Implicit Role Linking on Chinese Discourse: Exploiting Explicit Roles and Frame-to-Frame Relations

Ru Li[1,2]**, Juan Wu**[1]**, Zhiqiang Wang**[1] **and Qinghua Chai**[3]

[1]School of Computer and Information Technology

[2]Key Laboratory of Ministry of Education for Computation Intelligence and Chinese Information Processing

[3]School of Foreign Languages, Shanxi University, Taiyuan, China

{liru,charles}@sxu.edu.cn, {wujuan_0922,zhiq.wang}@163.com

Abstract

There is a growing interest in researching *null instantiations*, which are those implicit semantic arguments. Many of these implicit arguments can be linked to referents in context, and their discoveries are of great benefits to semantic processing. We address the issue of automatically identifying and resolving implicit arguments in Chinese discourse. For their resolutions, we present an approach that combines the information about overtly labeled arguments and frame-to-frame relations defined by FrameNet. Experimental results on our created corpus demonstrate the effectiveness of our approach.

1 Introduction

In natural discourse, only a small proportion of the theoretically possible semantic arguments of predicates tend to be locally instantiated. Other locally unrealized semantic roles are called *null instantiations* (NIs). Nevertheless, many of these implicit roles, while linguistically unexpressed, can often be bound to antecedent referents in the discourse context. What's more, capturing such implicit semantic roles and linking them to their antecedents can dramatically help text understanding.

Example (1) shows an analyzed result (Li, 2012) by employing Chinese FrameNet (Liu, 2011), which is a lexical semantic knowledge base based on the frame semantics of Fillmore (1982) and takes Berkeley's FrameNet Project (Baker et al., 1998) as the reference. In Chinese FrameNet, the predicates, called lexical units (LU), evoke frames which roughly correspond to different events or scenarios. Each frame defines a set of arguments called Frame Elements (FE). The set of FEs is further split into core FEs and non-core

FEs. Particularly, the core FEs are the essential components of a frame and can be defined by themselves. However, not all core FEs of a frame can be realized simultaneously in a sentence. These non-instantiated FEs are considered as *null instantiations of the frame elements*. Depending on the interpretation type of the omission, Chinese FrameNet divides the NIs into two categories: 1) Indefinite Null Instantiations (INIs), the missing element which can be understood given interpretational conventions and do not need resolution, and 2) Definite Null Instantiations (DNIs), the missing element which is something that can be understood in the linguistic or discourse context, and the fillers need to be inferred from the context through resolutions.

(1) [天葬卫星]$_{Entity}$也是$_{属于某类}$ [人造卫星]$_{Category}$，和侦察卫星、通信卫星、气象卫星等同属一类。

[The celestial burial satellite]$_{Entity}$ **is** *$_{Being\_in\_category}$* [artificial satellite]$_{Category}$, and belongs to the same category with reconnaissance, communications and meteorological satellite.

不同的是用途各异，是专门用来存放骨灰盒的；高度亦不同，一般发射$_{使位移}$ [到距地球表面3000多公里处的环球轨道上]$_{Goal}$。[Theme DNI] [Agent INI]

The purpose is different, specially used for storing the urn; due to the different heights, generally **launched** *$_{Cause\_motion}$* into [the orbit over 3000 kilometers away from the surface of earth]$_{Goal}$. [*Theme* DNI] [*Agent* INI]

Particularly, in example (1), lexical unit (or target) *launched*/发射 evokes the semantic frame Cause_motion, which has nine core FEs, namely *Agent*, *Theme*, *Source*, *Path*, *Goal*, *Area*, *Cause*, *Result*, *Initial_State*, but only one of them is instantiated, i.e. *Goal*, whose filler is [*the orbit over 3000 kilometers away from the surface of earth*/到距地球表面3000多公里处的环球轨道上]. For another core FE *Theme*, it is filled by [*The celestial burial satellite*/天葬卫星] that occurs in

the previous sentence.

Clearly, human beings have no problem to infer these uninstantiated roles and find the corresponding fillers based on the relevant context information, but this is beyond the capacity of state-of-the-art semantic role labeling systems.

