Geography Gaokao-oriented Knowledge Acquisition for Comparative Sentences Based on Logic Programming

Xuelian Li[1,2], Qian Liu[3], Man Zhu[3], Feifei Xu[2], Yunxiu Yu[1,2], Shang Zhang[1,2], Zhaoxi Ni[1,2], and Zhiqiang Gao[1,2]

[1] Key Lab of Computer Network and Information Integration (Southeast University), Ministry of Education, China
[2] School of Computer Science and Engineering, Southeast University, China
[3] School of Computer Science and Technology, Nanjing University of Posts and Telecommunications, China
{lixuelian,zqgao}@seu.edu.cn, {qianliu,mzhu}@njupt.edu.cn

Abstract. Multiple-choice questions of comparing one entity with another in a university's entrance examination like Gaokao in China are very common but require high knowledge skill. As a preliminary attempt to address this problem, we build a geography Gaokao-oriented knowledge acquisition system for comparative sentences based on logic programming to help solve real geography examinations. Our work consists of two consecutive tasks: identify comparative sentences from geographical texts and extract comparative elements from the identified comparative sentences. Specifically, for the former task, logic programming is employed to filter out non-comparative sentences, and for the latter task, the information of dependency grammar and heuristic position is adopted to represent the relations among comparative elements. The experimental results show that our system achieves outstanding performance for practical use.

Keywords: Knowledge Acquisition, Comparative Sentences, Logic Programming, Answer Set Programming

1 Introduction

The grand challenge of an AI robot, to pass entrance examinations at different levels of education has been approached. Much research has been devoted to AI robots on the education, such as the NII's Todai Robot Project [1] and the Allen Institute for Artificial Intelligence's Project Aristo [2]. Recently, China shared a similar motivation with them and has launched a similar project that would enable the computer to "learn" from textbooks and Web resources, and then pass the National Higher Education Entrance Examination (commonly known as Gaokao). This project is dedicated to four out of nine subjects in Gaokao, namely, Chinese, mathematics, geography and history. Cheng et al. [3] develop a three-stage approach including retrieving, ranking, and filtering concept and quote pages to automatically answer multiple-choice questions of history in Gaokao, which is a general solution to the question. Because of the diversity of questions in Gaokao, for example, some questions required to be more profound understanding and more fine-grained analysis, their approach may be not suitable for such questions. Therefore, answering complex multiple-choice questions of comparing one entity with another in Gaokao such as in Figure 1 requires specialized knowledge and is still far from being solved.

The 30th Summer Olympic Games are to be held in London, the United Kingdom from July 27 to August 12, 2012. Please read Figure and answer Q1.

Q1: In July and August, ______ in London than in Beijing.

A. the temperature is higher and covers a wider daily range B. it is less windy but more foggy and rainy

C. the sun rises later, days get shorter but nights longer D. the noontime altitude of the sun is smaller

Fig. 1. A multiple-choice item on the geographical subtest of the 2012 Gaokao held in Beijing, consisting of a stem and four possible options.

As a preliminary attempt to deal with this problem, we aim at building a geography Gaokao-oriented knowledge acquisition system for comparative sentences to solve real geography examinations. This problem consists of two challenging tasks. The former refers to a sentence classification problem while the latter refers to an information extraction problem.

Task 1. Identify candidate comparative sentences from a given collection of geographical texts like textbooks, reference books and relevant geographical web pages and then filter out non-comparative sentences.

Task 2. Extract comparative elements from the identified comparative sentences. For example, the sentence "伦敦的气温比北京高。(The temperature in London is higher than that in Beijing.)" is a comparative sentence, the word "伦敦(London)" is the subject entity (SE), the word "北京(Beijing)" is the object entity (OE), the word "气温(temperature)" is the comparative aspect (CA), and the word "高(higher)" is the comparative result (CR).

In recent years, the aforementioned tasks have been studied extensively. There are two main approaches: *machine learning approach* and *rule-based approach*. Some existing work [4] has shown that the rule-based approach is more appropriate because of the structural uniqueness of comparative sentences. The key idea of the rule-based approach is that it can describe the explicit domain information of comparative elements in a declarative way. Due to the fact that Logic Programming can offer detail-giving, natural-language explanations for its answers, it is more suitable for answering comparative questions in geographical domain. Specially, the non-monotonicity and scalability of answer set programming (ASP) - a variant of Logic Programming, provide an intuitively appealing way to address these issues.

