Pronominal Anaphora Resolution using a shallow meaning representation of sentences

Hilda Ho[1], Kyongho Min[1], Wai-Kiang Yeap[1]

[1] Language Research Group, Institute for IT Research,
Auckland University of Technology, Auckland 1020 New Zealand
{ hilda.ho, kyongho.min, wai.yeap}@aut.ac.nz

Abstract. This paper describes a knowledge-poor anaphora resolution approach based on shallow meaning representation of sentences. Within our representation, we define a new local domain which provides a powerful cue for resolving pronominal anaphora. Other information used included syntactic information, syntactic parallelism and salience weights. We collected 111 singular 3rd person pronouns from open domain resources such as children's novel and examples from several anaphora resolution papers. There are 111 third-person singular pronouns in the experiment data set and 94 of them demonstrate pronominal anaphora in domain of test data. The system successfully resolves 78.4% of anaphoric examples.

1 Introduction

Due to the variety of anaphors that could occur in text and the different knowledge sources needed to resolve anaphors, anaphoric resolution is an interesting and a difficult problem in language research. Some examples of the different types of anaphors are shown below:

1. intra-sentential (e.g. "Jack tried to start his car") versus inter-sentential (e.g. "Mary has a sister. Her name is Jane")),
2. according to a co-reference direction, anaphora (e.g. "Jack started his car") versus cataphora (e.g. "On his way to school, Tom met his friend")),
3. according to a syntactic category, definite noun phrase (e.g. "An intruder has stolen a vase. The intruder stole the vase from a cupboard"), pronominal (e.g. "Jack started his car"), reflexive (e.g. "John told us about himself"), and reciprocal (e.g. "Tom and Mary like each other")),
4. shared common knowledge between a speaker and a hearer within a context (extra-textual (e.g. "the sun")),
5. use of real world knowledge (associative (e.g. "John bought a new computer yesterday and he found that the keyboard was broken")),
6. pronoun "it" can mean one of a personal pronoun (e.g. "baby"), a non-personal pronoun (e.g. "company" or "cat"), a pleonastic pronoun (e.g. "it is raining"), or a referring event.

7.   others include: anaphoric dependencies in ellipsis [5], associative anaphors [1], [7], quantifier [13], and anaphors in captions of pictures [3].

Methods developed in the past for anaphoric resolution include the use of syntactic information [6], syntactic information with constraints and preferences based on structural binding dependency between anaphors and antecedents [12], [15], a machine learning approach with limited semantic information (e.g. ISA hierarchy) from WordNet [14], and knowledge-poor approach (e.g. MARS system [4]) using results of a POS tagger rather than syntactic parses [8], [9].

Lappin and Lease [6] used syntactic information and an attentional state of antecedent candidate to resolve inter-/intra-sentential anaphoric expressions restricted to third person pronouns and lexical anaphors (reflexives and reciprocals). The syntactic salience was measured by using syntactic structure produced by McCord's Slot Grammar Parser. They tested 360 pronoun occurrences in computer manual texts. However, the structural salience measurement affected their algorithm's performance of inter-sentential anaphoric cases.

Soon et. al. [14] studied coreference resolution of general noun phrases (e.g. definite/demonstrative noun phrases, proper names, and appositive) in terms of a machine learning approach. They used 12 features to resolve coreference problems including semantic class features such as a simple ISA hierarchy from WordNet. Among the 12 features, they found 3 features (ALIAS, APPOSITIVE, and STRING_MATCHING) were more highly informative than other features. Modjeska et. al. [11] employed lexico-syntactic pattern and semantic knowledge extracted from the WWW to resolve other-anaphora in a machine learning framework. They employed a Naïve Bayes approach with 9 different features and tested the system with/without web information.

Palomar et. al. [12] studied anaphora resolution in Spanish texts and their system was based on slot unification parser and constraints and preferences extracted from lexical, morphological, syntactic, and statistical knowledge. The preferential knowledge was retrieved from a training corpus to give more priority to resolve anaphoric coreference. Their system showed good performance to process a reflexive pronoun and a demonstrative pronoun in a prepositional phrase.

Stuckardt [15] studied anaphora resolution based on restrictions and preferences extracted from syntactic tree structures resulted from FDG (Functional Dependency Grammar of English). The system used binding constraints to extract rule patterns based on syntactic relations between anaphors and antecedents. In addition, a text-genre specific choice of preferred antecedent was applied. Among various preference factors, sentence recency was the most valuable factor in his system. Walker [16] studied anaphora resolution based on a centring model of discourse structure with a cache method. The centre in a hierarchical/linear discourse structure was stored in a cache for anaphoric resolution.

