

Resolving Co-reference Anaphora Using Semantic Constraints



Parma Nand

School of Computer and Mathematical Sciences

Auckland University of Technology

A thesis submitted for the degree of

Doctor of Philosophy (PhD)

2012 June

Abstract

Anaphora resolution forms a critical cornerstone of natural language computational systems and hence forms a large part of computational linguistics literature. Study of anaphora is as old as the origin of languages, however computational study of anaphora started in the late seventies. Since then, there has been substantial progress in terms of both our understanding of anaphora as well as their resolution by computational systems. A variety of anaphora exist in natural languages and correspondingly, a variety of strategies are needed to resolve them. While some proportion of anaphora can be resolved by the use of basic syntactical and pragmatic strategies, the leftover ones require the use of various types of semantics. This thesis presents an enhanced framework for resolving anaphora which integrates the existing syntactic and pragmatic based strategies with the use of semantics.

In this thesis, we present the use of anaphora in a new light; that is, we emphasize its role as a shortcut elaborative device in addition to its role as a co-reference pointer. The latter has been emphasized a lot in previous studies. We show that this elaborative function is achieved by the use of an alternative word or a combination of words to form a noun phrase that serves as an anaphor. This additional functionality can be seen across the whole range of anaphora, from simple pronouns to multi-word noun phrases or compound nouns. We show how the interpretation of knowledge embedded in them is used to identify the antecedent and this is the key strategy we used in developing our new approach.

The resulting anaphora resolution algorithm is directly based on anaphora functioning as a shortcut elaborative device. The algorithm extracts knowledge from the document itself and from WordNet, and uses it to uncover elaborative information embedded in the anaphor. The latter is then used to identify the antecedent entity. The implemented system, named *aCAR*, is written in java, and it takes as input, a shallow parsed clausal structure of each sentence found in newspaper articles. The latter is a LISP based, in-house, parser. The resolution algorithm uses the fact that information about an entity in a discourse is expressed sequentially. Hence, critical information that can be used to help resolve an anaphor may be expressed after the mention of the anaphor. This fact gives rise to two crucial aspects of our algorithm. Firstly, an anaphor that is difficult to resolve with the current amount of information is left in a semi-resolved state, to be resolved later, instead of immediately taking a decision based on the current information. Hence, our algorithm uses a multi-pass approach. Secondly, our algorithm resolves anaphora at the level of the discourse, not to a single antecedent at a local level. This approach uses the fact that an entity is referred by various noun phrases (NPs) in a document and the information required to resolve an anaphoric NP can be embedded in any one of the NPs referring to the same entity. This can be before or after the anaphor. Hence when an anaphor gets resolved, all the information pertaining to the anaphor-NP and the antecedent-NP is merged and this accumulated information is used for all subsequent resolutions as well as attempts to resolve an anaphor in a semi-resolved state.

The anaphora resolution system was tested using a corpus consisting of 35 online newspaper articles from The Press, The Dominion and The Herald. Out of these, 20 were used as training data and 15 as test data. These gave us a total of 915 (out of 2323) anaphoric NPs for training data and 723 (out of 1895) anaphoric NPs for test data. The results were eval-

uated both in terms of correct resolutions as determined by the author and also compared with the results obtained by several similar systems. Our system performed at similar or better levels compared to most of the systems evaluated against, however our system resolves a much wider range of anaphora. The resolution task that was most similar in terms of the range of anaphora attempted and the level of resolution was *Message Understanding Competition* tasks (*MUC-6*). The highest precision rates achieved by systems participating in this competition were 71% compared to a precision rate of 78% for our system.

In summary, this thesis delivers results on three levels. Firstly, it provides an enhancement of the theory on the natural language phenomena of anaphora usage. Secondly, it provides a relational framework to substantiate the theory. Thirdly, it provides the results of implementing the framework and resolving anaphora from naturally occurring discourses and the results include a careful evaluation of its performance.

Acknowledgements

It is my pleasure to acknowledge those individuals who have significantly impacted on the completion of this thesis.

To start with, I would like to thank Professor Wai Yeap for firstly accepting me as a PhD candidate, and secondly seeing me through to completion. I have gained immensely from his guidance, insights and intense discussions, both on the subject matter as well as research in general.

Secondly, I would like to thank my wife, Kalpana Nand, and only son, Pranav Nand for enduring with me coming home late and disappearing in the weekends. Additionally, Kalpana did the proof reading and Pranav helped by doing his homework by himself.

I would also like to express my utmost appreciation to The School of Computing and Mathematical Sciences at AUT for funding this study in entirety. This includes the fees as well as all other required resources. I would also like to thank the individual members of staff of the department for the feedbacks and words of encouragement.

The completion of this thesis was also made possible due to direct and indirect contributions by several other individuals from my circle of family and friends. They are so many, hence I would not name them individually, however I would like to express my deepest gratitude to each one of them for their well-wishes and having an unwavering faith in me.

Lastly, I would like to thank my extended family from Labasa in Fiji, consisting of my parents, five brothers and a sister. In particular, I would like to thank my eldest brother, late, Mr Nitya Nand for kindling the hunger for knowledge in my early school years. Without this hunger, perhaps pursuing this PhD to completion would have remained only a dream. With this PhD, I would like to kindle the same hunger to quest for knowledge in my son, Pranav Nand.

To

Pranav Nand

For

Liberation With Knowledge

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

The thesis work was conducted from Nov 2005 to May 2012 under the supervision of Professor Wai Yeap at Auckland University of Technology.

A handwritten signature in blue ink, appearing to read 'Parma Nand', with a stylized flourish underneath.

Parma Nand

Auckland, New Zealand

June 2012.

Contents

List of Figures	xiii
------------------------	-------------

List of Tables	xv
-----------------------	-----------

1	Introduction	1
1.1	An Introduction to This Thesis	1
1.2	An Introduction to Anaphora	2
1.3	Publications From This Research	12
2	Review of Works on Anaphora Resolution	14
2.1	Factors Used in Anaphora Resolution	17
2.1.1	Hard Constraints	17
2.1.1.1	Morphological Compatibility	17
2.1.1.2	Syntactic Binding Constraints	18
2.1.1.3	Semantic Consistency	18
2.1.2	Soft-Constraints or Preferences	18

CONTENTS

2.1.2.1	Parallelism	19
2.1.2.2	Focus	19
2.1.2.3	Segmentation	20
2.1.2.4	Coherence	22
2.2	Works Focussed on Pronominal Anaphora	23
2.2.1	Hobbs' Syntax Based System	23
2.2.2	Carter's Shallow Processing Hypothesis (SPAR)	24
2.2.3	Carbonell et al's Multi Strategy approach	25
2.2.4	Mitkov et al's Knowledge Poor and Genetic Algorithm Approach (MARS)	26
2.2.5	Lappin et al's Syntax Based Approach (RAP)	29
2.2.6	Kennedy et al's Parser-less Approach	31
2.2.7	Baldwin's High Precision Approach(CogNIAC)	32
2.2.8	A Comparative Summary of Pronominal Systems	33
2.3	Works that also Resolve NP Anaphora	33
2.3.1	Dahl's Discourse Structure Approach	33
2.3.2	Poesio et al's Multi Strategy Approach	37
2.3.3	Soon et al's Machine Learning Approach	40
2.3.4	Ng et al's Machine Learning Approach with Richer Features	43
2.3.5	Poesio's Neural Network Approach	44
2.4	Chapter Summary	44

3	Interpretation of Anaphora – Our Approach	47
3.1	Some Preliminaries on our Approach	48
3.2	Anaphora – An Elaboration of the Antecedent	50
3.3	Generation of Compound Nouns	56
3.4	Predicate Deletion Constraints	63
3.5	Normalization Constraints	65
3.6	Chapter Summary	69
4	Anaphora Resolution Implementation	71
4.1	Descriptions of Major Modules	72
4.1.1	Parsing	72
4.1.2	Other Pre-Processing Adjustments	75
4.1.3	Handling of Multi-Word Noun Phrases	76
4.1.4	Handling of Synonymous Nouns	79
4.1.5	Handling of Other Forms of Noun Phrases	80
4.1.6	Treatment of Word Senses	82
4.2	Noun Phrase Resolution Algorithm	86
4.2.1	Data Structures Used in the Implementation	86
4.2.2	Upper Level Algorithm	88
4.2.3	Treatment of Determiners	96
4.3	Extracting Relational Knowledge	101
4.3.1	The CAUSE Relation	102
4.3.2	The HAVE Relation	105
4.3.3	The MAKE Relation	106

CONTENTS

4.3.4	The USE Relation	107
4.3.5	The IN Relation	108
4.3.6	The FOR Relation	109
4.3.7	The FROM Relation	109
4.3.8	The ABOUT Relation	110
4.4	A Summary of Derivational Aspects of Relations	110
4.5	Hard Constraints Guiding Anaphora Usage	112
4.5.1	Relational Constraints	113
4.5.1.1	Predicate Deletion	113
4.5.1.2	Normalization	114
4.5.2	Morphological Constraints	116
4.5.3	Syntactical Constraints	117
4.5.3.1	Orthogonality	117
4.5.3.2	Pronouns	119
4.6	Soft Constraints Guiding Anaphora Usage	123
4.6.1	A Summary of the List of Constraints Used	125
4.7	Chapter Summary	126
5	Input Data and Experimental Results	127
5.1	Input Data	128
5.2	Results	129
5.2.1	Some Comparisons	130
5.2.2	Result Breakdowns	133
5.3	Some Identified Shortfalls of this Research	146
5.4	Chapter Summary	147

6	Conclusions and Future Work	148
6.1	The Contributions of this Research	148
6.2	Future Works	149
6.3	Concluding Remarks	151
A	Appendix	153
A.1	Local Lexicons used in the Algorithm	153
A.2	Sample Annotator Forms Used	160
A.2.1	Normalization Annotator Form Sample	160
A.3	Samples of Input Text	162
A.3.1	The Full “Trucks” Article	162
A.3.2	The Full “Taupo” Article	164
A.3.3	The Full “Levinman” Article	166
A.3.4	The Full “Zaoui” Article	168
A.3.5	The Full “Rugby” Article	170
A.3.6	The Full “Aaron” Article	172
A.3.7	The Full “Poisoning” Article	174
A.4	A Sample Evaluation File from aCAR	176
A.5	Sample Text Output from aCAR	188
	Bibliography	202

List of Figures

1.1	Top Level Schema of the Anaphora Resolution Framework	9
2.1	Decision Tree Describing Poesio et al's Algorithm	39
3.1	Diagram Illustrating the Use of Anaphoric NPs in a Discourse	49
3.2	Context Drawing from Downing (1977)	59
3.3	Table showing Approximate Relation Mappings of Tratz and Hovy (2010) to Other Relations Sets.	60
4.1	Lisp Style Sample Clausal Output from the Parser.	74
4.2	The Syntactical Order of Components for a Generic NP	77
4.3	Document Class Diagram	86
4.4	The Main Modules of aCAR	91
4.5	Top Level Anaphora Resolution Algorithm.	92
4.6	Diagram Showing the Inner Details of the Constraints Application Module	95
4.7	Diagram Illustrating the Complement Process	97
4.8	Diagram Illustrating the Qualifier Process	98

LIST OF FIGURES

4.9	Diagram Illustrating the Quantifier Process	99
4.10	Illustration of a Causative Relation between Two NPs with a Connect- ing Verb	104
5.1	Trucks Article.	136

List of Tables

2.1	Pronoun Resolution Comparison	34
3.1	List of Nine Deletable Predicates	62
4.1	List of Determiner Types with Examples	77
4.2	List of Of-construction Types with Examples	82
4.3	Criteria Used for Causative Verbs	105
4.4	Summary of Derivational Aspects of Relations	111
4.5	List of Verb Inflicted Suffixes	115
4.6	Morphological Properties of Nouns	116
4.7	Anaphora Resolution Relevant Categories of Pronouns	120
4.8	Possible Syntactical Roles of Third Person Pronouns	122
4.9	Salience Factors Used for Resolution	124
4.10	Summary of Hard Constrains Used in the Implementation	125
5.1	Raw Statistics for the Input Data	128
5.2	Overall Resolution Results.	129
5.3	Pronoun Resolution Comparison	132

LIST OF TABLES

5.4	Overall Evaluation of Heuristics	133
5.5	Relational Resolution Breakdown	134
5.6	Incorrect Resolution Results Breakdown	142
A.1	Object Types Used in aCAR	153
A.2	Pronouns Used in aCAR	154
A.3	Prepositions Used in CAR	155
A.4	Be-Verbs in aCAR	156
A.5	Demonstratives Used in aCAR	156
A.6	Time and Place Prepositions Used in aCAR	156
A.7	Qualifiers Used in aCAR	157
A.8	Quantifiers Used in aCAR	157
A.9	Oral Verbs Used in aCAR	157
A.10	Articles Used in aCAR	157
A.11	Titles Used in aCAR	158
A.12	Irregular Plurals Used in aCAR	158
A.13	Coordinating Conjunctions Used in aCAR	158
A.14	Subordinating Conjunctions Used in aCAR	159

NP	Noun Phrase.
NPs	Noun Phrases.
NLP	Natural Language Processing.
VP	Verb Phrase.
HN	Head Noun of a compound noun.
POS	Part of Speech.
XML	Extensible Markup Language.
AR	Anaphora Resolution.
aCAR	almost Complete Anaphora Resolver.
RU	Referrable Unit.
MUC	Message Understanding Competition.
P	Precision value.
R	Recall value.
F-measure	A measure that combines P and R.
RST	Rhetorical Structure Theory.
c-relation	Compound noun generation relation.
c-command	Aa theory which enforces two NPs in a clause to be different.

Chapter 1

Introduction

1.1 An Introduction to This Thesis

This thesis presents a novel framework for resolving anaphora which integrates the existing syntactic and pragmatic strategies with a new semantic framework resulting into a unified, wide-ranging, anaphora resolution algorithm. The semantic framework is based on an enhanced definition of anaphora which is able to better explain the use of complex noun phrases as referents. The new semantic framework was then formulated into an anaphora resolution algorithm which was implemented as an automated system and tested on corpus consisting of newspaper articles. The results obtained is analyzed for performance including a comparison with other similar implementations. Hence, the thesis proposes a new theoretical framework for the use of anaphora in natural languages including a a practical application of the framework in a language engineering task. Section 1.2 describes the range of examples of anaphora usage followed by the details of how it will be examined in this thesis.

1.2 An Introduction to Anaphora

The term **anaphora** originated from an ancient Greek word “*αναφορά*” which means “the act of carrying back upstream”. In the context of natural language processing, the term **anaphor** is a reference which points back to an object that has been mentioned previously in the text being processed. The referred object is called the **antecedent**. The anaphor can be the same noun as the antecedent, a variation of the noun or a completely different noun. This phenomenon of using another word (an anaphor) to refer to an object in the current context is exhibited in all language use, implying an underlying psychological or pedagogical theory. However to date, there is no single universally accepted theory to explain the phenomenon. Anaphora usage and its interpretation has been studied for decades by researchers from different disciplines with varying objectives pertaining to their respective disciplines. Psycho-linguists (for example, Butterworth, 1974; Clark and Sengul, 1979; Gordon et al., 1993; Hobbs, 1979; Johnson-Laird, 1983) have studied it from the perspective of the mental representations and processes used in its production and interpretation. Linguists (for example, Mann and Thompson, 1988; Marslen-Wilson et al., 1982; Passonneau, 1998; Tasmowski and Verluyten, 1985; Walker, 1997) have investigated it within the syntactic and semantic constraints of language. Computational Linguists (for example, Brennan et al., 1987; Kameyama, 1986; Kehler, 1993; Mccord, 1990; Reichman-Adar, 1984; Sidner, 1979) study how anaphora can be resolved by computers and they implement and test theories from the related disciplines. This thesis falls into this last category.

To illustrate some examples of English anaphora usage in the order of complexity, consider the sentences from (1.1) to (1.6) below:

(1.1) Jason loves **his** toy car.

(1.2) Peter and John love apples and **they** eat **them** often.

(1.3) A boy and a girl entered the room. **The boy** was tall and wore a hat.

(1.4) I bought a Honda Civic. **The car** is red in color.

(1.5) John bought a house. **The windows** are wooden.

(1.6) John drove into an electricity pole last night. **The accident** caused a blackout.