For the first task, we present an approach that integrates ASP and keyword-based method to implement comparative sentence identification by constructing a linguistic-based comparative keyword lexicon which takes into account the comparative keyword collocation principles.

For the second task, we propose to employ ASP to implement comparative element extraction by adopting dependency grammar and heuristic position to represent the relations among comparative elements.

Moreover, we manually construct a geographical dataset of comparative sentences to answer multiple-choice questions of comparing one entity with another. It achieves good results on a set of real multiple-choice questions collected from recent geography examinations. Also, this dataset, which covers more comparative relations of physical and human geographical topics, has been combined with linked data for the geographical domain, such as Clinga[4], GeoNames[5] and GeoLink[6] to help realize real artificial intellegence, enabling the computer to pass the geography exams in Gaokao.

2 Related Work

Linguistic researchers have studied the syntax and semantics of comparative constructs from the beginning of modern Chinese linguistic research [5, 6]. However, our focus is mainly on computational methods.

The most related works are comparative opinion mining [4, 7]. There are two main approaches: *machine learning approach* and *rule-based approach*. The former is mainly based on some of the most popular approaches such as conditional random fields (CRF) [8, 9] and support vector machine (SVM) [10], while the latter is mainly based on the combination between machine learning methods and rules [11–13].

Wang et al. [10, 9] build a SVM model for comparative sentence extraction and use CRF for comparative element extraction. Jindal and Liu [11, 12] apply Class Sequential Rules (CSR) and Label Sequential Rules (LSR) to extract comparative sentences and relations from English text documents. Varathan [4] gives a good survey of existing methods of comparative opinion mining. And it is shown that the pattern-based approach or rule-based approach are suitable for comparative opinion mining because comparative sentences follow a specific pattern or rule. In this paper, we show that our method obtains a better performance than these approaches for our tasks in geographical domain.

Our work is also related to information retrieval. Specifically, the most relevant work is by Cheng et al. [3] on multiple-choice questions in Gaokao. They propose a three-stage framework for answering multiple-choice questions in history tests. As their method is based on a set of Wikipedia pages, it is a general solution to the problem, thereby, it is not suitable for some questions required to be more profound understanding and more fine-grained analysis, such as questions about comparative sentence identification and comparative element extraction.

[4] http://w3id.org/clinga
[5] http://www.geonames.org/ontology
[6] http://www.geolink.org

3 System Architecture

Answer Set Programming originates from non-monotonic logic and logic programming. It is a logic programming paradigm based on the answer set semantics [14], which offers an elegant declarative semantics to the negation as failure operator in Prolog. An ASP program consists of *rules* of the form:

l_0 :- l_1, ..., l_m, not l_{m+1}, ..., not l_n.

where each l_i for $i \in [0..n]$ is a literal of some signature, i.e., expressions of the form $p(t)$ or $\neg p(t)$ where p is a predicate and t is a term, and *not* is called $negation\ as\ failure$ or $default\ negation$. A rule without body is called a $fact$.

An ASP based system architecture of knowledge acquisition for comparative sentences consists of the following steps:

1) Extract relevant parts of the knowledge base and represent the POS tags of words and collocation relations with comparative keywords as ASP facts;

2) Extract relevant parts from comparative sentences and represent the POS tags, dependency relations of words and heuristic position relations with comparative keywords as ASP facts;

3) Identify non-comparative sentence filtering rules and comparative element extraction rules, respectively, and represent them by ASP rules;

4) Compute the answer set of the logic program resulted from the first and the third steps to perform the first task, and then from the second and the third steps to perform the second task using an ASP solver like clingo[7]. Finally the non-comparative sentences and comparative elements are extracted respectively from the answer set.

