

# Employing Morphological Structures and Sememes for Chinese Event Extraction

*Li Pei Feng and ZHOU Guo Dong\**

School of Computer Science & Technology, Soochow University, Suzhou, China, 215006

{pfli, gdzhou}@suda.edu.cn

## ABSTRACT

Current Chinese event extraction systems suffer much from the low recall due to unknown triggers. To resolve this problem, this paper firstly introduces morphological structures to better represent the compositional semantics inside Chinese triggers and then proposes a mechanism to automatically identify the head morpheme (either verb or noun) as the governing sememe of a trigger. Finally, it proposes a mechanism of combining the morphological structures and sememes of Chinese words to infer unknown triggers to improve the recall of the Chinese event extraction system. Evaluation on the ACE 2005 Chinese corpus justifies the effectiveness of our approach over a state-of-the-art system.

## 形态结构和义原在中文事件抽取中的应用

由于存在大量未知的触发词，当前的中文事件抽取系统受限于它的低召回率。为了解决这个问题，本文首先引入形态结构来更好地表示隐含在中文触发词内部的组合语义，然后提出了一个自动识别触发词中作为支配义原的核心词素（动词或名词）的机制。最后，本文提出了一个结合了中文词语的形态结构和义原去推测未知触发词的方法，用于提高中文事件抽取系统的召回率。在ACE 2005中文语料上的实验验证了我们方法的性能超越了目前最好的中文事件抽取系统。

---

**KEYWORDS:** Chinese event extraction; Morphological structure; Governing sememe; Trigger identification; Head morpheme.

**Keywords in L2:** 中文事件抽取; 形态结构; 支配语义; 触发词识别; 核心语素

---

---

\* Corresponding author

## 1 Introduction

As a compromise to natural language understanding, Information Extraction (IE) aims to extract structured information (e.g., entities, relations and events) from a text. Event extraction, a classic subtask in IE, is to recognize event trigger mentions of a predefined event type and their participants and attributes. While most studies in the literature focus on English event extraction, there are few successful stories concerning Chinese event extraction due to the special characteristics and challenges in Chinese language. Even with ground truth entities, times and values, the performance of most Chinese event extraction systems is much lower than that of English ones.

For Chinese event extraction, unknown triggers (a trigger in the test set doesn't occur in the training set and otherwise, a known trigger.) and word segmentation errors are two major reasons for the low performance, particularly the recall. The statistics on the ACE 2005 Chinese and English corpora (Li et al., 2012) shows that these two cases cover almost 30% of Chinese trigger mentions while this figure reduces to only about 9% in English. Besides, given the same number of event mentions, there are about 30% more different triggers in Chinese than those in English. This amplifies the problem. Therefore, trigger identification becomes a key to the success of Chinese event extraction.

Currently, there mainly exist two major mechanisms to solve this problem. The first one is to expand the triggers using predefined or automatically-clustered synsets, a common mechanism widely used in various NLP applications. The problem with this mechanism is that it fails to consider the sense shifting of a word in difficult contexts and thus may introduce too many pseudo triggers and harm the precision. This largely limits the contribution of this mechanism (Chen and Ji, 2009b; Ji, 2009; Qin et al. 2010). For example, as a trigger of the *Start-position* event, “担任” has more than five senses (e.g., serve as, bear, engage, do, etc.) and only one of them (serve as) can trigger a *Start-position* event. Take following two sentences as samples:

(E1) 我们将**承担**所有本公司的费用。 (*We will bear all the expenses for our company.*)

(E2) 他将在IBM**从事**科学研究工作。 (*He will engage in scientific research in IBM.*)

Although “承担” (bear) and “从事” (engage) are two synonyms of “担任”, they do not trigger the *Start-position* event but any other events.

The second one is to expand the triggers using the compositional semantics inside Chinese words. The intuition is that if a Chinese word contains more than one character, and its meaning can be often inferred from the meanings of its component characters (Yuan, 1998). For example, Li et al. (2012) infer the semantics of a verb (most triggers in Chinese events are verbs) from its basic single-character verb (BV) and significantly improve the F1-measure, largely due to the dramatic increase in the recall. The problem with Li et al. (2012) is that they extract all single-character verbs contained in triggers as BVs (e.g., “担” (undertake, verb) and “任” (serve as, verb) are treated as two BVs for “担任” (serve as)). Therefore, pseudo triggers are much introduced. This severely harms the precision. Take the following sentence as a sample:

(E3) 所有的公司员工**信任**他们的董事长。 (*All employees trust their chairman.*)

Although “信任” (trust) and “担任” have the same BV (“任”) and the same verb structure (verb+BV), “信任” (trust) does not trigger the *Start-position* event but any other events.

Further analysis indicates that above two mechanisms are quite complementary. For example, we can find out that if we introduce the semantic similarity into the compositional semantics, “信任”(trust) in (E3) will not be expanded as a trigger for the *Start-position* event because of its different sense from “担任”(serve as), while if we introduce the compositional semantics into the semantic similarity, “从事”(engage) in (E2) will be filtered out from the trigger list of the *Start-position* event since it doesn’t have the same BV as “担任”(serve as). However, a more refined mechanism is required to filter out “承担”(bear) in (E1).