The simplest case is illustrated by the use of a pronoun in (1.1) where the anaphor *his* refers to the previously mentioned entity *Jason*. In this case the resolution is trivial, since there is no other suitable candidate for the anaphor. The complexity increases in example (1.2), where there is more than one candidate for the pronouns *they* and *them*. The reader has to decide that *they* refers to *Peter and John*, and *them* refers to *apples*. Apart from the use of pronouns we can use alternative noun phrases (referred to as noun phrase (NP) anaphora) to serve the same purpose as pronouns as illustrated by example (1.3), where *The boy* in the second sentence refers to *A boy* mentioned in the first sentence. Again, in this case there is no competitive candidate, and hence a computational system can do simple string matching to deduce that the two mentions of *boy* are coreferential. However this is not a strategy that can be used in example (1.4). This example requires knowing that *Honda Civic* is a type of car. Hence, *The car* in the second sentence refers to *Honda Civic*. Further, in example (1.5) the anaphor *The windows* refers to *a house*, however not in the same way as the anaphora *The boy* and *The car* used in examples (1.3) and (1.4). This has been described as an example of associative or bridging anaphora where the anaphor indirectly refers to the antecedent. The level of complexity is elevated even further in example (1.6) where the anaphor *The accident* refers to the whole sentence rather than a particular entity. In this case the resolution is dependent on interpreting the combined meaning of the sum of words from the whole sentence.

The above examples illustrate the wide-ranging and complex use of anaphora and significant research has been conducted into understanding how anaphora are resolved, both theoretically from a linguistic point of view and practically from a computational standpoint. From the above examples, it is clear that anaphora resolution requires

complex knowledge of word meanings (such as semantic, discourse, domain, commonsense, and others) and this has posed a major problem for developing practical systems. How do we encode and make the right piece of knowledge available? Early computational studies on anaphora resolution thus began by investigating the use of constraints (such as morphological compatibility, syntactical constraints, and others) that require little or no reasoning of word meanings and surprisingly, these constraints were found to be quite effective for filtering away some initial candidates and especially useful for resolving pronominal anaphora. Good reviews of these early studies could be found in Mitkov et al. (2002)

Since then, researchers have focused on how to extend such basic systems to incorporate the use of other knowledge for anaphora resolution. Works done on this problem are varied. Some have focused on developing various logics (such as the use of Dynamic Predicate Logic and Dynamic Quantifier Logic (Livia Polanyi, 1999), in an attempt to provide a better semantic machinery for anaphora resolution. Others have discussed various architectures needed to support a complete anaphora resolution system. For example, Lappin (2003) discussed how the different methods of resolution should be utilized within a single model. He suggested that one should begin with the most computationally efficient and inexpensive methods and then progress to the more costly ones if needed. Chan and T'sou (1999) discussed the use of a Bayesian model network to represent, and a neural network to generate, related knowledge to aid resolution. Yet, others have discussed how the much needed knowledge could be derived and used. In these studies, two popular tools used were databases (such as WordNet (Fellbaum, 1998)) that provide synonyms, hypernyms and other relations between words, and the Web where word relations could either be learned or found. Examples of work using WordNet are Strube and Hahn (1999), Vieira and Poesio (2000) and those using the web are Bunescu (2003) and Poesio et al. (2004a). While domain knowledge and word meanings are important, the focus or the attentional structure of a discourse also plays an important role for resolving anaphora. Examples of work

that have shown the importance of discourse focus¹ for anaphora resolution include Brennan et al. (1987), Mitkov (1993), Kasper et al. (1999), and Palomar et al. (2001).

With much work done on extending the basic system to utilize different knowledge sources in recent years, the problem of resolving bridging anaphora is drawing much attention among computational linguists interested in anaphora resolution (e.g. Abad et al., 2010; Markert and Nissim, 2005; Nedoluzhko et al., 2009; Vieira et al., 2006). Partly, this is an intriguing problem that needs to be solved, and secondly, rich corpora are now becoming available whereby different kinds of relations between words can be retrieved. Some of these relations are useful for identifying bridging anaphora and researchers are keen to demonstrate how this could be done computationally. However, in this thesis, we argue that the extension of the basic notion of an anaphor as a co-reference to that of bridging is both unnecessary and unhelpful. It is unnecessary because the bridging problem is a separate and a well-known problem referred to as the text cohesion problem². It is unhelpful because solving it does not help us to advance our understanding of the basic anaphora resolution problem, namely the co-reference problem. In fact, doing so distracts researchers from discovering further constraints needed to solve the fundamental co-reference problem and gives one the impression that the latter is largely solved. It is not.

The problem of anaphora resolution as a co-reference problem can easily be understood in examples involving the use of pronouns or pronominal anaphora such as (1.7) below:

(1.7) **John** bought a new camera for **himself**. **He** bought a **house** in Tauranga and **it** was beautiful.

The pronouns co-refer to a noun mentioned earlier. Note that this is a special and a unique problem that exists within the more general problem of text cohesion. If

¹referred to as *centre* in some studies, e.g. Grosz and Sidner (1986)

²An overview of text cohesion can be found in Halliday and Hasan (1976)

different words are used in the same text to refer to the same entity, then they must be identified as such. If not, the text would be interpreted incoherently.

The idea of a bridging anaphor is often introduced using sentences such as (1.8) below:

(1.8) John bought a new **camera** for himself. He was pleased because the **pictures** were perfect. He bought a **house** in Tauranga and it was beautiful. The **windows** were large.

In this case, “pictures” and “windows” were referred to as bridging anaphora for “camera” and “house” respectively. It was argued that these words refer to an entity whose existence is only inferred via some other words mentioned earlier and without which the two sentences could become incoherent. However, we argue, the generality of this problem makes it indistinguishable from the problem of text cohesion.

First, note that not all definite descriptions are anaphoric in the sense that they could only be understood via the finding of its antecedent in a previous sentence. In fact, Poesio and Vieira (1998) found that such first-mention definite descriptions are numerous in their corpus of newspaper articles. Second, bridging is not a problem confined to definite descriptions as the example (1.9) below shows (from Marie (2006)):

(1.9) The real estate agent recommended us **the house**. However, **the door** was rotten and **a window** was broken.

Thus, bridging is needed to properly interpret all noun terms. Finally, in the same study, Poesio and Vieira (1998) also showed that the antecedent for bridging anaphora is not necessarily unique. Different words or context mentioned earlier could provide the additional information to enhance one’s understanding of the noun terms used in the current sentence. Not surprisingly then, resolving bridging anaphora require all kinds of inferences, some of which include one’s belief that could later turn out to be incorrect. For example, consider the sentence (1.10) below (from Asher and Lascarides (1998)):

(1.10) Jack was going to commit suicide. He got a rope.

Asher and Lascarides (1988, p.84) noted that one infers that the rope is to be used in the suicide and without which there is no connection between the contents of the two sentences, leading to text incoherence. The assumption is that Jack is about to hang himself. However, this could change as in (1.11) below:

(1.11) Jack was going to commit suicide. He got a rope. He drove to the hospital and used the rope to climb into the drug store.

Bridging is thus a more general problem and it is about enriching noun terms with whatever contextual information one could bring to bear on their interpretation. Given the generality of the problem it is no wonder that the definitions of bridging anaphor is inconsistent. For example, Asher (1993) defined bridging anaphora from a linguistic point of view as:

“an inference that two objects or events that are introduced in a text that are related in a particular way that isnt explicitly stated...”

Poesio et al. (2004a) (based on Clark (1977)) defined bridging anaphora from a computational viewpoint as:

“anaphoric expression that cannot be resolved purely on the basis of string matching and thus require the reader to bridge the gap using common sense inference.”

For a start, the two definitions of a bridging anaphor are inconsistent. According to Poesio’s definition, when the noun “car” is referred to using the noun “vehicle”, it is a bridging anaphor, since it cannot be resolved by string matching. However according to Asher, it is not, since the anaphor explicitly functions to identify the entity “car”, hence it is not a bridging anaphor. We thus conclude that bridging anaphora are not part of the anaphora resolution problem.

Anaphora resolution problem is thus the process of identifying the same real world entity represented by different words in a discourse. In short it is a co-referencing problem. An entity such as “John” introduced in a discourse can be referred to by the words “he” “driver” “father” “employee” etc. Each of these words, or any other anaphora used expresses further meaning for the same entity. For example, when we use the pronoun “he” as an anaphor for “John” we are implicitly communicating that John is male, singular and animate. If the word “father” is used instead, then it implicitly communicates that the antecedent is animate, male, has an offspring and all the other connotations that go with the meaning of the word “father”. Similarly when a subset or a superset noun is used as an anaphor, it again implicitly brings in new information to the forefront without its explicit statement. For example, consider the subset-noun “tanker” used as a co-reference to “truck”. The word “tanker” implicitly states that the “truck” being referred to is used to carry some form of liquid. Compound nouns consisting of two or more words when used as an anaphor provides even more meaning for the antecedent. It consists of the meanings of the individual words as well as the relation between them. For example, when the anaphoric compound noun “diesel tanker” co-refers to the noun “truck”, it communicates the fact that the truck had “something to do with diesel, probably a truck meant to carry diesel”.

In the remainder of this thesis, we present an enhanced framework for the study of anaphora as a co-referencing problem. The basic goal of anaphora resolution is to find all the unique entities mentioned in a discourse. This framework is different from other earlier studies in that it considers all noun terms to be anaphoric potentials. As observed above, the meanings of a noun term could help us locate its antecedent, if any. In this study, we found new ways to exploit the use of word meanings to locate antecedents, in particular, we focus on how compound nouns are formed. Research into how compound nouns are formed has been a separate and active research area. Examples of such work include Kim and Nakov (2011), Abad et al. (2010), Hendrickx et al. (2010), Girju et al. (2009), Nakov (2008), Butnariu and Veale (2008), Kim and Baldwin (2006), Nastase et al. (2006) and Tratz and Hovy (2010). These researchers

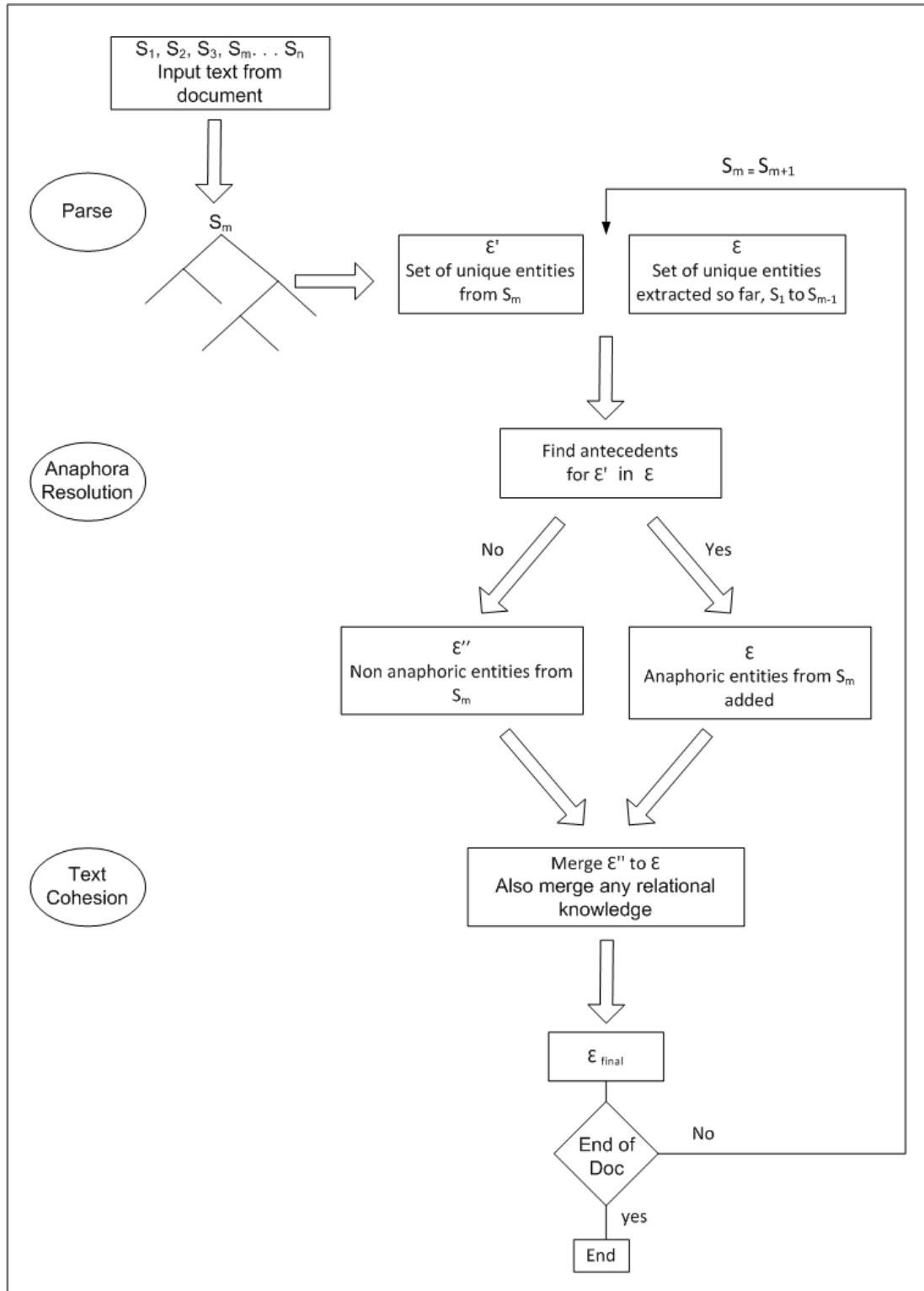


Figure 1.1: Top Level Schema of the Anaphora Resolution Framework - The diagram also shows the use of parsing and text cohesion in the framework

showed that compound nouns are not formed by combining *any* two or more words together. There are fixed rules that must be obeyed and our hypothesis is that these rules could be used to help locate the antecedents of anaphoric compound nouns and we have shown how to do this in an implemented system.

In the wider context of text processing, our treatment of all nouns as anaphoric potentials suggests that this is an important first step in the processing of nouns in a discourse. Unlike pronouns, each new noun is evaluated against the set of unique nouns encountered so far to establish if any of the unique nouns is an antecedent for it. If it is, its meanings and that of its antecedent are merged to enhance the overall meaning of the referred entity. If not, it becomes a new entity in the discourse and its meanings might then be enhanced in a later text cohesion process. Note that it is possible that a set of candidates can be found but a decision could not be made as to which is the antecedent. Such undecided cases are left alone and revisited whenever another noun term is encountered. As such, our anaphora resolution strategy is multi-passing. Figure 1.1 shows an outline of this overall process. In the implementation, the input is a set of shallow parsed clausal structures of all the sentences in the text and the output is a set of unique entities with all their respective references. Note that no text cohesion is performed, as this is not the concern of this study.

We have successfully implemented a new anaphora resolution system using the framework outlined above. The new system uses the following rules to help resolve anaphors: (i) rules for forming compound nouns, (ii) synonyms, (iii) traditional constraints such as preference rules (e.g. subject preference over objects), syntactic/semantic parallelism, c-command, and others as used in existing anaphora resolution systems (e.g. Lappin and Leass, 1994; Mitkov et al., 2002; Poesio et al., 2004a). The implemented system was tested on a corpus of 45 newspaper articles. The reason for choosing newspaper articles as the input data is that they contain some “difficult” as well as a relatively high usage of anaphora. For example our corpus consists of 4218 noun terms out of which 1638 (38.8%) were anaphoric. If we include the antecedents in the statistic for anaphora, this then gives us a figure of 2651 or 63%. This means only 37%

of the noun terms were first mentions, highlighting the extensive use of co-referential anaphora hence the importance of being able to resolve them for text interpretation. Our resolution system achieves an overall recall rate of 71% and a precision of 78% on the test data. This compares with recall values ranging from 35.69% and 62.78% and precision values ranging from 40% to 67% for the MUC-6 (van Deemter and Kib-ble, 1995) co-reference task competition. Details of evaluations with MUC6 and other systems is discussed in Chapter 5.

A synopsis of the thesis is as follows. Chapter 2 presents a review of the earlier works on anaphora resolution in order to present the range of strategies that have been employed and their success rates in resolving them. Chapter 3 presents a review of the works on compound noun formation and how these were used for our purpose of anaphora resolution. Chapter 4 discusses how the new anaphora resolution system was implemented and in particular, incorporating the compound noun rules identified in Chapter 3. Chapter 5 discusses the results obtained and how the system was evaluated. Finally, Chapter 6 presents the future directions and the concluding remarks.

1.3 Publications From This Research

Some partial results from this research have already been published in refereed conferences. The following are the list of these publications.