We now give an introduction on dependency grammar as it is useful to our proposed approach. Dependency grammar is adopted to describe the syntactic structure of a sentence by using a dependency tree to establish dependency relation between sentence components. The dependency relations among comparative elements contain *SBV* (subject-verb), *ATT* (attribute), *ADV* (adverbial), *POB* (preposition-object), *COO* (coordinate), *RAD* (right adjunct) and *HED* (head). Since we need part-of-speech (POS) tags throughout the paper, let us show the important POS tags of Language Technology Platform (LTP)[8]. n: general noun, nd: direction noun, nh: person name, ni: organization name, nl: location noun, ns: geographical name, nz: other proper noun, a: adjective, d: adverb, p: preposition, u: auxiliary, v: verb, c: conjunction, r: pronoun.

4 Identifying Comparative Sentences from Geographical Texts

In this section, we first introduce how to use keyword-based method to identify all candidate comparative sentences and then discuss in detail how to filter out non-comparative sentences using ASP.

[7] https://potassco.org/

[8] http://www.ltp-cloud.com/demo/

4.1 Identifying Candidate Comparative Sentences

According to the category of comparative questions in geographical tests, comparative sentences can be divided into two broad comparative types: gradable comparison such as “伦敦的气温比北京高。” and superlative comparison such as “中国是世界最大的稻米生产国。”.

As comparative keywords are an important symbol for comparative sentences, both the keyword-based method [13] and CSR-based method [11] use them to identify comparative sentences but show a relatively low precision. This is mainly because lexicon can not perfectly express the meaning or structure of the Chinese comparative sentences. Furthermore, some complicated comparative sentences tend to be more flexible in forms, comparative keywords need to pair with the words such as predicate verb, adjective and preposition to identify comparative sentences. Therefore, this paper makes some improvement and proposes an approach combining ASP with keyword-based method to recognize comparative sentences. Firstly, our strategy is to manually construct a linguistic-based comparative keyword lexicon containing a total of 202 common comparative keywords and their synonyms. The comparative keyword lexicon (CK) consists of a gradable keyword lexicon (CK_1) and a superlative keyword lexicon (CK_2). Subsequently, the lexicon CK is used for scanning the geographical corpus to identify all candidate comparative sentences S. Once the candidate comparative sentences are recognized, the next step is to filter out non-comparative ones using ASP.

4.2 Filtering out Non-comparative Sentences

To efficiently filter out these non-comparative sentences T from the identified candidate comparative sentences S, the collocation relations between pairs of words can be manually constructed, and the information about the POS tag of every word in S can be directly captured by LTP. We first represent them as ASP facts such as the form *keyword* (CK) representing that CK is a comparative keyword. Subsequently, the rules of filtering out non-comparative sentences are denoted by ASP rules.

For example, one filter-out rule of non-gradable comparisons could be “*if a word W has the direct collocation relation in the sentence T with CK_1, such as “于”, a preposition after W, and there is no reason to believe that W is an adjective, then the sentence T is a non-gradable comparison*”, which can be formulated by the following rule r_1:

non-gradableComparison (T) :- collocation (lb, W, CK$_1$, T), keyword (CK$_1$), pos (CK$_1$, p), not pos (W, a).

where *not* is used to exclude the fact that the POS of W is an adjective, *collocation* (lb, W, CK$_1$, T) means W and CK_1 have a collocation relation *lb*, namely, W is located on the left of CK_1 in the sentence T. For example, given the sentence, “南京位于江苏。(Nanjing is located in Jiangsu.)”, we can identify it as a non-comparative sentence using this rule.

To filter out non-comparative sentences, we group collocation rules (or filtering rules) $\mathcal{R}$ into two types ($\mathcal{R}^1$, $\mathcal{R}^2$) based on the classification of comparative sentences as below:

Type 1 rules ($\mathcal{R}^1$): using gradable comparative keywords to filter out non-gradable comparisons (based on some collocation relations between them), e.g., rule r_1. A gradable keyword lexicon CK_1 including the gradable comparative keywords and their POS tags is given a priori.

Type 2 rules ($\mathcal{R}^2$): using superlative comparative keywords to filter out non-superlative comparisons (based on some collocation rules between them). A set of superlative keyword lexicon CK_2 including the superlative keywords and their POS tags are the known seeds. The following is an example of such rules:

non-superlativeComparison (T) :- collocation (la, W, CK_2, T), keyword (CK_2), pos (CK_2, d), pos (W, v).

where *collocation* (la, W, CK_2, T) means W and CK_2 have a collocation relation *la*, namely, W is located on the right of CK_2 in the sentence T.