Next, we formalize the problem as follows: given a discourse $D = \{S_1, S_2, ..., S_n\}$, where S_k ($k \in [1, n]$) is the k-th sentence in D. The lexical unit set in S_k is $T_k = \{T_{k1}, T_{k2}, ..., T_{kp}\}$, and $F_k = \{F_{k1}, F_{k2}, ..., F_{kp}\}$ is relevant frame set. For a particular frame F_{ki} ($i \in [1, p]$), its core FE set is $E_{ki} = \{e_1, e_2, ..., e_m\}$, but it is possible that only part of core FEs C_{ki} appears in S_k, i.e. $C_{ki} \subseteq E_{ki}$. Apparently the set $E_{ki} - C_{ki}$ includes the uninstantiated core FEs. Thus, we need to determine which elements in $E_{ki} - C_{ki}$ are null instantiations. If e_m ($e_m \in E_{ki} - C_{ki}$) has been identified as a null instantiated FE, we should determine whether e_m is a DNI. If so, we need to find the corresponding antecedent d_m in context.

The major contributions of this paper can be summarized as follows:

(i) We have created a null instantiation (NI) annotations corpus, consisting of 164 Chinese discourses across different fields.

(ii) We use frame-to-frame relations to find antecedents from those explicit semantic roles.

2 Related Work

Among the researches of null instantiation on English, the most representative work is the task "Linking Events and Their Participants in Discourse" shared by the SemEval-2010 (Ruppenhofer et al., 2010). The two systems participated in the NI resolution task, VENSES++ and SEMAFOR, took very different approaches.

Tonelli and Delmonte (2010) develop a knowledge-based system called VENSES++, and describe two strategies depending on the predicate class (either nominal or verbal). For verbal predicates, they try to map the predicate argument structure extracted by VENSES with the valence patterns generated from FrameNet data, to identify missing arguments. And NIs are resolved by reasoning about the semantic similarity between an NI and a potential filler using WordNet. For nominal predicates, they resolve NIs by utilizing a common sense reasoning module that builds on ConceptNet (Liu and Singh, 2004). The final Precision and Recall are 4.62%

and 0.86% respectively.

Later on, Tonelli and Delmonte (2011) propose a simpler role linking strategy that based on computing a relevancy score for the nominal head of each potential antecedent. The intuition is that heads which often serve as role fillers and occur close to the target NI are more likely to function as antecedents for the NI. Finally they reported an F-score of 8% for role linking. However, being strongly lexicalized, their trained model seems heavily dependent on the training data.

The second system (Chen et al., 2010) is statistical based and extends an existing semantic role labeler (Das et al., 2010). Resolving DNIs is modeled in the same way as labeling overt arguments, with the search space being extended to nouns, pronouns, and noun phrases from the previous three sentences. When evaluating a potential filler, the syntactic features used in argument labeling of overt arguments are replaced by two semantic features: firstly the system checks whether a potential filler fills the null instantiated role overtly in at least one of the FrameNet sentences and train data, if not, the system calculates the distributional similarity between filler and role. While this system achieved 5% in F-score, data sparseness is a potential limiting factor.

Also closely related studies are as follows. Silberer and Frank (2012) cast NI resolution as a coreference resolution (CR) task, and employ an entity-mention model. They experiment with features of SRL and CR, and automatically expand the training set with examples generated from coreference corpus to avoid data sparseness, ultimately achieving F-score of 7.1%.

Gorinski et al. (2013) present a weakly supervised approach that investigates and combines a number of linguistically motivated strategies, which consist of four basic NI resolvers that exploit different types of linguistic knowledge, and achieve F-score of 12%.

Wang et al. (2013) conduct DNI resolution on SemEval2010 task10 data. They considered the task as a classified problem, by adding new features such as the information of head word and frame to traditional features, proposed a rule to choose the best candidate words set and combination of features, achieving F-score of 14.65% finally.

Laparra and Rigau (2013) present an attempt to apply a set of features that have been traditionally

used to model anaphora and coreference resolution tasks to implicit argument resolution, and got the best results: F-score of 18%.