In this paper, we first introduce the more general morphological structures in Chinese triggers, in place of verb structures in Li et al. (2102), to better represent the compositional semantics inside Chinese words and then propose a mechanism to automatically identify the head morpheme (either verb or noun) as the governing sememe of a trigger based on its morphological structure. The intuition behind is that the head morpheme can better represent the semantics of a Chinese word than the combination of all its component BVs. Finally, we propose a mechanism of combining the morphological structures and sememes of Chinese words to infer unknown triggers. Evaluation on the ACE 2005 Chinese corpus justifies the appropriateness of our approach.

To better understand the Chinese event extraction task as defined in ACE evaluations, where an event is defined as a specific occurrence involving participants, we list some ACE terminologies:

- **Event mention:** a phrase or sentence within which an event is described;
- **Trigger:** the main word that most clearly expresses the occurrence of an event, so recognizing an event can be recast as identifying a corresponding trigger;
- **Trigger mention:** a reference to a trigger.
- **Trigger type/Event type:** the type of an event;
- **Argument:** the entity mentions involved in an event;
- **Argument role:** the relation of an argument to an event where it participates.

In particular, the event extraction task is divided into four components:

- **Trigger identification:** to distinguish true trigger mentions from pseudo trigger mentions;
- **Event type determination:** to classify trigger mentions by event types;
- **Argument identification:** to distinguish true arguments from pseudo arguments;
- **Argument role determination:** to classify arguments by argument roles.

The rest of this paper is organized as follows. Section 2 overviews the related work. Section 3 describes various morphological structures in Chinese words and proposes a mechanism for determining the morphological structure and head morpheme in a Chinese trigger. Section 4 proposes an algorithm to infer unknown triggers on their morphological structures and sememes. Section 5 presents the experimental results. Finally, we conclude the paper with future work.

## 2 Related work

In the literature, most of existing studies on event extraction concern English and can be classified into either pattern-based (e.g., Riloff, 1996; Yangarber et al., 2000; Stevenson and Greenwood, 2005; Shinyama and Sekine, 2006; Patwardhan and Riloff, 2007; Chambers and

Jurafsky, 2011) or classifier-based (e.g., Grishman et al., 2005; Ahn, 2006; Hardy et al., 2006; Maslennikov and Chua, 2007; Ji and Grishman, 2008; Patwardhan and Riloff, 2009; Liao and Grishman, 2010; & 2011; Hong et al., 2011; Lu and Roth, 2012; Llorens et al., 2012). In particular, while earlier studies focus on sentence-level extraction, later ones turn to employ global information.

Compared with tremendous work on English event extraction, there are only a few studies on Chinese event extraction with focus on either feature engineering or trigger expansion, under the same framework as English event extraction.

On feature engineering, Tan et al. (2008) first employ a local feature selection method to ensure the performance of trigger classification and then apply multiple levels of patterns to improve the coverage in argument classification. Fu et al. (2010) apply a feature weighting scheme to re-weight various features for trigger identification and event type determination. Chen and Ji (2009b) apply various kinds of lexical, syntactic and semantic features to address the special issues in Chinese. Li et al. (2012) extend Chen and Ji (2009b) with more refined features and additional dependency and semantic role features.

On trigger expansion, Chen and Ji (2009a) propose a bootstrapping framework to exploit extra information captured by an English event extraction system. Ji (2009) first extracts some cross-lingual predicate clusters using bilingual parallel corpora and a cross-lingual information extraction system, and then employs the derived clusters to expand the triggers. Qin et al. (2010) employ a semantic dictionary “Tong Yi Ci Ciling (expanded version)” to expand triggers for Chinese event type determination. Li et al. (2012) propose an inference mechanism to infer new triggers by employing the verb structures to explore the compositional semantics inside Chinese triggers (verbs only) and achieve the state-of-the-art performance of 67.4% in F1-measure on the ACE 2005 Chinese corpus, ignoring the post-processing – discourse consistency.

### **3 Morphological structures and head morphemes inside Chinese triggers**

In this section, we introduce various morphological structures to better represent the compositional semantics inside Chinese triggers and then propose two mechanisms to identify the morphological structures and the head morpheme in Chinese triggers respectively.

#### **3.1 Compositional semantics and morphological structures in Chinese words**

Both in English and Chinese languages, a word is composed of one or more characters. However, a component character in English is just the basic unit to form a word instead of a semantic unit. In comparison, almost all Chinese characters have their own meanings and are called morpheme (or single-morpheme word), the minimal meaningful unit in Chinese language. If a Chinese word contains more than one character, its meaning can be often interpreted in terms of its composite characters/morphemes. This more fine-grained semantics are the compositional semantics inside Chinese words namely. Actually, it is also a normal way to understand a new Chinese word in everyday life for a Chinese native speaker.

Without doubt, a general method to represent the compositional semantics inside Chinese words is to systematically explore the morphological structures in Chinese words since it is the nature of compound words. Morphological structures in Chinese words are the word-building process to form the morphemes into words and are formulated by three major processes: compounding, affixation, and conversion. Compounding is a process, by which two or more morphemes are composed together to form a compound word. Affixation is a morphological process to add

grammatical or lexical information to a base form. By the conversion process, a word is changed from one part-of-speech (POS) into another without the addition or deletion of any morphemes. Compounding is the most productive way to compose a Chinese word while affixation is the most popular way to construct an English word. Affixation also is used widely in Chinese, but its prefix or suffix doesn't have the meaning and can be always omitted (e.g., “老虎” (tiger) and “虎” (tiger) have the same meaning.). As for conversion, it's really not a way to construct a word and just represents the fact that some words have more than one tense.

