Task Driven Coreference Resolution for Relation Extraction

Feiyu Xu and **Hans Uszkoreit** and **Hong Li** [1]

Abstract. This paper presents the extension of an existing mimimally supervised rule acquisition method for relation extraction by coreference resolution (CR). To this end, a novel approach to CR was designed and tested. In comparison to state-of-the-art methods for CR, our strategy is driven by the target semantic relation and utilizes domain-specific ontological and lexical knowledge in addition to the learned relation extraction rules. An empirical investigation reveals that newswire texts in our selected domains contain more coreferring noun phrases than prononimal coreferences. This means that existing methods for CR would not suffice and a semantic approach is needed. Our experiments show that the utilization of domain knowledge can boost CR. In our approach, the tasks of relation extraction and CR support each other. On the one hand, reference resolution is needed for the detection of arguments of the target relation. On the other hand, domain modelling for the IE task is used for semantic classification of the referring nouns. Moreover, the application of the learned relation extraction rules often narrows down the number of candidates for CR.

With respect to the minimally supervised learning of relation extraction grammars, we design and evaluate two integration strategies: (i) resolution after the complete pattern acquisition process and (ii) resolution embedded in the iterations of the learning process. The evaluation helps us to gain and substantiate a relevant insight: CR effectively improves recall in both strategies but it can hurt the precision because of its error spreading potential.

1 INTRODUCTION

Minimally supervised pattern acquisition methods for relation extraction such as [1] and [5] can be viewed as attempts to realize a major dream of machine learning: After receiving a few semantic examples of relevant information units, the machine autonomously learns from texts how humans express such kinds of information.

[20] show how this goal can be much better achieved if the machine is able to perform a syntactic analysis of the relevant sentences. The method does not need an annotated corpus for learning the extraction rules but it needs a small semantic domain model and general linguistic knowledge. For the method to be successful, only the relevant pieces of the text have to be analyzed. But it turns out that many instances of the targeted relation cannot be found because some required arguments are not directly contained in the learned pattern. They are indirectly present represented by a pronoun or another coreferring noun phrase. Yet all existing bootstrapping methods do not provide means for detecting the real arguments that usually follow or sometimes precede the detected relation pattern.

Most text sorts, among them newspaper texts, often make use of referential expressions which form the textual coherence(see [9] and [10]). Without effective coreference resolution such relation instances cannot be recognized. Coreference resolution is an important research area in general linguistics and computational linguistics (see [2], [3], [6], [9], [10], [11], [12], [15], [16], [17] and [21]).

Our novel method for coreference resolution is very restrictive. Its sole purpose is the improvement of relation extraction, thus it only attempts to resolve coreferences that matter for the recognition of relation arguments. To this end it utilizes the learned rules, the semantic domain model and generic linguistic knowledge sources such as WordNet (see [14]). In this way, our coreference resolution does not only support dedicated relation extraction, but it is also supported by the relation extraction.

Our main objective is to improve recall without doing too much damage to precision. The reasons for setting the priority on recall are: (i) recall is lower than than precision, (ii) in the framework of DARE, precision can still be further improved by exploiting additional linguistic components as filters, but instances not detected during the bootstrapping cycles cannot be used for learning new rules and (iii) we utilize relation extraction technology mainly in business intelligence and similar applications, where truly relevant information must not be missed but some manual filtering belongs to the work process.

The paper starts with a desription of the baseline, i.e., an existing minimally supervised bootstrapping approach to relation extraction (see [19] and [20]). Next a data analysis demonstrates the need for CR by quantifying the proportion of relation instances that involve coreferenced arguments. We then show how semantic domain modelling can provide valuable resources for CR including hierarchical noun classes and synonyms. In contrast to widely used CR methods, our approach treats noun phrases as complex descriptions thus also detecting subsequent references to elements of set denoting noun phrases. It is also suited for gathering information on arguments by aggregating the descriptions of more than one coreferring noun phrase. Our novel approach to CR is then described and demonstrated in action. Finally, the improvement of the baseline relation extraction system is measured. CR is shown to raise the recall in detecting prize award events and the management succession event. The paper closes with a summary and discussion of the gained insights.

