The main contributions of this paper are as follows (Baroni et al., 2009):

(1) We do not need to develop a corpus or other statistic resources for the system, but, instead employ the latent semantic similarity based on WordNet. The latent semantic similarity actually reflects the mapping of image schema between the source and target concepts.
(2) Compared with Shutova (2010), our system deals with, not only verbal, but also nominal metaphors.
(3) Our system achieves 85% accuracy for nominal metaphors and 86% accuracy for verbal metaphors.

2. Related work

2.1. Models based on inference

Martin (1994) described a metaphor comprehension system (MIDAS) and applied it to teaching software based on Unix.

Martin takes 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 extending system (MES). Firstly, the metaphor algorithm searches the 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 results of the interpretation. MIDAS relies on inference, and deals with novel metaphors without any corpus.

Veale and Hao (2008) described a “fluid knowledge presentation for metaphor interpretation and generation”, which is called Talking Points. Talking Points extracts the 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.

2.2. Models based on word paraphrasing

According to Shutova (2010), the result of interpretation should be directly embedded in other systems. They thus define metaphor interpretation “as a paraphrasing task” and describe a system that automatically derives 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.

Bollegala and Shutova (2013) 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.

In this paper, we apply the idea that views the interpretation of verbal metaphor as a paraphrasing task. Different from Shutova, when choosing a literal verb to paraphrase the metaphorical verb, we apply the semantic information of the source and target domains, not only statistic data or selectional preference.

2.3. Model based on term vector

Shutova et al. (2012) 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 the meaning list of target verbs, paraphrase candidates are extracted. After annotating the text with UKwac corpus (Baroni et al., 2009) and Stanford Part-of-Speech Tagger (Toutanova et al., 2003), 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 the interpretation, they score the obtained paraphrases with the supplementary target verb itself. The method using vector space model is also viewed as computing the similarity between the source and target domains. Compared to such similarity, this paper proposes the latent semantic similarity, which reflect how the source and target domains are similar, not only implying that the source and target domains are similar.

Overall, in terms of the previous studies on metaphor interpretation, a majority of work focus on verbal metaphors and some statistic methods are applied in metaphor interpretation. In this paper, we focus on Chinese nominal and verbal metaphor. Specially, we propose the idea of latent semantic similarity (See Section 3). Based on latent semantic similarity, this paper tries to explore the latent relation between the target and source domains. Such relation implies the semantic relation between the target and source domains, not only some statistic data.

3. Some definitions and assumptions

In this paper, we propose that the key of metaphor interpretation is to comprehend the “meaning” of the source and target concepts and to find the relationships between the two concepts. Ogden et al. (1946) gave 22 definitions of “meaning” in *The Meaning of Meaning*, three of which we focus on as follows: *“An Intrinsic property”*, which points out that the meaning of a word is its features; *“Emotion around by anything”*, which points out that sentiment is also a part of a word's meaning, and *“That which is actually related to a sign by a chosen relation”*, which points out that the meaning of a concept contains the relation to other concepts.

To build a system for metaphor interpretation, based on Ogden's definition to the “meaning”, we give some definitions and assumptions in this section.

Definition 1 (*The feature $f(C)$ of a concept C*). $f(C)$ is defined as a set of its features and formalized as $f(C)=\{p_1, p_2, \ldots, p_n\}$, among that p_i is one of the features belonging to C.

Definition 2 (*A concept's semantic $S(C)$*). The semantic of a concept is defined as a triple vector, $S(C) \cdot S(C)=\{concept\ name, f(C), Se(C)\}$, among that, $f(C)=\{p_1, p_2, \ldots, p_n\}$ is the feature of C and $Se(C)=\{se_1, se_2, \ldots, se_n\}$ is the feature's sentiment in which se_i is the sentiment relevant to the feature p_i. In this paper, the feature's sentiment is induced by the sentiment analysis tool of Stanford CoreNLP (Christopher et al., 2014) and is quantized as follows: Zero:Very Negative; One: Negative; Two: Neutral; Three: Positive; and Four: Very Positive.

Definition 3 (*A feature's semantic $S(P)$*). The semantic of a feature is defined as $S(P)=\{p, se\}$, where p is the feature and se is the feature's sentiment.

Definition 4 (*The shallow semantic similarity*). Given two features, A and B, if A is a synonym of B in WordNet, then A and B have shallow semantic similarity, formalized as $Lat^0(A, B)$. For example, the word "soft" is synonymous with "delicate" in WordNet, and it is represented as $Lat^0(soft, delicate)$.

Definition 5 (*The latent semantic similarity*). We define latent semantic similarity in the following situations (See Fig. 1):

Situation 1: Given two features, A and B, they do not have shallow semantic similarity. However, there is another feature, C. If $Lat^0(A, C) \wedge Lat^0(C, B)$, then A and B are latent semantic similar, formalized as $Lat^1(A, B)$. For example, the word "soft" has shallow semantic similarity with "delicate" and "delicate" has shallow semantic similarity with "tender", then, "soft" has latent semantic similarity with "tender", represented as $Lat^1(soft, tender)$.

Situation 2: Given three features, A, B, and C, if $Lat^1(A, B) \wedge Lat^1(B, C)$, A and C are latent semantic similar.

Situation 3: Given four features, A, B, C, and D, if $Lat^0(A, B) \wedge Lat^0(B, C) \wedge Lat^0(C, D)$, A and D are latent semantic similar, formalized as $Lat^2(A, D)$.

Given two features, E and F, they are latent semantic similar, formalized as $Lat^i(E, F)$, where ^{i} means that by having shallow semantic similarity with other i features, E and F are latent semantic similar. For example, in Situation 1, i equals one; in Situation 3, i equals two.

Hypothesis 1. A given metaphor is comprehended by the cognitive subject when the cognitive subject obtains the shallow semantic similarity or latent similarity between features of the source and target concepts.

Based on the above definitions and hypothesis, we propose an algorithm of metaphor interpretation using latent semantic similarity between the perceptual features of the source and target concepts. In this paper, the features' latent semantic similarity is represented as the synonymous extension in WordNet and we limit the level of such extension to six. If one of the target concept's features is latent semantic similar to one of the source concept's features, the source concept's feature is the metaphor's interpretation.