### 3.2 Morphological structures in Chinese triggers

Since almost all triggers in Chinese events are verbs and nouns, we focus on the morphological structures of Chinese verbs and nouns. Actually, statistics on the ACE 2005 Chinese corpus shows that 95% of triggers are either verbs or verbal nouns and just nearly 5% are pure nouns (e.g., “公开信” (open letter), “大会” (plenary session)). In ACE 2005 English corpus, there are some adjectives triggering an event of special type. However, no adjective acts as a trigger in the ACE 2005 Chinese corpus for the special characteristics in Chinese language. Besides, almost 95% of triggers in the training set just contain one or two morphemes, so this paper only considers the one-morpheme and two-morpheme triggers of verbs and nouns.

There are two type words in Chinese triggers: single-morpheme words and compound words. Single-morpheme word just contains one morpheme. Sometimes, a single-morpheme word maybe is composed by more than one character, such as the transliterated word. But it doesn't occur in Chinese triggers and we disregard them in this paper. So there is only one morphological structure concerning a single-morpheme trigger:

**Single-Morpheme Structure:** Single-morpheme trigger whose POS is a verb or a noun (e.g., “死” (die), “去” (go), “信” (letter), etc.).

Compounding is the most productive way to compose a Chinese trigger. In this paper we define five types (similar to (Chang, 1995)) of the morphological structures in Chinese triggers based on the relations between their morphemes.

**Coordinative Structure:** The two morphemes of a trigger play coordinative role. For example, “合” (combine) and “并” (merge) are coordinative in trigger “合并” (merge).

**Modifier-Head Structure:** The modified morpheme follows the modifying one in a trigger. For example, “婚” (marry) is modified by “新” (new) in trigger “新婚” (newly-married).

**Subject-Predicate Structure:** One morpheme is the subject and the other one tells something about the subject. This structure is like a subject-predicate sentence condensed in a trigger. For example, “身” (body) is a subject of predicate “亡” (die) in trigger “身亡” (die).

**Predicate-Object Structure:** The first morpheme (predicate) governs the second one (object) in a trigger. For example, “业” (business) serves as the object of predicate “开” (start) in trigger “开业” (start business).

**Predicate-Complement Structure:** The first morpheme is a predicate and the second one interprets the first one from different aspects (e.g., direction, result and tense) in a trigger. For example, morpheme “入” (into) expresses the direction of action “进” (go) in trigger “进入” (go into).

### 3.3 Determining the morphological structure in a Chinese trigger

A general method to determine the morphological structures in Chinese triggers is to first annotate some instances manually and then train a classifier. Alternatively, a simple way is employed in this paper to determine the morphological structures in Chinese triggers via their POS structures, due to our finding that the morphological structures in Chinese triggers can be inferred from their POS structures. Following are the inference rules employed in this paper for different morphological structures:

**Single-Morpheme Structure:** For a single-morpheme trigger whose POS is a noun or a verb, its morphological structure is *Single-Morpheme*. The statistics on the training set shows that this inference rule covers almost 100% of cases given correct POSs.

**Predicate-Complement Structure:** If the POS structure of a trigger is (*verb + preposition*) or (*verb + auxiliary*), its morphological structure is *Predicate-Complement*. The statistics on the training set shows that this inference rule covers almost 100% of cases given correct POSs.

**Predicate-Object Structure:** If the POS structure of a trigger is (*verb + noun*), its morphological structure is *Predicate-Object*. The statistics on the training set shows that this inference rule covers almost 100% of cases given correct POSs.

**Coordinative Structure:** If the POS structure of a trigger is (*verb + verb*) (e.g., “捐/VV 赠/VV” (donate), “购/VV 买/VV” (buy), etc.), its morphological structure is *Coordinative*. The statistics on the training set shows that this inference rule covers almost 98% of cases given correct POSs. The only exception to this inference rule is that it ignores those triggers whose POS structure is (*noun + noun*). This happens in Chinese triggers, though seldom. In such cases, i.e. if the POS structure of a trigger is (*noun + noun*), its morphological structure can be either *Modifier-Head* or *Coordinative* (e.g., “婚/NN 姻/NN” (marriage)).

**Modifier-Head Structure:** The morphological structure of a trigger is *Modifier-Head*, if its POS structure is one of following four structures: 1) (*adjective + verb*); 2) (*adjective + noun*); 3) (*noun + noun*); 4) (*noun + verb*). The statistics on the training set shows that this inference rule covers almost 96% of cases given correct POSs. The only exceptions to this inference rule are that if the POS structure of a trigger is (*noun + noun*) or (*noun + verb*), its morphological structure can also be *Coordinative* or *Subject-Predicate*, respectively.

**Subject-Predicate Structure:** Our exploration on the ACE 2005 Chinese corpus shows that only one trigger (i.e. “身亡” (die)) has the *Subject-Predicate* structure. Therefore, we ignore this structure.