As shown in Algorithm 1, by repeatedly implementing $\mathcal{R}$ based on the identified candidate comparative sentences $\mathcal{S}$ and comparative keyword lexicon $\mathcal{CK}$, a set of non-comparative sentences $\mathcal{T}$ are identified.

Algorithm 1 Non-Comparison($\mathcal{S}$, $\mathcal{R}$, $\mathcal{CK}$)

Input: Candidate comparative sentences $\mathcal{S}$, pre-defined filtering rules $\mathcal{R}$ ($\mathcal{R}^1$, $\mathcal{R}^2$), comparative keyword lexicon $\mathcal{CK}$ ($\mathcal{CK}_1$, $\mathcal{CK}_2$).
Output: A non-comparative sentence set $\mathcal{T}$.

$\mathcal{T} \leftarrow \{\}$; // initialize an empty non-comparative sentence set $\mathcal{T}$
for each sentence $s \in \mathcal{S}$ **do**
  if $\mathcal{CK}_1$ in s **then**
    implement $\mathcal{R}^1$;
    if s is a non-gradable comparison **then**
      insert s into $\mathcal{T}$;
    end if
  else if $\mathcal{CK}_2$ in s **then**
    implement $\mathcal{R}^2$;
    if s is a non-superlative comparison **then**
      insert s into $\mathcal{T}$;
    end if
  end if
end for
Output $\mathcal{T}$ as the final non-comparative sentence set.

5 Extracting Comparative Elements from the Identified Sentences

In our work, a gradable comparison can be defined as a quadruple $\langle SE, OE, CA, CR \rangle$ and a superlative comparison can also be expressed as a quadruple $\langle SE,\ CS,\ CA,\ CR \rangle$.

For example, given the comparative sentence "伦敦的气温比北京高。", our objective is to extract the following comparative elements (CE):

⟨伦敦(*London*), 北京(*Beijing*), 气温(*temperature*), 高(*higher*)⟩

To achieve the goal, a comparative keyword lexicon CK is given a priori. And a converted lexicon of comparative keywords CK' is manually constructed to address comparative sentences with the same meaning but different structures, for example, "伦敦的气温比北京高。" and "北京的气温不及伦敦高。". They are represented as facts $convertKeyword$ (CK') denoting that CK' needs to be converted to accomplish the change from a negative keyword "不及(lower than) " to a positive keyword "比(higher than)". The information about the POS of words and dependency relations that connect pairs of words or phrases in our corpus are automatically generated by LTP. The knowledge about heuristic position relations is generated according to the location between comparative elements and CK. According to the structure of comparative sentences in our corpus, the gradable comparisons are divided into five types and the superlative comparisons are divided into three types. We identify 35 extraction rules from gradable comparative sentences and 15 extraction rules from superlative comparative sentences. In the following, we will present some examples of the ASP based extraction rules for comparative elements.

$\mathrm{R1}_1$: Rule $\mathrm{R1}_1$ means "if a word CR, whose POS is an adjective, is located on the right of the comparative keyword CK_1, and directly depends on *Root* through dependency relation HED-Dep (namely, CR is the head of a comparative sentence), then CR is a gradable comparative result." This rule can be represented as follows:

result (CR) :- depends (HED-Dep, Root, CR),
location (la, CR, CK_1),
keyword (CK_1), pos (CR, a).

$\mathrm{R2}_1$: Rule $\mathrm{R2}_1$ means "if a word CA, whose POS is a noun, directly depends on a gradable comparative result CR through dependency relation SVB-Dep, and it can be depended by a subject entity SE through dependency relation ATT-Dep, then CA is a gradable comparative aspect." It is represented as follows:

aspect (CA) :- depends (SVB-Dep, CA, CR),
depends (ATT-Dep, SE, CA),
pos (CA, n).

$\mathrm{R3}_1$: Rule $\mathrm{R3}_1$ means "if a word SE, whose POS is a noun, is located on the left of the comparative keyword CK_1, and directly depends on a gradable comparative aspect CA through dependency relation ATT-Dep, then SE is a subject entity of gradable comparative sentence." This rule can be represented as follows:

subject (SE) :- depends (ATT-Dep, SE, CA),
location (lb, SE, CK_1),
keyword (CK_1), pos (SE, n).