4. Our method

The system takes metaphors as input, where the target and source concepts are annotated. After extending the perceptual features (Bhatt, 2011 discussed the effectiveness of this extension in the interpretation of pictorial metaphors) of the source and target concepts based on the inference mechanism of latent semantic similarity, the system

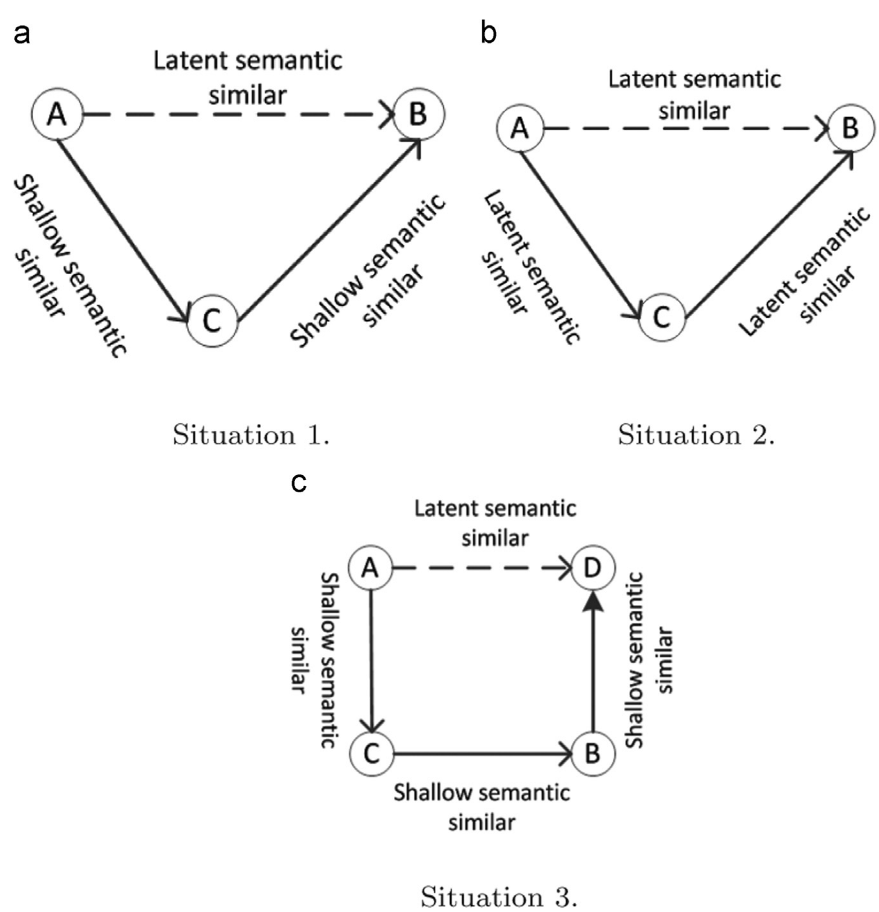

Fig. 1. Definition of latent semantic similarity. (a) Situation 1. (b) Situation 2. (c) Situation 3.

generates a list of possible interpretations that may be the literal sense of the given metaphors and ranks them by using Google Distance. The main algorithms are as follows:

Algorithm 1. Automatic interpretation for Chinese nominal metaphor.

Input: Nominal metaphors;
Output: The result of interpretation;
function NOMINAL ($Ajective1, Ajective2$)
Set six arrays, *Array*0, *Array*1, *Array*2, *Array*3, *Array*4, *Array*5;
if $Ajective1 = Ajective2$ **then return** True;
else
Extract the synonyms of *Ajective*1 by using WordNet, and save the synonyms, which have the same sentiment with *Ajective*1, in *Array*0;
Traverse *Array*0, if any element in *Array*0 equals *Ajective*2, then return “True”;
Extract the synonyms of each element in *Array*0 by using WordNet, and save the synonyms, which have the same sentiment with the corresponding element, in *Array*1;
Traverse *Array*1, if any element in *Array*1 equals *Ajective*2, then return “True”;
The iterative process continues, until *Array*5 is traversed;
end if
end function

function MAIN
Label the target, *T*, and the source, *S*;
Extract the features of *T*, p_i, and the features of *S*, c_j, by using *Sardonicus*; Record the total number of p_i as *N*, and the total number of c_j as *M*;
for $i = 0 \rightarrow N-1$ **do**
for $j = 0 \rightarrow M-1$ **do**
$mark \leftarrow Nominal(p_i, c_j)$;
if $mark = True$ **then**
Take *T* and c_j as one result of interpretation, and formalize the result as “*T* is c_j”;
end if
end for
end for
Rank all of the results by the Google Distance;
Take the result in rank 1 as the output result of interpretation;
end function

Algorithm 2. Automatic interpretation for Chinese verbal metaphor.

Input: Verbal metaphors: “Verb–Object” or “Subject–Verb”;
Output: The result of interpretation;
function VERBAL $Verb1, Verb2$
Set six arrays, *Array*0, *Array*1, *Array*2, *Array*3, *Array*4, *Array*5;
if $Verb1 = Verb2$ **then return** True;
else
Extract the synonyms of *Verb*1 by using WordNet, and save the synonyms, which have the same sentiment with *Verb*1, in *Array*0;
Traverse *Array*0, if any element in *Array*0 equals *Verb*2, then return “True”;
Extract the synonyms of each element in *Array*0 by using WordNet, and save the synonyms, which have the same sentiment with the correspond element, in *Array*1;
Traverse *Array*1, if any element in *Array*1 equals *Verb*2, then return “True”;
The iterative process continues, until *Array*5 is traversed;
end if
end function

function MAIN
Label the verb, *V*, and the subject or object, *T*;
Extract the literal verbs, V_i, with a relation of “$V_i - T$” or “$T - V_i$”, using *BNC*; Record the total number of V_i as *N*;
for $i = 0 \rightarrow N-1$ **do**
$mark \leftarrow Verbal(V, V_i)$
if $mark = True$ **then**
Query V_i in *WordNet*; Parse the senses of V_i with Stanford CoreNLP tools; Extract the direct nouns of V_i, N_k; Record the total number of N_k as *Z*;
for $k = 0 \rightarrow Z-1$ **do**
if *T is hypernym, hyponym, or latent semantic similar with* N_k **then**
Take *T* and V_i as one result of interpretation, and formalize the result as “$T - V_i$” or “$V_i - T$”;
end if
end for
end if
end for
Rank all of the results by the Google Distance;
Take the result in rank 1 as the output result of interpretation;
end function

4.1. *The extraction of perceptual features*

The perceptual features of the source and target concepts in our system are extracted from the simile-finder, *Sardonicus. Sardonicus* is an adjective taxonomy that knows the exemplary properties of different objects in the real world by sifting the contents of the web in search of meaningful comparisons. To test the reliability of the perceptual features in *Sardonicus*, we invited fifteen independent volunteers with sparse or no linguistic knowledge. All received the same part of our extracted perceptual features from Sardonicus and were asked to express their acceptability of the features on a scale from one to five where one means very low acceptability and five means very high acceptability. For example, Table 1 shows the average acceptability for the features of *Lawyer* and *Fox*(律师是狐狸。).

