Leveraging Eventive Information for Better Metaphor Detection and Classification

I-Hsuan Chen[1], Yunfei Long[2], Qin Lu[2], Chu-Ren Huang[1]
[1]Department of Chinese & Bilingual Studies, The Hong Kong Polytechnic University
{ihsuan.chen, churen.huang}@polyu.edu.hk
[2]Department of Computing, The Hong Kong Polytechnic University
{csylong, csluqin}@comp.polyu.edu.hk

Abstract

Metaphor detection has been both challenging and rewarding in natural language processing applications. This study offers a new approach based on eventive information in detecting metaphors by leveraging the Chinese writing system, which is a culturally bound ontological system organized according to the basic concepts represented by radicals. As such, the information represented is available in all Chinese text without pre-processing. Since metaphor detection is another culturally based conceptual representation, we hypothesize that sub-textual information can facilitate the identification and classification of the types of metaphoric events denoted in Chinese text. We propose a set of syntactic conditions crucial to event structures to improve the model based on the classification of radical groups. With the proposed syntactic conditions, the model achieves a performance of 0.8859 in terms of F-scores, making 1.7% of improvement than the same classifier with only Bag-of-word features. Results show that eventive information can improve the effectiveness of metaphor detection. Event information is rooted in every language, and thus this approach has a high potential to be applied to metaphor detection in other languages.

1 Introduction

Metaphors are a cross linguistic phenomenon in everyday language as shown in a great amount of corpus linguistic and experimental studies. The Conceptual Metaphor Theory (Lakoff, 1989; Lakoff and Johnson, 1981) shows how linguistic expressions reflect the mapping of two conceptual domains. For example, the expression *I see what you mean* instantiates the conceptual metaphor of KNOWING IS SEEING. The phrase is the result of mapping the source domain SEEING, which is embodied daily experience onto the target domain, KNOWING, as exemplified in the examples of *shed some light* on this, an *illuminating* article, and take a *close look*. Due to the pervasive use of metaphors, there is an enormous amount of studies in the techniques of detecting metaphors. Relevant studies of detecting metaphors primarily rely on contextual information.

This study provides a novel approach to detect and classify metaphors by analyzing eventive information. Concepts can be classified into a wide array of event types according to ontology, the organization of knowledge (Huang et al., 2007). Eventive information thus can be applied to the classification of metaphors, which concern mappings of conceptual structures from a source domain to a target domain.

The classification of metaphoric and literal senses has been approached by different methods such as vector-space models with distributional statistics (Hovy et al., 2013; Tsvetkov et al., 2014) and compositional distributional semantic models (CDSMs) (Kartsaklis and Sadrzadeh, 2013a). Most of the studies regarding metaphoric detection have been done in English, while the task in Chinese is at the incipient stage. The relevant studies such as clustering models and similarity computation in context (Fu et al., 2016; Wang, 2010) mainly focus on the metaphoric sense of each individual noun or adjectival phrase because the analyses are highly dependent on contextual information. However, metaphoric senses of verbs are less touched because it is difficult to define regularities of their contextual information. This study deals with the challenge of the verb category by including eventive information, which is the basis of the classification of metaphors.

Chinese is featured by its semantic-based orthography in the writing system. Specifically, Chinese characters are composed of radicals and components, which are ideographic or phonetic symbols. Radicals, which represent core conceptual properties, encode eventive information of the literal senses of characters (Huang 2009, Huang and Hsieh, 2015). For instance, the verb 踢 *ti* 'kick' contains the radical 足 'foot'; the verb 吃 *chi* 'eat' has the radical 口 'mouth'. The radicals clearly identify the body parts executing the actions. Chinese radicals, in particular, evoke the whole event structure such as the initiation, the process, and the termination of a kicking or an eating action. Also, radicals are good indicators of different types of events. For instance, radicals can encode the information of tools in the concept of separation. The radical 刀 *dao* 'knife' of the character 切 *qie* 'cut' implies that the action results in two pieces, while the radical 石 *shi* 'stone' of the character 破 *po* 'break' emphasizing that the action results into pieces. The radicals can thus provide detailed eventive information to identify the source domain in the task of metaphor detection.

Event information characterizes detailed properties such as the volition of the subject and the resulted status of the object. The properties can be accessed by their corresponding syntactic constructions. We propose 17 syntactic conditions which are appropriate to differentiate different event types. First, we implement the algorithm of metaphor detection based on a Support Vector Machine (SVM) classifier. The syntactic conditions serve as additional features using Bag-of-word features as the baseline. Second, we apply the SVM classifier to predict the senses, either literal or metaphoric, of each verb in *Baidu Baike* corpus, which has 1,543,669 million entries and 7.6 billion tokens.[1] We then measure the semantic similarities among different radical groups by the vector representation `according` to each sense of each character. The similarity of vectors based on word representation and sense representation proves that radicals can predict semantic groups of the literal senses. We delimit the syntactic environments where the literal senses tend to occur. When a sense does not occur in the defined set of syntactic conditions, it is highly possible to be metaphoric.

In this study, NLP technology is applied to two deeply culturally bound phenomena: (i) the Chinese writing system and (ii) the classification of metaphors. The Chinese character orthography is an ontological system organized based on the primitive concepts represented by radicals (Chou and Huang, 2010). Thus, the information represented by radicals is not only cultural specific but also available in all Chinese text without the need for processing. Metaphor detection, as another culturally based conceptual representation, has been proven to be both challenging and extremely valuable in natural language processing. Based on their shared event information, we hypothesize that sub-textual information can leverage the effectiveness to identify and classify different types of metaphoric events hidden in the Chinese text. Our experiments prove the effectiveness of eventive information in detecting metaphors. The approach of leveraging event type information by radicals increases both the precision and the recall in metaphor detection. Although this approach is especially effective for Chinese because of the information embedded in radicals, broader implications include the possibility of leveraging eventive information from different sources in other languages.