1. Nand, P. and Yeap, W. (2007). Challenges in anaphora resolution in the news media genre. In *The Proceedings of the fifth International Conference on Information Technology in Asia*, pages 156160, Sarawak, Malaysia.
2. Nand, P. and Yeap, W. (2008). On the use of salience weights in anaphora resolution. In *The Proceedings of the New Zealand Computer Science Research Student Conference, NZCSRSC*, Valencia, Spain.
3. Nand, P. and Yeap, W. (2010a). Generation of anaphoric noun phrases and their coreferential properties. In *The Proceedings of The Annual International Conference on Advanced Topics in Artificial Intelligence (ATAI)*, pages A75A80, Phuket, Thailand.
4. Nand, P. and Yeap, W. (2010b). Towards a discourse level coherence structure. In *The Proceedings of the 2nd International Conference on Agents and Artificial Intelligence*, pages 1119, Valencia, Spain.
5. Nand, P. and Yeap, W. (2011). A relational framework for interpreting anaphora. In *The Proceedings of the Annual International Conference on Advanced Topics in Artificial Intelligence (ATAI)*, pages 4449, Canning, Singapore.
6. Nand, P. and Yeap, W. (2012). Using relations to interpret anaphora. In *The Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, volume 1, pages 9299, Vilamoura, Algarve, Portugal.

1.3 Publications From This Research

(An extended version of this paper has been selected by Springer to be published in their book, *Lecture Notes In Computer Science (LNCS)*. We are in the process of editing this extended version to be submitted by 8 June 2012.)

Chapter 2

Review of Works on Anaphora

Resolution

Anaphora resolution is an area of computational linguistics that is as old as the discipline itself, hence there has been a considerable amount of work done in the past. However, new works are still being reported indicating that the problem is still not completely resolved. Anaphora in the English language can be divided into various types (Hirst, 1981), depending on the type of NP used for the anaphor. The most widespread types are:

- **Pronominal anaphora.** E.g. John likes apples and he eats them often.

This is a widespread category in which pronouns are used as the anaphoric NP. This category of anaphora is usually used to refer to antecedents which are in close proximity (Grosz and Sidner, 1986). Commonly used pronouns such as “he” and “she” are usually used to co-refer to a recently mentioned male or female, however sometimes, they can also be used to co-refer to entities mentioned after the anaphor. This is a rare

category and is referred to as *cataphora*. Apart from the forward and the backward co-reference, the pronoun “it” has also been shown to be sometimes non-coreferential. This category has been variously referred to as *pleonastic* and *zero* anaphora. An example of zero anaphora is the “It” in “It is going to rain today”. Hence, when resolving the pronoun “it” there is the added complication of determining whether a use of “it” is co-referential or not. Works have been reported, for example, Li et al. (2009) which employ various techniques to determine this.

- **NP anaphora.** E.g. John likes apples but the fruit is forbidden in his country.

The NP anaphora category includes cases in which an NP, apart from a pronoun, is used as the anaphor. This category has mostly been studied for definite NPs since these are easier to identify with the presence of the definite article “the”. However, works such as Abbott (2005) and Heim (1982) have argued that non-definite NPs can also be used to refer to entities in a discourse.

The pronominal and the NP anaphora represent the bulk of anaphora found in naturally occurring discourses. Apart from these, there are other minor categories of anaphora listed below:

- **One/Some anaphora.** E.g. If you cannot attend the morning lecture, you can attend the afternoon one.
- **This/That anaphora.** E.g. The stop button is located here. You can push this in an emergency.

Out of the above two categories, resolution of This/That anaphora is somewhat trivial, however resolution of the One/Some category is more complex as discussed in Dahl’s works Dahl (1986) and Dahl (1984).

Apart from the categories defined by the type of word used as anaphor, anaphora can also be categorized by the type of antecedent that it refers to. In addition to NPs,

antecedents can also be verb phrases, clauses, sentences or even whole paragraphs. Resolving to the latter targets is more complicated as it deals with extraction of semantics of the words comprising the discourse component. Some works, for example, Dahl et al. (1987) have attempted to solve a subset of anaphoric NPs (normalized NPs), that have verb phrases as their targets as these are less reliant on semantics. In our implementation we resolve such normalized NPs which have verb phrases as their targets. However, in addition we also examine normalized NPs in detail in order to resolve a wider range of normalized NPs.

Resolution of anaphora starts by first identifying the potential candidates in a search scope which typically are the NPs up to the previous sentence for pronouns and a couple of sentences for NP anaphora. The identification of the search scope is also critical as an over sized scope will introduce more competing candidates and an under sized one can potentially miss the antecedent. The search scope can also be dynamic so that when a suitably qualified candidate cannot be found in the initial scope, then it can be extended to include more candidates. The suitability of a candidate is based on various types of “factors”. Some examples are morphological compatibility, syntactic/semantic parallelism, semantic consistency and lexical distance. The factors have been referred to using various names such as “constraints and preferences” (Carbonell and Brown, 1988), symptoms (Mitkov, 1995) and indicators (Mitkov, 1998). The use of factors can either work in a discounting capacity to eliminate an NP as a candidate, or work to propose an NP as a higher preference candidate. A number of studies on anaphora resolution use a similar set of factors, however the computational strategy used to apply the factors may vary. In Section 2.1 I present a summary of these factors in order to give an overview of the common anaphora resolution factors independent of the strategies used to apply them in individual works. When searching for candidates for an anaphor, the factors are typically used as constraints applied to a set of previous NPs. Some of the factors can be applied in “absolute” capacity, that is, they have to be satisfied, while others can be applied only as a preference. We will use the term “hard constraints” for the “absolute factors” and these are discussed in Subsection 2.1.1. The

“soft constraints”, also called “preferences” in some works, will be used for constraints which can be occasionally violated. This is discussed in Subsection 2.1.2.

The strategies that can be used to resolve pronominal anaphora are characteristically different to those that can be used for NP anaphora. This has resulted in works that can be divided into two categories; those that focus on pronominal anaphora and those that focus on NP anaphora. The works reviewed in this chapter are divided accordingly. Section 2.2 reviews works that focuss on pronominal anaphora. Section 2.3 reviews works that focuss on NP anaphora. Finally Chapter 2.4 presents a summary of this chapter.

2.1 Factors Used in Anaphora Resolution

2.1.1 Hard Constraints

2.1.1.1 Morphological Compatibility

This is the most basic set of constraints which needs to be satisfied for both pronominal and NP anaphora. It includes gender, number and animacy agreement between an anaphor and its antecedent.

Examples:

John and **Jane_i** drove to town to buy **her_i** a dress.

The **driver_i** rammed the car into the pole. Police charged the **man_i** with dangerous driving.

2.1.1.2 Syntactic Binding Constraints

Government and Binding theory (GB) (Chomsky, 1969), formalized as *c-command* by Reinhart (1983) has also been used as constraints on anaphora and its antecedents. The c-command restrictions have mostly been used in a discounting capacity to eliminate NPs as candidates. It is based on the idea that an entity in a certain syntactical position has to be different to an entity in another syntactical position within a single clause. The following are illustrative examples from Ingria and Stallard (1989)

- (a) A non-pronominal NP cannot overlap in reference with any NP that c-commands it. E.g. **He**_{*i*} told them about **John**_{*j*}.
- (b) The antecedent of a bound anaphor c-commands it. E.g. **John**_{*i*} likes pictures of **himself**_{*i*}.
- (c) A personal pronoun cannot overlap in reference with an NP that c-commands it. **John**_{*i*} told **bill**_{*j*} about **him**_{*k*}.

2.1.1.3 Semantic Consistency

This is a knowledge dependent constraint which requires that both anaphor and antecedent satisfy the semantic possibility. The following example (from Mitkov (1999)) illustrates this.

- (a) John removed the diskette from the **computer**_{*i*} and then disconnected **it**_{*i*}.
- (b) John removed the **diskette**_{*i*} from the computer and then copied **it**_{*i*}.

2.1.2 Soft-Constraints or Preferences

As opposed to hard constraints, soft constraints are not obligatory, thus may occasionally not hold in some contexts. These are usually used to select an antecedent from

NPs that have already satisfied the hard constraints. Some illustrative examples are outlined in the following sections.

2.1.2.1 Parallelism

Parallelism can be either semantic or syntactic. In syntactic parallelism, the NP in the same syntactic function is given a higher preference. The following example from Mitkov (1999) illustrates syntactic parallelism:

- (a) The **programmer_i** successfully combined **Prolog_j** with C, but **he_i** had combined **it_j** with Pascal last time.
- (b) The **programmer_i** successfully combined Prolog with **C_j**, but **he_i** had combined Pascal with **it_j** last time.

In semantic parallelism, the NP in the same semantic role has a higher preference. An example is shown below:

- (a) John gave the disk to **Tom_i**. Kim also gave **him_i** a letter.
- (b) **John_i** gave the disk to Tom. **He_i** also gave Kim a letter.

2.1.2.2 Focus

The notion of “focus” in a discourse is also a widely used anaphora resolution factor. This factor again gives a higher preference to the NP which is currently the most salient entity in the discourse. The most salient entity is represented by the *Focus* (Sidner, 1979), or the *Center* (Grosz and Sidner, 1986), which are closely related but not exactly identical¹. In some instances it is difficult even for a human to resolve anaphora. For example, in “*Mary put the cup on the table and broke it*” the anaphor “it” is resolvable

¹Grosz et al. (1995) discusses the subtle difference between the two

to either “table” or “cup”. However, it becomes easier to resolve when the sentence appears in a context such as the following.

Mary got a cup for her birthday yesterday. It had a picture of a mountain on it. Immediately after receiving it, Mary put the cup on the table and broke it.

In the excerpt above, “cup” is the focus, hence has a higher preference to be the antecedent than “table”. The first requirement for an anaphora resolution algorithm that integrates focus is to devise a way to computationally track the focus in a sentence or clause. This may vary across computational strategies. Some examples are the use of *Centering Theory* (Sidner, 1979) in Poesio et al. (2004b) and *Recency and Frequency* in Akman and Ersan (1994).

2.1.2.3 Segmentation

Segmentation is a closely related concept to focus, which defines the accessibility of anaphora (esp. pronominal) restricted to only certain sentences in a discourse. The sum of these sentences are variously referred to as *segments* (Ide and Cristea, 2000) and *focus spaces* (Grosz and Sidner, 1986). The effect of segmentation on anaphora is effectively illustrated by the excerpt in (2.1), adapted from Vieira and Poesio (2000).

- (2.1) 2. *A deep trench now runs along its north wall, exposed when the **house_i** lurched two feet off its foundation during last week’s earthquake.*
 :
 19. *Others grab books, records, photo albums, sofas and chairs, working frantically in the fear that an aftershock will jolt the **house_i** again.*
 20. *The owners, William and Margie Hammack, are luckier than*

many others.

⋮

49. *When Aetna adjuster Bill Schaeffer visited a retired couple in Oakland last Thursday, he found them living in a **mobile home**_k parked in front of their yard.*

50. ***The house**_j itself, located about 50 yards from the collapsed section of double-decker highway Interstate 880, was pushed about four feet off its foundation and then collapsed into its basement.*

⋮

65. *As Ms. Johnson stands outside the **Hammack house**_i after winding up her chores there, the house begins to creak and sway.*

In sentence 49 a different instance of *house* is introduced as *mobile home*_i which is the antecedent of the anaphor *The house*_j in sentence 50. However the antecedent of *Hammack house*_i in sentence 65 does not co-refer to the most recent instance of house in sentence 50, but to the instance mentioned in sentences 2 and 19. According to segmentation theories, *The Hammack house*_i in 50 is able to be correctly resolved because sentences 49 and 50 form a segment and the entities within this segment are not accessible to the anaphor *Hammack house*_i. The segments are organized hierarchically and entities at a lower level are not visible to anaphora at a higher level (Fox, 1993; Grosz and Sidner, 1986; Reichman, 1985). Automatic determination of discourse segments is still largely an unresolved problem as it involves deeper level semantics dealing with intentions. An example of work using segmentation is reported in Vieira and Poesio (2000). In this study the authors experimented with segmentation by resolving anaphora using candidates from a fixed size window, indexed candidates consisting of head of nouns from the whole discourse, as well as a combination of these two methods. The results of these were not overly skewed towards any one segmentation scheme for it to be convincingly used or excluded in a system. We thus opted not to use any form of segmentation. We use all of the previous sentences as the candidate

window for NP anaphora and up to the previous sentence for pronominal anaphora, unless there are no candidates, in which case we progressively increase the window to the next previous sentence.

2.1.2.4 Coherence

An important component of discourse interpretation involves connecting clauses and phrases in order to establish a coherent discourse. Use of referring expressions is a critical tool which can be used to establish coherence between two clauses (Garnham, 2001; Halliday and Hasan, 1976; Johnson-Laird, 1983). A commonly used example (from Terry (1972)) used to illustrate the relation between anaphora and coherence is shown in (2.2)

- (2.2) (a) The city council denied the **demonstrators_i** the permit because **they_i** advocated violence.
- (b) The **city council_i** denied the demonstrators the permit because **they_i** feared violence.

In (2.2a) and (2.2b) the pronoun “they” co-refer to different antecedents which can be resolved only by establishing a *causal* inference relation between the two clauses Hobbs (1979). Although it is widely accepted that anaphora and coherence are closely related it is not clear what this relation is. Hobbs (1979) argues that anaphora resolution is a byproduct of establishing coherence, however Kehler (2002) argues that discourse coherence and pronoun resolution mutually constrain each other. He thus notes that resolution of inter-clausal anaphora is dependent on coherence relation between the clauses. In particular, three relations are proposed, *Causal*, *Similarity* and *Contrast*. He hypothesizes anaphora such as that in (2.3a) and (2.3b) is resolved as a result of establishing the relevant coherence relations.

- (2.3) (a) (Similarity) Gephardt organized rallies for **Gore_i**, and Daschle distributed pamphlets for **him_i**.
- (b) (Contrast) **Gephardt_i** supported Gore, but Arney opposed **him_i**.

The relation between anaphora and coherence is further re-enforced in Wolf and Gibson (2004). In this study the authors used self paced reading to test pronoun processing preferences in two clause sentences. The results showed that the preferences can be reversed by changing the coherence relation between the two clauses. In spite of these theoretical studies suggesting that coherence is an effective preference mechanism that can be used in anaphora resolution systems, it is still notably absent in implementations due to the challenge of computationally determining the coherence relations.

2.2 Works Focussed on Pronominal Anaphora

This section surveys a selection of works that focus on resolution of pronominal anaphora. The works are reviewed in approximate chronological order.

2.2.1 Hobbs' Syntax Based System

One of the simplest algorithms which is purely based on syntax is reported in Jerry (1978). It is described as a “naive” algorithm for finding the antecedents of pronouns which use a surface parsed flat tree structure identifying the subject, verb, object etc. of a clause. His algorithm uses simple syntactical rules and properties of pronouns to select antecedents. The algorithm was tested on selected chapters from two books and selected articles from *Newsweek*. Additionally, the algorithm assumes correctly parsed tree for the clauses hence manual corrections were done to achieve this. An overall resolution rate of 88.3% was obtained for the bare algorithm and 91.7% with some additional selectional constraints. Although these results look impressive, the authors

2.2 Works Focussed on Pronominal Anaphora

note that the numbers are somewhat deceptive since in over half the cases, there were no competing candidates. The algorithm and its achievement was a success at the time, however later it was regarded as inadequate (Hirst (1981)). The algorithm can not be expected to perform as well as a semantics based algorithm, however it is still a viable alternative due to its light weight hence computationally cheap.

Our system is an advancement of the Hobbs' system in two ways. It firstly resolves both pronominal as well as noun phrase anaphora hence can be more readily applied to a practical application. Secondly, it uses a semantic framework, hence is able to resolve the "difficult" category anaphora which can only be resolved by using knowledge.

2.2.2 Carter's Shallow Processing Hypothesis (SPAR)

In the work reported in Carter (1987), Carter uses a "shallow processing" approach to resolve anaphora. In this approach he proposes that since domain knowledge and reasoning tend to be expensive to implement, maintain and use, we should exploit general linguistic knowledge as much as possible. His proposal is based on the fact that natural language texts are relatively redundant and constructed considerably, hence it would be possible to interpret text (including resolving anaphora) by using linguistic and/or non-linguistic techniques based on information from the discourse itself. He further emphasizes that linguistic techniques should be preferred over non-linguistic ones since they are more general and less open ended. This idea is tested in a computational system called SPAR (Shallow Processing Anaphor Resolver) as part of his PHD thesis. The SPAR architecture firstly applies existing linguistic theories, most notably, Sidner's (Sidner, 1979) theory on local focusing to resolve anaphora. If the anaphor still cannot be resolved completely, then domain knowledge is sparingly used. In the system presented in this thesis we enhance the use of semantics to be used from the start rather than as a fall back mechanism. We also make use generic knowledge encoded in WordNet rather than domain knowledge which makes our system more generalizable.