This paper describes a knowledge-poor anaphora resolution approach based on surface meanings extracted from each sentences. The anaphoric resolution in this paper focuses on both intra- and inter-sentential anaphoric expressions of singular 3rd person pronouns. The system employed some syntactic (e.g. syntactic roles such as subject, object, indirect object, local domain, recency (position of antecedents)) and

semantic features (e.g. gender, number) rather than some syntactic dependency rules (e.g. c-command rule).

Section 2 briefly describes the input to the system and section 3 the basic algorithm developed. Section 4 presents some preliminary experimental results and section 5, the conclusion and future works.

2   Input

The anaphoric resolution system is developed as part of a text analysis system known as SmartINFO (Smart INFOrmation). SmartINFO processes paragraphs of text and generates a shallow meaning representation for each sentence. Such a representation captures the relationship between words in each sentence in a form that the system knows how to interpret. Some examples are shown below:

```
> (interpret '(The man who ordered the book foolishly paid the money))
[PAID* (:ACTOR (PERSON* (:SEX (MALE*))
                        (:NUMBER (SINGULAR*))
                        (:MODIFIER (THE*))
                        (:WHO* (ORDERED* (:WHAT (BOOK*
                                        (:MODIFIER (THE*))))))
                        (:MANNER (FOOLISHLY*))))
        (:WHAT (MONEY* (:MODIFIER (THE*))))]
```

```
> (interpret '(He got himself very dirty in the park)
[GOT** (:ACTOR (HE* (:SEX (MALE*))))
       (:WHAT (HIMSELF*))
       (:MODIFIER (VERY*) (DIRTY*))
       (:IN* (PARK** (:MODIFIER (THE*))))]
```

From the above input, all nouns and pronouns that appear in them are extracted and each is given a unique identification and a reference to the sentences in which they belong. As the focus of this paper is not on the output of SmartINFO, we will not give a formal definition of the above representation schema here.

3   The Anaphora Resolution Algorithm

The anaphora resolution algorithm works as follows:

1. Input – a paragraph of text which will be processed by SmartINFO and turned into a list of surface meanings representations.
2. For each of these representations, extract all the nouns and (third-person) pronouns that appear in them. When a pronoun appears in sentence, say, 5, then the

candidate list of nouns for this pronoun contains all nouns (only those with the correct gender and number agreement) appearing in the first 5 sentences.

3. Using the candidate list, the pronouns are resolved in the following manner: first we use local domain and syntactic constraints, next we apply syntactic parallelism between sentences, and finally we apply salience weightings.

3.1 Local Domain and Syntactic Constraints

Many existing systems uses Chomsky's Binding Theory [2] to define syntactic constraints for filtering invalid candidates with respects to c-command and local domain. The traditional definition of a local domain of a constituent C is defined as the set of constituents contained in the closest S or NP that contains C.
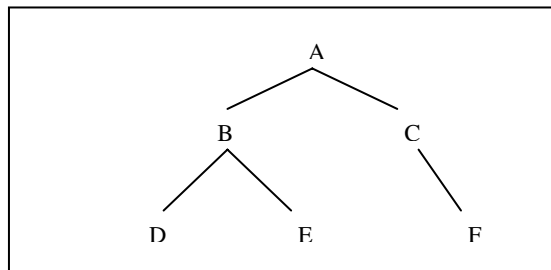

Fig. 3. Illustration of c-command

C-command is defined as follows ([10], pp 58): A node A c-commands a node B if and only if:

(i)   A does not dominate B.
(ii)   B does not dominate A.
(iii)   The first branching node dominating A also dominates B.

Such definition is illustrated in Figure 3: The node B c-commands node C and everything in C's sub-tree. The node C c-commands node B and everything in B's sub-tree. The node D only c-commands node E, but not any other node.

Three syntactic rules using the notions of c-command and local domain were defined [10] as follows:

(i)   A reflexive pronoun must co-refer with a NP that c-commands it and is in the same local domain.
(ii)   A non-reflexive pronoun cannot co-refer to a c-commanding NP within the same local domain.
(iii) A non-pronominal NP cannot co-refer with a NP that c-commands it.

Many existing anaphora resolution systems have applied these syntactic rules in their algorithms. Our approach is different. The local domain is already given explicitly as part of the structure in the shallow meaning representation (see Fig. 2).