2 DARE FRAMEWORK

[19] and [20] describe a minimally supervised machine learning framework for extracting relations of various complexity, called DARE (Domain Adaptive Relation Extraction). The bootstrapping starts from a small set of n-ary relation instances as "seed", in order to automatically learn pattern rules from parsed data, which can then

[1] the German Research Center for Artificial Intelligence, Germany, email: feiyu@dfki.de

extract new instances of the n-ary relation and its projections. The bootstrapping process stops when no new rules or instances can be detected. In DARE, the rule learning and the instance extraction interplays with each other. DARE presents a novel rule representation model which enables the composition of n-ary relation rules on top of the rules for projections of the relation. The compositional approach to rule construction is supported by a bottom-up pattern extraction method.

The first example relation comes from the prize award domain. The relation contains four arguments representing an event in which a person or an organization won a particular prize in a specific area and in a certain year:

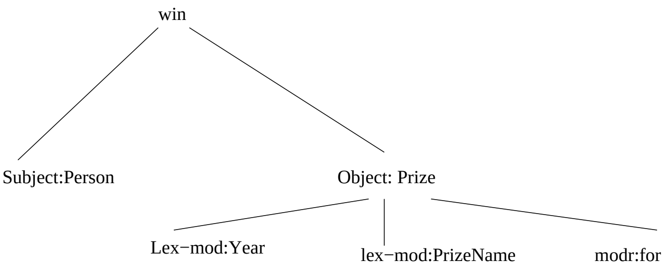
1. *<recipient, award, area, year>*
2. *<Mohamed ElBaradei, Nobel, Peace, 2005>*

(2) is an example relation instance of (1), referring to an event mentioned in the sentence (3).

3. *Mohamed ElBaradei, won the 2005 Nobel Prize for Peace on Friday for his efforts to limit the spread of atomic weapons.*

DARE learns three rules from the tree in (4), i.e., (5), (6) and (7).

4.



(5) extracts the semantic argument area from a prepositional phrase, while (6) extracts three arguments year, prize and area from the complex noun phrase and calls the rule (5) for the argument area.

5. `Rule name:: area_1`
`Rule body::`

$$\begin{bmatrix} \text{head} & \begin{bmatrix}\text{pos} & \text{preposition}\\ \text{lex-form} & \text{“for”}\end{bmatrix} \\ \text{daughters} & \left\langle \begin{bmatrix}\text{pcomp-n} & \begin{bmatrix}\text{head} & \boxed{1}\ \text{Area}\end{bmatrix}\end{bmatrix}\right\rangle \end{bmatrix}$$

`Output::` $\langle \boxed{1} Area \rangle$

6. `Rule name:: year_prize_area_1`
`Rule body::`

$$\begin{bmatrix} \text{head} & \begin{bmatrix}\text{pos} & \text{noun}\\ \text{lex-form} & \text{“prize”}\end{bmatrix} \\ \text{daughters} & \left\langle \begin{array}{l} \begin{bmatrix}\text{lex-mod} & \begin{bmatrix}\text{head} & \boxed{1}\ \text{Year}\end{bmatrix}\end{bmatrix}, \\ \begin{bmatrix}\text{lex-mod} & \begin{bmatrix}\text{head} & \boxed{2}\ \text{Prize}\end{bmatrix}\end{bmatrix}, \\ \begin{bmatrix}\text{mod} & \begin{bmatrix}\text{head} & \begin{bmatrix}\text{pos} & \text{preposition}\\ \text{lex-form} & \text{“for”}\end{bmatrix} \\ \text{rule} & \text{area\_1::}\ \langle \boxed{3} Area \rangle \end{bmatrix}\end{bmatrix} \end{array} \right\rangle \end{bmatrix}$$

`Output::` $\langle \boxed{1} Year, \boxed{2} Prize, \boxed{3} Area \rangle$

(7) is the rule that extracts all four arguments from the verb phrase dominated by the verb ”win” and calls (6) to handle the arguments embedded in the linguistic argument ”object”.