2.2 Works Focussed on Pronominal Anaphora

Carter reports a resolution rate of 93% (out of 242) for pronominal and 82% (out of 80 for non-pronominal anaphora. The high level of accuracy is impressive, however, the input text used for testing were specifically written for the project, albeit, not by people associated with SPAR. Nevertheless, the results provide evidence that shallow processing can be used to resolve anaphora with a high level of accuracy.

2.2.3 Carbonell et al's Multi Strategy approach

In Carbonell and Brown (1988), the authors report a pronominal anaphora resolution approach that is based on the assumption that anaphora resolution may be best accomplished by using a combination of multiple strategies rather than a single monolithic technique. The study mainly focusses on inter-sentential anaphora noting that intra-sentential is more tractable.

The framework applies a range of knowledge sources by categorizing them into *constraints* and *preferences*. The constraints are applied first to reduce the candidate list and then preferences are applied to the remaining candidates. In a situation in which conflicting knowledge of equal strength yields more than one candidate as potential antecedents, the list is presented as the solution, rather than selecting a unique, and possibly wrong, antecedent. This was driven by the fact that out of 70 examples, 17 cases were judged to be ambiguous even by human subjects.

The list of knowledge sources used in the framework consisted of Local Constraints, Case Role Semantic Constraints, Pre/Postcondition Constraints, Case Role Persistence, Inter-sentential Recency Preference, and Syntactic Topicalization Preference. The implementation uses a parsed input containing syntactic as well as semantic knowledge. The semantic information includes a reasonably deep knowledge, for example knowledge required for the example in (2.4). This is an example of *Semantic Alignment preference*, however the verbs “went” and “left” reverse the alignment hence the antecedent “club” in (2.4a) and “park” in (2.4b). The accuracy or the scope of the parser is not discussed in the paper.

- (2.4) (a) Mary drove from the park to the **club**_{*i*}. Peter went **there**_{*i*} too.
(b) Mary drove from the **park**_{*i*} to the club. Peter left **there**_{*i*} too.

The framework was evaluated against manual simulations using 70 examples, yielding 49 unique resolutions, 17 conflicting possibilities and 4 anomalous ones. In the 49 unique cases the system output and human judgements correlated very well. Furthermore, the majority of the conflicting possibilities from the system were also judged ambiguous by human subjects as well.

This system reported in this study is very similar to our system in terms of application of rules as a set of constraints and preferences. Carbonell et al's implementation is also similar to ours in they both make use of multiple strategies. However, our system makes more extensive use of semantics to resolve a much wider range of anaphora where knowledge from one resolution is used to resolve another in multi-passes.

2.2.4 Mitkov et al's Knowledge Poor and Genetic Algorithm Approach (MARS)

Perhaps the biggest bulk of work in computational anaphora has been done by Ruslan Mitkov (some examples are, Barbu and Mitkov, 2001; Mitkov, 1993, 1994, 1995, 1998, 1999; Mitkov et al., 2002), either alone or in conjunction with others. Among all the works, three papers report anaphora resolution results from implemented systems, starting with Mitkov (1994), followed by Mitkov (1998) and finally Mitkov et al. (2002). The three papers report results from continual improvement of the techniques used starting with the initial system in 1994 to a fully automatic system in 2002. The initial system described in Mitkov (1994) uses an integrated modular architecture consisting of the following key modules:

- Syntactic Knowledge

2.2 Works Focussed on Pronominal Anaphora

- number, gender and animacy agreement
 - c-command constraints
 - syntactic parallelism
 - syntactic topicalization
- Semantic Knowledge
 - semantic consistency
 - case roles
 - semantic parallelism
 - animacy
- Heuristical Knowledge
 - rating rules
 - recency
- Discourse Knowledge
 - center tracking
- Domain Knowledge
 - world knowledge (not implemented)

The algorithm is customized for the sub-language of Computer Science texts rather than general texts. Although the use of general world and common sense knowledge is not implemented in the initial version, it is still described as knowledge based because it uses semantic, syntactic and discourse knowledge. He uses an algorithm using probabilities to track the center of the discourse, which plays a decisive role in proposing the most likely candidate. Additionally, some sub-language based knowledge is also integrated, which for example, was able to eliminate the ambiguity between “storage

2.2 Works Focussed on Pronominal Anaphora

devices”, “tape cassettes” and “record discs” as antecedents for the anaphor “their”. The evaluation results of this implementation is reported in Mitkov (1996). The program was tested in two ways; with syntactic, semantic and domain modules activated and with all these modules in addition to discourse modules activated. The results show an improvement in resolving anaphora when traditional linguistic approaches (syntactic and semantic constraints) are combined with the statistical approach for tracking the center (e.g. 89.1% vs. 87.7% , in the second case 86.7% vs. 91.6% accuracy).

Mitkov went on to enhance the initial system, however instead of intensifying the implementation on the domain knowledge module, which was not implemented in Mitkov (1994), he shifted his efforts towards a knowledge poor or minimal knowledge implementation. In Mitkov (1998), Mitkov reports the results of an implementation in which he omits the use of sub-language specific knowledge that was used for tracking the center. In this implementation he emphasizes the use of mostly syntactical and heuristic anaphora indicators to resolve pronouns. The candidates are given salience scores between -1 and 2 based on each of the indicators. The indicators, which were identified empirically were *definiteness*, *givenness*, *indicating verbs*, *lexical reiteration*, *section heading preference*, *“nonprepositional” noun phrases*, *collocation*, *immediate reference*, *referential distance* and *term preference*. The algorithm assigns a score between 0 and 2, for each of the indicators and the entity which scores the largest sum was chosen as the antecedent. The system was tested using text from technical manuals in English, which contained 71 pronouns in 120 pages of text. The success rate on this is reported to be 95%. In a second evaluation using a larger corpus of 47-page Portable Style-Writer Users Guide, containing 223 pronouns, the system obtained a success rate of 83.6%, giving an overall average of 89.7%. A strength of the system claimed in the paper is that it can be adapted to different languages with minimal modification. The study reports some tests done on Polish texts with a success rate of 86.2%.

In Mitkov (2002), Mitkov et al report results of a revamped version of the previous implementation. This version, referred to as MARS, automates extraction of certain types of information such as identifying non-nominal anaphora and recognition of animacy.

2.2 Works Focussed on Pronominal Anaphora

However, the most notable enhancement is the use of genetic algorithm to determine the values for salience scores to attain the optimum performance for anaphora resolution. While in the previous version a score between 0 and 2 was allocated using heuristics, the revamped version uses genetic algorithm to find the optimum set of scores for each of the indicators that gives the best results. In addition to these MARS has two other notable enhancements. The first is the use of a richer parsed input instead of a POS tagger. This enables MARS to change the way in which some of the indicators are implemented. Secondly, MARS uses three new indicators, namely *Boost Pronoun*, *Syntactic Parallelism* and *Frequent Candidates*.

MARS was evaluated on a much larger corpus consisting of computer hardware and software technical manuals featuring some 247,401 words and 2,263 pronouns. The overall success rate obtained without the use of genetic algorithm was 59.2% and 61.55% with the use of genetic algorithm. These success rates include any errors introduced by pre-processing.

The performance results reported in Mitkov's works are impressive, especially with limited use of knowledge. However, the range of anaphora covered are restricted to only pronouns. In leftover pronouns that have been incorrectly resolved are the ones in the "difficult" category and would require knowledge for resolution. Although Mitkov reports limited use of knowledge, it was very domain specific, hand coded knowledge. The framework reported in this thesis provides an enhanced knowledge integration framework which is generic and can be mined from any lexicon including the Web. The generic nature of the knowledge framework enables it to be useful for not only pronominal anaphora but for all types of noun phrase anaphora as will be shown in chapter 5.

2.2.5 Lappin et al's Syntax Based Approach (RAP)

In Lappin and Leass (1994), Lappin et al report the results from an implemented system, RAP (Resolution of Anaphora Procedure), which is predominantly based on syn-

2.2 Works Focussed on Pronominal Anaphora

tactical constraints implemented as preference scores. RAP uses parsed input generated by McCord's Slot Grammar parser (McCord, 1990). The sentences are parsed on the fly and the clausal representations of the previous four sentences are retained in the workspace which are used to supply the list of candidates. Preference for candidates is achieved via salience scores corresponding to a set of anaphora factors. The scores get degraded as the discourse gets processed. A summary of the factors is given below.

- **Intra-sentential syntactic filter** - a set of six conditions for non-coreference within a sentence.
- **Morphological filter** - rules out candidates on the basis of person, number and person features.
- **Pleonastic pronouns** - A set of syntactic and semantic tests are used to determine if an occurrence of "it" is pleonastic.
- **Anaphor Binding** - The hierarchy subject > agent > object > i-object > pre-object is used to bind inter-sentential anaphora occurring in pre-determined syntactical forms.
- **Salience Weights** - A procedure for assigning salience weights based on grammatical role, parallelism of grammatical roles, frequency, proximity and recency. Higher salience weights are assigned to subject over non-subject, direct objects over other complements, arguments of a verb over adjuncts and objects of prepositional phrase over adjuncts of the verb, and head nouns over complements of head nouns.
- **Anaphorically linked NPs** - a procedure for identifying anaphorically linked NPs for which a global salience score is computed as the sum of the salience value of its constituents.

The system was tested on a corpus of five computer manuals containing a total of 560 third person pronouns including reflexives and reciprocals. The overall success rate

2.2 Works Focussed on Pronominal Anaphora

(which would be precision in this case) obtained was 85%. This is again a syntax based scheme designed to work on only pronominal anaphora hence cannot be expected to perform on the leftover “difficult” category of anaphora nor on noun phrase anaphora.

2.2.6 Kennedy et al’s Parser-less Approach

Kennedy and Boguraev (Kennedy and Boguraev, 1996), report the results of a modified and extended version of the algorithm from Lappin and Leass (1994), however the main strength of this work is that it does not require in-depth, full syntactic parsing of text. The algorithm works off a part-of-speech tagger (POS), similar to Mitkov (1994), but enriched with grammatical function of the lexical terms. The reason for using a simplistic pre-processing tool was to reduce the level of pre-processing errors. The POS tagger used in the study produces a flat analysis of sentences with extremely high levels of accuracy, over 95% reported in Voutilainen et al. (1993). This, according to the authors, makes the algorithm more adaptable to a variety of text genres.

The algorithm was tested on 27 texts taken from a random selection of genres including newspaper articles, product announcements, news stories, magazine articles and World Wide Web pages. The corpus consisted of 306 third person pronouns out of which 231 (75%) were correctly resolved. This compares with 85% for Lappin and Leass who used a similar set of anaphora factors but used a richer parser over a very specific domain of computer manuals. The results reported in this study shows that it is possible to attain comparable levels of accuracy for pronominal anaphora resolution, without the use of a complex parser. The strength of this approach is that it does not use a parser, hence removes a dimension of error, however it would be possible to integrate any form of semantics in this approach without any knowledge of the syntactical role of the words. This means that the “difficult” category of pronouns and a large proportion of noun phrase anaphora would not be resolvable. The framework presented in this thesis presents a unified structure that can be used to resolve all categories of anaphora including the “difficult” category of pronominal as well as noun phrase anaphora.

2.2.7 Baldwin's High Precision Approach(CogNIAC)

Baldwin (1997) reports the results of an implementation which is based on only resolving anaphora which can be resolved using very high confidence rules. The system, named CogNIAC, only resolves anaphora which can be resolved using a sequence of disjoint syntactical rules applied in a predetermined order. Pronouns that fall outside the scope of these rules are left unresolved. Interestingly, for the chosen corpora (narrative texts), the rules have a precision of 97% (121/125) and a recall of 60% (121/201) for training data. The set of 6 rules used in the order in which they were applied are:

- **Unique in Discourse:** If there is a single possible antecedent *i* in the read-in portion of the entire discourse, then pick *i* as the antecedent: 8 correct, and 0 incorrect.
- **Reflexive:** Pick nearest possible antecedent in read-in portion of current sentence if the anaphor is a reflexive pronoun: 16 correct, and 1 incorrect.
- **Unique in Current + Prior:** If there is a single possible antecedent *i* in the prior sentence and the read-in portion of the current sentence, then pick *i* as the antecedent: 114 correct, and 2 incorrect.
- **Possessive Pronoun:** If the anaphor is a possessive pronoun and there is a single exact string match *i* of the possessive in the prior sentence, then pick *i* as the antecedent: 4 correct, and 1 incorrect.
- **Unique Current Sentence:** If there is a single possible antecedent in the read-in portion of the current sentence, then pick *i* as the antecedent: 21 correct, and 1 incorrect.
- **Unique Subject/ Subject Pronoun:** If the subject of the prior sentence contains a single possible antecedent *i*, and the anaphor is the subject of the current sentence, then pick *i* as the antecedent: 11 correct, and 0 incorrect.

2.3 Works that also Resolve NP Anaphora

For each of the pronouns, the above six rules are applied in the given sequence and if an antecedent is not found after applying the sixth rule, then that pronoun is left unresolved. The paper reports a precision rate of 92% and a recall rate of 64%. The low recall rate indicates the the low generalizability ability of the system. Again, the system does not use any form of knowledge hence would not be able to make a reasonable attempt at resolving the full range of anaphora. Our framework is thus an enhancement of this work in being able to attempt resolution of a much wider range of anaphora, including a structure to represent the relations between the different types of anaphora.

2.2.8 A Comparative Summary of Pronominal Systems

Table 5.3 summarizes the reported resolution rates and other relevant factors from a selection of described studies. It should however be noted that the numbers should not be directly compared as there are numerous varying factors in the implementations which tilt the playing field eroding the usefulness of direct comparison.

2.3 Works that also Resolve NP Anaphora

This section reviews a selection of works that either resolve only NP anaphora or resolve both pronominal as well as NP anaphora.

2.3.1 Dahl's Discourse Structure Approach

This subsection describes anaphora resolution works that use some form of higher level discourse structure such as Focussing and Segmentation.

Factor	Lappin and Leass (1994) - RAP	Kennedy and Boguraev (1996)	Baldwin (1997) - CogNIAC	Mitkov (1998) - MARS	Ge et al. (1998)
Genre of strategy	Mixed	Non-learning	Non-learning	Non-learning	Learning
Recall rate	–	–	64%	–	–
Precision rate	85%	75%	92%	89.7%	82.9%
Genre of data	Technical manual	Newspaper	Narratives and Newspaper	Technical manual	Journal
Input corpus	Computer Manuals	Randomly from newspapers	Narrative text about 2 persons	Manuals for VCR	Wall Street Journal
Sample size	5 Manuals, 82,000 words	27 texts	15 articles, 200 pronouns	140 pages, 71 pronouns, 48 considered	3975 sentences, 2477 pronouns, 1371 considered
Amount of knowledge used	Morphological info supplied	Morphological info from LINGSOFT	Morphological info supplied	Morphological info supplied	Morphological info supplied and mined
Pre-processing	McCords's Slot Grammar parser	Part-Of-Speech tagger (LINGSOFT)	Part-Of-Speech tagger	Part-Of-Speech tagger	None

Table 2.1: Table showing the pronominal anaphora resolution rates and some pertinent factors used in five implemented systems.

2.3 Works that also Resolve NP Anaphora

In the work published in Dahl (1986), Dahl describes the implementation of a system that resolves *one/some* anaphora as well anaphoric NPs due to normalization, which is later described in detail in Dahl et al. (1987). Dahl uses Sidner's (Sidner (1979)) Focussing Theory to process anaphoric noun phrases in a restricted domain related to technical discourses. A point of difference in Dahl's use of focussing is that her focussing algorithm is based on syntactic constituents rather than thematic roles as described by the originator of the theory. For instance, Sidner's *Agent* role is encoded as the *Subject* by Dahl and similarly, Sidner's *Verb Phrase* is represented by objects of prepositional phrases in Dahl's case. She justifies (Dahl, 1986, (page 19)) this use of surface structure using examples in which a sentence is represented using two different surface structures but constant thematic roles resulting in one-anaphora being interpreted incorrectly. The focussing algorithm works by arranging the focus elements of the current sentence in the order *Sentence*, *Direct Object*, *Subject* and *Objects of the Prepositional Phrases*, with the *Sentence* being the highest ranked. Updating of this focus list is based on the use of pronouns in a subsequent sentence, in which case the antecedent of the pronoun is moved up to the top of the list, else the focus list is untouched. This makes the focussing algorithm solely dependent on the use of pronouns. This is an implementation of the Centering Theory (Grosz and Sidner, 1986) which dictates that the focus of a clause be pronominalized in a subsequent clause. Hence "John went to town. He took his car" is preferred over "John drove to town. John took his car".

