

Chinese Zero Pronoun Resolution: A Chain to Chain Approach

Kong Fang and Zhou Guodong*

School of Computer Science and Technology
Soochow University, China
{kongfang, gdzhou}@suda.edu.cn

Abstract. Chinese zero pronoun (ZP) resolution plays a critical role in discourse analysis. Different from traditional mention to mention approaches, this paper proposes a chain to chain approach to improve the performance of ZP resolution from three aspects. Firstly, consecutive ZPs are clustered into coreferential chains, each working as one independent anaphor as a whole. In this way, those ZPs far away from their overt antecedents can be bridged via other consecutive ZPs in the same coreferential chains and thus better resolved. Secondly, common noun phrases (NPs) are automatically grouped into coreferential chains using traditional approaches, each working as one independent antecedent candidate as a whole. Then, ZP resolution is made between ZP coreferential chains and common NP coreferential chains. In this way, the performance can be much improved due to the effective reduction of search space by pruning singletons and negative instances. Finally, additional features from ZP and common NP coreferential chains are employed to better represent anaphors and their antecedent candidates, respectively. Comprehensive experiments on the OntoNotes corpus show that our chain to chain approach significantly outperforms the state-of-the-art mention to mention approaches. To our knowledge, this is the first work to resolve zero pronouns in a chain to chain way.

Keywords: Chinese zero pronoun resolution, chain to chain approach, zero pronoun Coreferential Chains, chain-level features

1 Introduction

As a gap in a sentence, a ZP exists when a phonetically null form is used to refer to a real world entity, and an anaphoric ZP (AZP) corefers with a preceding NP. In comparison with overt pronouns, it is more challenging to resolve ZPs due to the lack of syntactic variations and grammatical attributes such as number and gender. Unlike English, Chinese is a pro-drop language. According to our statistics on the Ontonotes corpus, while only 4% of subjects are ZPs in English, this percentage soars to 36% in Chinese. This indicates the necessity of ZP resolution in Chinese understanding. During the past few years, Chinese ZP resolution has been drawing more and more attention due to its importance to various natural language processing (NLP) applications.

* Corresponding author

In the literature, most studies on Chinese ZP resolution employ mention to mention approaches [13, 8, 4, 3, 1], which resolve each detected ZP and its single candidate independently and thus can only rely on local information. Although various kinds of lexical and syntactic features have been employed in ZP resolution to a certain extent, the performance of Chinese ZP resolution is still far from satisfactory. In this paper, we propose a chain to chain approach to Chinese ZP resolution. In particular, all possible ZPs are first identified and then clustered into coreferential chains in a consecutive way. Then common NPs are automatically grouped into coreferential chains using traditional approaches. Finally, ZP resolution is made between ZP coreferential chains and common NP coreferential chains.

In comparison with traditional mention to mention approaches, our chain to chain approach improves the performance of ZP resolution from three aspects.

- Cluster consecutive ZPs into coreferential chains, each working as one independent anaphor as a whole. In this way, those ZPs far away from their overt antecedents can be better resolved by bridging the coreferential relations via other consecutive ZPs.
- Only consider those mentions in the coreferential chains of common NPs as antecedent candidates and resolve them on a chain to chain basis. That is, those NPs occurring in the same coreferential chain are viewed as one antecedent candidate as a whole. In this way, the search space can be much reduced by pruning singletons and negative instances.
- Employ additional features from ZP and common NP coreferential chains.

2 Background Knowledge

In this section, we introduce the ZP resolution and overview some related work.

2.1 ZP Resolution

Example (1) shows an excerpt from article chtb_0001 in the Chinese part of the OntoNotes corpus. In this example, two ZPs are denoted by “ Φ ” and the mention “一系列规范建设市场的文件/a series of documents for standardizing the construction market” in the same coreferential chain with the given ZPs is shown in bold. Besides, the corresponding mentions in translated English provided by the corpus are shown in font style.