For nominal predicates, Gerber and Chai (2010) investigate the linking of implicit arguments using the PropBank role labeling scheme. In contrast to the SemEval task, which focuses on a verbs and nouns, their system is only applied to nouns and is restricted to 10 predicates with 120 annotated instances per predicate on average. They propose a discriminative model that selects an antecedent for an implicit role from an extended context window. The approach incorporates some aspects relating to CR that go beyond the SRL oriented SemEval systems: A candidate representation includes information about all the candidates' coreferent mentions (determined by automatic CR), in particular their semantic roles (provided by gold annotations) and WordNet synsets. Patterns of semantic associations between filler candidates and implicit roles are learned for all mentions contained in the candidate's entity chain. They achieve an F-score of 42.3%, which is noticeably higher than those obtained on the SemEval data.

And Gerber (2011) presents an extended model that incorporates strategies suggested in Burchardt et al. (2005): using frame relations as well as coreference patterns acquired from large corpora. This model achieves an F-score of 50.3%.

Lei et al. (2013) conduct DNI identification on SemEval2010 task10 data. They adopt the method of combining rules and machine learning. Different from them, we conduct two-level identifying for NI detection and use more features on Chinese data. Wang et al. (2013) take noun phrases and pronoun as candidate words for DNI filler. We use several similar features with them. The differences are that 1) we take the fillers of overt instantiated FE as candidate words and 2) we use Frame-to-Frame relations. And Gerber (2011) also used frame relations. Different from them, we limit relation paths to 2.

3 Null Instantiation Detection

Now, we are ready to address the first subtask, i.e. null instantiation detection.

3.1 Frame element relations

Not all core arguments of all frames can be realized simultaneously. Some frames involve core FEs that are mutually exclusive. In example (2), in the `Amalgamation` frame, there are four core FEs, namely *Part_1*, *Part_2*, *Parts* and *Whole*, in which the first two FEs are mutually exclusive with *Parts*, thus formed an *Excludes* relation (relation 1). At the same time, *Part_1* and *Part_2* are in a *Requires* relation (relation 2), which means that if one of these two core FEs is present, then the other must occur as well. FE *Whole*, the result of the `Amalgamation`, is only existentially bound within the discourse, annotated as NI.

CoreSet (relation 3) specifies that *at least* one of the set must be instantiated overtly, though more of them can also be instantiated. As shown in example (3), in the `Awareness` frame, the two FEs *Content* and *Topic* are in one *CoreSet*. As *Content* is overtly realized, we consider *Topic* is not annotated as NI. The frame owning this relation is complicated. Sometimes, if one FE of this set is explicit, the absence of the other FEs in the set is not annotated as NI, but sometimes it is not true.

(2) [旧体制]$_{Part\_1}$ 和[新体制]$_{Part\_2}$ **结合**$_{结合}$ 在一起。 [*Whole* INI]

   [The old system]$_{Part\_1}$ and [the new system]$_{Part\_2}$ are **combined** $_{Amalgamation}$ together. [*Whole* INI]

(3) [你老板]$_{Cognizer}$ **知道**$_{知道}$ [你的任务]$_{Content}$。

   [Your boss]$_{Cognizer}$ is **aware**$_{Awareness}$ [of your commitment]$_{Content}$.

3.2 Modeling Null Instantiation detection

As shown in example (1), given a frame F_{ki} (e.g. `Cause_motion` evoked by *launched*/发射), NI detector needs to determine whether core FEs in $E_{F_{ki}} - subE_{F_{ki}}$ are missing, relying on information about the three types of the relations among core FEs: $CoreSet_{F_{ki}}$, $Excludes_{F_{ki}}$, $Requires_{F_{ki}}$ (as discussed in Section 3.1). In `Cause_motion`, the core FEs *Initial_State*, *Goal*, *Path*, *Source* and *Result* belong to the same $CoreSet$, and *Goal* is instantiated, thus *Initial_State*, *Path*, *Source* and *Result* are not annotated as NIs. Meanwhile core FEs *Goal* and *Area* are connected by the $Excludes$ relation, so do *Cause* and *Agent*. Therefore, according to the context, *Area* and *Cause* are not annotated as NIs.