2 Related Work

The task of metaphor detection has been handled in a wide variety of approaches including clustering models (Birke and Sarkar, 2006; Shutova et al., 2010; Li and Sporleder, 2010), semantic similarity graphs (Sporleder and Li, 2009), topic modeling (Li et al., 2010; Heintz et al., 2013), and compositional distributional semantic models (CDSMs) (Gutiérrez et al. 2016). Feature-based classification, in particular, attracts most attention since a wide array of contextual information is included (Sporleder and Li, 2009; Dunn., 2013; Hovy et al., 2011; Mohler et al., 2013; Neuman et al., 2013; Tsvetkov et al., 2013; Tsvetkov et al., 2014). Since the studies regarding metaphor identification have primarily focused on English, there are more available datasets in English in both manually-tagged linguistic resources (Gedigian et al., 2006; Krishnakumaran and Zhu, 2007; Broadwell et al., 2013) and corpus-based approach (Birke and Sarker, 2007; Shutova et al., 2013; Neuman et al., 2013; Hovy et al., 2013). Metaphor detection in Chinese is at the incipient stage. Fu et al., (2016) uses hierarchical clustering

---

[1] https://en.wikipedia.org/wiki/Baidu_Baike

for Chinese noun phrases according to their contextual information to recognize metaphoric phrases. Zhou et al. (2011) use the Maximum Entrophy model to detect the metaphoric reading of verb phrases based on collocation with noun phrases, and point out that there is no mature syntactic and semantic tool for metaphor analysis in Chinese. Our study will close the gap by building a model of metaphor detection based on syntactic conditions.

Regarding metaphor detection, most papers emphasize on distinguishing metaphoric senses from literal senses in a polysemy network. Disambiguation of senses has been handled by DSMs based on the availability of contextual information (Baroni et al., 2014; Boleda et al., 2012; Erk and Padó, 2010; Kartsaklis and Sadrzadeh 2013). When more contextual information is incorporated, disambiguation would be more successful. It should be noted that the senses of one form have different degrees of transparency to be traced in semantics. The senses of a form which can be chained together via overlapping semantics, as in the case of polysemy (cut a new *window* in the wall vs. the ball broke a *window*), are more likely to be traced. On the contrary, when the senses of a linguistic form are discrete as in the case of *homonymy* (e.g. piano *keys* vs. *key* point), they may be problematic to DSM (Baroni et al., 2014). Gutiérrez et al. (2016) point out that the challenge arises from the highly context-dependent property of *homonymies* since the relations of senses are not unsystematic. In contrast, the senses of a polysemy form a systematic system, and thus CDSM has a better chance to detect metaphoric senses (Gutiérrez et al. 2016). Nevertheless, how to group a variety of senses including metonymic and metaphoric senses as a polysemy has been a challenge in Chinese (Fu et al., 2016). In this paper, the use of Chinese radicals for grouping senses can avoid the confusion of polysemy and homonym because Chinese radicals stand for semantic classification, reflecting the structure of our ontological knowledge structure (Huang 2009).

Contextual information has been regarded as an important determinant in identifying metaphors. Previous studies thus primarily focus on the adjectives or nouns as in the studies of English metaphors due to the abundant contextual information from these categories. This study, instead, focuses on the verb category and shows the literal and metaphoric senses of a verb can be predicted by their syntactic conditions. The event structure evoked by a verb offers reliable information for metaphor detection.

3 Methodology

Our task is to define the syntactic environments where the metaphoric sense of a verb would be more likely to occur. Each verb corresponds to a type of event structure. Chinese radicals denote the most profiled element in an event structure. For example, the literal sense of 灌 *guan* 'pour', which has a water radical 氵, specifies the material of this action is water. Based on the properties of water, the verb emphasizes dynamic flows. Thus the verb tends to appear in non-passive constructions for expressing the dynamics. The literal meaning of 墊 *dian* 'pad', which has the mud radical 土, profiles mud as a loctum, and therefore it tends to appear with a locative phrase in order to specify the object to be padded. The literal sense of 切 *qie* 'cut', which has a knife radical 刀, specifies the instrument of the separation. The verb occurs mostly in the VO word order, as in 切蛋糕 *qie diangao* 'cut cakes' to emphasize on transitivity. In summary, each verb has its own event structure, which can be observed in the syntactic environments where the verb frequently occurs. Since a metaphoric sense describes a concept different from that of a literal sense, it should have a different event structure from that of a literal sense. According to corpus data, it can be observed that the literal senses of a verb tend to occur under a set of syntactic conditions, while the metaphoric senses of the same verb tend to occur in the environments deviating from the standards. For instance, the metaphoric sense of 灌 *guan* 'pour' frequently appears in passive constructions, while the literal sense generally occurs in non-passive constructions. The metaphoric sense of 墊 *dian* 'pad' is more likely to occur without a locative phrase, whereas the literal senses normally occur with a locative phrase. The metaphoric sense of 切 *qie* 'cut' as in 'cannot cut the relationship' occurs more frequently in the OV word order, while the literal sense tends to occur in the VO word order. The change of event types is expected since the source domain and the target domain refer to different settings although their underlying conceptual structures are organized in a similar way. For instance, both the literal and metaphoric senses of