- (1) 为规范建筑行为，防止出现无序现象，新区管委会根据国家 and 上海市的有关规定，结合浦东开发实际，及时出台了一**系列规范建设市场的文件**，其中 Φ 包括工程施工招投标管理办法、拆迁工作若干规定、整治违章建筑实施办法、通信设施及管线配套建设意见、建设工地施工环境管理暂行办法等， Φ 基本做到了每个环节都有明确而又具体的规定。
- (` In order to standardize construction procedures and to guard against the emergence of disorderly phenomena , the new region 's management committee promptly announced *a series of documents for standardizing the construction market* in accordance with the relevant national regulations and the regulations of Shanghai Municipality , while accommodating the realities of Pudong 's development.

The documents include: management methods for bidding on construction projects; a certain number of regulations for demolition and removal work; implementation methods for fixing construction that violated regulations; construction suggestions for communications installations and cable setups; provisional methods for environmental management at construction work sites; etc.

Essentially *they* are worked out to the point where every single link has clearly defined and specific regulations.")

Just as illustrated in Example (1), in comparison with its English translation, Chinese has much more ZPs and these ZPs can be translated into English common NPs, demonstrative NPs or pronouns. In the literature, ZP resolution contains three subtasks, i.e., **ZP detection**, which extracts all possible ZPs, e.g., the two ZPs in Example(1). **Anaphoricity determination**, for detected ZP, which determines its anaphoricity, i.e., whether the given ZP corefers with a preceding NP. In Example(1), both detected ZPs are anaphoric. In fact, just as noted in Kong and Zhou [8], about 15% of Chinese ZPs on the OntoNotes corpus are non-anaphoric. **Antecedent identification**, which finds the antecedent for every AZP. In Example(1), the same mention "一系列规范建设市场的文件" should be identified as the antecedent of both the first and the second ZP.

2.2 Related Work

Although ZPs are prevalent in Chinese, there is only a few studies, which can be classified into two categories: supervised learning (Zhao and Ng [13], Kong and Zhou [8], Chen and Ng [4, 3, 5]) and unsupervised learning (Chen and Ng [2, 1]).

For supervised learning, Zhao and Ng [13] proposed a feature-based method to ZP resolution. However, they only focused on the sub-task of antecedent identification. For ZP detection, a simple heuristic rule was employed, suffering from very low precision by introducing too many false ZPs. Kong and Zhou [8] proposed a unified framework for ZP resolution and extracted different kinds of syntactic parse tree structures for three sub-tasks. However, only the performance on gold parse trees was reported. Chen and Ng [4] built the first end-to-end ZP resolver. In particular, they proposed various kinds of syntactic and contextual features and allowed coreference links between multiple zero pronouns. Chen and Ng [5] further proposed an approach to AZP resolution based on deep neural networks to reduce feature engineering efforts involved in exploiting lexical features. In order to eliminate the reliance on annotated data, Chen and Ng [3, 2] presented a generative model for unsupervised Chinese ZP resolution. Chen and Ng [1] further proposed a probabilistic model for this task, which tried to jointly identify and resolve zero pronouns. However, both supervised and unsupervised learning strategies described above employ mention to mention approaches to Chinese ZP resolution, which resolve each detected ZP independently and thus can only rely on local information. This is contrary to the basic fact that ZP is a discourse phenomenon. That is ZPs normally don't exist independently.

Besides, for common NP coreference resolution, although mention-pair models have been successfully employed, their weaknesses have drawn much greater attention in recent research. Some improvements are proposed from different perspectives. Mention-ranking models attempt to rank preceding candidates for a given anaphor (Yang et

al. [12], Denis and Baldrige [6]), and entity-mention models attempt to determine whether a preceding cluster is coreferent with a given mention (Yang et al. [11], Lee et al. [9]).

Inspired by the above research for common NP coreference resolution, in this paper, we improve the performance of ZP resolution in a chain to chain way. Different from their work, we focus on ZP resolution. There are two obvious differences.