Example1: Mary gave a book to Jane for her to do the work for her.

```
[gave (:actor (person (:name (mary))))
     (:what (book (:modifier (a))
                     (:to (jane (:person (:name (jane)))
                             (:for (her (:sex (female))
                                     (:to (do (:what (work (:modifier (the))
                                                     (:for (her (:sex (female))))
                                                     ))))))))))]
```

Example2: Mary told her that she loves the picture of her.
```
[that (:ms2 (loves (:actor (she (:sex (female))))
                     (:what (picture (:modifier (the)) (:of (her (:sex (female))))))
                     (:number (singular))))
     (:ms1 (told (:actor (person (:name (mary))))
                     (:what (her (:sex (female))))))]
```

Fig. 2. Identifying local domains: In example 1, her2 is embedded inside the local domain of the first her1 whereas this is not the case in example 2.

It is interesting to compare the output for example 1 in Fig. 2 with the syntactic parse tree generated by link grammer:

```
(S (NP Mary)
   (VP gave
      (NP a book)
      (PP to
         (NP Jane))
      (PP for
         (NP her))
      (S2 (VP to
            (VP do
               (NP the work)
               (PP for
                  (NP her))))))
   )
```

The above output shows that the local domain for her2 is S2 but the local domain of her1 is the top S. Thus, according to c-command rules, her1 does not c-command her2 and they're in different local domains, implying that they could co-refer.

Once the local domain of the pronoun is identified, several syntactic constraints are applied to filter out invalid antecedent candidates. Two sets of rules are defined: one for reflexive pronouns and the other for non-reflexive pronouns.

The syntactic constraints for reflexive pronouns state that:

- If the local domain of the pronoun is a VP or a NP that is composed of possessive noun/possessive form of noun, the reflexive anaphor can only co-refer to the candidates that are located in the same local domain.

- If the local domain of the pronoun is an infinitive phrase or PP-infinitive phrase, the reflexive anaphor cannot co-refer to any candidates that are located in the same local domain.

The syntactic constraints for non-reflexive pronouns state that:

- The pronoun cannot co-refer to any candidates that are located in the same local domain, except when the anaphor is a possessive pronoun such as "Peter loves *his* car".

Another strength of SmartINFO's anaphora resolution system is that it can filter out invalid pronoun candidates as well as the NP candidates. To illustrate, consider the following example:

Peter met John and he$_1$ told him that he$_2$ would leave early.

Although this sentence is very ambiguous, one can still be sure that the pronoun "he$_1$" does not co-refer to the pronoun "him". When SmartINFO is trying to resolve the pronoun "him", it searches for any pronoun within its same local domain (i.e. the pronoun "he$_1$") and filter it out as invalid antecedent candidate. Furthermore, the algorithm also filters the antecedent of that invalid pronoun from the candidate list of the current anaphor. In the example above, the pronoun "he$_1$" is resolved to co-refer to "Peter", then "Peter" is filter out from the candidate list of the pronoun "him", leaving only "John" as a valid antecedent candidate.

2.2.2 Salience Weights

Six salience factors are applied to select the most preferred antecedent from the set of competing candidates. The salience depends solely on simple syntactic and statistic information that are provided by SmartINFO. Additional information such as word class or syntactic tags (e.g. head nouns) is not used as salience in this system. As the SmartINFO system aims to be applied to an open domain, all of the salience factors are genre-free. However, two of them are specific to the local domain of infinitive phrase and PP-Infinitive phrase. The antecedent candidates obtain a score between -2 to 2 for each salience factor.

Infinitive Phrase

If a local domain of a current pronoun (anaphor) is an infinitive phrase and the pronoun has a prepositional phrase (PP) as its parent (e.g. "Mary asks Jane [to do it for *her*.]"), then the object of a sentence is not preferred. A candidate with its syntactic role as object is given a score of -2. On the other hand, if a pronoun is not within a prepositional phrase (e.g. "Mary asks Jane [to tell *her* about Peter.]"), the subject of a sentence is more preferred as antecedent. A candidate with its syntactic role as subject is given a score of 2.

PP-Infinitive Phrase

If a local domain of a current pronoun (anaphor) is a PP-Infinitive phrase and the pronoun does not have a prepositional phrase as its parent (e.g. "Mary is desperate [for Jane to depend on *her*.]"), then a candidate with its syntactic role as subject is more preferred and a score of 1 is assigned. On the other hand, if a pronoun has a prepositional phrase as its parent, which

defines the pronoun's local domain (e.g. "Mary gave a book [to Jane for *her* to do the homework.]"), then a candidate with its syntactic role as indirect object is more preferred (scores 1).

Syntactic Roles

This salience factor simply says that subject is more preferred than object, which is more preferred than indirect object. A candidate with its syntactic role as subject is given a weight of 2, object scores 1, and indirect object scores 0.5.