4.2. *Generating interpretations using WordNet*

Our system aims to extract perceptual features of the source and target concepts from *Sardonicus* and use them to generate interpretations of a metaphorical expression with WordNet. We think that there must be some similarities between the source's features and the target's features in a metaphorical expression. In some cases, the similarities are shallow semantic similarities. In some metaphors, they are latent semantic similarities that are found through the inference mechanism of latent semantic similarity (See Section 3). Thus, the synonyms in WordNet suggest some relationships between the source and the target. The goal of this paper is to discover this relationship to generate the metaphor's interpretations.

4.2.1. *Generating interpretations with WordNet*

We present an example which represents the process of our synonymous extension to generate a metaphor's interpretation with the synsets in WordNet (See Fig. 2). We take the metaphorical expression *Lawyer is a fox* as an example. In this example,

Table 1
The average acceptability to the features of *Lawyer* and *Fox*.

Subjects	Features	Acceptability
Lawyer律师	professional专业	3.667
	prestigious有名望	2.667
	smart机智	4.333
	expert熟练	4.000
	knowledgeable知识渊博	3.667
	crooked不正当	1.333
	discreet谨慎	2.667
	unethical不道德	2.667
Fox狐狸	dishonest不诚实	3.000
	canny精明	4.000
	red红色	4.667
	sly淘气	3.333
	shy害羞	3.000
	wily狡猾	3.333
	cunning狡猾	4.333
	quick快	4.667

The degree of acceptability is from 5 (easy to be accepted) to 1 (hard to be accepted) and the degrees showed are an average degrees.

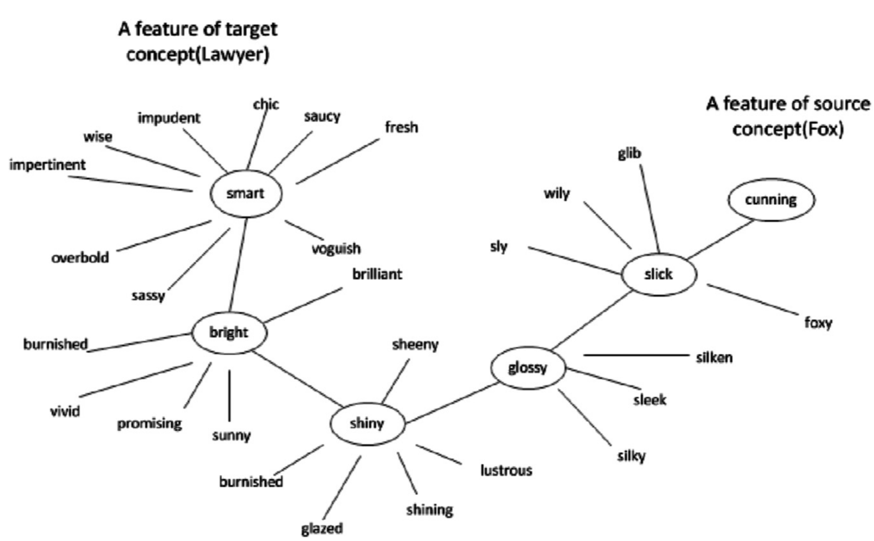

Fig. 2. The synonyms extension from the source's feature, *smart*, to the target's feature, *cunning*.

smart is one of the perceptual features of the target concept, *Lawyer*, and *cunning* is a perceptual feature of the source concept, *Fox*.

Via an iterative process, our system seeks to find a synonymous relationship between the perceptual feature of target concept, *smart* and the perceptual feature of source concept, *cunning*.

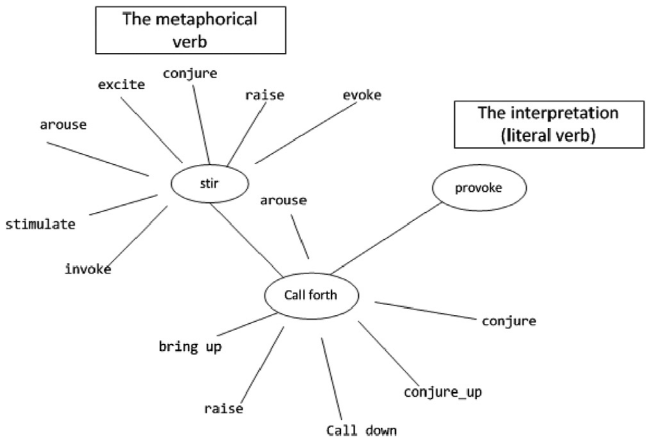

Fig. 3. The extension from “stir” to the literal verb “provoke” using WordNet.

Iteration 1: In the first iteration, we discover all the synonyms of *smart*, some of which (impudent, chic, saucy, fresh, voguish, *bright*, sassy, overbold, impertinent, and wise) are shown in Fig. 2.

Iteration 2: In the second iteration, we randomly select one of the example synonyms listed in the first iteration (*bright*) and discover all of its synonyms, some of which (burnished, vivid, promising, sunny, *shiny*, and brilliant) are shown in Fig. 2.

Iteration 3: In the third iteration, we randomly select one of the example synonyms listed in the second iteration (*shiny*) and repeat the discovery process given for Iteration 2.

This iterative process continues (as shown in Fig. 2) until, after five iterations, we reach *cunning*. In this paper, the maximum number of iterations is limited to six.

If within the iterative process discussed above, a path from the feature of the target concept to the feature of the source concept is discovered, the system views the source's feature (“cunning” in Fig. 3) as the feature common to the source and target. The common feature (“cunning” in Fig. 2) is available to describe the target concept. The system outputs the interpretation of any given metaphor with a form “The TARGET is FEATURE.”, where TARGET is the target concept and FEATURE is the common feature shared by the source and target. For example, the interpretation of the metaphorical expression “Lawyer is a fox. (律师是狐狸)” is generated as “The lawyer is cunning. (律师是狡猾的)” Table 2 shows all the interpretations of the metaphor, “Lawyer is a fox (律师是狐狸)”.

4.2.2. A heuristic generation of interpretations by word's sentiment

There are other iterative approaches that might lead to words (which we do not want to consider) that are unrelated to the features of the source concept. Thus, taking “sentiment” as the constraint, we restrict the direction of the iterative process. In this paper, using the sentiment analysis tool of Stanford CoreNLP, we apply the sentiment of each feature to constrain the direction of synonymous extension, according to a scale from zero to four as follows: zero (very negative), one (negative), two (neutral), three (positive), and four (very positive). For example, Table 3 shows the sentiment of each synonym of the feature “smart” whose sentiment is four.