7. `Rule name:: recipient_prize_area_year_1`
`Rule body::`

$$\begin{bmatrix} \text{head} & \begin{bmatrix}\text{pos} & \text{verb}\\ \text{mode} & \text{active}\\ \text{lex-form} & \text{“win”}\end{bmatrix} \\ \text{daughters} & \left\langle \begin{array}{l} \begin{bmatrix}\text{subject} & \begin{bmatrix}\text{head} & \boxed{1}\ \text{Person}\end{bmatrix}\end{bmatrix}, \\ \begin{bmatrix}\text{object} & \begin{bmatrix}\text{head} & \begin{bmatrix}\text{pos} & \text{noun}\\ \text{lex-form} & \text{“prize”}\end{bmatrix} \\ \text{rule} & \begin{array}{l}\text{year\_prize\_area\_1::} \\ \langle \boxed{4} Year, \boxed{2} Prize, \boxed{3} Area \rangle\end{array} \end{bmatrix}\end{bmatrix} \end{array} \right\rangle \end{bmatrix}$$

`Output::` $\langle \boxed{1} Recipient, \boxed{2} Prize, \boxed{3} Area, \boxed{4} Year \rangle$

All DARE rules are extracted from sentences in which the arguments of the seed example such as (2) occur. The arguments are named entities (or other selected entity types) recognized by a named-entity-recognition system called SProUT (see [7]). SProUT also resolves variants of names, e.g., *Dr. Mohamed ElBaradei* and *Dr. ElBaradei* are recognized as the same person. However, the current learning system cannot cope with sentences that mention the target relation but contain anaphoric references to their actual arguments. If a learned rule such as (7) matches the parsed tree of (8), DARE will not be able to extract a new instance from (8) because the subject is not a person name.

8. *He/The scientist won the 2005 Nobel Prize for Peace on Friday for his efforts to limit the spread of atomic weapons.*

It is known that arguments belonging to a relation instance are often distributed over several sentences. These sentences are usually linked by coreferences, semantic chains or various discourse relations, e.g., (9).

9. a. *Three of the Nobel Prizes for Chemistry during the first decade were awarded for pioneering work in organic chemistry.*

b. *In 1902 Emil Fischer (1852-1919), then in Berlin, was given the prize for his work on sugar and purine syntheses.*

c. *Another major influence from organic chemistry was the development of the chemical industry, and a chief contributor here was Fischer's teacher, Adolf von Baeyer (1835-1917) in Munich, who was awarded the prize in 1905.*

In example (9), two concrete Nobel Prize winning event instances in Chemistry are mentioned, one in the year 1902 for Emil Fischer and another one in 1905 for Adolf von Baeyer. However, the linking between the Nobel Prize winners with the Nobel Prize is expressed indirectly via the anaphoric NP *the prize*. The two arguments (prize name and area) shared by the two event instances are located in the first sentence. The two winners and their prize award years can be found in sentence (b) and (c), respectively. If we consider sentence (b) and (c) independently from the context, we cannot tell that they are about the Nobel Prize events, without resolving the anaphoric reference *the prize* as the Nobel Prize.

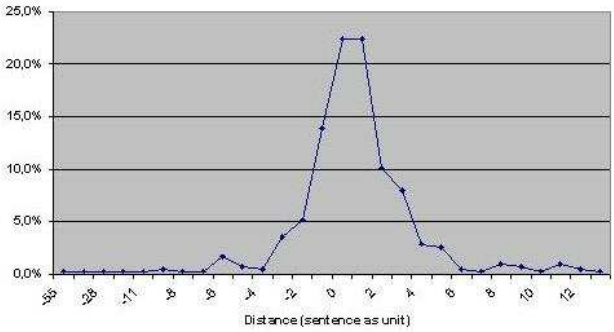
An evaluation in [19] shows us that more than 10% of the relation instances in the Nobel Prize award domain can only be detected with the help of coreference resolution. Therefore we can expect coreference resolution to improve the learning performance by detecting more relation instances as seed.