Since Dahl's system is geared to be used in a restricted domain, she uses a domain-specific knowledge base instead of a generic one. The knowledge base contains association relationships such as object properties and part-whole relations hence relations between nouns such as keyboards, motors disk drives can be established. The NP anaphora are resolved using a combination of the entities in the focus list and *current context*, which contains information from the knowledge base as well as the previous resolutions. Exactly how these are used is not detailed in the paper. Additionally, no statistics about the success rate is reported in the paper.

2.3 Works that also Resolve NP Anaphora

Palomar et al. (2001) is an example of work that integrates two different aspects of a discourse structure in their anaphora resolution algorithm. The paper describes an algorithm applicable to a dialogue oriented discourse in which they use a form of focus and segmentation. The focus element is referred to as *Discourse Topic* and the segmentation as *Accessibility Space*. The topic is determined using a weighted function of frequency which identifies NPs occurring with the highest frequency in the shortest distance as the Topic. Accessibility Segments are determined by using a function of *Adjacency Pair* which is a pair of sentences with two different speakers. They define the Accessibility Space for an anaphor as NPs taken from the sum of:

- the adjacency pair containing the anaphor
- the preceding adjacency pair
- any adjacency pair including the adjacency pair containing the anaphor
- the NP representing the main topic of the dialogue

The above segmentation policy was tested on 204 dialogues and the study reports that 95.9% of the antecedents were located in the defined Accessibility Space. They also report 81.3% precision rate for pronominal anaphora and an impressive 81.5% for NP anaphora. There are two limitations of this approach that prevents it from being applied to more general cases. Firstly, the authors used manual annotation of Adjacency Pairs, which needs to be somehow automated at the parsing level. Secondly, it is not clear, how the segmentation concept could be extended to a multi-party dialogue or to discourses which have direct speeches inserted in between author written sentences.

Apart from the discussed studies other studies such as Brennan et al. (1987), Mitkov (1993) Kasper et al. (1999) and Walker (1998) have attempted to integrate various forms of discourse structure in anaphora resolution algorithms with mixed success rates. However, the main challenge remains; how do we determine the discourse structure? The determination of most forms of discourse structure such as focus, coherence and rhetorical structure requires deeper semantic and inference mechanism which is

still an evolving discipline in computational linguistics. The framework proposed in this thesis makes use of semantic features which are more directly computable from surface structures hence are more practically usable compared to the frameworks described in this section which use discourse related semantics.

2.3.2 Poesio et al's Multi Strategy Approach

A significant number of works on NP anaphora resolution from a computational viewpoint were reported by Poesio et al in Poesio et al. (1997), Vieira and Poesio (2000), Poesio et al. (2002) and Poesio et al. (2004a). With the exception of Poesio et al. (2004a), the other studies are based on syntactic and pragmatic rules and use either WordNet and/or the web as a lexicon. I will review these works in reasonable detail for two reasons. Firstly our implementation is also a rule based system and uses the same lexical resource, WordNet. Secondly, similar to Poesio et al's, our system is based on shallow processing of the text, hence does not make use of deeper level inference mechanisms based on discourse theories such as focussing, coherence and rhetorical structure. Poesio et al. (2000), reports, some rudimentary experimentation with segmentation, however the results did not indicate any major improvement in the resolution rate.

The scope of these studies is resolution of NPs that are definite descriptions¹ and include both direct and bridging² anaphora. Direct anaphora includes cases where the head noun is the same, e.g. "house/red house". Indirect anaphora includes co-reference relations such as "truck/tanker" while Bridging anaphora includes the cases where the anaphor and the antecedent are related by a relation other than co-reference, e.g. "car/tyre". The authors follow on from Fraurud's (Fraurud, 1990) claim that definite NPs should not be processed based on the premise that they are primarily anaphoric. Fraurud's examination of definite description NPs revealed that a large proportion of

¹NPs preceded by the article *the*

²referred to as associative anaphora by Hawkins (1978) and inferrables by Prince (1992).

them were discourse new. Hence we need to incorporate rules to process discourse new entities rather than relying on a “failing search for an already established discourse referent” Fraurud (1990). In other words, an NP anaphora system should have separate rules or methods for recognizing discourse new NPs running in parallel with the rules for resolving anaphoric NPs, rather than as a fall back default when no antecedent has been identified. Poesio et al implement this simultaneous discourse new NP identification process in their resolution algorithm which is described in Figure 2.1. As an example of Fraurud’s definition of a discourse new entity, Poesio et al straight away declare NPs which are special predicates and appositions (at the top of the decision tree in Figure 2.1) to be discourse new without any attempt at resolving them. The alternative strategy that does not implement Fraurud’s notion of rules for discourse new entities, would be to attempt a resolution, using all available heuristics and declare it as discourse new as a default corresponding the *Fail* branch in Figure 2.1. We adopt the former strategy in our implementation, since it is error prone to declare an NP as discourse new, before even considering it against a candidate. I later use an example to illustrate the violation of this rule.

An aspect of discourse theory tested in Vieira and Poesio (2000) study, is the effect of various types of segmentation on anaphora resolution. Segmentation divides a discourse into various text units and the accessibility of any anaphor in the discourse is restricted to nouns in only predetermined segments. Such an idea is similar to the discourse concept of *focus spaces* as described in Grosz and Sidner (1986). There are various other forms of segmentation, for instance, those discussed in Reichman (1985) and Fox (1993). The effect of segmentation on anaphora was illustrated using excerpt (2.1) in Section 2.1.2.3. The excerpt contained two instances of “house” with anaphoric references to each of the instance. Resolution to the correct instance could only be explained using segmentation. In Vieira and Poesio (2000), the authors experimented with segmentation using candidates from a fixed size window, indexed candidates consisting of head of nouns from the whole discourse, as well as a combination of these two methods. However, the results of these were not overly skewed

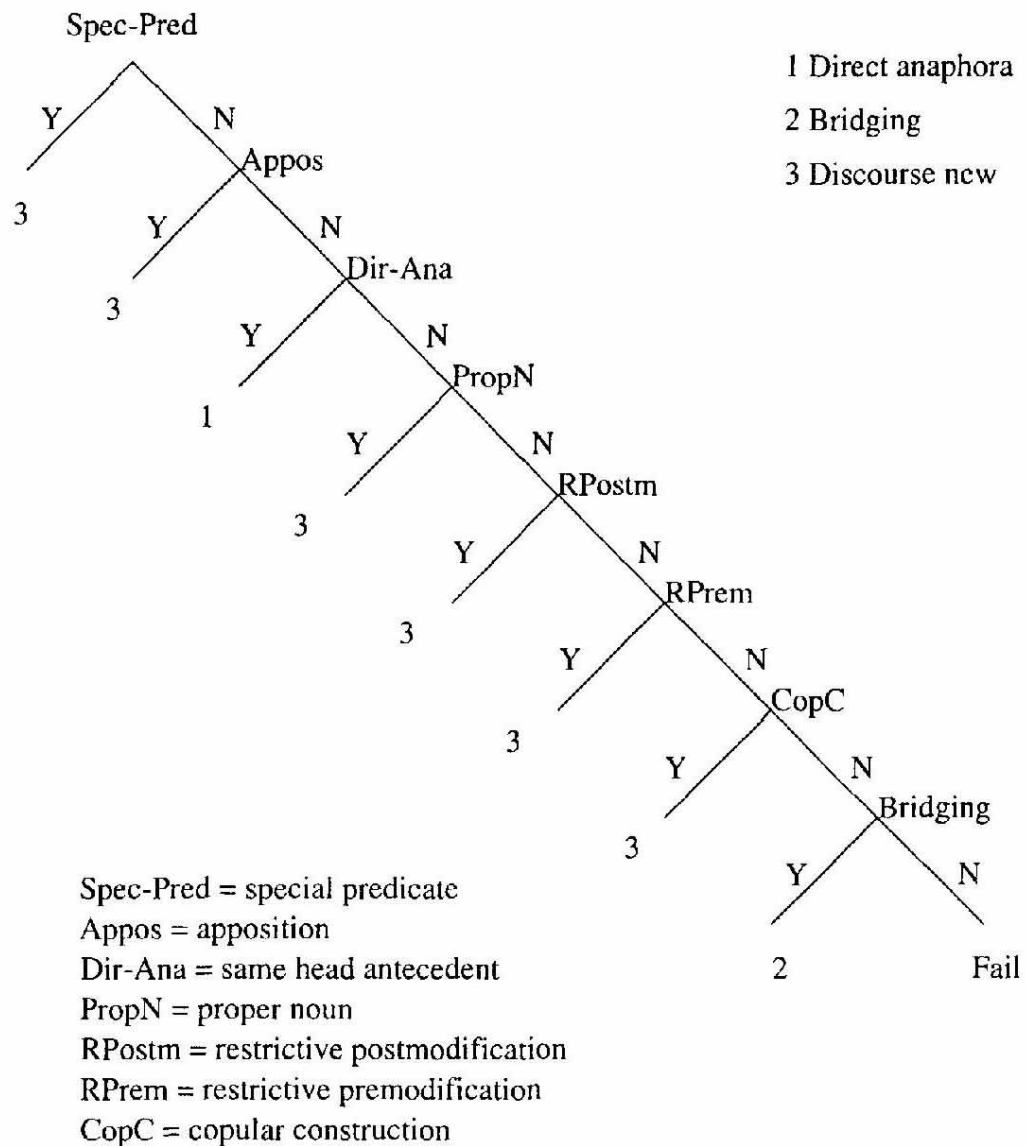


Figure 2.1: Decision Tree Describing Poesio et al's Algorithm - From Vieira and Poesio (2000, pg. 577)

towards any one segmentation scheme for it to be convincingly used or excluded in a system. As a comparison, our implementation is tailored for news oriented (not feature articles) newspaper articles. Most of the discourses in this genre are very quickly written with a lot of direct and indirect speeches from the participants making up the news. This makes newspaper articles more “erratic” and somewhat “less coherent” compared to a well thought out structured discourse with multiple proof-reads. Thus newspaper articles can be characteristically challenging in terms of analyzing it for discourse concepts such as focussing, coherence and segmentation structure. We thus opted not to use any form of segmentation. We use all of the previous sentences as the candidate window for NP anaphora and up to the previous sentence for pronominal anaphora, unless there are no candidates, in which case we progressively increase the window to the next previous sentence. We also considered a wider range of noun phrase anaphora compared to Poesio et al’s framework which only considered definite descriptions as anaphoric.

2.3.3 Soon et al’s Machine Learning Approach

This section discusses NP anaphora resolution systems that use some form of machine learning to resolve anaphora.

Soon et al. (2001) describes a corpus based machine learning approach which uses a decision tree learning algorithm, C5, described in Quinlan (1993). In this approach the “machine” is presented with an annotated corpus of anaphor-candidate pairs and a feature vector. The feature vector represents the various anaphora features used to determine the correct pairing based on values of the features in the feature vector. The feature values for correct pairs generates a decision tree which is then used by a clustering algorithm to create a partitioned cluster for each set of co-referent pairs as well as non-co-referent pairs. The content of the feature vector were derived from both linguistic as well as non-linguistic sources and contained core features such as number, gender, syntactical roles etc., which are also used in non-learning approaches.

2.3 Works that also Resolve NP Anaphora

The difference is that in non-learning approaches, the constraints associated with the features are applied as filters to completely rule out incompatible candidates. In the case of the learning approach, the constraints are implemented as features, hence are defeasible. This allows a learning system to base a classification, (or resolution in this case) collectively on a set of soft constraints represented by a cluster.

Soon et al. (2001) use a set of 12, mostly linguistic features summarized below: (i is the candidate and j is the anaphor):

- **Distance Feature** - This feature captures the sentence-distance between i and j .
- **i-Pronoun Feature** - If i is a pronoun, return true; else return false. Pronouns include reflexive pronouns (himself, herself), personal pronouns (he, him, you), and possessive pronouns (hers, her).
- **j-Pronoun Feature** - If j is a pronoun (as described above), then return true; else return false.
- **String Match Feature** - If the string of i matches the string of j , return true; else return false.
- **Definite Noun Phrase Feature** - If j is a definite noun phrase, return true; else return false.
- **Demonstrative Noun Phrase Feature** - If j is a demonstrative noun phrase, then return true; else return false.
- **Number Agreement Feature** - If i and j agree in number (i.e., they are both singular or both plural), the value is true; otherwise false.
- **Semantic Class Agreement Feature** - The defined semantic classes used were: female, male, person, organization, location, date, time, money, percent, and object. These semantic classes are arranged in a simple hierarchy. Agreement between i and j in the semantic class return true; else return false.

2.3 Works that also Resolve NP Anaphora

- **Gender Agreement Feature** - If the gender of either markable i or j is unknown, then the gender agreement feature value is unknown; else if i and j agree in gender, then the feature value is true; otherwise its value is false.
- **Both-Proper-Names Feature** - If i and j are both proper names, return true; else return false.
- **Alias Feature** - This feature accounts for aliases such as dates (e.g., *01-08/Jan. 8*), and cases such as *Mr. Simpson/Bent Simpson*.
- **Appositive Feature** - Accounts for cases such as *the chairman of Microsoft Corp* is an apposition to *Bill Gates* in the sentence *Bill Gates, the chairman of Microsoft Corp*.

The co-reference resolution algorithm works by generating a feature vector for each of the previous nouns (referred to as markables in the study) starting from the immediate previous one. The feature vector is passed to the classifier until it returns true (resolved) or there are no more markables in the document. An immediate shortfall of this approach is that the most recent candidate satisfying the classification criteria will be resolved to be the antecedent, while there might be a stronger candidate further away from the anaphor. This is amended by Ng and Cardie (2002) who use Soon et al's algorithm as basis to build a much richer set of feature vectors and use a score system to test all candidates in a window, and then choose the highest scoring candidate.

Soon et al process a compound noun by breaking it into a *head noun* and a *nested noun phrase*, where the nested noun phrase is the prenominal of the compound noun. They then use the union of the full noun phrase, name-entities and the nested noun phrases as candidates for co-reference resolution. In this approach there is an attempt to process the prenominal of a noun phrase by presenting the prenominal as a candidate for co-reference to an anaphor. This study has two similarities to this thesis; firstly it considers all nouns (as opposed to just the definite descriptions) as potentially anaphoric,

and secondly, it perceives the modifiers (referred to as prenominals) as discourse participants and makes them accessible as candidates. The authors of Soon et al. (2001) report recall values between 60 and 80 percent which, they claim is higher than any of the machine learning approaches¹ and comparable to some of the non-learning approaches, which are generally accepted to be better performing. Our framework also falls in the non-learning approach with a much richer knowledge integration compared to the single Semantic Class Agreement Feature used by Soon et al. The semantic feature used by Soon et al would also need a customized lexicon, classifying person, organization, location, time, etc., compared to our use of a generic lexicon such as WordNet.

2.3.4 Ng et al's Machine Learning Approach with Richer Features

Ng and Cardie (2002) take Soon et al's work a step further by expanding on the 12 features to an arguably deeper set of 53, although not all of them necessarily result in better resolution rates. However they report an improved performance compared to Soon et al's results when using a manually selected set of 22-26 features. An improvement from 62.6 to 70.4 is claimed on MUC-6 (Kaufmann, 1995) and MUC-7 (Kaufmann, 1998) data sets. Apart from the use of additional features Ng et al used a best-first clustering in which the antecedent chosen was the candidate with the highest co-reference likelihood value among preceding NPs with co-reference class values above a threshold of 0.5 (the values were between 0 and 1). This is in contrast to Soon et al, where the first candidate satisfying the clustering criteria was chosen as the antecedent. Some other examples of additional features include ancestor-dependent relations from WordNet, paragraph distance between the two anaphor-candidate pairs etc. Most of the features originate from Soon et al's set, except they are implemented in finer detail.