Structure	% Trigger mentions
Single-Morpheme	19.1
Coordinative	46.3
Modifier-Head	13.3
Predicate-Object	11.4
Predicate-Complement	8.7
Words (length>=3)	1.2

TABLE 1 – Distribution of different morphological structures in Chinese trigger mentions

To obtain the POS structures of Chinese triggers, we split all triggers into characters and employ a Chinese POS tool – ICTCLAS to tag their POSs. Table 1 shows the distribution of the morphological structures in Chinese triggers in the training set, extracted using above inference rules. Random manual evaluation of 1000 instances shows that our inference rules achieve the accuracy of more than 91% given automatically-tagged POSs.

### 3.4 Identifying head morpheme in Chinese triggers

Normally, almost all Chinese verbs or nouns contain one morpheme as the governing semantic element, called Head Morpheme (HM), to construct a word and the semantics of such a word thus can be inferred from its HM. Since the semantics of a Chinese trigger can be often inferred from its HM, it's natural to infer unknown triggers via HMs. For example, given verb “死” (die) as HM in trigger “烧死” (burn to death, trigger of the *Die* event) whose morphological structure is *Coordinative*, it is reasonable to infer “砸死” (crush to death), “炸死” (burst to death), “闷死” (stifle to death) to be triggers of the same event, due to their same HM and morphological structure as “烧死”.

Li et al. (2012) regards all single-character verbs contained in triggers as BVs and use them to infer unknown triggers. It may introduce many pseudo triggers into candidates and harm the precision for that loose constraint. For example, the morphological structure of “烧死” is *Coordinative*, and “烧” (burn) and “死” (die) are two single-morpheme verbs in it. Following Li's inference rule, all words including BV “烧” or “死” are regarded as triggers if their verb structures are (*BV + verb*) or (*verb + BV*). Hence, some pseudo triggers, such as “烧烤” (barbecue), “烧焊” (weld), “烧制” (fire), etc., would be expanded to be triggers.

Besides, a noun may be a HM to infer new triggers. For example, given “信” (letter) as the HM in trigger “私信”(private letter, trigger of *Phone-Write* event) whose morphological structure is *Modifier-Head*. It's correctly to infer those words (e.g., “贺信” (congratulatory letter), “密信” (secret letter), etc.) with the HM “信” (letter) and the morphological structure *Modifier-Head*, as triggers.

Therefore, how to identify the HM in a Chinese trigger becomes the key to infer unknown triggers. Table 2 shows our automatic mechanism to identify HM, where  $LM(w)$  and  $RM(w)$  are used to obtain the left and right morphemes from one-morpheme or two-morphemes word  $w$  respectively.

Structure	Inferences to select HM
Single-morpheme	$tr$
Coordinative	$LM(tr)$ : if $SSIM(tr, LM(tr)) > \alpha$ $RM(tr)$ : if $SSIM(tr, RM(tr)) > \alpha$
Modifier-Head	$RM(tr)$
Predicate-Object	$RM(tr)$
Predicate-Complement	$LM(tr)$

TABLE 2 – Inferences on different morphological structures to extract HMs

For a trigger whose morphological structure is *Single-morpheme*, *Predicate-Complement* or *Modifier-Head*, it's easy to identify its HM from the relationship between its two morphemes. If the structure of a trigger is *Predicate-Object*, we select the noun (object) as HM because it better represents the semantics of the trigger than the predicate, i.e. the governing semantic element

always comes from the object. However, without additional information, it's hard to select HM from a trigger whose morphological structure is *Coordinative*. For example, given the trigger “访问” (visit) whose morphological structure is *Coordinative*, its two component morphemes, “访” (visit) and “问” (ask), have their own semantics respectively. Fortunately, we can find out that morpheme “访” (visit) has the same meaning as trigger “访问” (visit). So an effective way to identify HM in a trigger with the *Coordinative* structure is via the semantic similarity (SemSim).

In this paper, we employ HowNet<sup>1</sup> (Dong and Dong, 2006) to obtain the semantics of Chinese words. Similar to Wordnet in English, HowNet is a structured Chinese lexical semantic resource. In HowNet, sememe is a basic semantic unit and represents the meaning of a word. In total, about 2200 sememes are used to define 95000 Chinese words. In this paper, the governing sememe is introduced to recognize HMs from those triggers with the *Coordinative* structure. That is, if a morpheme represents the governing sememe, it is recognized as HM of that trigger. Following Liu and Li (2002), function  $SemSim(x, y)$  is used to calculate the semantic similarity between the sememes of the trigger  $x$  and its morpheme  $y$  as follow:

$$SemSim(x, y) = \frac{\phi}{Dis(x, y) + \phi} \tag{1}$$

where  $Dis(x, y)$  is the distance between the sememe of  $x$  and  $y$  in HowNet's sememe hierarchical architecture, and  $\phi$  is an adjustable parameter and assigned 0.75 following Liu and Li (2002).