Radical	Sample Characters	Radical	Sample Characters
火 *huo* 'fire'	熬 *ao* 'simmer' 烤 *kao* 'grill'	糸 *mi* 'thread'	綁 *bang* 'tight' 織 *zhi* 'weave'
水 *shui* 'water'	灌 *guan* 'pour' 沖 *chong* 'flush'	力 *li* 'power'	動 *dong* 'move' 加 *jia* 'add'
土 *tu* 'mud'	墊 *dian* 'pad' 塞 *sai* 'pack'	扌 *shou* 'hand'	抱 *bao* 'hug' 推 *tui* 'push'
金 *jin* 'gold'	釘 *ding* 'pin' 鑽 *zuan* 'drill'	口 *kou* 'mouth'	吃 *chi* 'eat' 咬 *yao* 'bite'
石 *shi* 'stone'	砍 *kan* 'chop' 破 *po* 'break' 碰 *peng* 'clash'	辵/辶 *chuo* 'interval walk'	逃 *tiao* 'escape' 追 *zhui* 'chase'
刀 *dao* 'knife'	刷 *shua* 'brush' 切 *qie* 'cut'	足 *zu* 'foot'	跳 *tiao* 'jump' 踢 *ti* 'kick'
斤 *jin* 'ax'	斬 *zan* 'cut' 斷 *duan* 'snap'	走 *zou* 'walk'	走 *zou* 'walk' 趕 *gan* 'chase'

Table 1: Types of radicals and sample characters

切 *qie* 'cut' refer to the concept of separation which results in two entities, but the separation is employed to describe different contexts. The literal one refers to the separation of an entity with a specific instrument, while the metaphoric one refers to the discontinuation of a relationship. It is the change of event types that provides information of predicting which sense is in use.

3.1 Syntactic conditions and Radicals

Radicals: The advantages of radical-based analysis are the transparency and traceability of semantic relations among different senses in a polysemy network. The current experiments include 14 types of radicals as listed in Table 1. Each type of radicals has two to three verbs which have high frequency in *Chinese Gigaword* (Huang 2009) as the representatives.

Syntactic conditions: We hypothesize that the literal senses of a verb tend to appear in a set of syntactic conditions whereas the metaphoric senses tend to deviate from those conditions due to the

change of event types. To test this hypothesis, we propose a variety of syntactic conditions to characterize each sense and its relevant event structure. The conditions are selected based on the frequency of where the literal senses of these verbs occur.

(i)   Word order (VO): If a verb can take an object, the verb and its object may occur in either VO or OV word order.

(ii)   Compounding (VV): The verb may form a compound with another verb in VV form. The target verb is the second one.

(iii)   Transitivity (Vt): The verb may be transitive or intransitive.

(iv)   Passivity (Pass): The verb may occur in a passive construction. The indicators are the occurrences of passive markers.

(v)   Disposal constructions (Disposal): The verb may occur with the disposal markers to foreground the semantic patient or the direct object.

(vi)   Aspectual markers (Asp): The verb may appear with aspectual markers to specify the status of the process.

(vii)   Double-object construction (DO): The verb may take both a direct object and an indirect object.

(viii)   Relative clauses (RC): The verbs may occur with a relative clause. This feature is indicated by the markers of a relative clause.

(ix)   Numeral phrases (Num): Amounts relevant to the event are specified by numeral-classifier phrases.

(x)   Locative phrases (Loc): Location of the event is specified. The locative phrase can occur either before or after the verb.

(xi)   Negation (Neg): Negative markers appear in the main clause which contains

(xii)   Postpositions (Post): The verb may take a postposition phrase.

(xiii)   Prepositions (Prep): The verb may occur with a preposition phrase. . The indicators are the occurrences of a variety of prepositions.

(xiv)   Instrumental 用 *yong* 'use' (yong): The instruments are profiled.

(xv)   對 *dui* 'to/ toward' (dui): The goal of the verb is profiled by this marker.

(xvi)   Beneficiary/ maleficent marker 給 *gei* (gei): The affectiveness of the event relevant to the target verb is specified.
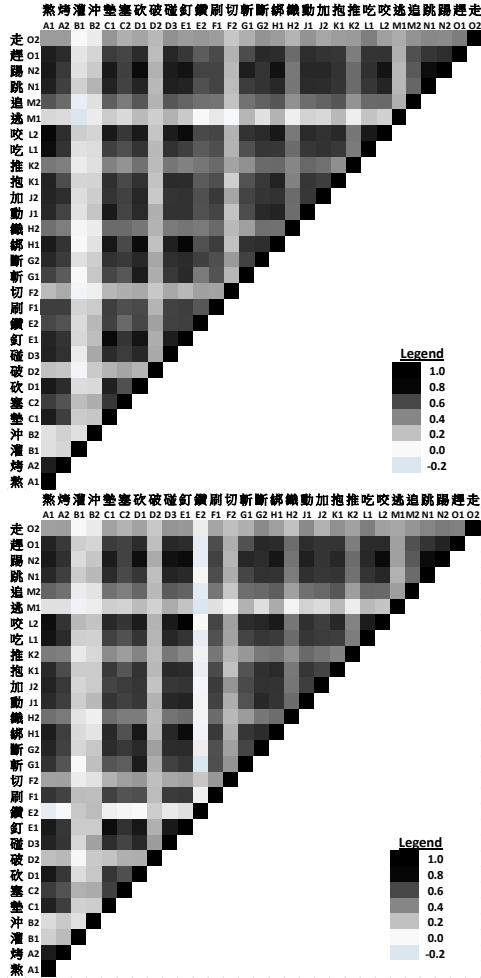
Figure 1: Semantic closeness among different verbs [upper graph: literal sense; lower graph: metaphoric sense]

(xvii) Postverbal adverbs (Vadv): The verb may be followed by an adverb which specifies degrees or durations of time.

