

Mining activation force defined dependency patterns for relation extraction



Chunyun Zhang^a, Yichang Zhang^b, Weiran Xu^{b,*}, Zhanyu Ma^b, Yan Leng^c, Jun Guo^b

^a School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan, China

^b Pattern Recognition and Intelligent System Laboratory, Beijing University of Posts and Telecommunications, Beijing, China

^c School of Physical and Electronic Science, Shandong Normal University, Jinan, China

ARTICLE INFO

Article history:

Received 21 March 2014

Received in revised form 19 June 2015

Accepted 20 June 2015

Available online 24 June 2015

Keywords:

Relation extraction

Pattern learning

Trigger word

Activation force

ABSTRACT

Relation extraction is essential for most text mining tasks. Existing approaches on relation extraction are generally based on bootstrapping methodology which implies semantic drift problem. This paper presents a new approach to learn semantic dependency patterns, which can significantly alleviate this problem. To this end, a unique representation of activation force defined dependency pattern is presented. It is a trigger word mediated relation between an entity and its attribute value, and the trigger word is extracted by using the statistics of word activation forces between those words. The adaptability and the scalability of the framework are facilitated by the recursive and compositional bootstrap learning of patterns and seed pairs. To obtain insights of the reliability and applicability of the method, we applied it to the English Slot Filling task of Knowledge Base Population track at Text Analysis Conference 2013. Experimental results show that the proposed method has good performance in the implementation of English Slot Filling 2013 with the overall F1 value significantly higher than the best automatic result reported. The experimental results also demonstrate that the activation force based trigger word mining method plays an essential role in improving the performance.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Relation extraction is a very important task, which can potentially benefit a wide range of natural language processing (NLP) tasks such as question answering, ontology learning, and summarization [1].

In decades, many related tasks, e.g., Message Understanding Conference (MUC) [2], Automatic Content Extraction (ACE) [3], and Knowledge Base Population (KBP) [4], arose and facilitated the development of relation extraction technology. The English Slot Filling task (ESF) of KBP track in Text Analysis Conference (TAC), which involves mining information about entities from text, has been very attractive for its potential application in dealing with big data in the web. ESF systems determine from a large source collection of documents the values of specified attributes ('slots') of an entity, such as the *age* and the *birthplace* of a person or the *top employees* of a corporation [5]. Entities investigated in ESF are generally either a person or some type of organization. Extracting predefined attributes of such kind of entities is considered as a key technology to realize a practical system that can automatically mine the Web data.

* Corresponding author.

E-mail address: xuweiran@bupt.edu.cn (W. Xu).

Reliable relation extraction which can be used in the task mentioned above is still an open problem. One of the major challenges is how to learn selective patterns that have high coverage to represent relations [6]. In 2010, the overview of TAC-KBP slot filling showed that the bootstrapping system [7] proposed by NYU with F1 value 28.3% was much better than the CRF system [8] proposed by BUPT with F1 value 14.4%. In addition, the slot filling report [9] of CUNY in the same year also showed further experimental results that the system only using MaxEnt method gave the performance with F1 value 17.9%, which was much lower than the result of bootstrapping system. All this verified the bootstrapping method, as an effective method, is superior to the classical MaxEnt and CRF classifiers.

Recent years, more studies show that bootstrap learning methods [6,10–16] are capable for learning relation patterns starting from a small number of seed examples. But they are limited by inflexible pattern representation and semantic drift. Many patterns [6,10–12] using shallow syntactic features have poor performance in extracting relations that are ambiguous or lexically distant in their expression. Dependency patterns [17–21,15,22,16,23] have been shown to be better patterns, since they are more informative. Among these dependency patterns, the shortest dependency pattern (SDP) and the subject-verb-object (SVO) pattern are two mostly used ones [17,1,19,20]. However, because there is less

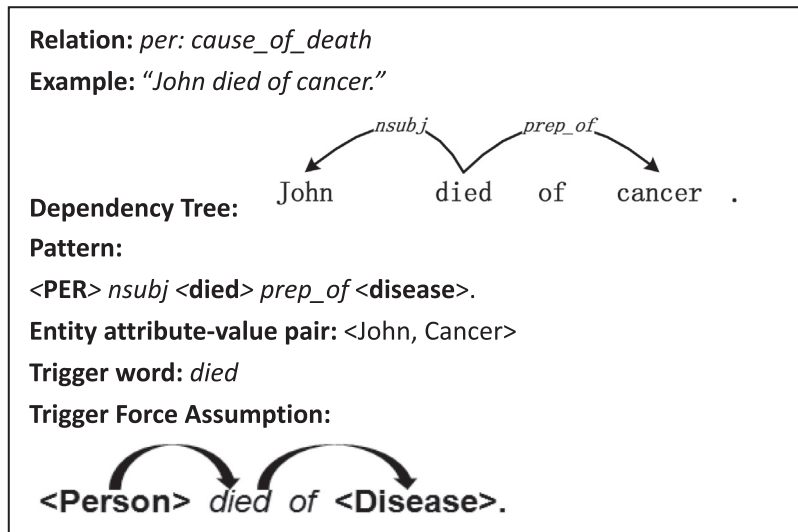


Fig. 1. An example of trigger word based dependency pattern.