$\mathrm{R4}_1$: Rule $\mathrm{R4}_1$ means "if a word OE, whose POS is a noun, is located on the right of the comparative keyword CK_1, and directly depends on a gradable comparative aspect CA through dependency relation $ATT\text{-}Dep$, then OE is an object entity of gradable comparative sentence." This rule can be represented as follows:

object (OE) :- depends (ATT-Dep, OE, CA),
location (la, OE, CK_1),
keyword (CK_1), pos (OE, n).

$R5_1$: Rule $R5_1$ outputs gradable comparative template of comparative elements. This rule can be represented as follows:

gradableTem (SE, OE, CA, CR) :- subject (SE), object (OE),
aspect (CA), result (CR),
not convertKeyword (CK′).

where *not* is used to exclude the comparative keywords that need to be converted. For example, the comparative keyword "比" in "伦敦的气温比北京高", is a positive keyword with no exceptions.

$R5_2$: Rule $R5_2$ outputs gradable comparative template of comparative elements. It is represented as follows:

gradableTem (OE, SE, CA, CR) :- object (OE), subject (SE),
aspect (CA), result (CR),
convertKeyword (CK′).

where CK' belongs to the converted lexicon, thereby, it needs to be converted to accomplish the change from a negative keyword to a positive keyword.

The proposed algorithm of comparative element extraction is called CE-Extraction, short for comparative element extraction. As shown in Algorithm 2, given the identified comparative sentence set $\mathcal{S}$, the extraction rules for comparative elements $\mathcal{R}$, a comparative keyword lexicon $\mathcal{CK}$, and a converted lexicon of comparative keywords $\mathcal{CK}'$, we apply extraction rules $\mathcal{R}$ to extract all possible comparative elements according to the type of comparative keywords existing in each sentence s. Finally the comparative element set $\mathcal{CE}$ is generated.

Algorithm 2 CE-Extraction($\mathcal{S}$, $\mathcal{R}$, $\mathcal{CK}$, $\mathcal{CK}'$)

Input: Comparative sentences $\mathcal{S}$, pre-defined extraction rules $\mathcal{R}$, comparative keyword lexicon $\mathcal{CK}$, converted lexicon of comparative keywords $\mathcal{CK}'$.
Output: A comparative element set $\mathcal{CE}$.

```
1: CE ← {};     // initialize an empty comparative element set CE
2: for each sentence s ∈ S do
3:    if CK in s then
4:       implement R1 - R4 and R5_1;
5:       insert results of R5_1 into CE;
6:    end if
7:    if CK′ in s then
8:       implement R1 - R4 and R5_2;
9:       insert results of R5_2 into CE;
10:   end if
11: end for
12: Output CE as the final comparative element set.
```

6 Experiment

6.1 Datasets

For the identifying task of comparative sentences, 2500 comparative sentences were collected from geographical texts like textbooks, reference books and relevant geographical web pages such as Baidu Baike[9], the largest collaboratively-built Chinese wiki encyclopedia, to represent different types of data. For the extracting task of comparative elements, 500 comparative sentences (Gradable: 400, Superlative: 100) are randomly selected from a population of 2500 comparative sentences as the development dataset to identify extraction rules and the remaining 2000 comparative sentences (Gradable: 1600, Superlative: 400) are used for evaluating the effectiveness of the proposed approach as the testing dataset, in which multiple-word comparative elements are very common. The related resources are partially available[10] for research purposes.

6.2 Labeling

The datasets were all annotated manually. Word segmentation and POS tagging were firstly conducted by using LTP and then double-checked by human labelers to guarantee the quality. IOB tags used in text chunking [16] and named entity recognition [17] tasks were employed for annotating the comparative elements in the 2500 comparative sentences with the corresponding CE labels by four trained human annotators. The sentence below has been labeled with IOB tags corresponding to phrases that should be extracted as comparative elements. Table 1 lists the process of representing comparative elements by using IOB tags. Our work was double-checked by one another, and any disagreement between two annotators was resolved by discussion among the four annotators before reaching an agreement.

Table 1. Feature examples for labeling.