Our approach for performing this detection is described as follows. For the first-level of detection, we make full use of the three types of relations, and adopt a rule-based strategy proposed

by Lei et al. (2013) to detect NIs. As for *CoreSet* relation, in particular, as long as one of the FEs in this set is expressed overtly, NIs are not annotated for the absence of the other FEs in the set. If none of *CoreSet* is expressed, the contextually most relevant one should be annotated as a NI. However, this is difficult for automatic detector, which inevitably introduces some false detected NIs.

Thus, we conduct a second-level identifying. To be specific, for the current lexical unit, i.e. the target word, we collect its frame element patterns from the training dataset. Frame element patterns are annotated semantic roles, which include the roles annotated as NIs. Taking lexical unit *launched/发射* as an example, Table 1 shows its frame element patterns in our data. Depending on this kind of patterns, we are able to filter out some false NIs effectively.

Patte1	Time	Agent$_{INI}$	Theme	Goal$_{INI}$
Patte2	Agent	Theme	Goal$_{INI}$	

Table 1: Frame element patterns for the target 发射/*launched* in our data

4 Definite Null Instantiation Identification

In this section, we focus on our second task of definite null instantiation (DNI) identification.

Before performing the implicit argument resolution in discourse, we have to decide which null instantiated frame elements should be selected, i.e. which null instantiations are definite. As shown in example (1) above, assuming one detected null instantiated FE in the previous step is e_m (e.g. *Theme*), we should determine whether e_m needs to be filled or not, that is, we should determine e_m as DNI or INI.

Num	Feature names	Feature Descriptions
T1	Target	Target predicate
T2	Pos	The part of speech of target
T3	Frame	The frame that target evokes
T4	FENI	NI of frame elements
T5	FE	Overtly expressed FEs

Table 2: Features description in DNI Identification

We treat this issue as a classification problem, and build a binary maximum entropy model to predict the null instantiation type of e_m. Table

2 lists all features used for training our models. In addition, we employ some similar features that were used in Lei et al. (2013). Meanwhile, we choose to learn a SVM classifier for comparison purpose.

5 Definite Null Instantiation Resolution

In this section, we tackle the last subtask, namely definite null instantiation resolution.

5.1 Frame-to-Frame Relations

The relations of Frame-to-Frame and FE-to-FE in FrameNet, serve as important information sources, to be leveraged for DNI resolutions.

FrameNet arranges frames into a net by defining frame-to-frame relations, including Inheritance, Inchoative Of, Subframe, Causative Of, Precedes, Using, See_also and Perspective On. In the case of Inheritance relation, it defines two frames, i.e. one more general frame and the other more specific frame. The specific frame `Commerce buy`, for example, is inherited from the general frame `Getting`.

As Figure 1 shows, the inheritance relation allows a general frame (e.g., `Getting`) to be specialized with a particular semantic interpretation (e.g., `Commerce buy`). Also the inheritance relation exists between the *frame elements* of two related frames. Each of the inheriting FEs contains all semantic properties of the inherited general frame elements and also owns its additional private properties.
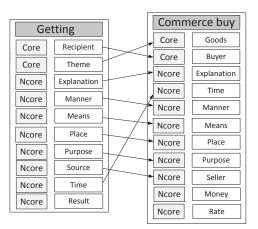

Figure 1: FE-FE relations of frame `Getting` and `Commerce buy`

Number		Features Name	Features Description
T1	F1	DistCT	The number of sentences between candidate FE content and target
	F2	CanFEcon	Candidate frame element content
	F3	CanFEpt	The phrase type or POS of candidate frame element content
T2	F4	Frame	The frame that target predicate evokes
	F5	FEDNI	DNI frame element
	F6	Target	Target predicate
	F7	TargetPOS	The part of speech of target

Table 3: Features description in Overt Frame Elements Based Resolver

5.2 Modeling Definite Null Instantiation Resolution

After accomplishing the previous processes, we can perform DNI resolutions. If the uninstantiated FE e_m (e.g., *Theme* in example (1)) has been identified as DNI previously, we need to find the corresponding antecedent mention d_m (e.g., [*The celestial burial satellite*/天葬卫星] in example (1)). Due to having fine-grained frame semantic role labeled for each sentence, we think the filler of DNI maybe also instantiates the FE of other annotated frames in the context. Therefore, we collect the overt FE content set φ instantiated in the discourse, and this set forms the overall set of candidates for DNI linking. Then, for DNI e_m, a subset of candidates φ_m ($\varphi_m \subseteq \varphi$) is chosen as candidate search space for resolving e_m.