semantic constraint, they gain the generality at the cost of lacking specific information and thus may produce semantic drift in bootstrapping iterations.

To address the problem of semantic drift, this paper presents an activation force based relation extraction model which can integrate the semantic constraints into patterns. This model adopts a unique representation of activation force defined dependency pattern (AFDDP). It is the shortest dependency path from an entity to its attribute value with a trigger word as the semantic anchor. The trigger word is extracted by using the statistics of word activation forces [24,25] between those words. Compared with the SDP and the SVO, the AFDDP can maintain necessary semantic information and largely suppress semantic drift of bootstrapping. During iterations, the AFDDP learning method evaluates qualities of these patterns and discards the unreliable ones.

Trigger words are words with semantic information. They activate patterns of specific relations and act as the patterns' conceptual anchor points [26]. "John died of cancer" is an example sentence of the relation of person *per:cause_of_death*. Its dependency tree is shown in Fig. 1. The sentence head is the main verb *died* which is modified by its nominal subject (*nsubj*) *John* and the *of* prepositional modifier (*prep_of*) *cancer*. The main verb *died* is a trigger word of the relation *per:cause_of_death* and the trigger word based dependency pattern is obtained as < PER > nsubj < died > prep_of < disease >. By matching the pattern, the entity-value pair <John, Cancer> can be extracted from the sentence. Obviously, how well the trigger word based dependency patterns perform largely depends on how well the trigger words are identified. Unfortunately, existing trigger word identification methods are rigid and poorly portable because they are simply based on dictionary or term frequency [27,28,26,29–31]. To tackle this challenge, we proposed a unique trigger word identification method by using the statistics of word activation force, which have proved powerful in defining word relations in context [24,25]. Through word activation forces, we define a trigger force (TF) metric to identify trigger words. The TF of a word is predicted by a weighted sum of the activation force that an entity exerts on the trigger word and the activation force that the trigger word exerts on the attribute-value of the entity.

To obtain insights of reliability and applicability of this method, we applied it to the KBP-ESF task [32] to extract values of specified attributes ('slots') of entities. Experimental results show that we achieved a good performance with the overall precision of

46.55%, recall of 33.04%, and F1 value of 38.65%, which is significantly higher than the best automatic result reported in ESF2013.

The main contributions of this paper are as follows:

- (1) We define a unique trigger force metric on trigger word identification, which identifies trigger words by using the statistics of word activation force.
- (2) A unique representation of activation force defined dependency pattern is proposed. It is a trigger word mediated relation between an entity and its attribute value. They form the shortest dependency path with a trigger word as the semantic anchor. Compared with the SDP and the SVO pattern, the new pattern is more relation oriented, and thus more noisy robust.
- (3) Two-level semantic drift suppressing strategies are adopted. At the first level, we set trigger word as the semantic anchor of a pattern to maintain the pattern's semantic information. At the second level, we used a specialized pattern evaluation method to select the most relation revealing patterns during iterations.

The rest of this paper is organized as follows: Section 2 reviews related work. Section 3 describes activation force defined dependency patterns via trigger words. Section 4 presents the algorithm of mining activation force defined dependency patterns in large scale text. The experimental results are shown and discussed in Section 5. Finally, we conclude the paper in Section 6.

2. Related work

2.1. Trigger word mining

WordNet [33] has played an important role in trigger words extraction methods. For example, the predicates of WordNet were used as trigger words to establish relations between words and concepts of a language independent ontology [27]. A trigger word list was gathered from WordNet by checking whether a word had the semantic class "person|...|relative" to personal social relation subtypes [28]. In practice, these dictionary dependent methods generally have poor domain portability.

Alternatively, event trigger word extraction was defined as a multiclass classification problem in [26]. The trigger word extraction method [29] of social tag computed the trigger power of a

Download English Version:

<https://daneshyari.com/en/article/6862473>

Download Persian Version:

<https://daneshyari.com/article/6862473>

[Daneshyari.com](https://daneshyari.com)