¹a detailed comparison is made in McCarthy and Lehnert (1995)

2.3.5 Poesio's Neural Network Approach

An NP resolution system using machine learning approach is reported in Poesio et al. (2004a). This study uses Multi-layer perceptrons (MLPs) with back-propagation. The study uses relatively fewer features (compared to Soon et al. (2001) and Ng and Cardie (2002)), however restrict the application to a subset of NP anaphora, namely mereological (part-of) relations . They use a set of features associated with local focus and lexical distance between anaphor-candidate pairs. The lexical distance is found from two sources, **google distance** and **WordNet distance**, the details of both are described in the paper. The other features used are based on the interaction of global and local focus (described in Grosz and Sidner, 1986). For example, **local first mention**(whether the entity has been realized¹ in the previous 5 sentences) and **global first mention** (whether it had been realized anywhere else). The learning system was trained using both positive and a randomly chosen set of negative instances with multi-layers of perceptrons including back-propagation. The recall rate reported from this study is between 84% and 86% (for a small dataset of 58 instances) with slight variations depending on the combination of features used. Use of perceptrons for anaphora resolution in this study looks promising, however a lot more work is needed in order to scope in the other categories of anaphora. The framework proposed in this thesis currently uses a much richer set of features and uses constraints or rules for resolution. It would be interesting to investigate the effectiveness of the features in a neural network system such as that described in Poesio et al. (2004a).

2.4 Chapter Summary

In the first part of the chapter I discussed the various types of anaphora and how they can be categorized into groups depending on the type of word used as the anaphor. We observed that anaphora resolution is based on common set of factors, however the

¹realize in Grosz et al's terminology is used as an anaphor

strategy with which these factors are applied varies between works. We also noted that some of the factors such as coherence and focus are deeper level discourse concepts, determination of which presents a challenge for computational systems. The latter part of the chapter outlines a sample of works on anaphora resolution, including what has been described as bridging anaphora. The works on pronominal anaphora indicated that relatively a high success rate is achievable using none or very limited integration of knowledge, however, for the leftover pronouns integration of knowledge becomes essential. The last section reviewed a range of works on NP anaphora including what has been referred to as bridging anaphora. The works on bridging anaphora illustrated that they have predominantly concentrated on *part-of* relations although most of them indicate that there are other, possibly infinite, types of bridging relations. The review of works also illustrate, that there are also a wide range of co-reference anaphora and hence a corresponding range of approaches to resolve them. The last part of the chapter also reviewed some works that have used machine learning approaches which have achieved success rates comparable to the traditional rule based approaches. With the easy availability of computational power and large corpora such as the web, we are sure to see a lot more work using machine learning approaches for anaphora resolution.

An overall conclusion drawn from all of the reviewed works is that the task of anaphora resolution is a complex process which is dependent on a large number of variables. Firstly, most of the approaches are based on preprocessing. The level of preprocessing can vary between different works and the way the preprocessing errors get handled might also be different. Secondly, the level of knowledge integration can also widely vary. Even the most trivial information such as the gender information can vary where some works specify it while others mine it from the text itself. Thirdly, the categories of anaphora being resolved also varies. Although reporting of both the precision and recall values somewhat mitigates this variance, it still becomes difficult to compare a system which resolves both pronominal and NP anaphora with one that only resolves pronominal anaphora. A related aspect to this is the level of resolution, that is, whether an anaphor is resolved locally or at a discourse level, and how this statistics is reported.

Lastly, the size and the genre of the corpus used for testing can also vary widely. Table 5.3 shows that the test data between works can vary from 15 articles to approximately 4000 sentences. In a situation like this it becomes difficult to compare the success rates to determine the robustness of the system with open texts.

The next chapter describes the related works on the generation of compound nouns and how this was utilized in the proposed framework for resolving anaphora.

Chapter 3

Interpretation of Anaphora – Our Approach

This chapter describes the theoretical framework which forms the basis of the anaphora resolution algorithm which will be described in Chapter 4. In Chapter 1 we proposed a novel theory on anaphora; that is, anaphora is also used as a shortcut elaborative device, in addition to its known function as a co-reference pointer. This elaborative theory forms a core part of the new anaphora resolution framework which uses knowledge integration to resolve what has been conceived as “difficult” anaphora.

In order to develop this new theory I firstly outline some preliminaries required to support the theory. This is done in Section 3.1. Then in Section 3.2 I use real examples from the input corpus to further demonstrate the additional elaborative function of anaphora which was introduced in Chapter 1. Using these examples I also show how part of the elaborative information can also be used to identify the antecedent entity. The discussion of examples in 3.2 also include the elaborative role of anaphora represented by putting two or more words together. This phenomenon is called compound

3.1 Some Preliminaries on our Approach

noun generation. This is discussed next in Section 3.3, which presents an overview of the linguistic theories that describe the construction of the meaning of an NP by putting two words together in forming a compound noun. After discussion of a range of theories on compound nouns, I examine the work reported in Levi (1973) in greater detail and explain why the set of relations from this work was used for our purpose of anaphora resolution. Note that the relations described in these works describe the *production* of compound nouns, which is a process done by the producer of a discourse. The process of anaphora resolution is the reverse, done by the consumer of a discourse. Hence, in order to use the relations from compound noun generation for resolving anaphora, they had to be formulated into reverse-constraints so that they could be integrated into the anaphora resolution algorithm. The formulation of these constraints is discussed in Sections 3.4 and 3.5. Finally, Section 3.6 presents a summary of this chapter.

3.1 Some Preliminaries on our Approach

As discussed in the introduction, we consider anaphora resolution to be a purely co-referencing problem. This co-referencing problem has traditionally been referred to as a reference pointer from a source NP known as the anaphor to the destination NP, known as the antecedent. The antecedent NP is the most recent one and if that NP is also an anaphor, then this has been referred to as chain reference. Our approach to co-reference is slightly different. We don't consider an **NP** to be the antecedent, rather, the **entity** referred by it is the antecedent. Hence, the task of resolving anaphora involves identifying the real world entity referred by the various NPs in a discourse. The real world entity can be referred to by an already used NP or a different one. The different NP used can be a pronoun, an alternative noun or a compound noun. When a different NP is used, it communicates additional information about the entity pertaining to the

3.1 Some Preliminaries on our Approach

meaning of the NP used, enriching the consumer's knowledge about the entity. This perception of anaphora is illustrated in Figure 3.1.

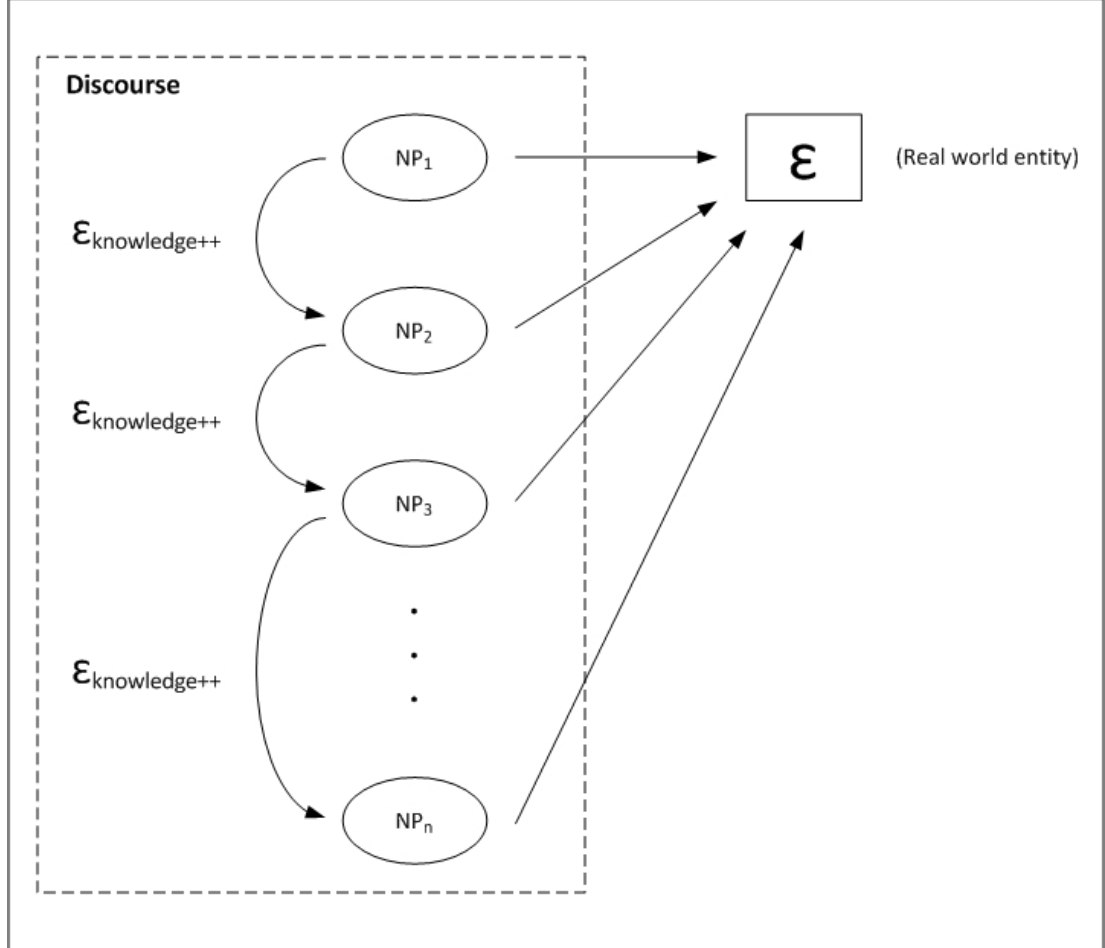


Figure 3.1: Diagram Illustrating the Use of Anaphoric NPs in a Discourse - The diagram firstly shows that the antecedent is the referred entity rather than an NP in the discourse. It also illustrates how a subsequent use of an NP works in accumulating knowledge about the referred entity.

The figure shows how multiple NPs are used to refer to the same entity, however, note that none of the NPs is the antecedent. The antecedent is ϵ , which is an entity existing in the real world which has been referred using various NPs. With this perception, the process of anaphora resolution can be stated as:

Finding the minimum set of entities in a discourse that are referred to

by the NPs therein.

Hence the difference between the minimum set of entities and the number of NPs in a discourse is the number of repeat references to entities already referred to by an NP. In Chapter 4, we will see that this perception of anaphora forms a crucial part of the proposed anaphora resolution framework by making use of knowledge from *any* of the NPs referring to the antecedent entity. The illustration in Figure 3.1 shows that the knowledge about the referred entity is cumulatively accumulated with each new mention of the entity with an NP. In processing a discourse, when a new NP is encountered, its meaning is interpreted in relation to the accumulated knowledge for all of the previously encountered entities. In doing so, one is able to resolve the NP to either one of the previous entities, or determine that it refers to a new entity. Thus the goal for an anaphora resolution system is to use the knowledge embedded in the meaning of an NP to determine if it could be referring to the same entity as referred to by one of the previous NPs, or is it referring to a new entity. Hence, the framework proposed in this study considers all NPs as potentially anaphoric and compares the meaning of each newly encountered NP with the accumulated knowledge of all of the previously encountered entities to determine if it is co-referential. If the knowledge available does not suggest that it might be referring to an existing entity, it is declared as referring to a new entity.

3.2 Anaphora – An Elaboration of the Antecedent

The production of a discourse is dependent on many factors such as the language competency of the producer, discourse goal, audience and the genre of discourse. These factors account for the variety of discourses, and hence a variety of anaphora, that can be produced describing the same event. As already discussed, the choice of the noun used to refer to an entity depends on the semantic aspects that is most relevant in the view of the producer. The producer chooses an NP that best emphasizes the aspect of

3.2 Anaphora – An Elaboration of the Antecedent

the semantics that is most consistent with the intention of the discourse. Consider the following excerpt to illustrate this.

The [owner]_{i1} of a 4WD vehicle seized in connection with the Birgit Brauer murder case says it was stolen by [his]_{i2} employee six weeks ago.

The Toyota Hilux is registered to [Palmerston North man Brent Cleverley]_{i3}, who confirmed to the Herald yesterday that [his]_{i4} vehicle had been taken by a [man]_{j1} who worked for [him]_{i5}.

The Herald understands the [worker]_{j2}, who used the vehicle in [his]_{j3} job cutting firewood, failed to turn up to work one day last month.

The excerpt shows the first three paragraphs from a discourse. The referents to two entities are highlighted and labeled with subscripts *i* and *j*. In the first paragraph, the noun [owner]_{i1} was chosen instead of the NP [Palmerston North man Brent Cleverley]_{i3}, used in the second paragraph. The choice of the noun “owner” brings into picture the semantics related to the “owner of the vehicle” in question, rather than the specific person. This was deemed to be consistent with the intention of the first paragraph by the producer. In the first paragraph, other details like the name, gender, and where the owner lived was considered to be of secondary importance, hence included in the second paragraph using [Palmerston North man Brent Cleverley]_{i3}. Note that, the anaphor [Palmerston North man Brent Cleverley]_{i3} contains elaborative information about the antecedent. In particular, it says that the antecedent is a male, lives in Palmerston North and the name is Brent Cleverley. All this information has been embedded in a shortcut way in the NP, in addition to it functioning as an anaphor.

The second entity represented initially by the NP [man]_{j1} is co-referred with the noun [worker]_{j2} in the last paragraph. The anaphor [worker]_{j2} uses the semantics associated with the noun, “worker”, to achieve the intention of the third paragraph. Note that in this case, the pronoun “he” could also have been used as in sentence, “The Herald understands that he used the vehicle in . . .”. If the pronoun was used instead, it would

3.2 Anaphora – An Elaboration of the Antecedent

still be resolvable to “man” instead of “him” in the previous clause, since “man” is the subject hence the focus as stipulated by the Centering theory (Gordon et al., 1993). However, the author chose to use the noun [worker]_{j2} since it brings to the forefront additional semantic information such as he gets paid, he probably had associations with the employer etc.. When the meaning of the noun “worker” is interpreted in the context in which it is used, one is also able to resolve it. That is, the previous use of the verb “worked” enables one to identify the agent of the verb as the same entity as the noun “worker”. This makes [worker]_{j2} resolvable to [man]_{j1}, which in turn further elaborates that the gender of [worker]_{j2} is male. This information helps in resolving the subsequent pronoun [his]_{j3}. The example illustrates the motivation for an aspect of our algorithm which is very different to the previous approaches, that is, the use of multiple passes instead of “resolve as you encounter” approach. In this approach, relevant knowledge is accumulated from other, easier resolutions which is used to help resolve more difficult anaphora in a subsequent pass. Details of how this was implemented is presented in Chapter 4.

When choosing an NP to use in a co-reference capacity a producer is obligated to also ensure that there is enough identificatory information attached to the anaphor for the consumer to be able to identify the antecedent. The identificatory information can often be contained in the use of additional words as modifiers which are in turn connected via synonyms, hypernyms and hyponyms. Consider the following excerpt to illustrate this.