We implement two semantic resolvers based on different methods. For either of these two resolvers, if two or more candidates score equally well, the one closest to the target predicate is chosen.

OvertFE is based on machine learning, and FFR is an inference method. As the inherent difficulty of task, it's difficult to find all fillers for DNIs only using one of them. Thus finally we simultaneously employ OvertFE and FFR to find as many fillers for DNIs as possible.

Overt Frame Elements Based Resolver (OvertFE)

This resolver is based on the assumption that the filler of DNI can be found among the overt FE content set in context. Given a DNI e_m, DNI linking can be treated as a classification problem to judge whether a candidate overt FE content d ($d \in \varphi_m$) could be taken as filler of a DNI. Therefore, we employ a classification method to solve the problem. Clearly, the performance of classifiers largely depends on constructed features. Since corresponding antecedent of DNI is not overtly expressed, it is difficult to get some information from context to describe them. What we take as features is the information of candidate frame element contents and frame information. Table 3 lists all features used for training our models. Some similar features were employed by Wang et al. (2013) where they also considered DNI linking as a classification problem.

Then maximum entropy models, widely used in natural language processing (such as Chinese word segmentation and machine translation), are employed to predict whether a candidate FE content is the filler of DNI.

Frame-to-Frame Relations Based Resolver (F-FR)

Another way of finding the correct filler is through searching Frame-to-Frame relations in a given context window. This is because Frame-to-Frame relations and FE-to-FE relations can provide relevant information for finding DNI filler among candidate frame element contents. Specifically, for one frame f_1 that contains a DNI, firstly we need to find related frame f_2 with it from context. Then, if DNI frame element in f_1 has relation with the frame element (marked with fe_2) of f_2, the filler of fe_2 is the corresponding filler of this DNI. The detailed steps are reported in Algorithm 1.

If frame names are the same, we think they are related, and Figure 2 illustrates this case. As the frames evoked in two sentences are both `Arriving`, we link the antecedent of *Goal* in the second sentence to [*Tiananmen Square*/天安门广场], which is the content of *Goal* in the first sentence.

For other cases, we use the related frames which at most contain two relation paths (e.g., the paths from `Event` to `Process_start` to `Activity_start` in Figure 3). As shown in Figure 3, the target *initiated*/发起 in the first sentence evokes the `Activity_start` frame,

in which the two frame elements (*Agent*, *Place*) is expressed in a single constituent [*our country*/我国], i.e. the phenomenon of frame element fusion arises. Frame `Event` is evoked by the target *happened*/出现 in the second sentence, where *Time* and *Event* FEs are expressed overtly, except the core FE *Place*. In the net of FrameNet, frame `Activity_start` inherits from the frame `Process_start` which further inherits from the `Event` frame. These inheritance relationships also hold between the frame elements of the related frames. According to the FE-to-FE relations, the content of FE *Place* in the first sentence, [*our country*/我国], is the corresponding filler of implicit FE *Place* in the second sentence.

Algorithm 1 : Frame-to-Frame Relations Based Resolver

Input: The frame set in discourse is $F = \{f_1, f_2, ..., f_n\}$; overt core frame element set for frame f_i is $E_i = \{e_1, e_2, ..., e_m\}$, its corresponding filler set is $A_i = \{a_1, a_2, ..., a_m\}$; one frame that contains DNI e^* is f^*, target t evokes the frame f^*; $dis\,(a_i, t)$ is the distance between DNI filler a_i and target t; $relationpath\,(f_i, f^*)$ are the relation paths from f_i to f^*; A^{temp} is temporary DNI filler set