#### 4 Inferring unknown triggers on HMs and sememes

To better represent the compositional semantics inside Chinese words and filter out more pseudo triggers, we introduce the morphological structures and sememes of Chinese words to infer unknown triggers. The methodology is shown as follows: 1) following the principle of compositional semantics, we extract these one-morpheme or two-morpheme words in the test set as candidates when they contain at least one HM and their POS are nouns or verbs; 2) according to the morphological structure of each candidate word, we applied different inferences to choose unknown triggers. We implement an algorithm to determine whether a candidate is an unknown trigger and the input and output are shown as follows:

**Input:**  $HMs \leftarrow$  the set of all HMs extracting from the training set

$$candidates \leftarrow \{w \mid \begin{cases} (LM(w) \in HMs \vee RM(w) \in HMs) \\ \wedge (POS(w) = noun \vee POS(w) = verb) \\ \wedge MPRO(MORPH(w), HM(w)) \geq 0 \end{cases}\}$$

$triggerwords \leftarrow \phi$

**Output:**  $triggerwords$  : the set of unknown triggers accepted by our algorithm

$POS(w)$  and  $HM(w)$  are applied to get the POS of word  $w$  and obtain the HM in word  $w$  respectively.  $MPRO(ms, hm)$  is defined to compute the conditional probability of a trigger when it contains a HM  $hm$  and its morphological structure is  $ms$ .  $MORPH(w)$  is used to get the morphological structure of word  $w$ .

For each candidate word  $w$  in  $candidates$ , we apply following inferences to distinguish the true unknown triggers from the pseudo ones according to the morphological structure and sememe.

<sup>1</sup> <http://www.keenage.com>



**Single-Morpheme:** These expanding single-morpheme words are those HMs in two-morpheme triggers. So we apply a simple constraint to determine whether or not it's an unknown trigger:

$$\underset{m_i \in S}{MAX}(SemSim(w, m_i)) = 1 \quad (2)$$

where  $S$  is the set of triggers in the training set which contain word  $w$ . If the maximum score of the semantic similarity between these triggers and word  $w$  is equal to 1, we accept it.

**Predicate-Complement:** The first morpheme is usually a verb, so the sememe of word  $w$  always is similar to the sememe of its first morpheme. The constraint for *Predicate-Complement* structure is:

$$LM(w) \in S_{sm} \cup S_{pc} \quad (3)$$

where  $S_{sm}$  is the set of triggers in the training set whose structures are *Single-morpheme* while  $S_{pc}$  is the set of left morphemes of triggers in the training set whose structures are *Predicate-Complement*.

**Predicate-Object:** for a word  $w$  whose morphological structure is *Predicate-Object*, we regard it as the unknown trigger following two conditions to constrain its two morphemes:

$$RM(w) \in HMs \quad (4)$$

$$\underset{m_i \in SW}{MAX}(SemSim(LM(w), m_i)) \geq \beta \quad (5)$$

where  $SW$  is the set of predicates in the similar triggers<sup>2</sup> of word  $w$ . For example, if there are two triggers “离职” (resign) and “辞职” (resign), and their HMs are “职” (job) too. For a candidate “免职” (resign), its morphological structure is as same as the above two and its HM also is “职” (job). We call them similar triggers and calculate the similarities between “免”(dismiss) and the predicates (“离”(leave), “辞”(dismiss)) in its similar triggers in the training set.

**Modifier-Head:** The first morpheme of word  $w$  modifies the second one, so that the semantics of word  $w$  comes from its second morpheme. We apply following rules based on POS consistency and semantic similarity.

$$RM(w) \in HMs \quad (6)$$

$$POS(LM(w)) \in \{POS(l) | COM(l, b) \in S_{mh} \wedge b \in HMs\} \quad (7)$$

$$\underset{m_i \in S}{MAX}(SemSim(w, m_i)) = 1 \quad (8)$$

where  $S_{mh}$  is the set of triggers in the training set whose structures are *Modifier-Head* and  $COM(l, b)$  is to combine morpheme  $l$  and  $b$  to be a two-morpheme word. Otherwise,  $S$  is the set of those triggers which contain word  $w$ .

**Coordinative:** Since the two composite morphemes of word  $w$  are homogeneous and its semantics is flexible and maybe comes from the combination of its two morphemes or one of its morpheme. We calculate the average score of the similarities to infer trigger of this type:

---

<sup>2</sup> Similar triggers are those triggers with the same morphological structure and the same BM in the training set.

$$\frac{\sum_{co \in \{w, LM(w), RM(w)\}} \sum_{m_i \in SC} MAX(SemSim(co, m_i))}{3} > \lambda \quad (9)$$

where

$$SC \in \{s | MORPH(s) = MORPH(w) \wedge (LM(w) \in HMs \wedge LM(s) = LM(w) \vee RM(w) \in HMs \wedge RM(w) = RM(s))\} \quad (10)$$

where  $SC$  is the set of triggers in the training set with following two constraints: 1) their morphological structures are *Coordinative*; 2) their left/right morphemes and the left/right morpheme of word  $w$  are the same HM.

## 5 Experimentation and discussion

In this section, we evaluate our mechanism of combining the morphological structures and sememes of Chinese words in inferring unknown triggers and report the experimental results on trigger identification and its application to overall Chinese event extraction.

### 5.1 Experimental setting and baseline

We use a state-of-the-art Chinese event extraction system (Li et al., 2012) as one of our baselines which consists of four typical components (trigger identification, event type determination, argument identification and argument role determination) in a pipeline way. During testing, each word in the test set is first scanned for instances of known triggers from the training set and then scanned by employing the compositional semantics inside Chinese triggers to infer instances of unknown triggers. When an instance is found, the trigger identifier is applied to distinguish those true trigger mentions from pseudo ones. If true, the event type determiner is then applied to recognize its event type. For any entity mention in a sentence which is identified as an event, the argument identifier is employed to assign its possible arguments afterwards. Finally, the argument role determiner is introduced to assign a role to each argument.