3.2 Design of classification model

In order to evaluate the effectiveness of our proposed syntactic conditions, we have to extract the syntactic features which are relevant to the literal and metaphoric senses. The task of detecting the metaphor/literal senses is modeled as a binary class classification. The proposed syntactic conditions are implemented as additional features in this model.

SVM are well performed in higher dimension, particularly when targeted instances only hold a small portion in a dataset. Since our design focuses on the effectiveness of syntactic conditions in metaphor detection rather than on a classifier, we choose SVM with linear kernel as our classifier

for its linear binary classification and use LibSVM (Chang and Lin, 2011) as the SVM tool.

3.3 Word embedding for word similarity

Word embedding is known as a special form of word vectors which represents a word through a low dimensional dense vector and has been used in different lexical tasks, such as semantic similarity, word analogy, word synonym detection, and concept categorization (Baroni, 2014), (Levy, 2015). Our goal is to increase the precision of metaphor detection with the aid of the semantic classification of radicals. Thus we conduct word embedding to show how different concepts are categorized in terms of their semantic similarities. Based on the similarity from word embedding, we can infer semantic distance among verbs with different radicals and further quantify the differences between the metaphoric and literal senses of the same verb.

Various models are proposed to learn the dense vector representation of words, which are all based on the distributional hypothesis that words occur in similar context have similar meanings (Harris, 1954). Among those models, the most widely used one is the Skip-Gram model with negative sampling (Mikolov, 2013). In our task, word embedding is trained through the Skip-Gram model with default parameters on the *Baidu Baike* corpus[2] with word segmentation performed by the HIT LTP tool[3].

Since Chinese radicals encode semantic categorizations, verbs which belong to the same radical group are expected to be close semantically. In order to capture the predictive power of radicals in semantics, we use multi-dimensional vector space to show the distribution of verbs when they are used in their literal senses and metaphoric senses respectively (Baroni, 2014, Levy, 2015). First, we use our proposed classifier to predict the senses of 29 selected verbs, and treat metaphor /literal sense of each word as an individual word. And we calculate the cosine similarity between different senses. Figure 1 shows that verbs having the same radical are relatively similar to each other compared to verbs which belong to different radical groups. However, the grouping by radicals does not work well in the metaphoric senses, as shown in the lower graph. The sharp contrast supports the claim that the metaphoric senses of a verb have a
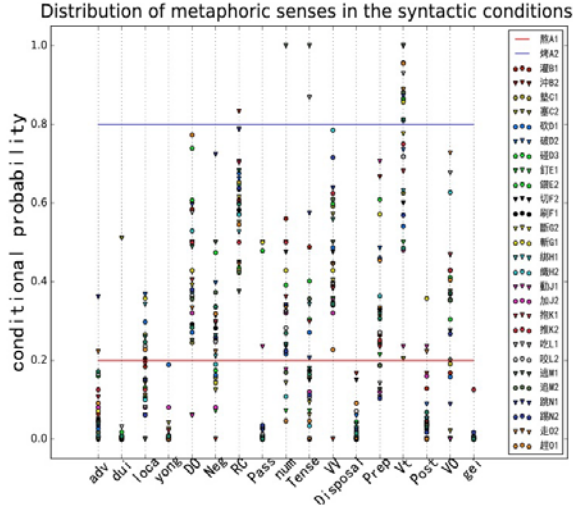
Figure 2: Probability of metaphoric senses in each syntactic condition

different event structure from that of the literal senses.

4 Experiments

Experiments of feature analysis are conducted to show whether our proposed syntactic conditions can improve the model of metaphor detection.

4.1 Dataset

The dataset is structured based on the 29 verbs from 14 radical groups introduced in Section 3.1. For each verb, a random sample of 200-300 sentences are collected from the *Chinese Gigaword* corpus (Huang, 2009), a comprehensive archive of newswire text data. Two Chinese native speakers manually annotated the metaphoric and literal senses of each token based on *Hantology* (Chou and Huang, 2006), a character-based Chinese language resource in which each character is sense-tagged. In the 6,047 tokens, 1,738 of them are labeled as a metaphoric sense and 4,309 are labeled as a literal sense. Our annotation task has kappa statistics (Banerjee, 1999) over 0.81 indicating strong inter-annotator consistency.

4.2 Model and analysis

We evaluate the 17 syntactic conditions using the SVM classification model in the dataset introduced in Section 4.1. In order to avoid overfitting, we perform 10-fold cross validation. To test the efficiency of our proposed syntactic conditions, the 17 conditions are divided into 3 feature groups.

Type	Precision	Recall	F score
Basic	0.8824	0.8559	0.8689
All features	0.8952	0.8768	0.8859
Feature group 1	**0.8925**	**0.8821**	**0.8872**
Feature group 2	0.8752	0.8631	0.8691
Feature group 3	0.8705	0.8521	0.8612

Table 2: Performance in different condition groups

- **Base group**: Using Bag-of-word features only
- **Group1:** transitivity (Vt), numeral phrases (Num), relative clauses (RC), compounding (VV), tense, word order (VO), and double-object construction (DO).
- **Group 2:** negation (Neg), prepositions (Prep), locative phrases (Loc), postverbal adverbial (Vadv), passivity (Pass), and aspectual markers (Asp).
- **Group 3:** disposal constructions (Disposal), postpositions (Post), instrumental 用 *yong* 'use' (yong), 對 *dui* 'to/ toward' (dui), and beneficiary/maleficient marker 給 *gei* (gei).