3 INVESTIGATION OF COREFERENCE RELATIONS IN THE EXPERIMENT DOMAIN

The phenomena of coreference has been investigated intensively in the literature (see [9], [10], [11] and [15]). They are complex linguistic phenomena influenced by lexical, syntactic, semantic and discourse constraints. In recent years, a number of computational approaches attempted to map these constraints to features of computational models, e.g., features of some classifiers (see [6], [13] and [16]). The constraints shared by many approaches are

- Distance: coreference expressions are often close to each other in the surface structure;
- Syntactic: pronominal resolution constraints within sentence
- Semantic: same semantic category, agreement in number, gender and person;
- Discourse: parallelism, repetition, apposition, name alias

We did a study to investigate the coreference relations in our experiment domain Nobel Prize award. The texts in the corpus are Nobel Prize related articles from New York Times, online BBC and CNN news reports. It contains 3328 documents with a size of 18.4 MB. We only consider documents mentioning the target relation. Figure 1 shows the distribution of distance of coreference links in the corpus. The target relation is located in sentence 0. We call this sentence our *anchor*. The context around the anchor where the antecedents can be found is within three sentences before or after the anchor. The distribution result confirms the closeness as indicator.



Furthermore, we calculate the distribution of the pronominal and the nominal coreference links. 25.08% of the links are pronominal, while 74.92% are nominal. Most of the anaphora expressions are singular (96.19%) and only 3.81% are plural. To our surprise, the forward links make up for 40%, but the backward links for 60%. Among the noun phrases, the definite phrases account for 19.92% of occurrences and indefinite NPs thus for 80.08%.

Let us look at the following two examples (10) and (11):

10.

a. *Two Americans have won the 2002 Nobel Prize in Economic Sciences.*

b. *The two scientists, Daniel Kahneman and Vernon L. Smith, received the honour on Wednesday for their work using psychological research and laboratory experiments in economic analysis.*

11.

a. *Egypt honours its Nobel Prize chemist.*

b. *President Hosni Mubarak of Egypt has awarded the country's most prestigious prize - the Nile Necklace - to the Egyptian-born chemist Ahmed Zewail.*

Many approaches emphasize semantic similarity and semantic consistency between the coreference expressions, e.g., ISA and synonym relations and make use of the cohesion indicator *repetition*. Although there are repetitions in both (10) and (11), e.g., the number *two* occurs both in (a) and (b) in (10) and the word *chemist* is mentioned in the two coreference noun phrases in (11). However, neither of the coreference expressions can be simply resolved by ISA and synonym relations. In both cases, the second phrase adds additional semantic information, corresponding to the elaboration phenomena discussed by [10]:

> $S1$ *is an Elaboration of* $S0$ *if a proposition* P *follows from the assertions of both* $S0$ *and* $S1$*, but* $S1$ *contains a property of one of the elements of P that is not in* $S0$.

Elaboration is an important feature of newspaper reports. In our experiment domain, we observe that various aspects of a person are mentioned in a report to describe a Nobel Prize winner. The most frequent properties are:

- Nationality/origin/inhabitant: e.g., two Americans, the Egyptian-born, a Dutch
- Profession/occupation: e.g., novelist, chemist, scientist, researcher
- Title/position: e.g., professor, president
- Domain description: e.g., recipient, winner, Nobel Laureate
- General description: e.g., the man, a woman, the team

The most frequently mentioned property in our domain for a Prize winner is profession or occupation. The second one is the title and position. Nationalities are ranked in third position. (10) and (11) also show that a noun phrase often describes more than one property of a person. *Egyptian-born chemist Ahmed Zewail* mentions not only the person name but also the origin and the profession of the person. Therefore, it is important to treat a referent as a complex semantic object. The antecedents of the anaphora noun phrase in both examples are in the second sentence. Both backward and forward search are important in our domain.