Besides, we adopt the same experimental setting as Li et al. (2012). The ACE 2005 Chinese corpus (only the training data is available) is used in all our experiments. The corpus contains 633 Chinese documents annotated with 8 predefined event types and 33 predefined event subtypes<sup>3</sup>. We randomly select 567 documents as the training set and the remaining 66 documents as the test set. Besides, we reserve 33 documents in the training set as the development set and follow the setting of ACE diagnostic tasks and use the ground truth entities, times and values for our training and testing. As for evaluation, we also follow the standards as defined in Li et al (2012):

- A trigger is **correctly** identified if its position in the document matches a reference trigger;
- An event type is **correctly** determined if the trigger’s event type and position in the document match a reference trigger;
- An argument is **correctly** identified if its involved event type and position in the document match any of the reference argument mentions;
- An argument role is **correctly** determined if its involved event type, position in the document, and role match any of the reference argument mentions.

<sup>3</sup> Similar to previous studies, we treat these subtypes simply as 33 separate event types and do not consider the hierarchical structure among them.

Finally, all the sentences in the corpus are divided into words using a word segmentation tool (ICTCLAS) with all entities annotated in the corpus kept. Besides, we use Berkeley Parser and Stanford Parser to create the constituent and dependency parse trees. We use N-gram features and employ the ME model<sup>4</sup> to train individual component classifiers.

## 5.2 Results on identifying HMs and unknown triggers

As the key to infer unknown triggers, Table 3 shows the performance of HM identification. For evaluation, the HMs of all the known triggers in the ACE 2005 Chinese corpus are manually labeled by three annotators and we accept those morphemes as HMs when at least two annotators agree on them. The threshold  $\alpha$  is fine-tuned to 0.85 using the development set. Compared to Li et al. (2012), our approach can improve the F1-measure by 6.9%, largely due to the dramatic increase in Precision of 15.8%. Li et al. (2012) extracted all single-character verbs as BVs, so their Recall is higher than that of ours. Otherwise, we extract 30 single-morpheme nouns as HMs and 73% of them occur in the gold set while this number in Li et al. (2012) is 0.

System	#BV/HMs	P(%)	R(%)	F1
Li et al. (2012)	361	64.3	88.5	74.5
Ours	266	80.1	82.1	81.4

TABLE 3— Performance of the HM identification (#Gold: 262)

We apply the mechanism of combining the morphological structures and sememes of Chinese words (CMS) to infer unknown triggers. The thresholds  $\beta$  and  $\lambda$  are fine-tuned to 0.7 using the development set. Following Li et al. (2012), we also apply the non-trigger filtering rule in our system and just filter out those candidates which occur as pseudo triggers more than 5 times in the training set. So we obtain a candidate set of words including known triggers in the training set and those unknown triggers identified by our mechanism. Manual inspection shows that 62 words are inferred as unknown triggers, among which 69.4% are true triggers.

To verify the effectiveness of our mechanism, we extract those trigger mentions from the test set when they are instances of known triggers from the training set or unknown triggers extracted by CMS. Table 4 shows the results of our CMS and two baseline systems in inferring unknown trigger mentions. Here, Baseline-1 (Chen and Ji (2009b)) just extracts those trigger mentions occurring in the training data while Baseline-2 (Li, et al., 2012) infers unknown trigger mentions based on the compositional semantics and verb structures of Chinese words.

System	#True trigger mentions	#Pseudo trigger mentions
Baseline-1	266	629
Baseline-2	302	444
CMS	326	508
Gold	367	-

Table 4 – Impact of combining the morphological structure and sememe of Chinese words in inferring unknown triggers

Compared with Baseline-1 and Baseline-2, our mechanism recovers 16.3% (60) and 6.5% (24) of true trigger mentions respectively. This improvement mainly comes from two factors. The first one is that we introduce those nouns to be HMs and almost 20% of the true unknown triggers

<sup>4</sup> <http://mallet.cs.umass.edu/>

(e.g., “失业” (lose one’s job), “出境” (leave the country)) are extracted. The second one is that our mechanism filters out more pseudo trigger mentions due to the contribution of combining the morphological structures and sememes of Chinese words. For example, Baseline-2 will infer “调频” (frequency adjustment) “妨害” (impair) to be triggers due to “调” (adjust) and “害” (harm) are BVs and their syntactic structures are *(BV+noun)* and *(verb+BV)* respectively. On the contrary, our mechanism will filter out “调频” since its structure is *Modifier-Head* and the head morpheme “频” (frequency) doesn’t appear in HMs while “妨害” will also be ignored because its sememe is not similar to any known triggers with the same HM “害” (harm). It justifies the effectiveness of our mechanism to combine the morphological structures and sememes of Chinese words in recovering true triggers.

Otherwise, some triggers in the training set are seldom used as trigger mentions. We also applied above mechanism to filter out those triggers. Table 4 shows that almost 28% of pseudo trigger mentions is filtered out, so the number of pseudo trigger mentions is reduced to 508.

### 5.3 Results on trigger identification and overall Chinese event extraction

There are too many pseudo trigger mentions showed in Table 4 by using our mechanism to infer unknown triggers and extract trigger mentions from the test set, so we introduce a ME-based trigger identifier to distinguish the true trigger mentions from the pseudo ones as previous works.