The three groups are defined based on two principles: (i) the probability of the occurrence of the metaphoric senses in the syntactic condition in question; (ii) the clusters of the verbs. As shown in Figure 2, the metaphoric senses frequently occur in a few syntactic conditions, such as Vt, VO, and relative clauses. Regarding the principle of the clusters, the condition which has less overlapping data points is more effective in distinguishing different senses.

The results of the experiment given in Table 2 show that the proposed syntactic conditions have improved the performance of the model. The incorporation of all the 17 features does improve the classification model by 1.70% in F-score. However, Group 1 has the best performance, outperforming the result when all the 17 features are used. However, when Group 2 and Group 3 are used alone, they do not contribute to improving the model. In fact, Group 3 decreases the effectiveness of the model. The decrease in performance is on both Precision and Recall. However, while the model incorporates Group 1, the precision is improved at the expense of a slight decrease in recall. This increase in precision indicates that the features of Group 2 and Group 3 still provide useful information in metaphor detection.
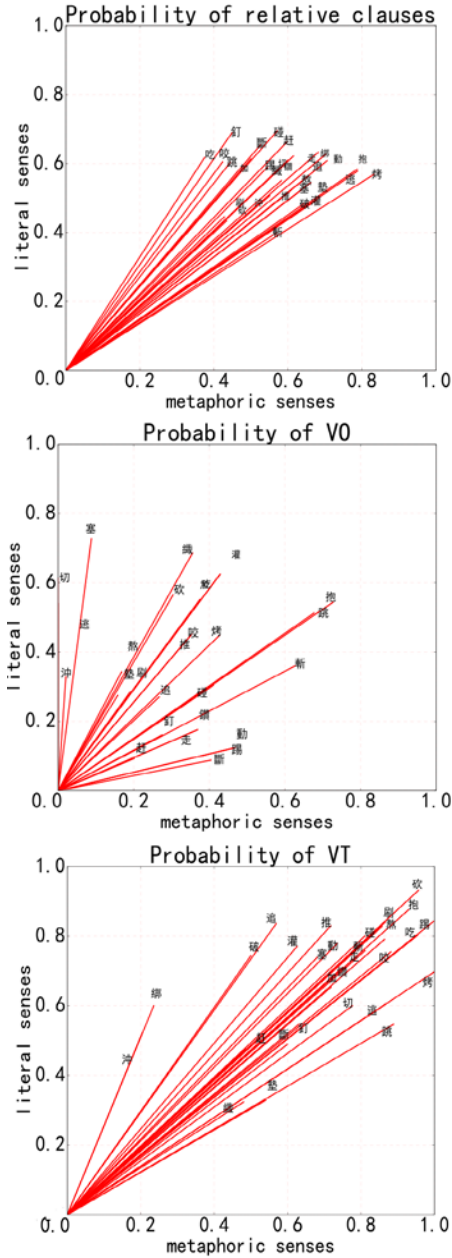
Figure 3: Distribution of example literal-metaphoric pair of verbs under individual syntactic conditions.

Figure 4: Examples of metaphoric and literal senses of verbs characterized by a core set of syntactic conditions.

5 Discussion

Our experimental result shows that the proposed syntactic conditions can predict where the literal and metaphoric senses of the same verb occur. This is because the two senses tend to be used in different event structure. For example, the literal sense of the verb 灌 *guan* 'pour' as in 灌良田 *guan liang tian* 'irrigating good farms' specifies the location right after the verb, while the metaphoric sense as in 灌水 *guan shui* pour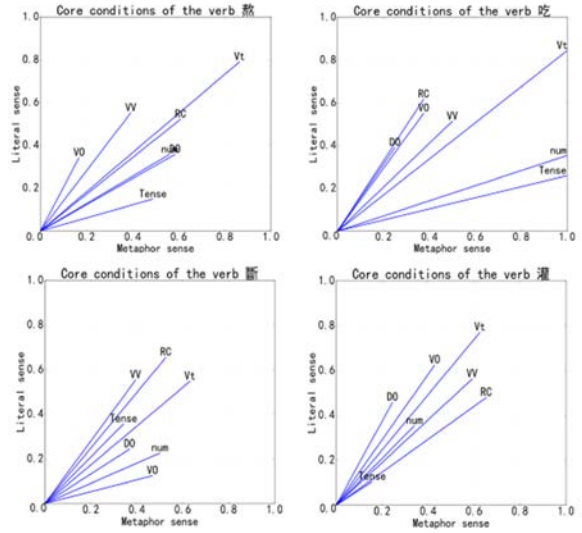 water 'artificially increasing the amount' has water as the material without specifying the location. The literal sense of 走 *zhou* 'walk' appears as an intransitive verb as in 他走了 *ta zhou le* he-walk-ASP 'he left', while the metaphoric sense tends to have a noun phrase following it as in 走好運 *zhou hao yun* walk-luck 'being lucky'. Since the metaphoric sense describes an event different from that of the literal sense, the syntactic properties of the metaphoric sense should differ from those of the literal sense. Among our proposed syntactic conditions, seven of them, transitivity, relative clauses, double objects, compounds, word order, aspectual markers, and numeral phrases, are the most effective conditions in detecting metaphors. Figure 3 shows examples from these syntactic conditions including transitivity, word order, and relative clauses. The horizontal axis shows conditional probabilities in metaphoric sense. The vertical axis shows the conditional probabilities in literal sense. A condition with a stronger predictive power has a bigger difference in the probability between the literal and metaphoric senses. For example, the literal sense and metaphoric sense of the verb 切 *qie* 'cut' have saliently different probabilities in the feature of word order. The literal sense is frequently found in the VO word order, while the metaphoric sense seldom occurs in the VO word order. It is the difference that can serve to predict which sense, literal or metaphoric, is in use.