The guidelines of sentiment's constraint are as follows: (1) if a positive/negative target word is extended to its synonyms with WordNet, the synonyms with positive/negative sentiment (rather than all of the synonyms) are put into the next iteration; (2) if the target word is neutral, then all of its synonyms are put into the next iteration.

Table 4 shows the interpretation result of the metaphor “Lawyer is a fox” when constrained by sentiment. According to Table 4, when the feature “smart” extends to the next iteration, the synonym whose sentiment is positive and very positive will be taken into account. That is, the first iteration in Fig. 2 will contain only “fresh”, “wise”, “sassy” and “bright”.

Comparing the result from Table 4 with those shown in Table 2, some unreliable interpretations are removed by using sentiment to constrain the synonymous extension.

Table 2
The all interpretations of the metaphor *Lawyer is a fox. (律师是狐狸).*

Lawyer is quick. 律师是快的	Lawyer is smart. 律师是机智的
Lawyer is red. 律师是红色的	Lawyer is sly. 律师是淘气的
Lawyer is lively. 律师是活泼的	Lawyer is crafty. 律师是灵巧的
Lawyer is alert. 律师是机警的	Lawyer is shy. 律师是害羞的
Lawyer is wily. 律师是狡猾的	Lawyer is crazy. 律师是疯狂的
Lawyer is cunning. 律师是狡猾的	

Table 3
The sentiment of each synonym of “smart”(the sentiment of smart is 4).

chic 别致	2	fresh 新鲜	4
wise 明智	3	impudent 无耻	2
overbold 鲁莽	2	sassy 活泼	3
bright 聪明	3	voguish 时	2

Table 4
The interpretation result of metaphor “Lawyer is a fox” when supervised by sentiment.

Lawyer is cunning. 律师是狡猾的	Lawyer is smart. 律师是机智的
Lawyer is lively. 律师是活泼的	Lawyer is crazy. 律师是疯狂的
Lawyer is dishonest. 律师是不诚实的	Lawyer is crafty. 律师是灵巧的

4.2.3. Ranking based on the Normalized Google Distance

The results shown in Table 4 contain some irrelevant interpretations (e.g., *Lawyer is lively*.) each of which has a different relevancy to the target concept. Therefore, the results must be re-ranked. This paper uses Normalized Google Distance (NGD) (Cilibrasi and Vitányi, 2007) to measure the relevancy of the target concept and the listed features. The Normalized Google Distance is defined as follows:

$$NGD(x,y) = \frac{\max\{\log f(x), \log f(y)\} - \log f(x,y)}{\log M - \min\{\log f(x), \log f(y)\}} \tag{1}$$

where M is the total number of web pages searched by Google; $f(x)$ and $f(y)$ are the number of hits for search terms x and y, respectively; and $f(x,y)$ is the number of web pages on which both x and y occur. If the two search terms, x and y, never occur together on the same web page, but do occur separately, the normalized Google distance between them is infinite. If both terms always occur together, their NGD is zero, or equivalent to the coefficient between x^2 and y^2.

To measure the relevancy of the target concept and the feature, we define Google Relevancy (GR) as a normalized standard to measure the relevancy between two words as follows:

$$GR(x,y) = \frac{\lambda}{NGD(x,y) + \lambda} \tag{2}$$

where λ is a regulation parameter from zero to one. In our experiment, we choose λ as 0.6. The ranking is demonstrated in Table 5.

The expectation is that the interpretation in the first rank (i.e. the target with which the feature in question has the highest association) represents the literal interpretation.

4.3. The method with Type 2 metaphors

Given a verbal metaphor with a form of “Verb–Object” or “Subject–Verb” (e.g., *stir excitement, hold back truth* Shutova, 2010), we think that if the verbs acting on the Object or Subject in literal usages are latent semantic similar to the verb in the metaphor, then the literal verbs are the interpretations of this metaphor. Then, taking the metaphor “stir excitement” as an example, we apply an iterative process (similar to that shown in Fig. 2) as follows:

Step 1: *The extraction of literal verbs*. We extract literal verbs with a relation pattern of “Verb [a/the/an] Object” and “Subject+Verb” from the BNC (British National Corpus) (Aston et al., 1998). For example, in the metaphor “stir excitement”, the literal verb of the object “excitement” extracted from BNC are “add, tingle, create, suppress, rise, provoke, produce, make, keep, contain, conjure”, etc.

Step 2: *Extension of verbs*. Using this iterative process, we extend the metaphorical verb “stir” using WordNet to decide whether “stir” is latent semantic similar to those literal verbs. If the “stir” is latent semantic similar to one of those literal, then the literal verb is the interpretation. For example, Fig. 3 shows iterative process to extend from “stir” to the literal verb “provoke”. As before, the iterative process is constrained by the sentiment of every verb.

Step 3: *Filtering the interpretations*. Among the interpretations, some are irrelevant to the “Object” or “Subject”. This paper uses WordNet to measure a verb's selectional preference, thus filtering the irrelevant interpretations. In WordNet, a verb has some senses and acts on different objects(or co-occur with different subjects) in each sense. The majority of a verb's senses in WordNet are literal. In some, but not all cases, it is effective to use these senses to filter the interpretation results. We extract the senses of the candidate verbs and parse these senses with Dependency Parse of Stanford CoreNLP (Christopher et al., 2014). Then we extract these objects as the verb's direct objects. Fig. 4 shows the extracted direct objects of “provoke” and “add”. We take the extracted direct objects as the objects that have selectional preference with the verb. If the object is a hypernym, a hyponym, or latent semantic similar with the target noun (In the metaphor “stir excitement”, “excitement” is the target noun), then the verb has selectional preference with the target noun and is an interpretation of the metaphor.

Table 5
The list of interpretations ranked by the Google Distance.