Table 5 shows the contribution of our mechanism to trigger identification on the held-out test set. Compared to Baseline-1, our approach can dramatically improve the F1-measure by 10.0%, with a big gain of 17.8% in Recall and a small loss of 1.8% in Precision. It further proves the effectiveness of the compositional semantics in inferring Chinese unknown triggers. Compared to the state-of-the-art system (Baseline-2), our approach also enhances F1-measure by 4.1%, largely due to a dramatic increase of 7.7% in Recall. It also justifies that the morphological structures of Chinese words are more effective than the verb structures when they are employed to infer unknown triggers. Besides, these results also show that introducing sememes of Chinese words into our mechanism is a helpful way to filter out those pseudo triggers.

We also employ the mechanism of discourse consistency (Li et al., 2012) to improve the Precision and our results show that our approach achieves 79.4%, 69.2% and 73.9% in F1-measure, Precision and Recall respectively and it outperforms Li et al. (2012) by 3.4% and 5.7% in F1-measure and Recall, with a small loss of 0.1% in Precision.

System	Trigger identification		
	P(%)	R(%)	F1
Baseline-1	75.2	52.0	61.5
Baseline-2 (Li et al. (2012))	73.5	62.1	67.4
CMS	73.4	<b>69.8</b>	71.5
Baseline-2+ Discourse consistency	79.3	63.5	70.5
CMS + Discourse consistency	<b>79.4</b>	69.2	<b>73.9</b>

Table 5 – Contribution to Chinese trigger identification

Table 6 shows the contribution of trigger identification to overall event extraction on the held-out test set. Compared to Baseline-2, we can find that our approach can improve the F1-measure for event type determination by 4.0%, argument identification by 3.3% and argument role determination (i.e. overall event extraction) by 2.9%, largely due to the dramatic increase in

Recall of 7.4%, 6.1% and 5.6%. These results also ensure the importance of trigger identification in Chinese event extraction.

System	Event type determination			Argument identification			Argument role determination		
	P(%)	R(%)	F1	P(%)	R(%)	F1	P(%)	R(%)	F1
Baseline-1	70.3	49.0	57.8	58.4	42.7	49.3	55.2	38.6	45.4
Baseline-2	70.2	59.1	64.2	58.0	48.9	53.0	54.7	44.5	49.1
CMS	<b>69.9</b>	<b>66.5</b>	<b>68.2</b>	<b>57.6</b>	<b>55.0</b>	<b>56.3</b>	<b>54.1</b>	<b>50.1</b>	<b>52.0</b>

Table 6 – Contribution to Overall Chinese event extraction

## 5.4 Discussion

Through manual inspection, we find that many remaining errors are related to three aspects. The first one is that almost 4.7% of trigger mentions in the test set doesn't have a morpheme appeared in the set of HMs. For example, there are so many ways to hurt a human to express an *injure* event and just a few of triggers or its HMs occurred in the training set. The second one comes from the errors in POS tagging in the verb structures of triggers and constituent parse tree. Almost all errors in determining morphological structures are come from those wrong POSs, especially those single-morpheme triggers, with the wrong POS in the parse tree will be ignored in inferring unknown triggers. The last one is the low quality of the annotated event corpus and many event mentions are missed. Those un-annotated true mentions would make the classifier confuse to distinguish true event mentions from pseudo ones. We look into those pseudo trigger mentions which are classified as true ones by the ME classifier and find out almost 20% of them maybe are true ones by our knowledge.

In order to evaluate the effect of the training set size on the performance, we modify the proportion of the training set to the test set from 9:1 to 1:9. Fig. 2 shows the percentages of true trigger mentions extracted by our baseline and our CMS. From Figure 1, we can find out that our mechanism can extract much more true trigger mentions than that of the baseline, especially for a smaller training set. When the proportion of the training set to the test set is set to 1:9, our mechanism can extract 67.5% of true trigger mentions while the figure drops to 43.3% in our baseline. This justifies that our mechanism can be well applied to minimally-supervised event extraction.

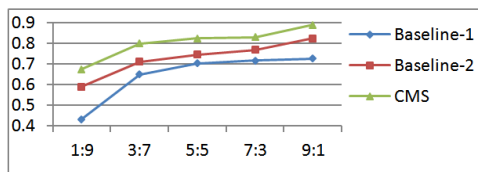


FIGURE 1 – The percentages of extracting true trigger mentions on different proportions of the training set to the test set

Compared to Li et al. (2012), There are three contributions in our work: 1) we use the morphological structure to better represent the compositional semantics inside Chinese triggers; 2) we introduce a mechanism to identify HMs in triggers automatically and those HMs can be verbs or nouns; 3) we propose a mechanism of combining the morphological structures and semantics of

Chinese words to extract unknown triggers. The results show that our mechanism outperforms the state-of-the-art system.

## Conclusion

To address the special characteristics of Chinese event extraction and extract more true trigger mentions, this paper presents a novel approach to Chinese trigger identification which combines the morphological structures and sememes of Chinese words to infer unknown triggers. The experimental results show that our approach can significantly improve the performance of the Chinese event extraction system, especially Chinese trigger identification in Recall. In future work, we will focus on how to apply the mechanism of compositional semantics to unsupervised or minimally supervised event extraction system and improve their performance.

## Acknowledgments

The authors would like to thank three anonymous reviewers for their comments on this paper. This research was supported by the National Natural Science Foundation of China under Grant Nos. 61070123 and 61273320, the National 863 Project of China under Grant No. 2012AA011102.