4 ACQUISITION OF DOMAIN KNOWLEDGE DURING BOOTSTRAPPING

The observation in section 3 tells us that it is important to acquire domain knowledge for the coreference resolution in our application domain. However, manual modeling is too time consuming and not easily adaptable to new domain (see [2] and [9]). The general principle of the DARE framework is to start with some relation instances as their semantic seed to acquire linguistic pattern rules. The learned extraction rules are applied to the parsed sentences in order to extract relation instances as new seed for the next iteration. The current relation instances contain only entity instances as their arguments. In order to acquire the knowledge about the semantic roles in the target relation, we add a knowledge mining step in each iteration during the bootstrapping process:

Given a target relation R with n arguments and a set of relation instances Γ detected by DARE:
for all argument arg_i in r such that $r \in \Gamma \wedge i \in [1, n]$ **do**

```
   1. collection appositions of arg_i in the
whole corpus
   2. extract adjectives and nouns from the
appositions and assign the frequency to the
words
   3. retrieve direct hypernyms, inherited
hypernyms and its sister terms of the
adjective and noun terms from WORDNET
end for
build a domain-specific ontology from the
extracted WORDNET relations for each
argument
```

In our approach, the domain ontology contains a list of subontologies for each semantic role in the target relation. For the prize award event, we have one for recipient, one for prize award, one for area. Each subontology models the domain knowledge of each semantic argument. The words mentioned in the corpus are marked and assigned with their frequencies in the domain ontology. Furthermore, we store all descriptions for each entity in the database. These descriptions will be used for validation when extracted entities are referred to again in later iterations.

5 TASK DRIVEN COREFERENCE RESOLUTION

Our coreference resolution is driven by the target relation. We consider only anaphora expressions that are potential semantic arguments of the mentioned relation instances embedded in the sentences matched against the learned pattern rules such as (7).

As confirmed by our domain investigation, antecedents in the nearest context are preferred. The search for an antecedent stops when an entity instance as a coreference candidate can be found. Therefore, we construct coreference chains during the backward and the forward search. The end of the chain is an entity instance.

In addition to the general features such as distance, agreement with number and gender and discourse parallelism, we introduce a novel unification method to verify the compatibility among the referential expressions. This method is suitable to handle relations among noun phrases, in particular, the complex noun phrases, e.g., the coreference relation between the infinitive plural noun phrase *two Americans* and the complex noun phrase *the two scientists, Daniel Kahneman and Vernon L. Smith* in (10).

Given the domain ontology learned from the wordNet, we construct a template for each semantic argument. For example, the recipient of the Nobel Prize has the following properities: *nationality, profession, title, domain-description, general-description* and *name*. Then we develop mappings between the wordNet concepts and the properties. For example, concepts inherited by the wordNet concept *inhabitant* or *native* are values of *nationality*.

Let us step through (10) with our solution. A DARE rule matches sentence (a) where the recipient role is an indefinite plural noun phrase. The features of this noun phrase and their values are described in (12):

12.
$$\begin{bmatrix} \textit{sentence\_id:} & i \\ \textit{number:} & \begin{bmatrix} \textit{type:} & \textit{plural} \\ \textit{amount:} & 2 \end{bmatrix} \\ \textit{definite:} & - \\ \textit{grammarrole:} & \textit{subject} \\ \textit{semantics:} & \begin{bmatrix} \textit{nationality:} & \textit{american} \end{bmatrix} \end{bmatrix}$$

The system looks forwards and finds a complex noun phrase in the next sentence (b) where the feature structure of this noun phrase is as follows:

13.
$$\begin{bmatrix} \textit{sentence\_id:} & i+1 \\ \textit{number:} & \begin{bmatrix} \textit{type:} & \textit{plural} \\ \textit{amount:} & 2 \end{bmatrix} \\ \textit{definite:} & + \\ \textit{grammarrole:} & \textit{subject} \\ \textit{semantics:} & \begin{bmatrix} \textit{profession:} & \textit{scientist} \end{bmatrix} \\ \textit{names:} & \langle\ \textit{name1, name2}\ \rangle \end{bmatrix}$$

These two feature structures are compatible because of their close distance, agreement in number and the parallel grammar function. (13) can be regarded as an elaboration of (12). It adds the profession information and instantiates the two persons with their names. The unified feature structure for (12) and (13) is (14):

14.
$$\begin{bmatrix} \textit{number:} & \begin{bmatrix} \textit{type:} & \textit{plural} \\ \textit{amount:} & 2 \end{bmatrix} \\ \textit{semantics:} & \begin{bmatrix} \textit{nationality:} & \textit{american} \\ \textit{profession:} & \textit{scientist} \end{bmatrix} \\ \textit{names:} & \langle\ \textit{name1, name2}\ \rangle \end{bmatrix}$$