Rank	Interpretation	Value of NGD
1	Lawyer is smart. 律师是机智的	0.72437656
2	Lawyer is cunning. 律师是狡猾的	0.62153876
3	Lawyer is dishonest. 律师是不诚实的	0.55879956
4	Lawyer is crafty. 律师是灵巧的	0.47348243
5	Lawyer is lively. 律师是活泼的	0.29129228
6	Lawyer is crazy. 律师是疯狂的	0.18534878

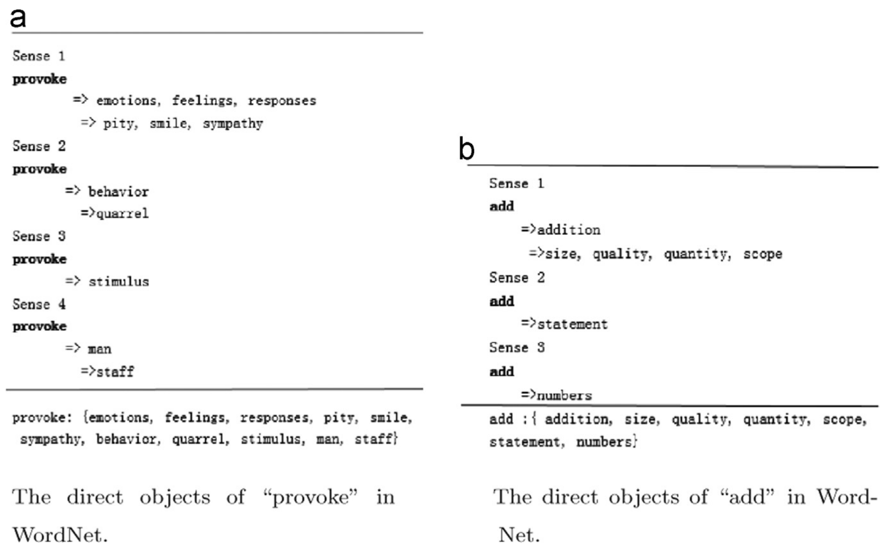

Fig. 4. The verb's direct objects in WordNet. (a) The direct objects of “provoke” in WordNet. (b) The direct objects of “add” in WordNet.

Table 6
The interpretation result of metaphor “stir excitement”.

	Metaphor: stir excitement(搅拌兴奋)	
Rank	Interpretation	Value of NGD
1	stimulate excitement 刺激兴奋	0.74238260
2	provoke excitement 引起兴奋	0.64392614
3	make excitement 制造兴奋	0.61325574
4	crate excitement 产生兴奋	0.49265650
5	suppress excitement 抑制兴奋	0.42063612

Step 4: *Ranking the results*. At last, we rank the interpretation results using Google Distance (see Table 6). In this paper, we take the interpretation in the first rank as the interpretation of the target metaphor.

5. Evaluation and discussion

5.1. Test data

We focus on Type 1 and Type 2 metaphors. Our test data is annotated from the web, newspapers, blogs, and the “Reader”.[1] The test data contains 85 nominal metaphors and 35 verbal metaphors within the following genres: news/ journal articles, politics, finance, essays, fiction and speech. Appendices A and B show a part of our test data and interpretation results.

Using multiple annotators on the corpus, we tested reliability of this annotated data.

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 evaluate the reliability of our annotation scheme by assessing the inter-annotator agreement in terms of κ (Siegel and Castellan, 1988). The classification was performed with the an agreement of 0.66(κ), which is considered reliable. The main source of disagreement was that some conventional metaphor had been viewed as literal expressions.

[1] A Chinese magazine. URL: http://www.duzhe.com

Table 7
The accuracy and four baselines.

Type		Random word (%)	Shallow similarity (%)	Without NGD (%)	Without sentiment (%)	Our's (%)
Accuracy	Nominal	41	58.33	50	70.83	85
	Verbal	45.24	61.87	52.66	66.33	86

5.2. Evaluation

5.2.1. Accuracy and baseline

To determine accuracy, we invited five volunteers to evaluate the interpretation results ranked one by NGD. The volunteers rated the acceptability of the interpretation results; and, 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 divide the acceptability into five levels instead of simple binary decision (accept/decline), because the five-levels method makes the evaluation finer-grained, compared with a binary decision. We then evaluate the system performance according to their judgments in terms of accuracy. We take the results with average acceptability less than three as incorrect results and those greater than three as correct results. The system achieves 85% accuracy for nominal metaphors and 86% accuracy for verbal metaphors.

In order to evaluate the quality of the system, we compare our approach against four baselines: (1) using shallow semantic similarity; (2) without sentiment constraints; (3) using a random adjective/verb: to nominal metaphors, we randomly select one feature of the source and output the interpretation with the target, "target is feature"; to verbal metaphors, we select a random verb from the candidates and output the interpretation; (4) without ranking by NGD: we choose a random result, rather than choosing the result ranked one with NGD. After evaluating the annotators' results, we compare the accuracies with the algorithm proposed in this paper (see Table 7).

Compared against the four baselines, our system achieves better accuracy for both nominal and verbal metaphors. According to the baselines, the method using WordNet expansion brings some noise, and the application of sentiment constraints and ranking by NGD is significant. Especially, while using only shallow semantic similarity (Baseline 1), the system did not output the results in 20% for the test instances.

5.2.2. Analysis result

The performance of our system is promising. In particular, the sentiment-based filter performs with considerable improvement in accuracy. However, the results still have problems, some of which are as follows:

1. Some metaphors have become conventional to human beings, e.g., "the net of love (情网)". It is difficult to extract the features of the source "net", causing a mistaken result.
2. Some metaphors contain a mapping of many aspects of the source and target concepts, e.g., "atom is solar system (原子是太阳系)". The features in this paper are adjectives; however, in the metaphor, the mapping is the structure of the two domains.
3. The semantics defined in this paper contains only features and sentiments. In some cases, the interpretation of metaphors is influenced by their context. The features extracted from the context may be more useful in some metaphors than in others.

6. Conclusion and future work

We presented a novel approach to metaphor interpretation. Our system achieves 85% accuracy for nominal metaphors and 86% accuracy for verbal metaphors. Distinguished from other work, our method does not develop any statistical corpus or any hand-crafted knowledge for the system, but, in contrast, relies on latent semantic similarity induced by WordNet. We propose that the latent semantic similarity reflects the mapping of the intentional graphical representation between the source and target concepts. By the extension of synonyms, our method represents the process of how humans understand metaphors. We evaluated this measure relative to human judgment and found it to be reasonably significant.

In this paper, we propose two algorithms that deal with Chinese nominal and verbal metaphors, by applying existing software tools (e.g., WordNet, Sardonicus and Stanford CoreNLP). We extract the features manually by using Sardonicus. Thus, an automatical algorithm to extract the features of the source and target domains is one of the future work. And, the experiment shows that the two algorithms in this paper are implemented.

According to the results, our method attains good performance for Chinese nominal and verbal metaphors without context, but is somewhat weak for metaphors with context. In some cases, the context contains important information for metaphor interpretation and also provides some constraints to the interpretation. Furthermore, a metaphor in different cultural backgrounds has different interpretations (e.g., "dragon" represents "nobility" and "stateliness" in Chinese but represents "evil" in western culture). Thus, an intelligent metaphor interpretation system must take the cultural background into account.

This paper, by using WordNet, took a synonym's extension to represent latent semantic similarity by using WordNet. In future work, a semantic inference mechanism is planned to replace this simple extension. Furthermore, a more precise method of feature extraction is also a part of our future work.