Original sentence	“伦敦的气温比北京高。” (The temperature in London is higher than that in Beijing.)
After word segmentation	伦敦 的 气温 比 北京 高。
After POS	伦敦/ns 的/u 气温/n 比/p 北京/ns 高/a 。/wp
After IOB tags	伦敦/ns 的/u 气温/n 比/p 北京/ns 高/a 。/wp B-SE O B-aspect keyword B-OE B-result

[9] http://baike.baidu.com
[10] http://www.corpora.com.cn/GaoKaoGeographyComSen/

6.3 Identifying Comparative Sentences

In the experiments for Task 1, we compare our approach with other representative approaches: the CSR-based approach [11, 15] and the keyword-based approach [13] for performance evaluation on comparative sentence identification. The comparison results are presented in Table 2, showing that our approach based on the combination of ASP and keywords achieves a higher precision. The keyword-based approach shows that these comparative keywords are good indicators, but the precision is low, which indicates that many sentences that contain comparative keywords are not comparative sentences. However, our method can properly deal with this issue and filter out those non-comparative sentences correctly. Although the results of the CSR-based approach are competitive to our approach, we can employ some collocation relation to filter out non-comparative sentences that the CSR-based approach can not address. Moreover, we also analyzed the incorrectly identified comparative sentences, and found that there are inherently ambiguity, which conforms to Huang et al.'s [15] analysis.

Table 2. Final results in comparative sentence identification (%).

Systems	Gradable			Superlative		
	P	R	F_1	P	R	F_1
keyword	90.8	97.9	88.1	91.7	97.3	90.4
CSR	93.4	97.2	92.3	95.3	96.8	94.2
keyword+ASP	96.2	97.4	94.8	97.9	97.2	96.4

6.4 Extracting Comparative Elements

In order to demonstrate whether our proposed logic programming approach is effective, we first evaluate our approach with different number of comparative sentences and then compare our approach to conditional random fields (CRF). Table 3 and Table 4 show comparisons of precision, recall and F-score results of our approach and CRF in gradable and superlative sentences separately. For CRF, the POS tags, dependency relations of words and heuristic position relations with comparative keywords were used as features for the element extraction, and the ratio of the training set and test set in both gradable and superlative comparative sentences was 2:1. Obviously, with all of the four sets of different number of comparative sentences, our method still reaches good performance and outperforms CRF.

For SE, our approach gave a precision of more than 98% for different number of gradable comparative sentences and a precision of more than 92% for different number of superlative comparative sentences, because subject entities have nice characteristics, e.g., a noun or noun phrase, occurring at the start of a sentence, before a comparative keyword. For CR, our approach gave a precision of more than 97% for different number of both gradable and superlative comparative sentences, because comparative results also have nice characteristics, e.g., an adjective, occurring at the end of a sentence, after a comparative keyword. For CA, it could appear after a subject entity or an object entity in gradable sentences, as long as it is extracted once, we think it is successfully extracted, therefore, it has a relatively higher precision. However, our approach showed bad performance on CA in superlative sentences, because they are omitted frequently and quite fuzzy, sometimes it is not easy for human to identify them. Moreover, a number of omitted comparative scopes and the high multiple-word portion in comparative scopes caused relatively lower recall in superlative sentences.

Table 3. Precision, Recall and F_1-score of our method with different number of gradable sentences and CRF (%).

	SE			OE			CA			CR		
	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1
400	98.67	82.36	87.97	92.11	75.81	83.17	95.23	79.46	88.56	97.81	95.71	96.75
800	98.58	82.25	87.89	92.02	75.69	83.03	95.04	79.23	88.43	97.70	95.54	96.61
1200	98.43	81.16	87.73	91.99	75.54	82.94	95.97	79.12	88.32	97.58	95.41	96.45
1600	98.31	80.99	87.54	91.67	75.42	82.87	95.84	79.01	88.16	97.40	95.32	96.25
CRF	82.41	80.18	81.28	69.81	66.67	68.20	74.55	67.21	70.69	97.20	93.69	95.41

Table 4. Precision, Recall and F_1-score of our method with different number of superlative sentences and CRF (%).