Output: the filler a^* of DNI e^*

$A^{temp} = \phi$

for each $f_i \in F$ **do**

  if f_i has frame relation with f^* **AND** $relationpath\,(f_i, f^*) \le 2$ **then**

    for each $e_i \in E_i, a_i \in A_i$ **do**

      if e_i has relation with e^* **then**

        $a_i \in A^{temp}$

      end if

    end for

  else if $f_i = f^*$ **then**

    $a_i \in A^{temp}$

  end if

end for

if $A^{temp} \ne \phi$ **then**

  for $a_i \in A^{temp}$ **do**

    if $dis\,(a_i, t)$ is minimum **then**

      $a^* = a_i$

    end if

  end for
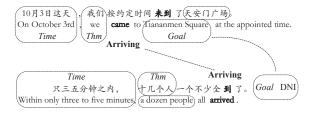
end if

return a^* ;

Figure 2: Two consecutive sentences owning the same frame. Bold fonts represent lexical units or frames. Dashed boxes represent FEs.

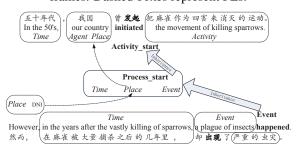Figure 3: Two consecutive sentences owing related frames. Bold fonts represent lexical units or frames. Dashed boxes represent FEs.

6 Experiments

6.1 Experimental Settings

Data: Experimental data set comes from Semantic Computing and Chinese FrameNet Research Centor of Shanxi University[1]. Because of the current low performance of CFN automatic semantic analysis systems, all discourses are labeled semantic roles manually, and the process is similar with the FrameNet annotation.

First, the ICTCLAS are used for part-of-speech tagging (omitted in examples), and we treat verbs, adjectives and nouns in each sentence as potential targets. As not all potential targets can be annotated, it is necessary to identify those targets which can evoke frames.

Then, we choose corresponding frames for those targets. For one verb target *launched*/发射 in example (1), we find its evoked frame `Cause_motion`.

Then annotate semantic roles for those constituents which share syntactical relations with this target, so the span [*the orbit over 3000 kilometers away from the surface of earth*/到距地球表面3000多公里处的环球轨道上] is annotated as role *Goal*, which is, however, the only one instantiated, out of nine `Cause_motion`'s core frame elements. So according to the context and frame element relations, we need to determine whether each

---

missing frame element should be annotated as DNI or INI.

Next, we generate the XML format for our annotated corpus, which is similar to the data format in SemEval-10 Task 10.

Our 164 discourses had been annotated by one person (to make it consistent), and they consist of 57 discourses from People's Daily and 107 discourses from Chinese reading comprehension, which cover technology, health care, social, geography and other fields. Each discourse contains 10 sentences in average. The data set contains about 37526 words in 1618 sentences; it has 175 frame types, including 2283 annotated frame instances. Table 4 shows the detailed statistics of our data set. we'll share our data in the website(http://sccfn.sxu.edu.cn/).

discourses	sentences	frame inst.	frame types	INIs	DNIs
164	1618	2283	175	213	212

Table 4: Corpus Statistics

Definite Null Instantiation identification and resolution model: Our maximum entropy classification model uses the toolkit from Zhang (2005) with the default parameter values. The SVM classifier for comparison was trained via SVM toolkit LIBSVM with the default parameter values too.

6.2 Experimental Results

Based on the experimental methods described in the previous section, we have systematically evaluated our approach on the constructed Chinese null instantiation corpus. Note all the performances are achieved using 5-fold cross validation.

Null Instantiation Detection
Table 5 gives the performance of NI detection, which achieves 72.71%, 86.12% and 78.84% in precision, recall and F-score, respectively. Here, the relatively lower precision is mainly due to the heuristic rules used to detect NIs. However, it is worth to point out that lower precision and higher recall is highly beneficial, as higher recall means less filtering of true NIs.

	P%	R%	F%
Ours	**72.71**	86.12	**78.84**
Lei et al.	56.18	90.57	69.34

Table 5: Performance of NI Detection

To illustrate the effectiveness of our method, we compare it with the Lei et al.'s method on our data, as shown in the Table 5. The F-score of our method is 78.84%, which is 9% higher than that of Lei et al.'s method. Clearly, these experimental results further prove that our second-level identification is very effective.