Syntactic Parallelism

A syntactic parallelism mechanism is different from the other salience factors because it is performed earlier in the AR process. It has a special constraint, which states that an anaphor and its antecedent are dominated by the same ACTION node (i.e. verb) in the parse tree (i.e. mental sketches). However, the constraint is removed when the syntactic parallelism is used as a salience factor in this system. The syntactic roles of the pronoun and its candidate are compared and if they match each other (e.g. both are subjects), a score of 1 is assigned to the candidate.

Frequency of Occurrence in the paragraph

Preference is given to candidates that appear more frequently in a paragraph in which an anaphor is located. Candidates that appear more then five times in the paragraph are given a score of 2. Candidates that appear between three to five times in the paragraph score 1 and those that appear twice in the paragraph are given a weight of 0.5.

Recency

The most recent candidate is given a score of 0.5.

When all preferences are taken into account, each candidate has an accumulated salience weight. Candidates that have a negative score are removed from the candidate list. The most highly scored candidate is then selected as the antecedent. In case of a tie, the most recent candidate will be selected.

3 Preliminary Experimental Results

The SmartINFO's anaphora resolution system is tested against a set of experiment data, which is collected from open domain resources such as children's novel and examples form several anaphora resolution papers. The data is slightly modified to be parsed by SmartINFO. There are 111 singular third-person pronouns in the experiment data set and 17 (15%) of them do not have any antecedents within the scope of the test data. The application of salience factors to resolve anaphoric expression in the system is able to filter out 90% of invalid candidates and resolve 78.4% of anaphoric examples used to test the AR system. By using local domain information and

syntactic constraints solely, the AR system is considered very successful in filtering invalid antecedent candidates.

Table 1 illustrates the evaluation of pronominal anaphors in this experiment. The pronominal anaphors that are examined in this experiment are the 3rd person singular pronouns such as she, her, he, his, him, and it. The success rate of masculine pronouns is much higher than that of feminine pronouns because the syntactic structure of sentences containing the masculine pronouns is simpler than those with feminine pronouns. In the case of pronoun "it", if real world text were tested, then its success rate would be worse.

Table 1. Evaluation of Pronominal Anaphors in Experiment

Type of Pronominal Anaphor	Total number of occurrences in data set	Number of correctly resolved anaphors	Correctly resolved anaphors (%)
She	34	28	82.3%
Her (objective)	39	31	79.5%
Her (possessive)	5	0	0%
He	12	12	100%
His	5	0	0%
Him	12	12	100%
It	5	4	80%
TOTAL	**111**	**87**	**78.4%**

Table 2. Evaluation of Intra-sentential & Inter-sentential Anaphors in Experiment

Type of Anaphor	Total number of occurrences in data set	Number of correctly resolved anaphors	Correctly resolved anaphors (%)
Intra-sentential	43	33	76.7%
Inter-sentential	68	54	79.4%
TOTAL	111	87	78.4%

Table 2 shows the evaluation of intra-sentential and inter-sentential anaphors tested. Of the 111 pronouns, 68 of them illustrate inter-sentential anaphors. SmartINFO correctly resolves 81% of inter-sentential anaphors and 77% of intra-sentential anaphors. As mentioned before, 15% of the pronouns in the experiment data do not have any antecedent in the test domain (e.g. "Peter asks Jane to tell Mary about *her*." – the pronoun does not co-refer to any of the NPs in the sentence.). By using the specially defined local domain and syntactic constrains, SmartINFO's anaphora resolution system can successfully identify 82% of these pronouns. Overall, SmartINFO is able to resolve 78% of the anaphors successfully.

In the test data 20.7% of the third-person pronouns are incorrectly resolved by the SmartINFO's anaphora resolution system. The causes of errors are shown in Table 3. The top two reasons of failure are parsing errors by SmartINFO (30%) and possessive pronouns (30%) that is not implemented in this system. As the SmartINFO system is still being developed, it has not reached maturity and some errors are produced during parsing. Such parsing errors therefore affect the performance of anaphora resolution. As mentioned in section 2, the SmartINFO extracts all the nouns and pronouns from

the parse tree. However, it is currently unable to extract any possessive pronoun concepts. Hence, all of the possessive pronouns in the testing data are not resolved.

Salient weights results in 17% of the errors. A salient factor based on syntactic roles states that a candidate as a subject role is preferred, and the salient factor based on syntactic parallelism states that a candidate having the same syntactic role as an anaphor is preferred. These two salient factors work well with sentences/paragraph in which the theme is consistent. However, if the theme changes suddenly, these two salient factors will incorrectly assign more weight to an unapproached candidate. For example, consider the paragraph:

"Mary lives in Auckland. She is ten years old. She loves singing. Mary has a sister. *She* is Jane."