	SE			CS			CA			CR		
	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1
100	92.59	83.33	87.72	87.50	70.00	77.78	71.74	66.67	68.56	98.89	97.74	98.68
200	92.38	83.16	87.69	87.46	70.09	77.53	70.04	66.23	68.43	98.70	97.57	98.64
300	92.33	82.99	87.63	87.23	69.94	77.34	69.87	66.09	68.37	98.56	97.46	98.58
400	92.31	82.87	87.59	87.17	69.85	77.27	69.84	66.02	68.19	97.99	97.39	98.51
CRF	91.30	70.01	79.25	83.33	50.03	62.58	63.16	40.56	48.98	96.26	92.79	94.51

7 Conclusion

This paper has studied a Chinese knowledge acquisition system for comparative sentences in geographical domain, including two important tasks, namely, comparative sentence identification and comparative element extraction. For the first task, both final precision and F_1-score rates are over 94%, and higher than the baseline, indicating the proposed method is effective and performs well in identifying comparative sentences. For the second task, we compare our approach with the state-of-the-art statistical method CRF. The proposed approach is much more effective in extracting comparative elements. These results demonstrated that Answer Set Programming could be used effectively and concisely in practical applications. Since multiple-choice questions of comparing one entity with another in Gaokao are very common, our study can contribute greatly to Chinese geographical data. In future work, we will investigate more extraction rules and increase the amount of geographical data to help answer a question correctly in Gaokao.

Acknowledgement. This work is partially funded by the 863 Program under Grant 2015AA015406 and the National Natural Science Foundation of China under Grant 61702279, 61170165, 61602260, 61502095.

References

1. Fujita, A., Kameda, A., Kawazoe, A., Miyao, Y.: Overview of Todai Robot Project and evaluation framework of its NLP-based problem solving. In Proceedings of the 9th International Conference on Language Resources and Evaluation, (2014)
2. Clark, P.: Elementary school science and math tests as a driver for AI: take the aristo challenge!. AAAI, pp. 4019-4021 (2015)
3. Cheng, G., Zhu, W., Wang, Z., Chen, J., Qu, Y.: Taking up the Gaokao challenge: An information retrieval approach. IJCAI, pp. 2479-2485 (2016)
4. Varathan, K. D., Giachanou, A., Crestani, F.: Comparative opinion mining: A review. Journal of the Association for Information Science and Technology, (2016)
5. Ma, J.:Mashi Wentong[M], Shanghai: The Commercial Press, (1898)
6. Chen, J., Zhou, X. B.: The Selection and Arrangement of Grammatical Items concerning Comparative Sentences. Language Teaching and Linguistic Studies, (2005)
7. Liu, B. Sentiment analysis: mining opinions, sentiments, and emotions. Cambridge University Press, (2015)
8. Liu, C., Xu, R., Liu, J., Qu, P., Wang, H., Zou C.: Comparative opinion sentences identification and elements extraction. In: Proceedings of the ICMLC, IEEE, (2013)
9. Wang, W., Zhao, T. J., Xin, G. D., Xu, Y. D.: Extraction of comparative elements using conditional random fields. Acta Automatica Sinica, 41(8), pp.1385-1393 (2015)
10. Wang, W., Zhao, T. J., Xin, G. D., Xu, Y. D.: Exploiting machine learning for comparative sentences extraction. International Journal of Hybrid Information Technology, (2015)
11. Jindal, N., Liu, B.: Identifying comparative sentences in text documents. In Proceedings of SIGIR'06, pp. 244-251 (2006)
12. Jindal, N., Liu, B.: Mining comparative sentences and relations. AAAI, (2006)
13. Yang, S., Ko, Y.: Extracting comparative entities and predicates from texts using comparative type classification. In Proceedings of HLT'11, pp. 1636-1644 (2011)
14. Gelfond, M., Lifschitz, V.: The stable model semantics for logic programming. Proceedings of the Fifth International Conference on Logic Programming (ICLP), pp.1070-1080 (1988)
15. Huang, X., Wan, X., Yang, J., Xiao, J.: Learning to identify Chinese comparative sentences. Journal of Chinese Information Processing, pp. 30-38 (2008)
16. Zhai, F., Potdar, S., Xiang, B., Zhou, B.: Neural Models for Sequence Chunking. In AAAI, pp. 3365-3371 (2017)
17. Jie, Z., Muis, A. O., Lu, W.: Efficient Dependency-Guided Named Entity Recognition. In AAAI, pp. 3457-3465 (2017)