First, in our chain to chain approach, each achieved ZP cluster works as one independent anaphor as a whole. While in mention-ranking and entity-mention models, the anaphor is a mention. Second, for ZP resolution, all complete common NP coreferential chains can be achieved in advance. Only the mentions (NPs) in common NP coreferential chains need be considered as the antecedent candidates. Even after linking some ZP coreferential chains to the specific common NP coreference chains, we only need to consider the NPs in common NP coreferential chains. That is to say, the search space of antecedent is static. While in mention-ranking and entity-mention models, all the partial coreference chains, some singletons and other non-anaphoric mentions preceding the given anaphor should be considered. So the search space of antecedent is dynamic. To some extent, our chain to chain approach is much simpler.

3 Motivation

In this section, we motivate our chain to chain approach by analyzing the problems of traditional mention to mention approaches to ZP resolution.

Just as described above, after extracting possible ZPs, traditional mention to mention approaches view each ZP as an independent anaphor and identify its overt antecedent independently. There are three issues in such a way.

- In Chinese, a long sentence can contain multiple ZPs (e.g., Example (1) has two ZPs). While these consecutive ZPs may have coreferential relations (e.g., the given two ZPs in Example (1)), those ZPs far away from their overt antecedents need to bridge their coreferential relations via other consecutive ZPs. For example, the second ZP in Example(1) refers to the overt antecedent "一系列规范建设市场的文件/a series of documents for standardizing the construction market" of the first ZP by treating the first ZP in Example(1) as bridge. In fact, just as noted in Chen and Ng [4], 22.7% of ZPs in the OntoNotes corpus appear two or more sentences away from their closest overt antecedents. Moreover, similar to common NP resolution, it is much more difficult to resolve ZPs over longer distances. It is not appropriate to view every ZP as an independent anaphor. Instead, those relevant coreferential relations among multiple ZPs should be exploited to improve ZP resolution.
- Similar to common NP resolution, the search space much influences the resolution performance. Previous research on common NP resolution shows that there is a significant performance gap between gold mentions and automatic mentions due to a large portion of mentions are non-coreferent. Just as noted in Moosavi and Strube [10], more than 80% of mentions are singletons in the OntoNotes English development set. This proportion on Chinese part climbs up to 89.7%. On the one hand, traditional mention to mention approaches normally extract all NPs

preceding the given ZP in current and previous two sentences to form the search space for the given ZP. This introduces many non-coreferent mentions. On the other hand, there exist various relations among these mentions, e.g., coreferential relation. Since common NP resolution is much easier than ZP resolution, it is expected that these achieved coreferential relations among common NPs can effectively prune the search space and thus much improve the performance of ZP resolution.

- Traditional mention to mention approaches only consider a pair of mentions and the local information between them. Due to its limitation to the discourse of ZPs, local information is always not enough for correct decision. This motivates the incorporation of chain-level information, i.e. various features defined over clusters of mentions to improve the performance of ZP resolution.

Motivated by above observations, we propose a chain to chain approach to ZP resolution. After detecting all possible ZPs, we firstly cluster consecutive ZPs into coreferential chains and view each ZP coreferential chain as an unified anaphor. In this way, those ZPs far away from their overt antecedents can be bridged via other consecutive ZPs in the same coreferential chains. Then, we only consider those mentions in common NP coreferential chains as the antecedent candidates and resolve them on a chain to chain basis. In this way, singletons and negative instances (i.e., non-coreferent mentions) can be better pruned to reduce the search space and thus improve the performance effectively. Finally, we introduce a set of additional features from common NP and ZP coreferential chains to better represent ZPs and their antecedent candidates.