## REFERENCES

- Ahn, D. (2006). The stages of event extraction. In *Proceedings of the Workshop on Annotating and Reasoning about Time and Events (ARTE 2006)*, pages 1-8.
- Chambers, N. and Jurafsky, D. (2011). Template-based information extraction without the templates. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*, pages 976-986.
- Chang, S. M. (1995). A study on Mandarin compounds. Hsinchu: National Tsing Hua University.
- Chen, Z. and Ji, H. (2009a). Can one language bootstrap the other: a case study on event extraction. In *Proceedings of the NAACL HLT 2009 Workshop on Semi-Supervised Learning for Natural Language Processing (SemiSupLearn 2009)*, pages 66-74.
- Chen, Z. and Ji, H. (2009b). Language specific issue and feature exploration in Chinese event extraction. In *Proceedings of Human Language Technologies: the 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2009)*, pages 209-212.
- Dong Z. D. and Dong Q. (2006). HowNet and the computation of meaning. World Scientific Pub Co. Inc.
- Finkel, J., Grenager, T. and Manning, C. (2005). Incorporating non-local information into information extraction systems by Gibbs sampling. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (ACL 2005)*, pages 363-370.
- Fu, J. F., Liu, Z. T., Zhong, Z. M. and Shan, J. F. (2010). Chinese event extraction based on feature weighting. *Information Technology Journal*, 9: 184-187.
- Gupta, P. and Ji, H. (2009). Predicting unknown time arguments based on cross-event propagation. In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers (ACLShort 2009)*, pages 369-272.
- Grishman, R., Westbrook, D. and Meyers, A. (2005). NYU's English ACE 2005 system description. In *Proceedings of ACE 2005 Evaluation Workshop (ACE workshop 2005)*.
- Hardy, H., Kanchakouskaya, V. and Strzalkowski, T. (2006). Automatic event classification

using surface text features. In *Proceedings of the Twenty-first National Conference on Artificial Intelligence Workshop on Event Extraction and Synthesis (AAAI workshop 2006)*, pages 36-41.

Hong, Y., Zhang, J., Ma, B., Yao, J. M., Zhou, G. G. and Zhu, Q. M. (2011). Using cross-entity inference to improve event extraction. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT 2011)*, pages 1127-1136.

Ji, H. (2009). Cross-lingual predicate cluster acquisition to improve bilingual event extraction by inductive learning. In *Proceedings of the Workshop on Unsupervised and Minimally Supervised Learning of Lexical Semantics (UMSLLS 2009)*, pages 27-35.

Ji, H. and Grishman, R. (2008). Refining event extraction through cross-document inference. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT 2008)*, pages 254-262.

Li, J. H., Zhou, G. D. and Ng, H. T. (2010). Joint syntactic and semantic parsing of Chinese. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, pages 1108-1117.

Li, P. F., Zhou G. D., Zhu Q. M. and Hou L. B. (2012). Employing compositional semantics and discourse consistency in Chinese event extraction. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012)*, pages 1006-1016.

Liao, S. S. and Grishman, R. (2010). Using document level cross-event inference to improve event extraction. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, pages 789-797.

Liao, S. S. and Grishman, R. (2011). Using prediction from sentential scope to build a pseudo co-testing learner for event extraction. In *Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP 2011)*, pages 714-722.

Llorens, H., Saquete, E., Navarro-Colorado, B. (2012). Applying semantic knowledge to the automatic processing of temporal expressions and events. *Information Processing & Management*.

Lu, W. and Roth, D. (2012). Automatic event extraction with structured preference modeling. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012)*, pages 835-844.

Maslennikov, M. and Chua, T. (2007). A multi resolution framework for information extraction from free text. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics (ACL 2007)*, pages 592-599.

Patwardhan, S. and Riloff, E. (2007). Effective information extraction with semantic affinity patterns and relevant regions. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-coNLL 2007)*, pages 717-727.

Patwardhan, S. and Riloff, E. (2009). A unified model of phrasal and sentential evidence for information extraction. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing (EMNLP 2009)*, pages 151-160.

Qin, B., Zhao, Y. Y., Ding, X., Liu, T. and Zhai, G. F. (2011). Event type recognition based on trigger expansion. *Tsinghua Science and Technology*. 15(3): 251-258.

- Riloff, E. (1996). Automatically generating extraction patterns from untagged text. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, pages 1044–1049.
- Shinyama, Y. and Sekine, S. (2006). Preemptive information extraction using unrestricted relation discovery. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2006)*, pages 304–311.
- Stevenson, M. and Greenwood, M. (2005). A semantic approach to IE pattern induction. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005)*, pages 379–386.
- Tan, H., Zhao, T., Zheng, J. (2008). Identification of Chinese event and their argument roles. In *Proceedings of the 2008 IEEE 8th International Conference on Computer and Information Technology Workshops (CITWORKSHOPS 2008)*, pages 14-19.
- Yangarber, R., Grishman, R., Tapanainen, P. and Huttunen, S. (2000). Automatic acquisition of domain knowledge for information extraction. In *Proceedings of the 18th International Conference on Computational Linguistics (ACL 2000)*, pages 940–946.
- Yuan, M. L. (1998). Studies on valency in modern Chinese. Chinese Commerce and Trade Press, Beijing, China.