Each syntactic condition is regarded as a measurement. The syntactic conditions then can be grouped to precisely identify the event types of the literal and metaphoric senses for each verb as

Category	Radicals
Materials	火 *huo* 'fire' 水 *shui* 'water' 土 *tu* 'mud'
Body parts	扌 *shou* 'hand' 口 *kou* 'mouth'
Instruments	走 *zou* 'walk' 辵/辶 *chuo* 'interval walk' 足 *zu* 'foot' 力 *li* 'power'
Movements	石 *shi* 'stone' 刀 *dao* 'knife' 斤 *jin* 'ax' 糸 *mi* 'thread' 金 *jin* 'gold'

Table 3: Higher-level Ontological Categories of Radicals

shown in the examples of Figure 4. The horizontal axis shows conditional probabilities in metaphoric sense. The vertical axis shows the conditional probabilities in literal sense. Each condition has different relevancy to a verb because each verb belongs to a different event type. For example, the condition of word order (labeled as VO) has higher effectiveness in the verb 吃 *chi* 'eat' than in the verb 斷 *duan* 'snap'. In other words, the senses of each verb can be identified by the most relevant syntactic conditions. Therefore, the syntactic environments of where a verb occurs can be used to predict whether it is metaphoric.

Furthermore, grouping verbs by Chinese radicals can offer generalizations of the event types associated with a particular semantic group. A group of relevant radicals denote a higher-level category in the ontological structure, which refers to the organization of knowledge structure and the representation of knowledge system in terms of relations between concepts (Prévot et al., 2010). For example, the radicals discussed in this paper can be classified into four larger semantic categories, which are instruments, body parts, materials, movements, as given in Table 3.

The differences in the distribution of the literal and metaphoric senses of the four semantic groups can be characterized by the rankings of the syntactic conditions. As shown in Figure 5, the group of the material radicals and the group of the movement radicals have different arrangement of the
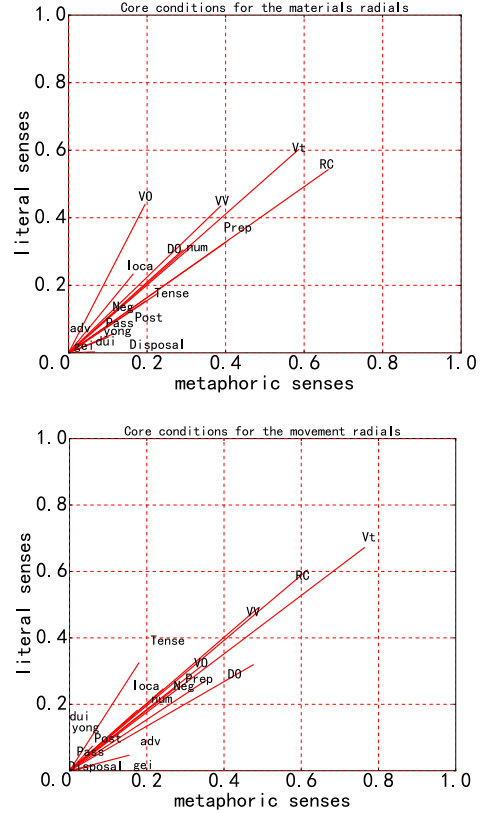

Figure 5: Examples of the syntactic conditions characterizing a higher-level semantic category

conditions. In other words, the literal sense of a larger semantic group can also be identified by its syntactic distribution. When a verb belonging to a larger semantic group does not occur in the set of syntactic conditions where the literal senses generally occur, it is highly possible to be metaphoric. Our design shows that syntactic conditions can offer informative clues in detecting metaphoric senses based on the fact that each sense of a verb has its own preferred syntactic environments.

The syntactic conditions can be further classified based on their effectiveness. As discussed in Section 4.2, the syntactic conditions of Group 1 in our model, transitivity (Vt), numeral phrases (Num), relative clauses (RC), compounding (VV), word order (VO), double-object construction (DO), are proven to be more efficient. The effectiveness of the conditions reflects three generalizations of where metaphoric senses tend to occur. First, a sense tends to be non-metaphoric when a numeral phrase is involved. The involvement of numeral phrases specifies the exact numbers of the object. Since the object has concrete details, the verb is more likely to be a literal. Second, a metaphoric sense is generally used to modify a concept. Due to this modification property, meta-

43

phoric senses tend to occur when there is a presence of a relative clause a relative clause, which serves the purpose of modification. Third, due to the changes of event types, the inherent properties of a verb are likely to change. More specifically, the transitivity of a verb changes when the verb is used in its metaphoric sense. For example, when a transitive verb becomes intransitive, the verb is likely not to be in its literal sense. Regarding the occurrence of compounding, the addition of another verb provides additional information and thus creates an event structure which differs from the original one. Similarly, when a verb which does not have two objects in its argument structure appears in the double-object construction, it is a sign of changing event types because the additional object cannot be accommodated in the original event structure. As for word order, it is associated with the information structure, which is a key component of an event structure. The change of word order therefore indicates the change of an event structure. Since each of the syntactic conditions links to a particular aspect of a conceptual event, its change is an informative indicator of which sense, literal or metaphoric, is in use.