Definite Null Instantiation Identification
Table 6 provides the performance of DNI identification on our automatic NI detection results. It shows that DNI identification based on maximum entropy model achieves the performance of 67.86%, 69.93% and 68.88% in terms of precision, recall and F-score respectively, which are better than the results using SVM classifier, as well as the results employing Lei et al.'s method on our data.

We observe, from Table 6, that the performance of DNI identification is not high, possibly due to the poorer results of NI detection in the previous step. Moreover, because of the diversity of NI distribution, the difference of frames, and target words or missing core frame elements, the interpretation of NI types may be quite different. Thus it is difficult to build a suitable and accurate uniform classification model.

	P%	R%	F%
DNI IdenME	**67.86**	**69.93**	**68.88**
DNI IdenSVM	67.25	62.02	64.53
Lei et al.	64.58	67.73	66.12

Table 6: Performance of DNI Identification

Resolution on golden Definite Null Instantiation
In order to select the most effective features for OvertFE resolver and choose the best search space, we assume perfect results for the first two steps, that is, we perform DNI resolution experiment just with the correct DNIs in discourse. After extensive experiments employing different sets of features in different window sizes, we conclude that combining all features can achieve the best performance. Table 7 shows the results on correct DNIs using the best feature set in the window of 2, 3 and 4 sentences containing and before the target predicate (Win2, Win3, Win4 for short).

For OvertFE resolver, it shows that the F-score with Win2 is higher than that in other windows, because the bigger the window size, the more the candidate fillers for DNI, and the more difficult for

	Win2			Win3			Win4		
	P%	R%	F%	P%	R%	F%	P%	R%	F%
OvertFE	45.22	18.20	25.95	43.04	17.64	25.02	38.63	15.23	21.84
FFR	65.56	16.29	26.11	63.50	16.81	26.58	58.53	17.32	26.72
OvertFE+FFR	51.17	31.59	**39.06**	52.41	32.02	**39.75**	45.88	31.10	**37.07**

Table 7: Results on golden DNI

OvertFE classifier to find right fillers.

For FFR resolver, it needs to find related frames, and we find that its resolved DNIs are less than that by OvertFE resolver, thereby resulting in the lower precision of OvertFE than FFR.

Though performances of OvertFE and FFR both are relatively low, FFR can resolve several DNIs that OvertFE can not. Figures 2 and 3 both are such cases. So when combining the two resolvers, the final result of OvertFE+FFR outperforms that of each individual resolver. Meanwhile, as shown in Table 7, for the combined resolver OvertFE+FFR, the F-score is the highest when the window size is 3 (i.e. Win3).

Overall: Null Instantiation Resolution

Table 8 gives the performance of overall null instantiations resolution with automatic NI detection and automatic DNI determination. It shows that our resolver OvertFE+FFR achieves 40.53%, 21.54% and 28.13% in terms of precision, recall and F-score. In comparison with the results (52.41%, 32.02% and 39.75% in P, R and F) in Win3 of Table 7, it shows that the errors caused by automatic NI detection and automatic DNI determination decrease the performance of overall NI resolution by about 11% in terms of F-score.

	P%	R%	F%
OvertFE	33.28	13.78	19.49
FFR	52.95	9.71	16.41
OvertFE+FFR	**40.53**	**21.54**	**28.13**
Wang et al.	31.93	12.76	18.23

Table 8: Performance of NI resolution for our models and comparative systems

For comparison, we also conduct DNI resolution on our constructed corpus employing the method proposed by Wang et al. (2013). Since our corpus does not contain annotation of head words, the results are obtained by using their features without head word information. As the last line of Table 8 shows, the performance behaves similarly with our OvertFE resolver. In addition, we

notice current state-of-the-art approach of Laparra and Rigau (2013) employs coreference models, although our corpus does not contain coreference annotation information. As such, we are not able to conduct experiments on our dataset using their method for comparison purpose.

Overall, the relatively low performance of resolution reflects the inherent difficulty of this task, also reveals that further research is needed.

7   Conclusion and Future Work

Apparently, linking implicit participants of a predicate is a challenging problem. We have presented a study for identifying implicit arguments and finding their antecedents in Chinese discourse.