*The **Welsh rugby fan** who killed a young Waikato woman when his camper-van smashed into her oncoming car has been ordered to pay almost \$10,000 in fines and reparation.*

***James Berry**, 23, had previously pleaded guilty to careless driving causing injury and careless driving causing the death of 18-year-old Cambridge woman **Liz Neels**.*

3.2 Anaphora – An Elaboration of the Antecedent

*This morning in the Hamilton District Court **he** stood looking scared and vulnerable as **he** was handed down a \$500 fine, plus costs, for the injury charge.*

*Judge Anne McAloon said she took into account that **Berry** had been held in custody for three days after the June 23 crash and the volunteer work **he** had been doing since his initial court appearance.*

*The **accounting graduate** from Swansea had come to New Zealand to follow the Lions tour but had not seen any of the games.*

The highlighted NPs in the excerpt all represent the same entity. The first NP, “Welsh rugby fan” is more general than the second mention, “James Berry”, followed by the use of three pronouns and a use of the surname “Berry”. All these anaphora are relatively easy to resolve and have already been done using various techniques in previous works. However, use of the NP “accounting graduate” in the last paragraph poses a challenge. It cannot be resolved by recency based systems which would have “Judge Anne MacAloon” and “Berry” from the previous paragraph as two equally competing candidates. The NP “accounting graduate” is gender independent and even if we assume that the gender for the candidates “Judge Anne McAloon” and “Berry” are known, we cannot use gender compatibility to resolve the anaphor “accounting graduate”. Firstly, note that the phrase “accounting graduate” is an elaboration of “Berry”. The NP was chosen over the alternative, the pronoun “he”, in order to emphasize that “Berry” was also an accounting graduate. There is no identificatory information in the NP itself, however the sentence in which the anaphor is used contains the required identificatory knowledge for it to be resolved to “Welsh rugby fan”. Specifically, the following pieces of information is available from the sentence:

- accounting graduate was from Swansea (Swansea \approx Welsh)
- accounting graduate had come to New Zealand (New Zealand \approx Waikato)
- accounting graduate follow Lions tour (Lions \approx rugby)

3.2 Anaphora – An Elaboration of the Antecedent

- accounting graduate had not seen any of the games(games \approx rugby)

Although all of the pieces of information is relevant for resolving the anaphor, in this study we focus on the category of semantics illustrated by the first example. The knowledge that the “accounting graduate” was from “Swansea”, the relation “from” enables us to form the compound noun “Swansea accounting graduate”. If we assume that “Swansea” is an alternative name for “Welsh” and if we have the additional reinforcement knowledge that “fan” and “graduate” are both animate/human than we can resolve “accounting graduate” to “Welsh rugby fan” in the first paragraph. The illustration highlights the motivation for two features of our algorithm. Firstly, it highlights one of the types of knowledge encoding used for resolving anaphora which was integrated in our algorithm. In particular this was the ‘FROM’ relation between “accounting graduate” and “Swansea”. Secondly, it highlights the motivation for using a search space extending to the beginning of the document. Even though the NPs “he”, “Berry” and “James Berry” are more immediate antecedents for the anaphor “accounting graduate”, the meaning embedded in the anaphor could only be used to resolve to the far antecedent “Welsh rugby fan” which appears in the first sentence. In other words, the identificatory information contained in an anaphor can “match” with *any* of the previous antecedent NPs, not necessarily the most recent one. A human consumer, incrementally accumulates all the information about the entity and resolves an antecedent based one or more pieces of information about the entity. It is not necessarily based on the information contained in the NP used in the most recent mention of the entity. This is reflected in our algorithm which accumulates the relevant information in a data structure referred to as Reference Unit or RU, described in Section 4.2.1.

The previous two examples illustrated how anaphoric compound nouns function to also elaborate on the antecedent apart from identifying the antecedent entity. This elaborative function can also be accomplished by use of even the simplest of the nouns, the pronouns. Consider the case in (3.1) to illustrate this.

(3.1) Jowenzo went roaring.

3.2 Anaphora – An Elaboration of the Antecedent

- (a) She was fast.
- (b) He was fast.
- (c) It was fast.

Assume that the name “Jowenzo” in (3.1) is an unfamiliar name which could be the name of a person or a non-animate entity such as a car. The statements in (3.1a), (3.1b) and (3.1c) are elaborations of (3.1). The pronouns “He”, “She” and “It” co-refer to the entity “Jowenzo”. Note that the referents by themselves also contain additional information about the entity “Jowenzo”. For example, the statement in (3.1a) asserts that “jowenzo” is female, animate, singular as well as fast. The first three assertions are implicit, embedded in the meaning of the pronoun “She” and the last assertion is explicit, expressed as the clause of the anaphor. Even the use of the simplest pronoun “It” in (3.1c) asserts that “Jowenzo” is non-animate and singular and if interpreted in the context of the whole clause, it probably represents some type of vehicle since it is fast.

Now consider the statement “The girl was fast” after (3.1). In this case the anaphor “The girl” achieves the same purpose as the pronoun “she”, however use of the NP “The girl” also implies that “Jowenzo” is a girl, as opposed to a woman, which is not the same as the use of the pronoun “she”. Going even further, the statement “The Auckland girl was fast” has the anaphor “The Auckland girl”, which contains even more semantic information. The full predicate “the girl lives in Auckland” or “the girl plays for Auckland” is implied in the anaphor, however it has been deleted and expressed as a shorthand by use of only the nouns in the predicate. This has been described as *predicate deletion* by Levi (1978) and Lees (1966) in the context of compound noun generation. An anaphor can also exist as an alternative derived noun, such as “Jowenzo” referred to with “runner” if “Jowenzo” has been asserted to be running in a previous clause. In this case the alternative noun is used for expressing the action that the antecedent was involved in. This is referred to as *normalization* (Levi, 1978). Both, predicate deletion and normalization are examined in detail later in this chapter as they are formulated as constraints used in the anaphora resolution algorithm.

In summary, use of anaphora is an implicit way of communicating the intended information in the most concise way possible. NPs are used as part of subsequent clauses which provide further information on the antecedent entity. Apart from the anaphor-clause itself, the choice of the NP used as the anaphor may also contain implicit additional information on the antecedent. The choice of the anaphoric NP is determined by what needs to be elaborated about the antecedent and the dictionary meaning of the word. For example, use of pronouns is constrained by the number, animacy and gender values while use of alternative nouns is constrained by the sortal organization of the nouns in the ontology in use. When the anaphoric NP is composed of more than one word, the semantics of the NP as a whole is governed by the dictionary meanings of the individual words as well as the relation between them. The next section details the theories on the formation of compound nouns in general. This is followed by Section 3.4, which discusses how the theory was used in our anaphora resolution algorithm.

3.3 Generation of Compound Nouns

In the following discussion we will restrict the composition of a compound noun to consist of two nouns in the form *noun* + *noun* where the first noun is referred to as the modifier and the second one as the head noun. Note that this is a simplified version as we could also have multiple modifiers, which can be made up of determiners, adjectives and even verbs. When a compound noun is used as an anaphor, the head noun is the noun that acts as the anaphor, however the total meaning of the compound noun may contain the knowledge required for identifying the antecedent. One form of knowledge contained in compound nouns is the existence of a relation between the modifier and the head noun. For example, in the compound noun “battle fatigue” there is a causal relation between the nouns “battle” and “fatigue”. However, the compound noun “olive oil” does not have the same relation between the nouns “olive” and “oil”. Studies on compound nouns have tried to understand how compound nouns such as

3.3 Generation of Compound Nouns

the above get generated and how a consumer is able to interpret it. That is, when we see the compound noun “battle fatigue”, we are able to recreate the predicate “battle causes fatigue” while in the case of “olive oil” we recreate that “oil comes from olive”. On the other hand we are not able to generate a compound noun on the basis of any predicate. For example, “war man” can not be formed on the basis of the relation “man who hates war” or similarly “house tree” can not be formed from “tree between two houses” (Zimmer, 1971). In both these examples there does exist a relation between the nouns, however it is not one that can be used to generate a compound noun. Studies on compound nouns is decades old and there are plenty of works (e.g. Chomsky, 1969; Downing, 1977; Lees, 1966; Levi, 1973; Li, 1971; Warren, 1978; Zimmer, 1971) that have proposed sets of relations that can give rise to compound nouns. Although the actual set of relations between studies vary slightly, there is a common core set defined in Levi (1973) which has been used in this study. These particular set of relations are later specified in this section, however from here on we will use the term c-relation for any relation that can be used to generate a compound noun.

The studies on compound nouns further suggest that their use is highly *productive* rather than *lexical*. Some compound nouns, for example, “police station” and “cable car”, have become part of the lexicon because of the common existence of the composite nouns in the same c-relation. However, a large number of compound nouns also get generated on the fly as a discourse is being produced rather than recalled and used from a lexicon. In generating a compound noun, the producer looks for the existence of a c-relation among the nouns within the discourse, or even from outside the discourse Levi (1973). When the compound noun is generated using nouns from within the discourse, the nouns in the c-relation have to exist prior to the use of the compound noun. Hence, for the purpose of anaphora resolution, the task boils down to identifying the modifier and the head noun in a c-relation in the discourse. A compound noun can also be generated on the basis of only one of the nouns from the c-relation existing in the discourse, and assuming the other one to be part of the consumer’s knowledge base. For example, “battle fatigue” can be used without the mention of “battle”, and the con-

3.3 Generation of Compound Nouns

sumer can be expected to draw the causal relation between “battle” from outside the discourse and “fatigue” mentioned in the discourse. In terms of a computational implementation, a lexicon encoding c-relations is required to bridge the gap to the second noun. Although there are no specific lexicons encoding this, they can be mined, with some degree of success, from generic lexicons such as WordNet which was used in this study. Note that compound nouns don’t only get generated on the basis of permanent c-relations. They can also be generated on the basis of “temporary or fortuitous” relations from the immediate context. In this case the c-relation is expressed between nouns, which otherwise would not be related. A compound noun gets generated on the basis of the expressed c-relation which is only valid in the context in which it is mentioned. This aspect was highlighted in Downing (1977) in which the author did experiments with human subjects to study the generation of novel NPs to describe situations given as text as well as pictures. For example subjects were asked to create noun compounds from drawings, such as in Figure 3.2, of the entities with no conventionalized name. She found that even when the subjects were hard-pressed to generate a compound name, they typically based their names on “generic or habitual” relations, referred to as c-relations in this discussion. For example, *barrel lamp*, *barrel phone* but not *barrel sofa* or *barrel cushion*. Even though *barrel lamp* and *barrel phone* are not “universally interpretable”, the subjects tended to generate them based on the ON relation which is the basis of NPs such as conventionalized names “table lamp” and “sofa cushion”. Overall Downing found that 90% of the relations were based on generic relations between entities (c-relations) and 10% on fortuitous relations. These results prove two aspects of compound noun generation; that they are highly based on a closed set of relations (c-relations), and secondly they can be generated from objects only in a temporary c-relation in the immediate context. Hence, for the purpose of anaphora resolution, one has to identify both permanent as well as temporary c-relations in the discourse and present them as candidates for anaphora which are compound nouns. This means extraction of the relations from discourse surface structures. Extraction of c-relations for the purpose of interpreting compound nouns is an already progressed area in computational linguistics which can also be utilized for our purpose of anaphora

resolution. I next examine a selection of the computational frameworks on compound nouns in order to explain the reason for choosing the particular framework used in this study.



Figure 3.2: Context Drawing from Downing (1977)

The set of c-relations used by computational linguists (e.g. Barker and Szpakowicz, 1998; Butnariu and Veale, 2008; Girju et al., 2005; Kim and Baldwin, 2005; Kim and Nakov, 2011; Nakov, 2008; Nastase and Szpakowicz, 2003) vary along similar lines as those proposed by the linguists. Most of these studies have used the relation set originating from either Levi (1978) or Warren (1978). Some of them have arrived at a slightly different variation while others have defined a finer grained set of relations dictated by the data sets used for the study. For example, Tratz and Hovy (2010) reports a set of 43 relations grouped into 10 upper level categories. Table 3.3 replicated from Tratz and Hovy (2010) outlines the set of relations posited in the study as well as a mapping of the relations proposed in a selection of other studies.

From Table 3.3 it can be seen that most of the relations from different studies can be mapped to an equivalent relation in the other studies. For this study we chose the coarse set of relations proposed in Levi (1978) for two reasons. Firstly, the objective for extraction of these relations is not for a fine grained interpretation of a compound noun. We only make an indirect use of these relations to help resolve the compound noun used as an anaphor. Hence, we only need to ascertain if ANY of these relations exist, and not specifically, WHICH of the relations exist. Thus for the purpose at hand,

3.3 Generation of Compound Nouns

Category Name	% Example	Approximate Mappings
Causal Group		
COMMUNICATOR OF COMMUNICATION	0.77 court order	\supset BGN:Agent, \supset L:Act _a +Product _a , \supset V:Subj
PERFORMER OF ACT/ACTIVITY	2.07 police abuse	\supset BGN:Agent, \supset L:Act _a +Product _a , \supset V:Subj
CREATOR/PROVIDER/CAUSE OF	2.55 ad revenue	\subset BGV:Cause(d-by), \subset L:Cause ₂ , \subset N:Effect
Purpose/Activity Group		
PERFORM/ENGAGE_IN	13.24 cooking pot	\supset BGV:Purpose, \supset L:For, \approx N:Purpose, \supset W:Activity \cup Purpose
CREATE/PROVIDE/SELL	8.94 nicotine patch	∞ BV:Purpose, \subset BG:Result, ∞ G:Make-Product, \subset GNV:Cause(s), ∞ L:Cause ₁ \cup Make ₁ \cup For, \subset N:Product, \supset W:Activity \cup Purpose
OBTAIN/ACCESS/SEEK	1.50 shrimp boat	\supset BGNV:Purpose, \supset L:For, \supset W:Activity \cup Purpose
MODIFY/PROCESS/CHANGE	1.50 eye surgery	\supset BGNV:Purpose, \supset L:For, \supset W:Activity \cup Purpose
MITIGATE/OPOSE/DESTROY	2.34 flak jacket	\supset BGV:Purpose, \supset L:For, \approx N:Detraction, \supset W:Activity \cup Purpose
ORGANIZE/SUPERVISE/AUTHORITY	4.82 ethics board	\supset BGNV:Purpose/Topic, \supset L:For/About _a , \supset W:Activity
PROPEL	0.16 water gun	\supset BGNV:Purpose, \supset L:For, \supset W:Activity \cup Purpose
PROTECT/CONSERVE	0.25 screen saver	\supset BGNV:Purpose, \supset L:For, \supset W:Activity \cup Purpose
TRANSPORT/TRANSFER/TRADE	1.92 freight train	\supset BGNV:Purpose, \supset L:For, \supset W:Activity \cup Purpose
TRAVERSE/VISIT	0.11 tree traversal	\supset BGNV:Purpose, \supset L:For, \supset W:Activity \cup Purpose
Ownership, Experience, Employment, and Use		
POSSESSOR + OWNED/POSSESSED	2.11 family estate	\supset BGNVW:Possess*, \supset L:Have ₂
EXPERIENCER + COGNITION/MENTAL	0.45 voter concern	\supset BNVW:Possess*, \approx G:Experiencer, \supset L:Have ₂
EMPLOYER + EMPLOYEE/VOLUNTEER	2.72 team doctor	\supset BGNVW:Possess*, \supset L:For/Have ₂ , \supset BGN:Beneficiary
CONSUMER + CONSUMED	0.09 cat food	\supset BGNVW:Purpose, \supset L:For, \supset BGN:Beneficiary
USER/RECIPIENT + USED/RECEIVED	1.02 voter guide	\supset BNVW:Purpose, \supset G:Recipient, \supset L:For, \supset BGN:Beneficiary
OWNED/POSSESSED + POSSESSION	1.20 store owner	\approx G:Possession, \supset L:Have ₁ , \approx W:Belonging-Possessor
EXPERIENCE + EXPERIENCER	0.27 fire victim	\approx G:Experiencer, ∞ L:Have ₁
THING CONSUMED + CONSUMER	0.41 fruit fly	\supset W:Obj-SingleBeing
THING/MEANS USED + USER	1.96 faith healer	\approx BNV:Instrument, \approx G:Means \cup Instrument, \approx L:Use, \subset W:MotivePower-Obj
Temporal Group		
TIME [SPAN] + X	2.35 night work	\approx BNV:Time(At), \supset G:Temporal, \approx L:Inc, \approx W:Time-Obj
X + TIME [SPAN]	0.50 birth date	\supset G:Temporal, \approx W:Obj-Time
Location and Whole+Part/Member of		
LOCATION/GEOGRAPHIC SCOPE OF X	4.99 hillside home	\approx BGV:Locat(ion/ive), \approx L:In _a \cup From _b , B:Source, \approx N:Location(At/From), \approx W:Place-Obj \cup PlaceOfOrigin
WHOLE + PART/MEMBER OF	1.75 robot arm	\supset B:Possess*, \approx G:Part-Whole, \supset L:Have ₂ , \approx N:Part, \approx V:Whole-Part, \approx W:Obj-Part \cup Group-Member
Composition and Containment Group		
SUBSTANCE/MATERIAL/INGREDIENT + WHOLE	2.42 plastic bag	\subset BNVW:Material*, ∞ GN:Source, ∞ L:From _a , \approx L:Have ₁ , ∞ L:Make _{2b} , ∞ N:Content
PART/MEMBER + COLLECTION/CONFIG/SERIES	1.78 truck convoy	\approx L:Make _{2ac} , \approx N:Whole, \approx V:Part-Whole, \approx W:Parts-Whole
X + SPATIAL CONTAINER/LOCATION/BOUNDS	1.39 shoe box	\supset B:Content \cup Located, \supset L:For, \supset L:Have ₁ , \approx N:Location, \approx W:Obj-Place
Topic Group		
TOPIC OF COMMUNICATION/IMAGERY/INFO	8.37 travel story	\supset BGNV:Topic, \supset L:About _{ab} , \supset W:SubjectMatter, \subset G:Depiction
TOPIC OF PLAN/DEAL/ARRANGEMENT/RULES	4.11 loan terms	\supset BGNV:Topic, \supset L:About _a , \supset W:SubjectMatter
TOPIC OF OBSERVATION/STUDY/EVALUATION	1.71 job survey	\supset BGNV:Topic, \supset L:About _a , \supset W:SubjectMatter
TOPIC OF COGNITION/EMOTION	0.58 jazz fan	\supset BGNV:Topic, \supset L:About _a , \supset W:SubjectMatter
TOPIC OF EXPERT	0.57 policy wonk	\supset BGNV:Topic, \supset L:About _a , \supset W:SubjectMatter
TOPIC OF SITUATION	1.64 oil glut	\supset BGNV:Topic, \approx L:About _e
TOPIC OF EVENT/PROCESS	1.09 lava flow	\supset G:Theme, \supset V:Subj
Attribute Group		
TOPIC/THING + ATTRIB	4.13 street name	\supset BNV:Possess*, \approx G:Property, \supset L:Have ₂ , \approx W:Obj-Quality
TOPIC/THING + ATTRIB VALUE CHARAC OF	0.31 earth tone	
Attributive and Coreferential		
COREFERENTIAL	4.51 fighter plane	\approx BV:Equative, \supset G:Type \cup IS-A, \approx L:BE _{bcd} , \approx N:Type \cup Equality, \approx W:Copula
PARTIAL ATTRIBUTE TRANSFER	0.69 skeleton crew	\approx W:Resemblance, \supset G:Type
MEASURE + WHOLE	4.37 hour meeting	\approx G:Measure, \subset N:TimeThrough \cup Measure, \approx W:Size-Whole
Other		
HIGHLY LEXICALIZED / FIXED PAIR	0.65 pig iron	
OTHER	1.67 contact lens	