On the other hand, the conditions in Group 3 do not contribute much to detecting senses. Although they provide additional information, the information is proven to be peripheral in indicating changes of event types. In brief, our experiments can successfully rank the relevancy of syntactic conditions with event types. The syntactic conditions which are related to the core elements of an event structure can improve the model of detecting metaphors.

6 Conclusion

This study offers an effective and precise way of detecting metaphoric and literal senses by including eventive information encoded in radicals. A set of syntactic conditions core to the event structure of a verb can define where its literal senses tend to occur. When a verb appears in the environments deviating from the defined set, it has a higher chance to be metaphoric. Instead of focusing on individual lexemes, we offer larger generalizations by event types encoded by radicals. Event types correspond to larger conceptual categories. Thus verbs of the same group have similar syntactic distribution. The generalizations can increase the efficiency of the model for metaphor detection.

Our study shows that other eventive information parsed in the existing platforms such as WordNet, FrameNet, and Tongyici Cilin should also have a high potential to be leveraged in the detecting of metaphors. The tools relevant to eventive information such as aspectual markers and word order can be applied to determine event types. This new approach refocuses metaphor detection in the inherent eventive information of metaphors instead of its contextual information, and thus it is more reliable. Our algorithm of modeling eventive information can provide a pathway to incorporate analysis of event types in deep learning as future studies.

In summary, our study show that by leveraging the Chinese writing system, culturally bound eventive information can facilitate processing of metaphor. This method is not only applicable to all Sinitic languages and a small sub-set of languages sharing Chinese orthography as their cultural heritage, such as Japanese and Korean. Huang and Chou (2015) already showed that lexical processing in Japanese and others based on Chinese orthography can be automatically bootstrapped. This study suggests the potential applications for the use of eventive information to conceptual processing such as automatic classification of metaphor. Eventive information in many languages can be automatically or semi-automatically extracted through the OntoLex interface approach (Huang et al. 2010). Eventive information in turn will be a powerful tool in the extraction of event types for studies based on eventive structures such as sarcasm and sentiment detection.

References

Marco Baroni, Raffaela Bernardi, and Roberto Zamparelli. 2014. Frege in space: A program of compositional distributional semantics. *Linguistic Issues in Language Technology*, 9.

Marco Baroni, Georgiana Dinu, and German´ Kruszewski. 2014. Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, pages 238-247. https:doi.org/10.3115/v1/P14-1023

George Aaron Broadwell, Umit Boz, Ignacio Cases, Tomek Strzalkowski, Laurie Feldman, Sarah Taylor, Samira Shaikh, Ting Liu, Kit Cho, and Nick Webb. 2013. Using imageability and topic chaining to locate metaphors in linguistic corpora. In *Social Computing, Behavioral-Cultural Modeling and Prediction.* Springer, pages 102-110.

Julia Birke and Anoop Sarkar. 2007. A clustering approach for nearly unsupervised recognition of non-literal language. In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, pages 329-336.

Gemma Boleda, Eva Maria Vecchi, Miquel Cornudella, and Louise McNally. 2012. First-order vs. higher order modification in distributional semantics. In *Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Association for Computational Linguistics, pages 1223-1233. http://www.aclweb.org/anthology/D12-1112

Yaming Chou and Chu-ren Huang. 2006. Hantology - A linguistic resource for Chinese language processing and studying. Paper presented at the *5th International Conference on Language Resources and Evaluation*, Genoa, Italy.

Ya-Min Chou and Chu-Ren Huang. 2010. Hantology: conceptual system discovery based on orthographic convention. In *Ontology and the Lexicon: A Natural Language Processing Perspective*. Cambridge University Press, page 122-143.

Jonathan Dunn. 2013. Evaluating the premises and results of four metaphor identification systems. In *Computational Linguistics and Intelligent Text Processing*. Springer, pages 471-486.

Katrin Erk and Sebastian Padó. 2010. Exemplar-based models for word meaning in context. In *Proceedings of the ACL 2010 Conference Short Papers*. Association for Computational Linguistics, pages 92-97.

Jianhui Fu, Shi Wang, Ya Wang, Cungen Cao. 2016. A Practical Method of Identifying Chinese Metaphor Phrases from Corpus. In *International Conference on Knowledge Science, Engineering and Management*. Springer, pages 43-54.

Matt Gedigian, John Bryant, Srini Narayanan, and Branimir Ciric. 2006. Catching metaphors. In *Proceedings of the Third Workshop on Scalable Natural Language Understanding*. Association for Computational Linguistics, pages 41-48.

Gutiérrez Darıo, Ekaterina Shutova, Tyler Marghetis, and Benjamin Bergen. 2016. Literal and metaphorical senses in compositional distributional semantic models. In *Proceedings of the 54th Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, pages 160-170.

Harris, Zellig, 1954. Distributional structure. *Word*, 10(23):146-162.

Ilana Heintz, Ryan Gabbard, Mahesh Srinivasan, David Barner, Donald S Black, Marjorie Freedman, and Ralph Weischedel. 2013. Automatic extraction of linguistic metaphor with lda topic modeling. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 58-66.

Dirk Hovy, Shashank Srivastava, Sujay Kumar Jauhar, Mrinmaya Sachan, Kartik Goyal, Huiying Li, Whitney Sanders, and Eduard Hovy. 2013. Identifying metaphorical word use with tree kernels. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 5257.