Acknowledgement

Funding was provided by the National Natural Science Foundation of China under Grant no. 61075058. The authors wish to thank Mr. Michael McAllister for his valuable assistance in proofreading this paper.

Appendix A. The acceptability of the interpretation results (Nominal Metaphors)

Metaphors	Interpretations	Acceptability					
		1	2	3	4	5	Average
Petroleum is gold. 石油是黄金	Petroleum is expensive 石油是昂贵的	5.0	5.0	5.0	5.0	5.0	5.0
Pour the flood of anger to him. 把愤怒的洪水向他倾泻	Anger is strong. 愤怒是强烈的	4.0	4.0	5.0	4.0	4.0	4.2
Appreciation is nectar. 欣赏是琼浆	Appreciation is helpful. 欣赏是有用的	2.0	2.0	2.0	2.0	2.0	2.0
Appreciation is sunshine. 欣赏是阳光	Appreciation is warm. 欣赏是温暖的	2.0	3.0	4.0	3.0	4.0	3.2
Argument is war. 争论是战争	Argument is serious. 争论是严重的	2.0	1.0	2.0	1.0	1.0	1.4
Atom is solar system. 原子是太阳系	Atom is complex. 原子是复杂的	4.0	2.0	3.0	4.0	3.0	3.2
Benediction is tent. 祝福是帐篷	Benediction is warm. 祝福是温暖的	3.0	3.0	2.0	4.0	3.0	3
Brain is a machine. 人的大脑是一台机器	Brain is precise. 大脑是精密的	4.0	5.0	5.0	4.0	5.0	4.6
Child is the flower of motherland. 儿童是祖国的花朵。	Child is innocent. 儿童是天真的	2.0	2.0	1.0	2.0	3.0	2.0
In the ocean of crowd. 在茫茫人海里	Crowd is massive. 人是大量的	5.0	5.0	5.0	5.0	5.0	5.0
Dew is pearl. 露水是珍珠	Dew is brilliant. 露水是晶莹的	5.0	5.0	5.0	5.0	5.0	5.0
Expectation is a spring breeze 期盼是一阵春风	Expectation is tender. 期盼是温和的	2.0	3.0	2.0	3.0	3.0	2.6
Her eyes are stars. 她那双眼睛是星星	Eye is bright. 眼睛是明亮的	5.0	5.0	5.0	5.0	5.0	5.0
Family life becomes hell. 家庭生活变成了地狱	Family life is horrible. 家庭生活是恐怖的	4.0	5.0	3.0	4.0	5.0	4.2
Fireworks blossoms into a poem. 烟火瞬间绽放成诗	Firework is beautiful. 烟火是漂亮的	4.0	4.0	2.0	4.0	2.0	3.6
Flood is beast. 洪水是猛兽	Flood is cruel. 洪水是残忍的	4.0	5.0	2.0	4.0	5.0	4.0

Full moon is a cake. 满月是块大饼	Full moon is round. 满月是圆的	5.0	5.0	5.0	5.0	5.0	5.0
New moon is a banana. 新月是只香蕉	New moon is curved. 新月是弯的	5.0	5.0	5.0	5.0	5.0	5.0
Happiness is goddess. 幸福是女神	Happiness is sacred. 幸福是神圣的	3.0	5.0	2.0	2.0	4.0	3.2
Intellect is a groom. 理智是一个马夫	Intellect is constrained. 理智是约束的	3.0	5.0	2.0	2.0	4.0	3.2
Kapok is a burning torch. 木棉是一支燃烧的火炬	Kapok is red. 木棉是红色的	4.0	5.0	5.0	4.0	5.0	4.6
Is this lake the tear of Aphrodite? 这湖水，是美神遣落的泪珠么？	Lake is clear. 湖水是清澈的	5.0	4.0	4.0	4.0	4.0	4.2
Landscape is poem. 长诗短歌般的山水	landscape is magnificent. 山水是壮丽的	5.0	3.0	4.0	3.0	4.0	3.8
Lawyer is a fox. 律师是狐狸	Lawyer is smart. 律师是机智的	4.0	4.0	4.0	4.0	4.0	4.0
Life is a bronco. 生命是一匹野马	Life is Lively. 生命是活泼的	4.0	4.0	5.0	4.0	4.0	4.2
Life is a cup of tea. 生命是一杯清茶	Life is fragrant. 生命是芬芳的	5.0	4.0	5.0	4.0	5.0	4.6
The flower of life. 生命之花	Life is fresh. 生命是新鲜的	2.0	4.0	4.0	3.0	4.0	3.4
Life is a spring. 生命是一眼清泉	Life is sweet. 生命是甜美的	2.0	2.0	2.0	2.0	3.0	2.2
The flame of love. 爱情之火	Love is intense. 爱情是热烈的	4.0	5.0	5.0	4.0	5.0	4.6
Love is sugar. 爱情是糖	Love is sweet. 爱情是甜蜜的	5.0	5.0	5.0	2.0	5.0	4.4
The net of love. 情网	Love is tight. 爱情是紧密的	1.0	2.0	2.0	1.0	5.0	1.8
Marriage is a war. 爱情是战争	Marriage is acute. 婚姻是激烈的	4.0	2.0	5.0	3.0	2.0	3.2
Mind is building. 理念是建筑	Mind is grand. 理念是宏伟的	2.0	4.0	2.0	3.0	3.0	2.8
Mind is building. 理念是建筑	Mind is grand. 理念是宏伟的	2.0	4.0	2.0	3.0	3.0	2.8
Missing is a ray of sunshine. 思念是一缕阳光	Missing is warm. 思念是温暖的	4.0	4.0	5.0	4.0	4.0	4.2