As shown in this paper, we split the difficult task into three subtasks: null instantiation detection, definite null instantiation identification and definite null instantiation resolution. Among the three subtasks, the third is our major focus. For the third subtask, we build two different resolvers: 1) OvertFE resolver, which represents that the filler of a DNI can be found among those overt FE content set in context, by employing classification methods; 2) FFR resolver, which is the frame-related search, leverages rich network of frame-frame relations to find antecedents. We have proved that these two resolvers are very useful for the third subtask, and a combination of two resolvers produced the best results.

In the near future, we plan to create and release a larger null instantiation corpus. As null instantiation detection and definite null instantiation identification are the foundation of resolving definite null instantiation, it is critical to improve the performance of both subtasks. Moreover, as different information sources have been used in our study, we cannot directly compare with some of the existing methods. For our future work, we plan to manually annotate coreference information so that we can compare with more methods. Finally, we hope to exploit some additional knowledge resources, such as HowNet, which could

potentially further improve the performance of our proposed method.

Acknowledgments

References

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The berkeley framenet project. In *Proceedings of COLING/ACL*.

Aljoscha Burchardt, Anette Frank, and Manfred Pinkal. 2005. Building text meaning representations from contextually related frames–a case study. *Proceedings of the 6th International Workshop on Computational Semantics (IWCS-6)*, pages 66–77.

Desai Chen, Nathan Schneider, Dipanjan Das, and Noah A Smith. 2010. Semafor: Frame argument resolution with log-linear models. In *Proceedings of the 5th international workshop on semantic evaluation*, pages 264–267.

Dipanjan Das, Nathan Schneider, Desai Chen, and Noah A Smith. 2010. Probabilistic frame-semantic parsing. In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pages 948–956.

Charles J. Fillmore. 1982. Frame semantics. *Linguistics in the morning calm*, pages 111–137.

Matthew Gerber and Joyce Y. Chai. 2010. Beyond nombank: a study of implicit arguments for nominal predicates. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1583–1592.

Matthew Steven Gerber. 2011. *Semantic Role Labeling of Implicit Arguments for Nominal Predicates*. Ph.D. thesis, Michigan State University.

Philip Gorinski, Josef Ruppenhofer, and Caroline Sporleder. 2013. Towards weakly supervised resolution of null instantiations. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)–Long Papers*, pages 119–130.

Egoitz Laparra and German Rigau. 2013. Sources of evidence for implicit argument resolution. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)–Long Papers*, pages 155–166.

Zhangzhang Lei, Ning Wang, Ru Li, and Zhiqiang Wang. 2013. Definite null instantiation recognizing in framenet. *Journal of Chinese Information*, 27(3):107–112.

Ru Li. 2012. *Research on Frame Semantic Structure Analysis Technology for Chinese Sentences*. Ph.D. thesis, ShanXi university.

Hugo Liu and Push Singh. 2004. Conceptnet: a practical commonsense reasoning tool-kit. *BT technology journal*, 22(4):211–226.

Kaiying Liu. 2011. Research on chinese framenet construction and application technologies. *Journal of Chinese Information Processing*, 6:006.

Josef Ruppenhofer, Caroline Sporleder, Roser Morante, Collin Baker, and Martha Palmer. 2010. Semeval-2010 task 10: Linking events and their participants in discourse. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 45–50.

Carina Silberer and Anette Frank. 2012. Casting implicit role linking as an anaphora resolution task. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*, pages 1–10.

Sara Tonelli and Rodolfo Delmonte. 2010. Venses++: Adapting a deep semantic processing system to the identification of null instantiations. In *Proceedings of the 5th international workshop on semantic evaluation*, pages 296–299.

Sara Tonelli and Rodolfo Delmonte. 2011. Desperately seeking implicit arguments in text. In *Proceedings of the ACL 2011 workshop on relational models of semantics*, pages 54–62.

Ning Wang, Ru Li, Zhangzhang Lei, Zhiqiang Wang, and Jingpan Jin. 2013. Document oriented gap filling of definite null instantiation in framenet. In *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*, pages 85–96.

Le Zhang. 2005. Maximum entropy modeling toolkit for python and c++. `http://homepages.inf.ed.ac.uk/lzhang10/maxent_toolkit.html`.