4 ZP Resolution: Chain to Chain Approach

Similar to previous studies on ZP resolution, our chain to chain approach contains two components, i.e., ZP detection and AZP resolution which combines anaphoricity determination and antecedent identification. For ZP detection, a clause-based approach is employed to generate ZP candidates [7]. Firstly, a simplified semantic role labeling (SRL) framework is adopted to determine clauses from a parse tree. ZP candidates are then generated for each clause in a bottom-up way. Particularly, for upper clauses, all the resolved sub-clauses are viewed as an inseparable ``constituent''. After generating the ZP candidates, a learning-based classifier is adopted to identify whether a given candidate is a true ZP, with the help of following features.

- Lexical: two words and their POSs before or after the given candidate, and their various combinations.
- Syntactic: whether the lowest clause covering the given candidate has a subject; whether the given candidate is the first gap of the clause; whether the clause is a terminal clause or non-terminal clause; whether the clause has a sibling immediately to its left; whether the left siblings of the clause contain an NP; whether the clause has a sibling immediately to its right; whether the right siblings of the clause contain a VP; whether the syntactic category of the immediate parent of the clause is an IP or VP; whether the path from the clause to the root of the parse tree contains an NP or VP or CP; whether the clause is a matrix clause, an independent clause, a subordinate clause, or others.

- Semantic: whether the clause has an agent or patient argument.

For AZP resolution, our chain to chain approach consists of three steps, i.e., ZP coreferential chain generation, ZP coreferential chain linking, and additional chain-level features incorporation.

4.1 ZP Coreferential Chains Generation

In ZP coreferential chain generation stage, we only consider consecutive ZPs in current and previous two sentences. For multiple consecutive ZPs, we first establish coreferent links between two consecutive ZPs, and then build the coreferential chain by merging the coreferent links using the transitivity principle. Obviously, if only having two consecutive ZPs, just as shown in Example(1), the achieved coreferent link is also the final ZP coreferential chain. For the case containing more than two consecutive ZPs¹, e.g., three ZPs denoted as $ZP1$, $ZP2$ and $ZP3$ respectively, we first pair two consecutive ZPs to generate two instances $(ZP1, ZP2)$, $(ZP2, ZP3)$, then determine the coreferential relation of every instance, and finally merge the two links to achieve the coreferential chain $ZP3 - ZP2 - ZP1$.

Here, establishing coreferent links between two consecutive ZPs is the key point. Fortunately, the involved two ZPs are always close to each other with similar context. In particular, we extract following contextual features to compute the contextual similarity of the two ZPs.

- The pair of grammatical roles of the two given ZPs. We only consider three grammatical roles, i.e., subject, object and other.
- The pair of clause categories of the clauses governing the two given ZPs. Clause category can be independent, subordinate or none.
- The pair of root nodes of the clauses comprising the two ZPs.
- Whether the clauses governing the two ZPs are siblings.
- The path from the root of one clause to the root of the other clause.
- Punctuation list between the two clauses governing the ZPs.

4.2 ZP Coreferential Chains Linking

After achieving ZP coreferential chains, we try to link every ZP coreferential chain to the common NP coreferential chains. Since ZP coreferential chains are always short and the ZPs in the same coreferential chain tend to have similar context, we only consider linking the first ZP of a ZP coreferential chain to a common NP coreferential chain.

Different from the mention to mention approaches, our chain to chain approach retrieves a complete common NP coreferential chain as the antecedent of a given ZP, or none which means the given ZP is non-anaphoric. Therefore, there exist three obvious differences between our chain to chain approach and traditional mention to mention approaches in instance generation.

¹ A corpus study of the OntoNotes corpus reveals that the case containing more than three consecutive ZPs occupies less than 1%. The proportion of three consecutive ZPs is about 5.8%. The case of two consecutive ZPs occupies about 14.3%. The case of just one ZP occupies the most.