Estrangement is a barrier. 隔阂,是一层厚障壁	Estrangement is impassable. 隔阂是无法逾越的	5.0	5.0	5.0	5.0	5.0	5.0
Motherhood is a big flame. 母爱是一种巨大的火焰	Motherhood is passionate. 母爱是炽热的	5.0	5.0	5.0	5.0	5.0	5.0
New York is the heart of America. 纽约是美国心脏	New York is important. 纽约是重要的	5.0	5.0	4.0	5.0	5.0	4.8
Nurse is an angel. 护士是天使	Nurse is sweet. 护士是甜美的	2.0	4.0	2.0	2.0	4.0	2.8
Organs on her face are elfin. 她脸上的器官是小精灵	Organ is delicate. 器官是精致的	5.0	2.0	4.0	3.0	2.0	3.2
The people are monkeys. 那些人简直是一群猴子	People is ridiculous. 人是滑稽的	4.0	3.0	4.0	3.0	3.0	3.4
Prairie is a picture. 草原是一幅画	Prairie is good. 草原是美丽的	5.0	5.0	5.0	5.0	4.0	4.8
Skin is snow. 肌肤如雪	Skin is white. 皮肤是白的	5.0	5.0	5.0	5.0	5.0	5.0
This smile is sunshine. 笑容是阳光	Smile is bright. 笑容是明亮的	3.0	4.0	4.0	3.0	4.0	3.6
Stars in the sky are gems. 天上的星星是一颗颗宝石	Star is brilliant. 星星是灿烂的	5.0	5.0	5.0	5.0	5.0	5.0
Sun becomes a fireball. 太阳变成了火球	Sun is hot. 太阳是热的	5.0	5.0	5.0	5.0	5.0	5.0
The swan is a prince. 天鹅是王子	Swan is courtly. 天鹅是庄严的	2.0	2.0	3.0	2.0	3.0	2.4
The swan is a princess. 天鹅是公主	Swan is graceful. 天鹅是优雅的	5.0	5.0	5.0	5.0	5.0	5.0
Trust is a weapon. 信任是一种武器	Trust is powerful. 信任是强有力的	5.0	5.0	5.0	5.0	5.0	5.0
Tree of youth grows more frondent. 青春的树越长越葱茏	Youth is strong. 青春是茁壮的	4.0	4.0	2.0	4.0	4.0	3.6
Life is grass. 生命是小草	Life is weak. 生命是脆弱的	3.0	5.0	2.0	2.0	5.0	3.4
Life is cloud. 生命是云朵	Life is pure. 生命是纯洁的	3.0	5.0	2.0	2.0	5.0	3.4
Missing is a cup of tea. 思念是一杯浓浓的茶	Missing is bitter. 思念是痛苦的	2.0	1.0	3.0	3.0	2.0	2.2
Life is a game. 人生是一场游戏	Life is virtual. 人生是虚幻的	1.0	1.0	4.0	4.0	5.0	3.0

Her life is a feast. 她的人生是一场盛宴	Life is wonderful. 人生是精彩的	4.0	5.0	5.0	4.0	5.0	4.6
Life is a cup of wine. 生活是一杯酒	Life is bitter. 生活是苦的	3.0	2.0	2.0	1.0	4.0	2.4
Meteor is a lantern in their hand. 那朵流星,是他们提着灯笼在走	Meteor is brilliant. 流星是闪亮的	4.0	5.0	2.0	4.0	5.0	4.0
Essay is flower. 散文这枝花	Essay is beautiful. 散文是美丽的	5.0	5.0	4.0	3.0	4.0	4.2
the fetter of racism 种族歧视的枷锁	racism is oppressive 种族歧视是压迫的	5.0	5.0	4.0	4.0	5.0	4.6
shackle of apartheid 种族隔离的镣铐	apartheid is restrained 种族隔离是束缚的	5.0	5.0	5.0	4.0	5.0	4.8
the tidewater of my sentiment 我思想感情的潮水	sentiment is powerful 思想感情是强有力的	5.0	4.0	2.0	4.0	5.0	4.0
Courtyard is a box. 四合院是一个盒子	Courtyard is quadrate. 四合院是方的	5.0	5.0	4.0	4.0	4.0	4.4
the flower of human's thought 人类思维的花朵	thought is gorgeous 思维是华丽的	4.0	5.0	2.0	3.0	5.0	3.8
Life is a cup of water. 生活是一杯白开水	Life is insipid. 生活是平淡的	5.0	5.0	5.0	5.0	5.0	5.0
Word is flower in heart. 文字，是心上的花	Word is beautiful. 文字是美丽的	5.0	5.0	3.0	4.0	5.0	4.4
Happiness is blanket. 幸福是一条毛毯	Happiness is warm. 幸福是温暖的	5.0	5.0	5.0	5.0	5.0	5.0
Friendship is the cloud. 友谊是天空中的云朵	Friendship is pure. 友谊是纯洁的	5.0	5.0	3.0	5.0	5.0	4.6
Friendship is flame. 友谊是焰火	Friendship is intense. 友谊是强烈的	3.0	4.0	4.0	4.0	5.0	4.0
Friendship is sunshine. 友谊是一片阳光	Friendship is warm. 友谊是温暖的	5.0	5.0	5.0	4.0	3.0	4.4
Opportunity is epiphyllum. 机会是昙花	Opportunity is transitory. 机会是短暂的	5.0	5.0	5.0	5.0	5.0	5.0
Father is tree. 父亲是大树	Father is strong. 父亲是强壮的	5.0	5.0	4.0	4.0	5.0	4.6
Client is god. 顾客是上帝	Client is exalted. 顾客是尊贵的	5.0	5.0	5.0	3.0	5.0	4.6
Time is arrow. 光阴是利箭	Time is fast. 时间是飞快的	5.0	5.0	5.0	5.0	5.0	5.0

Inspiration is spark. 灵感是火花	Inspiration is instantaneous. 灵感是瞬间的	4.0	2.0	3.0	5.0	5.0	3.8
Rumor is plague. 流言是瘟疫	Rumor is horrible. 流言是可怕的	5.0	5.0	3.0	5.0	4.0	4.8
Disaster is nightmare. 灾难是噩梦	Disaster is terrible. 灾难时恐怖的	5.0	5.0	5.0	5.0	4.0	4.8
Skin is satin. 肌肤是绸缎	Skin is smooth. 皮肤是光滑的	5.0	5.0	5.0	5.0	5.0	5.0
Hair is snow. 头发如雪	Hair is white. 头发是白的	5.0	5.0	5.0	4.0	5.0	4.8
Lake is jadeite. 湖水是‰翠	Lake is green. 湖水是绿的	5.0	5.0	5.0	4.0	4.0	4.6

Appendix B. The acceptability of the interpretation results (Verbal Metaphors)[2]

[2] The metaphorical verb is labeled as ***v***, the subject is labeled as ***s*** and object is labeled as ***o***.