SmartINFO's anaphora resolution system incorrectly resolved the last pronoun 'she' as Mary, which has been the focus of the paragraph in the previous four sentences.

Table 3. Major Error Types in SmartINFO's Anaphora Resolution System

Error Type	Number of Errors	%
Parsing Errors	7	30.4%
Possessive Pronouns	7	30.4%
Salience Weights	4	17.4%
Anaphor co-reference to pronoun	4	17.4%
Pleonastic It	1	4.2%
TOTAL	**24**	**100%**

Although the current system is able to filter out invalid pronoun candidates as well as their antecedents when solving an anaphor, it is currently unable to resolve any anaphors that co-refer to another pronouns. Such cases have caused 17% of the errors. The remaining 4% of errors are caused by pleonastic 'it' as it has not been implemented yet.

4 Future works

Further research is needed to improve the accuracy of SmartINFO's anaphora resolution system. The current system can only handle singular third-person pronouns such as he, him, she, her, and it. Once the system is improved to extract possessive pronouns (e.g. 'Peter loves his photos.'), new syntactic constraints with respects to local domains can be added to the anaphora resolution algorithm for resolving possessive pronouns. Apart from this, more studies on the distribution of salience weights and discovery of new salience factors can enhance the chance for selecting the correct antecedent from the set of competing candidates. A pronoun-pronoun co-reference algorithm will also be investigated in future for resolving any anaphors that have another pronoun as their antecedents (e.g. "Mary gave a book to Jane. *She* told her that *she* love the story very much.") In addition, when the anaphora resolution system for singular third-person pronouns is mature enough, the system can be improved to handle plural pronominal anaphors and pleonastic 'it'.

Reference

1. Bunescu, R.: Associative Anaphora Resolution: A Web-based Approach. Proceedings of the EACL-2003 Workshop on the Computational Treatment of Anaphora. Budapest Hungary (2003) 47-52
2. Chomsky, N.: Lectures on government and binding. Foris, Dodrecht (1981)
3. Denber, M.: Automatic Resolution of Anaphora in English. Imaging Science Division, Eastman Kodak Co. (1998)
4. Evans R.: Mitkov's Anaphora Resolution System (MARS Demo). University of Wolverhampton. Available: http://clg.wlv.ac.uk/MARS/
5. Kehler, A., Shieber, S.: Anaphoric Dependencies in Ellipsis. Computational Linguistics. 23(3) (1997) 457-466
6. Lappin, S., Leass, H.: An Algorithm for Pronominal Anaphora Resolution. Computational Linguistics. 20(4) (1994) 535-561
7. Meyer, J., Dale, R.: Using the WordNet Hierarchy for Associative Anaphora Resolution. Proceedings of Building and Using Semantic Networks. Taipei Taiwan (2002)
8. Mitkov R.: Robust pronoun resolution with limited knowledge. Proceedings of COLING'98/ACL'98. Montreal Canada (1998) 869-875
9. Mitkov R.: 1999. "Anaphora resolution: The state of The Art". *Working paper (Based on the COLING'98/ACL'98 tutorial on anaphora resolution).* University of Wolverhampton UK (1999)
10. Mitkov R: Anaphora Resolution. Pearson Education Limited. Longman. Great Britain. (2002)
11. Modjeska, N., Markert, K., Nissim, M.: Using the Web in Machine Learning for Other-Anaphora Resolution. Proceedings of EMNLP-2003. Sapporo Japan (2003) 176-183
12. Palomar, M., Ferrandez, A., Moreno, L., Martinez-Barco, P., Peral, J., Siz-Noeda, M. Munoz, R.: Am Algorithm for anaphora Resolution in Spanish Texts. Computational Linguistics. 27(4) (2001) 544-567
13. Polany, L., van den Berg, M.: Logical Structure and Discourse Anaphora Resolution. Proceedings of ACL99 Workshop on the Relation of Discourse/Dialogue Structure and Reference. College Park USA (1999) 110-117
14. Soon, W. Ng, H., Lim, D.: A Machine Learning Approach to Coreference Resolution of Noun Phrases. Computational Linguistics. 27(4) (2001) 521-544
15. Stuckardt, R.: Design and Enhanced Evaluation of a Robust Anaphora Resolution Algorithm. Computational Linguistics. 27(4) (2001) 481-50514.
16. Walker, M.: Centering, Anaphora Resolution, and Discourse Structure. In: Walker, M., Joshi, K. Prince, E. (eds.): Centering in Discourse. Oxford University Press, UK (1998)