- Elementary processing unit. Traditional mention to mention approaches conduct instance generation for every ZP. That is, traditional approaches first extract antecedent candidates for every ZP, and then pair the achieved candidates with the given ZP to generate the learning instances. In comparison, our chain to chain approach extracts antecedent candidates for every ZP coreferential chain (i.e., every ZP coreferential chain is viewed as an anaphor as a whole), and pair the achieved candidates with the first ZP of the ZP coreferential chain (i.e., using the first ZP to represent the whole ZP coreferential chain) to generate instances.
- Antecedent candidate extraction. Traditional mention to mention approaches extract the NPs in a search space with some heuristic rules as antecedent candidates. While our chain to chain approach only considers the mentions in coreferential chains. In this way, those non-anaphoric NPs and singletons can be well ignored.
- Pairing strategy. From chain to chain perspective, we do not need to pair every antecedent candidate with the anaphor. For the antecedent candidates in the same coreference chain, only the nearest one need to be considered. That is, we only pair the first ZP of one ZP coreferential chain with the nearest mentions (NPs) in different coreferential chains to generate the instances.

| Feature | Description | Value |
|----------|--|---------|
| ZPNum | the number of ZP occurring in the given ZP coreferential chain. Intuitively, the given ZP coreferential chain having multiple ZPs more likely to be anaphoric. | 2 |
| ZPRSft | If the given ZP coreferential chain has multiple ZPs, whether all the ZPs have the same grammar role. | True |
| ZPRPt | The grammar roles of every ZP in the given ZP coreferential chain. | Sub-Sub |
| SemClass | The semantic class of the coreferential chain containing the antecedent candidate. | Object |
| Number | Singular or plural of the coreferential chain containing the antecedent candidate. | Pluar |
| Gender | Male, female or unknown, determined by the coreferential chain containing the antecedent candidate. | Unknown |
| ZPCDst | the number of other ZP coreferential chains between the anaphor and current candidate. | 0 |
| NPCDst | the number of other coreferential chains of common NPs between the anaphor and current candidate. | 0 |

Table 1. Additional chain-level features employed in our chain to chain approach

4.3 Incorporating Additional Chain-level Features

For every instance, we extract various kinds of lexical and syntactic features which have been proven useful in previous work.

- Features on ZP: whether the path of nodes from the ZP to the root of the parse tree contains NP, IP, CP, or VP; whether the ZP is the first or last ZP of the sentence; whether the ZP is in the headline of the text.
- Features on antecedent candidate (CA): whether the CA is a first person, second person, third person, neutral pronoun, or others; whether the CA is a subject, object, or others; whether the CA is in a matrix clause, an independent clause, a subordinate clause, or none of the above; whether the path of nodes from the CA to the root of the parse tree contains NP, IP, CP, or VP.
- Features between ZP and CA: their distance in sentence²; whether the CA is the closest preceding NP; whether the CA and the ZP are siblings in the parse tree.

Besides, although we only pair the first ZP of a ZP coreferential chain with its antecedent candidates occurring in different coreferential chains of common NPs to generate the learning instances, additional features describing the ZP coreferential chain and the NPs coreferential chain containing the antecedent candidate are introduced to better represent the ZP and the antecedent candidate.

| Source | Chen and Ng [1] | | | Chen and Ng [5] | | | Our System | | |
|---------|-----------------|------|------|-----------------|------|------|------------|------|------|
| | R(%) | P(%) | F | R(%) | P(%) | F | R(%) | P(%) | F |
| NW | 11.9 | 14.3 | 13.0 | 11.9 | 12.8 | 12.3 | 17.2 | 24.1 | 20.1 |
| MZ | 4.9 | 4.7 | 4.8 | 9.3 | 7.3 | 8.2 | 10.3 | 18.4 | 13.2 |
| WB | 20.1 | 14.3 | 16.7 | 23.9 | 16.1 | 19.2 | 21.2 | 27.4 | 23.9 |
| BN | 18.2 | 22.3 | 20.0 | 22.1 | 23.2 | 22.6 | 19.4 | 29.1 | 23.3 |
| BC | 19.4 | 14.6 | 16.7 | 21.2 | 14.6 | 17.3 | 25.3 | 21.2 | 23.1 |
| TC | 31.8 | 17.0 | 22.2 | 31.4 | 15.9 | 21.1 | 35.7 | 29.4 | 32.2 |
| Overall | 19.6 | 15.5 | 17.3 | 21.9 | 15.8 | 18.4 | 25.8 | 26.9 | 26.3 |

Table 2. Performance of three end-to-end Chinese ZP resolution approaches.