6 EVALUATION

The general motivation of this approach is to improve the recall of the DARE framework. We conduct two strategies: (i) resolution after the complete pattern acquisition process and (ii) resolution embedded in the iterations of the learning process. In the first strategy, we apply the learned pattern rules again to the experiment corpus and extract sentences with anaphora experssions. In the second strategy, the coreference resolution belongs to the relation extraction in each iteration. We experiment with the two strategies in two domains: the Nobel Prize award and the management succession (MUC-6) (see [8]). In MUC-6, our target relation contains also four arguments. It is about a person (person_in) taking over a position in an organisation and a person (person_out) leaving the position.

15. *<person_in, person_out, position, organisation>*

The data set and the DARE performance before the coreference resolution is given in Table 1: Given the same data set and the same initial seed examples, we apply the two strategies to the two experiment domains. The final performance is listed in Table 2.

Table 1. relation extraction without coreference resolution

domain	data size	initial seed no.	precision	recall
Nobel Prize	18.4MB	1	86.5%	50.7%
MUC-6	1 MB	55	62%	48%

In [19] we describe why the data of the MUC-6 management succession task are not well suited for our RE method. Our hope that CR might improve recall for this task without ruining precision was not fulfilled. Minor improvements of recall were outweighed by a drastic drop of precision. To a large part the disappointing performance of the CR extension for this task can be attributed to shortcomings of the person-name recognition of our NEE system. The missing precision

Table 2. relation extraction with coreference resolution

domain	strategy (1) after pattern acquisiton		strategy (2) during pattern acquisition	
	precision	recall	precision	recall
Nobel Prize	82.76%	53.47%	83.9%	54.21%
MUC-6	48.89%	51.55%	33.5%	52.85%

in NEE interacts in bad ways with the confusion between person-in and person-out in coreference resolution.

In the Nobel prize domain, the contributions of the CR component were much more promising. Through a manual data analysis, we found 42 relation instances that could not be detected by DARE because of missing coreference resolution. In these cases at least one participant of the relation instance could not be found because the identifiying NP occured outside the relation pattern. Our CR method found 11 of these instances. In addition, CR also correctly resolved 29 cases of coreference that did not contribute to RE recall because they occurred in mentions of instances that could be detected without CR. Besides the 40 correctly resolved coreferences, our method also returned 13 false coreference resolutions. Thus the precision of the coreference resolution was 75.5 %. Because of the tight coupling of CR and RE, the false positives of CR also turned into false positives of RE. Thus the overall precision of the RE system slightly decreased.

In the experiment with a tight integration of CR and rule learning, recall improved by 3.5 %. Although precision decreased by 2.6 % even the unweighted F-measure gained 1.9 %. The measured performance gain would be higher for any F-measure variant reflecting the stronger relevance of recall. The performance gain was slightly lower for the experiment in which CR was applied after rule learning.

7 CONCLUSION

The coreference resolution approach proposed by us is driven by the relation extraction task. The investigation of the coreference relations in the application domain shows us that coreferential nominal phrases do not only share the same semantic category (repetition), but there also often exists elaboration relationship between them. We make use of the general bootstrapping strategy to learn and extract subontologies for the relation arguments from WordNet. The domain ontology reflects the domain-specific properties of the relation arguments and helps on the one hand the validation of the semantic compatibility of the coreferences and on the other hand the construction of the information content of the individual referents. In our experiments, we show that integration of coreference resolution generally improves the recall value. However, precision can be hurt at different degrees.

Our experiments have shown that a low base-line performance in both CR and RE precision can be aggravated by a combination of CR and RE. For the Nobel prize domain, the decrease in precision was outweighed by the improvement in recall. We expect that an improvement of CR precision by an enhanced NEE component will lead to even better overall effects of the integration of CR into semi-supervised RE. In our method, the result of coreference resolution is not a simple yes or no such as in classification-based methods. It is an aggregation of semantic descriptions about the referents. These descriptions can be reused for further coreference resolution and even for identity resolution across documents (see [4] and [13]).