Chu-Ren Huang. 2009. Tagged Chinese Gigaword Version 2.0, LDC2009T14. Linguistic Data Consortium.

Chu-Ren Huang. 2009. Semantics as an Orthography-Relevant Level for Mandarin Chinese. In *The 17th Annual Conference of the International Association of Chinese Linguistics*.

Chu-Ren Huang, Siaw-Fong Chung, and Kathleen Ahrens, 2007. An ontology-based exploration of knowledge systems for metaphor. In *Ontologies* Springer, page 489-517.

Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, and Laurent Prévot. 2010 (Eds.) *Ontology and the lexicon: A natural language processing perspective*. Cambridge: Cambridge University Press,

Chu-Ren Huang and Chou Ya-Min. 2015. Multilingual conceptual access to lexicon based on shared orthography: an ontology-driven study of Chinese and Japanese. In *Language Production, Cognition, and the Lexicon,* ed. Núria Gala, Reinhard Rapp, and Gemma Bel-Enguix. Springer, pages 135-150.

Chu-Ren Huang and Shu-Kai Hsieh. 2015. Chinese lexical semantics: from radicals to event structure. In *The Oxford Handbook of Chinese Linguistics*, ed. William S.-Y. Wang and Chao-Fen Sun. Oxford University Press, pages 290-305.

Dimitri Kartsaklis and Mehrnoosh Sadrzadeh. 2013. Prior disambiguation of word tensors for constructing sentence vectors. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1590-1601.

Dimitri Kartsaklis, Mehrnoosh Sadrzadeh, and Stephen Pulman. 2013. Separating disambiguation from composition in distributional semantics. In *Proceedings of the 2013 Conference on Computational Natural Language Learning*, pages 114-123.

Saisuresh Krishnakumaran and Xiaojin Zhu. 2007. Hunting elusive metaphors using lexical resources. In *Proceedings of the Workshop on Computational approaches to Figurative Language*. Association for Computational Linguistics, pages 13-20.

George Lakoff and Mark Johnson. 1981. *Metaphors we live by*. University of Chicago Press, Chicago, IL.

George Lakoff. 1989. Some empirical results about the nature of concepts. *Mind & Language*, 4(1-2): 103-109.

Prévot, Laurent, Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, and Alessandro Oltramari. 2010. Ontology and the lexicon: A multi-disciplinary perspective. In *Ontology and the lexicon: A natural language processing perspective*, eds. Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, and Laurent Prévot. Cambridge: Cambridge University Press, pages 3-24.

Omer Levy, Yoav Goldberg, Ido Dagan, and Israel Ramat-Gan. 2015. Improving distributional similarity with lessons learned from word embeddings. *Transactions of the Association for Computational Linguistics*, 3.

Linlin Li, Benjamin Roth, and Caroline Sporleder. 2010. Topic models for word sense disambiguation and token-based idiom detection. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, pages 1138-1147.

Linlin Li and Caroline Sporleder. 2010. 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, pages 297-300.

Michael Mohler, David Bracewell, David Hinote, and Marc Tomlinson. 2013. Semantic signatures for example-based linguistic metaphor detection. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 27–35.

Yair Neuman, Dan Assaf, Yohai Cohen, Mark Last, Shlomo Argamon, Newton Howard, and Ophir

Frieder. 2013. Metaphor identification in large texts corpora. *PLoS ONE*, 8:e62343.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of NIPS*, pages 3111-3119.

Yulia Tsvetkov, Leonid Boytsov, Anatole Gershman, Eric Nyberg, and Chris Dyer. 2014. Metaphor detection with cross-lingual model transfer. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, pages 248-258

Ekaterina Shutova. 2010. Models of metaphor in NLP. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, pages 688-697.

Ekaterina Shutova, Lin Sun, and Anna Korhonen. 2010. Metaphor identification using verb and noun clustering. In *Proceedings of the 23rd International Conference on Computational Linguistics*. Association for Computational Linguistics, pages 1002-1010.

Ekaterina Shutova, Simone Teufel, and Anna Korhonen. 2013. Statistical metaphor processing. *Computational Linguistics*, 39(2):301-353.

Caroline Sporleder and Linlin Li. 2009. Unsupervised recognition of literal and non-literal use of idiomatic expressions. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 754-762.

Tomek Strzalkowski, George A. Broadwell, Sarah Taylor, Laurie Feldman, Boris Yamrom, Samira Shaikh, Ting Liu, Kit Cho, Umit Boz, Ignacio Cases, and Kyle Elliot. 2013. Robust extraction of metaphors from novel data. In *Proceedings of the First Workshop on Metaphor in NLP*. Association for Computational Linguistics, pages 67-76.

Yulia Tsvetkov, Elena Mukomel, and Anatole Gershman. 2013. Cross-lingual metaphor detection using common semantic features. In *Proceedings of the First Workshop on Metaphor in NLP*. Association for Computational Linguistics, pages 45-51.

Yulia Tsvetkov, Leonid Boytsov, Anatole Gershman, Eric Nyberg, and Chris Dyer. 2014. Metaphor detection with cross-lingual model transfer. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, page 248-258. http://www.aclweb.org/anthology/P14-1024

Zhao Hongyan, Qu Weiguang, Zhang Fen, and Zhou Junsheng. 2011. Chinese verb metaphor recognition based on machine learning and semantic knowledge. *Journal of Nanjing Normal University (Engineering and Technology)* 11(3):59-64.