Table 1 shows additional chain-level features introduced in our chain to chain approach. In Example (1), there is a ZP coreferential chain having two ZPs. The third column in Table 1 lists the feature values viewing the ZP coreferential chain as the anaphor, and the mention “一系列规范建设市场的文件/a series of documents for standardizing the construction market” as the candidate.

5 Experimentation and discussion

We evaluate our proposed approach in this section.

² if the CA and the ZP are in the same sentence, the value is 0; if they are one sentence apart, the value is 1; and so on

5.1 Experimental Setup

Following Chen and Ng [1, 5], we employ the Chinese portion of the OntoNotes 5.0 corpus, which was used in the official CoNLL-2012 shared task, in all our experiments. Since only the training and development sets in the CoNLL-2012 data contain ZP coreference annotations, we train our models on the training set and perform evaluation on the development set. Besides, we employ the automatic parse trees provided by the CoNLL-2012 shared task as the default one and report our performance using traditional precision, recall and F1-score. In addition, maximum entropy is employed as our learning-based algorithm. All maximum entropy classifiers are trained using the OpenNLP maximum entropy package³ with the default parameters (i.e. without smoothing and with 100 iterations). For end-to-end performance, automatic common NP coreference chains are achieved using the Stanford Deterministic Coreference Resolution System⁴ with the default Chinese models (i.e. without additional training). Its performance of Chinese common NP resolution can be learned from Lee etc. [9]. To see whether an improvement is significant, we conduct significance testing using paired t-test.

5.2 Experimental Results

Table 2 shows the performance of our chain to chain approach. For comparison, Table 2 also includes two state-of-the-art mention to mention approaches, where Chen and Ng [1] is the representative of unsupervised mention to mention approach, and Chen and Ng [5] is the representative of supervised mention to mention approach. These two approaches achieved by far the best performance. We can find that,

- For overall performance, our approach beats the state-of-the-art unsupervised approach, i.e., Chen and Ng [1] by 9.0% in F1-score. In comparison with the state-of-the-art supervised approach, i.e. Chen and Ng [5], our chain to chain approach still outperforms it significantly by 7.9% in F1-score.
- Over different sources, our approach significantly outperforms Chen and Ng [1, 5] on all 6 sources by 7.1%/7.8% (NW), 8.4%/5.0% (MZ), 7.2%/4.7% (WB), 3.3%/0.7% (BN), 6.4%/5.8% (BC), and 10.0%/11.1% (TC) in F1-score, respectively. This suggests that our approach works well across different sources.

Similar to traditional ZP resolution, our approach contains two components, ZP detection and AZP resolution. With traditional ZP detection adopted, our chain to chain approach focuses on improving the performance of the second stage. For ZP detection, our ZP detector achieves 59.4%, 70.1% and 64.3% in precision, recall and F1-score respectively when gold parse trees are employed, and the performance drops to 40.2%(P), 60.2%(R) and 48.2%(F) using automatic parse trees.

Just as described above, after identifying all possible ZPs, our chain-to-chain approach divides AZP resolution into three steps, i.e., ZP coreferential chain generation, ZP coreferential chain linking and additional chain-level feature incorporation. In the following, we evaluate these three steps one by one.