Our approach is quite in line with our philosophy of information extraction, i.e., the view that truly systematic approaches to information extraction may turn into controlled gradual approximations of text understanding [18] .

ACKNOWLEDGEMENTS

The work presented here was partially supported by a research grant from the German Federal Ministry of Education, Science, Research and Technology (BMBF) to the DFKI project Hylap (FKZ: 01 IW F02) and the international project RASCALLI supported by the European Commission Cognitive Systems Programme (IST-27596-2004).

REFERENCES

[1] Eugene Agichtein and Luis Gravano, 'Snowball: extracting relations from large plain-text collections', in *ACM 2000*, pp. 85–94, (2000).

[2] D. Bean and E. Riloff, 'Unsupervised learning of contextual role knowledge for coreference resolution', in *Proceedings of HLT/NAACL*, pp. 297–304, (2004).

[3] David Bean and Ellen Riloff, 'Unsupervised learning of contextual role knowledge for coreference resolution', in *Proceedings of HLT-NAACL 2004*, pp. 297–304, Boston, Massachusetts, USA, (May 2 - May 7 2004).

[4] R. Bekkerman and A. McCallumb2Golub, 'Disambiguating web appearances of people in a social network', in *Proceedings of the 14th international conference on World Wide Web, 2005, Chiba, Japan*, (2005).

[5] S. Brin, 'Extracting patterns and relations from the World Wide Web', *Lecture Notes in Computer Science*, **1590**, (1999).

[6] C. Cardie and K. Wagstaff, 'Noun phrase coreference as clustering', in *Proceedings of EMNLP/VLC*, pp. 82–89, (1999).

[7] Witold Drozdzynski, Hans-Ulrich Krieger, Jakub Piskorski, Ulrich Schäfer, and Feiyu Xu, 'Shallow processing with unification and typed feature structures - foundations and applications', *KI*, **18**(1), 17, (2004).

[8] Grishman and B. Sundheim, *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, Morgan Kaufmann, 1995.

[9] S. Harabagiu, R. Bunescu, and S. Maiorano, 'Text and knowledge mining for coreference resolution', in *Proceedings of NAACL-2001*, pp. 55–62, (2001).

[10] J. Hobbs, 'Coherence and coreferen'ce', in *Cognitive Science 3*, pp. 67–90, (1979).

[11] A. Kehler, 'Probabilistic coreference in information extraction', in *Proceedings of EMNLP*, pp. 163–173, (1997).

[12] Shalom Lappin and Herbert J. Leass, 'An algorithm for pronominal anaphora resolution', *Computational Linguistics*, **20**(4), 535–562, (1994).

[13] Andrew McCallum and Ben Wellner, 'Conditional models of identity uncertainty with application to noun coreference', in *NIPS*, (2004).

[14] George A. Miller, 'Wordnet: A lexical database for english', *Commun. ACM*, **38**(11), 39–41, (1995).

[15] Ruslan Mitkov and Wolverhampton Wv Sb. Factors in anaphora resolution: they are not the only things that matter. A case study based on two different approaches, July 24 1997.

[16] Vincent Ng, 'Machine learning for coreference resolution: From local classification to global ranking', in *Proceedings of ACL 2005*. The Association for Computer Linguistics, (2005).

[17] M. Strube, S. Rapp, and C. Muller. The influence of minimum edit distance on reference resolution, 2002.

[18] Hans Uszkoreit, 'Methods and applications for relation detection', in *Proceedings of the Third IEEE International Conference on Natural Language Processing and Knowledge Engineering*, (2007).

[19] Feiyu Xu, *Bootstrapping Relation Extraction from Semantic Seeds*, Saarbruecken dissertation series in Computational Linguistics and Language Technology, phd disseration, volume xxiv edn., 2007.

[20] Feiyu Xu, Hans Uszkoreit, and Hong Li, 'A seed-driven bottom-up machine learning framework for extracting relations of various complexity', in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pp. 584–591, Prague, Czech Republic, (June 2007). Association for Computational Linguistics.

[21] D. Zelenko, C. Aone, and Tibbetts J., 'Coreference resolution for information extraction', in *Proceedings of the ACL Workshop on Reference Resolution and its Applications*, pp. 9–16, (2004).