Metaphors	Interpretations	Acceptability					
		1	2	3	4	5	Average
bank stock$_S$ dive$_V$ 银行股跳水	stock plummet 股票下跌	5.0	5.0	5.0	4.0	5.0	4.8
economy$_S$ pick up$_V$ 经济回暖	economy recover 经济恢复	5.0	4.0	5.0	5.0	5.0	4.8
internationalization$_S$ take off$_V$ 国际化进程起飞	internationalization begin 国际化进程开始	5.0	4.0	5.0	3.0	5.0	4.4
kill$_V$ economic creativity$_O$ 扼杀经济创新力	reduce creativity 降低创新力	5.0	5.0	4.0	4.0	2.0	4.0
investment demand$_S$ ebb$_V$ 投资性需求退潮	demand reduce 需求减少	5.0	5.0	4.0	5.0	5.0	4.8
multilateralism$_S$ die$_V$ 多边主义死亡	multilateralism disappear 多边主义消失	5.0	5.0	5.0	4.0	5.0	4.8
economic index$_S$ landslide$_V$ 经济指标滑坡	index reduce 指标降低	5.0	5.0	5.0	4.0	5.0	4.8
company$_S$ turn back$_V$ 公司转身	company transform 公司转型	5.0	5.0	5.0	3.0	5.0	4.6
stock$_S$ sink$_V$ and float$_V$ 股票沉浮	stock rise and plument 股票上升和下跌	4.0	1.0	4.0	2.0	4.0	3.0
thunderclap$_S$ howl$_V$ ruthlessly 雷？无情地怒吼	thunderclap thunder 雷电轰隆	3.0	2.0	4.0	2.0	1.0	2.4
Do you see the wave$_S$ dancing$_V$? 你看到海浪在舞蹈吗?	wave surge 海浪汹涌	5.0	2.0	4.0	4.0	5.0	4.0
All cedars$_S$ wave$_V$ to me. 所有冷杉在向我招手。	cedar swing 杉树摆动	5.0	5.0	5.0	4.0	5.0	4.8
carve$_V$ life$_O$ 镌刻人生	live life 度过人生	3.0	2.0	3.0	2.0	2.0	2.4
boat$_S$ twitch$_V$ suddently 船身突然一阵痉挛	boat rock 船身摇晃	5.0	5.0	4.0	4.0	4.0	4.4
weave$_V$ relationship$_O$ 编织人际关系	establish relationship 建立人际关系	3.0	2.0	4.0	3.0	2.0	2.8

warm sunshine$_S$ embrace$_V$ them 温煦的阳光拥抱着他们	sunshine irradiate 阳光照耀	5.0	5.0	4.0	4.0	5.0	4.6
sow$_V$ moral$_O$ 播种道德	promote moral 弘扬道德	4.0	5.0	5.0	5.0	5.0	4.8
nightingales$_S$ sing$_V$ in the evening 夜莺在夜晚歌唱	nightingale tweet 夜莺鸣叫	5.0	5.0	4.0	4.0	5.0	4.6
embrace$_V$ life$_O$ 拥抱生活	face life 面对生活	2.0	2.0	3.0	3.0	3.0	2.6
branch$_S$ struggle$_V$ in the wind 枝条在狂风中挣扎	branch swing 枝条摆动	5.0	5.0	4.0	4.0	3.0	4.2
rose spill$_V$ perfume$_O$ 玫瑰花蕊溢出芳香	give off perfume 散发芬芳	5.0	5.0	5.0	5.0	5.0	5.0
Earth$_S$ shivers$_V$ under the wheels. 大地在车轮下颤栗	earth shake 大地震动	5.0	5.0	5.0	5.0	5.0	5.0
radiate$_V$ tender$_O$ using language 用语言辐射温柔	express tender 表达温柔	5.0	5.0	4.0	4.0	5.0	4.6
chew$_V$ joke$_O$ 咀嚼笑话	understand joke 理解笑话	5.0	5.0	5.0	3.0	3.0	4.2
enthusiasm$_S$ burns$_V$ 热情燃烧	enthusiasm spurt 热情迸发	5.0	5.0	3.0	4.0	5.0	4.4
brew$_V$ darkness$_O$ 酿造黑暗	make darkness 制造黑暗	5.0	5.0	4.0	4.0	5.0	4.6
My brain$_S$ a little rusts$_V$ today. 我的大脑今天有点生锈	brain dull 大脑迟钝	5.0	5.0	5.0	5.0	5.0	5.0

References

Aston, G., Burnard, L., 1998. The BNC Handbook: Exploring the British National Corpus with Sara, Designing A Corpus of Cameroonian English.

Baroni, M., Bernardini, S., Ferraresi, A., Zanchetta, E., 2009. The wacky wide web: a collection of very large linguistically processed web-crawled corpora. Lang. Resour. Eval. 43 (3), 209–226.

Bhatt, R., 2011. Interpretation of metaphors with perceptual features using Wordnet, Engineering Psychology and Cognitive Ergonomics. Springer, Berlin, Heidelberg, pp. 21–27.

Bollegala, D., Shutova, E., 2013. Metaphor interpretation using paraphrases extracted from the web. PloS One 8 (9), e74304.

Christopher, M., Mihai, D.S., John, B., 2014. The stanford corenlp natural language processing toolkit. In: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations. Association for Computational Linguistics, Baltimore, Maryland, USA, pp. 55–66.

Cilibrasi, R.L., Vitányi, P.M.B., 2007. The google similarity distance. IEEE Trans. Knowl. Data Eng. 19 (3), 370–383.

Gandy, L.M., Atallah, M., Al, E., 2013. Automatic identification of conceptual metaphors with limited knowledge. In: Twenty-Seventh AAAI Conference on Artificial Intelligence.

Krishnakumaran, S., Zhu, X., 2007. Hunting elusive metaphors using lexical resources. In: Proceedings of the Workshop on Computational approaches to Figurative Language. Association for Computational Linguistics, Rochester, NY, pp. 13–20.

Lakoff, G., Johnson, M., 1980. Metaphors we live by. In: Philosophical Books, 2nd Ed., vol. 23, pp. 111–116.

Martin, J.H., 1994. A corpus-based analysis of context effects on metaphor comprehension. In: Corpus-Based Approaches to Metaphor and Metonymy.

Miller, G.A., 1995. Wordnet: a lexical database for english. Commun. ACM 38 (11), 39–41.

Ogden, C.K., Richards, I.A., Malinowski, B., Crookshank, F.G., 1946. The Meaning of Meaning. Harcourt, Brace & World, New York.

Shutova, E., Van de Cruys, T., Korhonen, A., 2012. Unsupervised metaphor paraphrasing using a vector space model. In: COLING (Posters), Citeseer, 2012, pp. 1121–1130.

Shutova, E., 2010. 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, Los Angeles, California, pp. 1029–1037.

Siegel, Castellan, N.J., 1988. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill Book Company, New York.

Toutanova, K., Klein, D., Manning, C.D., Singer, Y., 2003. 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, pp. 173–180.

Veale, T., Hao, Y., 2007. Comprehending and generating apt metaphors: a web-driven, case-based approach to figurative language. In: AAAI, pp. 1471–1476.

Veale, T., Hao, Y., 2008. 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, Manchester, pp. 945–952.