³ <http://maxent.sourceforge.net/>

⁴ <http://stanfordnlp.github.io/CoreNLP/>

| | Gold ZPs | | | Auto ZPs | | |
|------|----------|------|------|----------|------|------|
| | R | P | F | R | P | F |
| GS | 92.4 | 92.4 | 92.4 | 53.7 | 64.7 | 58.7 |
| Auto | 89.7 | 89.7 | 89.7 | 36.2 | 56.4 | 44.1 |

Table 3. Performance of ZP coreferential chain generation(GS - using gold standard parse trees, Auto - using automatic parse trees)

Table 3 shows the performance of ZP coreferential chain generation. In our evaluation, we say that a ZP coreferential chain is correct only when all the ZPs in the chain are same as the annotated ZP coreferential chain. We can find that,

- Under gold ZPs, ZP coreferential chain generation achieves the performance of 92.4% and 89.7% in F1-score using both gold and automatic parse trees. In comparison with using gold parse trees, the performance of ZP coreferential chain generation reduces 2.7% in F1-score using automatic parse trees.
- Under auto ZPs, ZP coreferential chain generation achieves 58.7% and 44.1% in F1-score using gold and automatic parse trees, respectively. In comparison with the performance under gold ZPs, we can see the significant performance drop of 14.6% in F1-score in ZP coreferential chain generation due to the errors introduced in ZP detection.

After achieving the ZP coreferential chains, we link each ZP coreferential chain into common NP coreferential chains. Table 4 lists the results under gold common NP coreferential chains. We can find that, under gold ZPs, our chain-to-chain approach to AZP resolution can achieve 90.4% and 86.7% in F1-score using gold and automatic parse trees, respectively. And under auto ZPs, our approach also achieves 55.4% and 42.0% in F1-score using gold and automatic parse trees, respectively.

| | Gold ZPs | | | Auto ZPs | | |
|------|----------|------|------|----------|------|------|
| | R | P | F | R | P | F |
| GS | 90.4 | 90.4 | 90.4 | 50.2 | 61.7 | 55.4 |
| Auto | 86.7 | 86.7 | 86.7 | 33.9 | 55.1 | 42.0 |

Table 4. Performance of Chinese AZP resolution with gold coreference chains of common NPs

While knowing the standard common NP coreferential chains is an ideal case, evaluating AZP resolution using auto coreferential chains of common NPs is more practical and thus meaningful. Table 5 shows the performance of our chain to chain approach to AZP resolution when automatic coreferential chains of common NPs are considered. We can find that,

- Under gold ZPs, our chain-to-chain approach achieves 61.8% and 52.0% in F1-score using gold standard and automatic parse trees, respectively. In comparison

| | Gold ZPs | | | Auto ZPs | | |
|------|----------|------|------|----------|------|------|
| | R | P | F | R | P | F |
| GS | 60.8 | 62.9 | 61.8 | 37.9 | 46.8 | 41.9 |
| Auto | 51.9 | 52.1 | 52.0 | 25.4 | 23.2 | 24.3 |

Table 5. Performance of AZP resolution with automatic common NP coreference chains

with the results using gold coreferential chains of common NPs shown in Table 4, the F1-score drops by 28.6% and 34.7%, respectively. This suggests the performance of common NP resolution significantly influences ZP resolution.

- Under auto ZPs, our approach achieves 41.9% and 24.3% in F1-score using gold and automatic parse trees, respectively. In comparison with the results using gold common NP coreferential chains, the performance of our chain-to-chain approach drops 13.5% and 17.7% using gold and automatic parse trees, respectively, due to the performance of common NP coreference resolution. In comparison with the results under gold ZPs, the performance drops by 19.9% and 27.7% using gold standard and automatic parse trees, respectively. This suggests the significant influence of ZP detection. Ignoring the influence of ZP detection and other factors, we further compare our results with Chen and Ng [5] under both gold ZPs and gold parse trees. In this setting, our approach achieves 61.8% in F1-score, while their approach achieved 52.2%.

Table 6 shows the performance after additional chain-level features as shown in Table 1 are incorporated when automatic coreference chains of common NPs are adopted. In comparison with Table 5, we can find that, under both gold and auto ZPs, additional chain-level features can improve the performance using both gold standard and automatic parse trees by about 2-3% in F1-score. This suggest the effectiveness of the incorporated chain-level features.

| | Gold ZPs | | | Auto ZPs | | |
|------|----------|------|------|----------|------|------|
| | R | P | F | R | P | F |
| GS | 64.5 | 63.5 | 64.0 | 39.7 | 48.3 | 43.6 |
| Auto | 54.4 | 57.2 | 55.8 | 25.8 | 26.9 | 26.3 |

Table 6. Performance of Chinese ZP resolution after additional chain-level features are incorporated under automatic coreference chains of common NPs

6 Conclusion and future work

In this paper, we improve Chinese zero pronoun resolution from chain-to-chain perspective, i.e., from ZP coreferential chains to common NP coreferential chains. The

experimental results on the OntoNotes corpus show that our approach significantly outperforms the state-of-the-art mention to mention approaches.

Although our chain-to-chain approach much improves the performance of AZP resolution, the evaluation suggests that both ZP detection and common NP resolution have heavy impact on the final ZP resolution performance. In future work, we will focus on ZP detection and jointly resolving common NPs and ZPs as a whole.

Acknowledgements

This work is supported by Key Project 61333018 under the National Natural Science Foundation of China, Project 61472264 and 61673290 under the National Natural Science Foundation of China.

References

1. Chen, C., Ng, V.: Chinese zero pronoun resolution: A joint unsupervised discourse-aware model rivaling state-of-the-art resolvers. In: Proceedings of ACL 2015. pp. 320--326. Association for Computational Linguistics
2. Chen, C., Ng, V.: Chinese zero pronoun resolution: A unsupervised approach combining ranking and integer linear programming. In: Proceedings of AAAI 2014. pp. 1622--1628
3. Chen, C., Ng, V.: Chinese zero pronoun resolution: An unsupervised probabilistic model rivaling supervised resolvers. In: Proceedings of EMNLP 2014. pp. 763--774. Association for Computational Linguistics
4. Chen, C., Ng, V.: Chinese zero pronoun resolution: Some recent advances. In: Proceedings of EMNLP 2010. pp. 1360--1365. Association for Computational Linguistics
5. Chen, C., Ng, V.: Chinese zero pronoun resolution with deep neural networks. In: Proceedings of ACL 2016. pp. 778--788. Association for Computational Linguistics
6. Denis, P., Baldridge, J.: Specialized models and ranking for coreference resolution. In: Proceedings of EMNLP 2008. pp. 660--669. Association for Computational Linguistics
7. Kong, F., Zhou, G.: A clause-level hybrid approach to Chinese empty element recovery. In: Proceedings of IJCAI 2013. pp. 2113--2119
8. Kong, F., Zhou, G.: A tree kernel-based unified framework for Chinese zero anaphora resolution. In: Proceedings of EMNLP 2010. pp. 882--891. Association for Computational Linguistics
9. Lee, H., Chang, A., Peirsman, Y., Chambers, N., Surdeanu, M., Jurafsky, D.: Deterministic coreference resolution based on entity-centric, precision-ranked rules. vol. 39, pp. 885--916 (2013)
10. Moosavi, N.S., Strube, M.: Search space pruning: A simple solution for better coreference resolvers. In: Proceedings of NAACL 2016. pp. 1005--1011. Association for Computational Linguistics
11. Yang, X., Su, J., Lang, J., Tan, C.L., Liu, T., Li, S.: An entity-mention model for coreference resolution with inductive logic programming. In: Proceedings of ACL 2008. pp. 843--851. Association for Computational Linguistics
12. Yang, X., Zhou, G., Su, J., Tan, C.L.: Coreference resolution using competition learning approach. In: Proceedings of ACL 2013. pp. 176--183. Association for Computational Linguistics
13. Zhao, S., Ng, H.T.: Identification and resolution of Chinese zero pronouns: A machine learning approach. In: Proceedings of EMNLP-CoNLL 2007. pp. 541--550. Association for Computational Linguistics