Figure 3.3: Table showing Approximate Relation Mappings of Tratz and Hovy (2010) to Other Relations Sets. - replicated from Tratz and Hovy (2010), representing semantic relations, frequency in their data set, examples, approximate relation mappings to previous relations sets. \approx -approximately equivalent; \supset/\subset -super/sub set; ∞ -some overlap; \cup -union; initials BGLNVW refer respectively to the works of (Barker and Szpakowicz, 1998; Girju, 2007; Girju et al., 2005; Levi, 1978; Nastase and Szpakowicz, 2003; Vanderwende, 1994; Warren, 1978)

extraction of the upper level coarse relations is sufficient. Secondly, there are already several works that computationally extract these relations from existing lexicons such as WordNet and the Web. In fact several very recent studies (e.g. Abad et al., 2010; Hendrickx et al., 2010; Kim and Nakov, 2011; Tratz and Hovy, 2010) are dedicated towards automatic derivation of such relations, most of them based on the relation set from Levi (1978). In future, this will result in improved techniques for extraction and availability of ontology describing the semantic relations between nouns, and with an anaphora resolution framework such as that proposed in this thesis, these will also be useful for the purpose of resolving anaphora. I will now describe the set of nine c-relations from Levi (1978) which were used for anaphora resolution. These were originally proposed as seven relations in Levi (1973).

Levi (1978) proposed that compound NPs are derived from the underlying clause or complement structures by the two processes of *predicate deletion* and *normalization*. Her work is based on similar framework as Lees (1966) except Less' transformational process is based on verb classifications. Levi proposed that, in the case of normalization, 'the underlying predicate survives overtly in the head noun, with the modifier derived from either the subject or the object of the underlying clause. This gives rise to subjective (e.g. industrial production) or objective (e.g. heart massage) normalization. In the case of predicate deletion, Levi proposed that the number of deletable predicates is limited to only nine primitive relations. They are CAUSE, HAVE, MAKE, USE, BE, IN, FROM, ABOUT and FOR. With these relations she accounts for the range of example NPs shown in Table 3.1.

In the process of predicate deletion the subject and the object combine to form the NP if the underlying clause can be represented by any one of the relations in Table 3.1. Levi further specifies that the first three (CAUSE, HAVE AND MAKE) to be bidirectional while the rest of them are unidirectional. For bidirectional relations it is possible for NPs to be formed in either in *subject object*, or *object subject* order. For example the clause "battle causes fatigue" can generate the compound noun "battle fatigue" in the

3.3 Generation of Compound Nouns

Predicate	N + N	ADJ + N
CAUSE	battle fatigue	viral infection
	disease germ	malarial mosquito
HAVE	snake poison	reptilian scales
	apple cake	musical comedy
MAKE	cellblock	floral wreath
	silk worm	sebaceous glands
USE	steam engine	electric drill
BE	girlfriend	professorial friends
ON	fieldmouse	polar bear
	hay truck	sugary cake
FOR	bird sanctuary	avian sanctuary
FROM	olive oil	
ABOUT	travel story	oily glut

Table 3.1: Table showing the list of nine deletable predicates from Levi (1978)

subject object order. On the other hand, the clause “industry produces milk” can generate “milk industry” which is in the *object subject* order. In the case of unidirectional relations (USE, BE, ON, FOR, FROM, ABOUT) she maintains that a compound noun can only be formed in the *subject object* order. This gives a total of 12 way ambiguity for any compound noun (1 each for 9 relations and 3 more for the bidirectional ones). That is, given any compound noun the modifier and the head noun could be related by any one of the 9 relations, but in addition the order of the nouns could have been reversed for the three bidirectional relations. Translating this to resolving compound noun anaphora, each anaphor will need to be searched for the existence of any of the nine relations between the previously mentioned nouns and in addition the ones with the bidirectional relations will need to be searched for the reverse order as well. Additionally, the modifier noun can be converted into morphological adjectives where the adjective exists in the language lexicon, e.g. virus converted to viral in Table 3.1. For these cases the stem noun corresponding to the adjective needs to be identified before searching for the relations among the previously mentioned nouns.

Levi’s theory on predicate deletion and normalization describe two ways in which new noun terms get generated from relations between existing entities. We formulate the

two processes of normalization and predicate deletion as constraints on a discourse producer and used them in reverse for resolving anaphora used in the discourse. Section 3.4 describes the predicate deletion constraints and Section 3.5 describes the normalization constraints.

3.4 Predicate Deletion Constraints

Consider the excerpt below from one of the discourses used in this study to illustrate how the compound noun generation relations were used for resolving anaphora:

The southbound truck₁ collided with a big truck₂ laden with diesel.
:
The diesel truck₃ exploded in an inferno.

In the example above “truck₃” co-refers to “truck₂” and not truck₁. This co-reference cannot be established by any form of modifier or head noun matching, nor any of the higher level discourse structure techniques, such as focussing, since there is lexical distance of six sentences between the first and second sentence. However, the anaphor “diesel truck₃” can be described by Levi’s process of predicate deletion where the predicate “diesel is on the truck” can have the predicate deleted, producing the shorthand form “diesel truck”. This is described by the ON relation from Table 3.1, which is expressed as the surface structure “laden with”. Note that this relation supersedes the lexicalized relation “truck USE diesel” hence the compound noun “diesel truck”. In order to use the constraint on compound noun generation for NP anaphora resolution, we can search for surface structures that instantiate ANY of the relations in 3.1 and map the NPs to the extracted relations. For the specific purpose of anaphora resolution we do not need to determine which specific relation relates the two nouns, just that the

two nouns are related by one of the relations in the set. Extraction of the relations to describe noun compounds has already been discussed as an active research area.

The set of relations used in our implementation were directly taken from Levi (1978), shown in Table 3.1, with the exception of the BE relation. Levi’s BE relation is equivalent to *is-a* relation which itself is a co-reference. In the examples, “John is a teacher” and “car is a vehicle”, John/teacher and car/vehicle are anaphoric since one can be used to co-reference the other. Hence the BE relation was itself treated as anaphoric and not as a constraint like the rest of the relations in Table 3.1. However, Levi’s BE relation also includes the group-membership relations, for example, as in example, “southbound truck *is-a* truck”. When a plural form such as “both trucks” is mentioned followed by use of the NP “southbound truck”, there is a group-membership relation which needs to be recognized in order to correctly resolve “southbound truck”. This knowledge is crucial in determining if multiple instances of an entity exists so that subsequent singular mentions of the entity can be resolved to the appropriate instance. In order to distinguish between the group-membership relation from BE relations which are co-referential, we will use the term INSTANCE relation (short form BE-INST) to represent the former, while the latter was treated as anaphoric.

With this discussion we can now define and exemplify relation types that were used for the purpose of anaphora resolution. They are:

CAUSE - Includes all causal relations. For example, *battle/fatigue*, *earthquake/debris*

HAVE - Includes notions of possession. This includes diverse examples such as *snake/poison*, *house/window* and *cake/apple*.

MAKE - Includes examples such as *concrete house*, *tar/road* and *lead/pencil*.

USE - Some examples are *drill/electricity* and *steam/ship*.

IN - This relation captures grouping of things that share physical or temporal properties. For example *lamp/table* and *Auckland/New Zealand*.

FOR - This includes purpose of one entity for another. For example *pen/writing* and *soccer ball/play*.

FROM - This includes cases where one entity is derived from another. For example *olive/oil* and *wheat/flour*.

ABOUT - Describes cases where one entity is a topic of the other. For example *travel/story* and *loan/terms*.

INSTANCE - Describes a singular form of an entity that is a member of a plural form. For example *Northbound truck/both trucks* and *John/drivers*.

Apart from searching for the exact pair of nouns in a compound noun relation, we also need to do an expanded search for either of the nouns to be represented by a synonymous noun. For example the noun “diesel” can also be represented by the noun “fuel” and the noun “truck” by the nouns “rig” and “tanker”. This enables a producer to use any permutation of the above nouns in either the predicate expressing the relation, e.g. “big truck laden with fuel”, or the anaphoric NP, e.g. “fuel tanker”. Hence to be effective we integrated this expanded search and we used WordNet as the lexicon for the source of the synonymous knowledge.

The techniques and the set of surface structures used to extract the relations is described in Section 4.3. I next describe the other process, normalization, which can also be used as a source of constraints for anaphoric compound nouns.

3.5 Normalization Constraints

The process of normalization is the opposite of predicate deletion discussed in the previous section. In the normalization process an NP is generated by retaining predicate, instead of deleting it as in the case of predicate deletion. This is done by using a morpheme of the functioning¹ verb or adverb as either the head-noun or a modifier of the resulting noun term. Some examples of normalization for each of the categories are :

¹used as the main verb of the clause

verb to head-noun: **drive** normalized to NP **driver**.
adverb to head-noun: **accidentally** drove a car into ... normalized to **accident**.
adverb to modifier: **accidentally** drove a car into ... normalized to **accidental** act.
verb to modifier: **communicate** normalized to **communicative** act ...

In the first and the second cases the verb and adverb are normalized to the head noun, however the first one co-refers to the agent argument of the verb while the second one co-refers to the clause as a whole. In the last two cases, normalization results in a modifier rather than the head noun. In both these cases the resulting compound noun co-refers to the source clause as a whole. However, the co-reference is not constrained by the modifier. For example the NP “accidental person” is also an adverb to modifier normalization, however it now co-refers to the agent instead of the clause. For the purpose of anaphora resolution, we want to identify if there are any constraints on the process of normalization, which can be used in reverse for identifying the antecedent of a normalized NP.

There are only limited works which attempt to resolve anaphora from normalized NPs, two examples found were (Vieira and Poesio, 2000) and (Dahl et al., 1987). These works implement simple normalization cases in which the normalized NP taken to co-refer to the agent of the verb, as in the example “John drives a car”, the normalized NP “driver” or “car driver” refers to “John”. However, the agent of the verb is not the only candidate from the source clause. We want to know if it is possible for normalized NPs to also co-refer to other components of the source clause, for example the object and the clause. We will examine this starting from a theoretical perspective.

Levi (1978) proposed two forms of normalization; one which describes the previous example of “car driver”, where a morpheme of the subject survives as the modifier. This is referred to as **subjective infliction**. The second form is when the object survives as the modifier, referred to as the **objective infliction**. For example, the compound noun “heart massage” is an objective infliction from the clause “John massaged the

heart”. We will examine both forms of normalization to determine if it has any effect on co-reference.

NPs generated by both subjective and objective inflections can co-refer to different components of the source clause. Sometimes the antecedent may not be exactly definitive. Consider the phrase “John massaged the heart”. Does the normalized NP “heart massage” co-refer to the *action* “massage”, the *object* “heart”, the *predicate* “massage the heart” or the whole *clause*? We also want to examine if the antecedent for a normalized NP is influenced by the *type of inflection*, that is, either objective or subjective. In addition, we want to determine if the *type of morpheme* acts as a constraint which determines its antecedent. The type of morpheme (*morpheme type*) is determined by the suffix of the inflected noun.

Let us first examine if the inflection type has any influence on the co-reference target of the normalized NP. Consider the case in example (3.2).

(3.2) *The industry produces useful bacteria.*

- (a) The bacterial production ...
- (b) The industrial production ...

The morpheme type of the verb *produce* is same (ion-form) in both (3.2a) and (3.2b), however (3.2a) is an objective inflection, while (3.2b) is a subjective inflection. Firstly it is apparent that the two NPs are not referring to either the subject or the object of the clause but to some variation of the event represented by the source clause. The two forms do not have exactly the same semantics hence have to be referring to two different components of the clause determined solely by the modifier, objective in the first case and subjective in the second. In the case of (3.2a), the anaphor is referring to the “act of producing bacteria”. For the case in (3.2b), the modifier is the agent (industry) engaging in the action (produce) which is limited to only producing bacteria. Although at a fine level, the semantics associated with objective and subjective modifiers are slightly different, both forms co-refer to slight variations of the whole event

associated with the clause. We will ignore the subtle variations in the semantics and will classify both (3.2a) and (3.2b) as a co-reference to the clause as a whole. Hence we will rule out the use of inflection type as a constraint to determine the co-reference target. This effectively confirms that the co-reference target is entirely determined by the head noun of a compound noun. However, in the case of normalized NPs, the inflection type can be used to assist in the resolution process by searching for clauses which have the subject or the object which is equivalent to the modifier noun. If the modifier noun exists as either the subject or the object of a previously mentioned clause, then the morpheme type can be used to identify the actual antecedent within the clause. The morpheme types and their co-reference properties are discussed next.

The morphemes of a verb are the different inflected forms that can be derived from the root verb. For example the verb “produce” has the morphemes “producing”, “producer”, “producible” and “product”, and each of these have different semantics. The range of morphemes for a particular verb are determined by their suffixes allowed by the language lexicon. Not all suffixes can be used with all verbs. The range of suffixes that can be used with a particular verb is determined by the range of the allowed morphemes. The allowable morphemes originate from the semantics of the verb as well as the basic sounds like vowels and consonants that makes up the verb (Fabb, 1988). The range of suffixes are especially dependent on the ending sound of the verb so that associated suffix sound can be easily appended to the verb. The morphemes of the verb “produce”, for example, has the morpheme types “**producer**”, “**product**”, “**producible**” and “**producing**” while the verb “govern” has the forms “**governor**”, “**government**”, “**governable**” and “**governing**”. For the verb “produce” the morpheme “**product**” represents the object argument while “**producer**” represents the agent argument. However, this is not consistent across all verbs, that is, the morphemes ending with the suffix “er” can not be used to represent the agent for all the verbs. For example, in the case of “govern” the agent argument is communicated by the morpheme “**governor**” whereas the verb “produce” has the morpheme “**producer**” serving the same purpose. In yet another case, the verb “abortion” has the morpheme “**abortionist**” for the agent argu-

ment. The discussion of the examples illustrate that using the process of normalization, a producer has access to a wide range of verb morphemes that can be used to co-refer to different components of the source clause. The component of the clause co-referred by the normalized NP is dependent on the suffix of the morpheme. Hence, to be able to correctly resolve an anaphor generated from normalization, we need to extract the co-reference target from the morpheme type, which is determined by the suffix of the NP. As a subsidiary part of this study we studied the anaphoric uses of normalized NPs in the input corpora and examined their co-reference targets as a function of the suffix types. The range of suffix types and their corresponding targets were then implemented in the general anaphora resolution algorithm as constraints. The range of suffix types used are specified in the next chapter which describes the implementation details all of the constraints discussed in this chapter as well the other traditional constraints used in our anaphora resolution algorithm.

3.6 Chapter Summary

In this chapter I presented a new perspective on anaphora; as a shortcut means of elaboration of the antecedent rather than a purely co-referential device as has been viewed by previous works. The amount of elaboration can vary. When the same NP as the antecedent is used, the amount of elaboration is zero. Use of any other NP adds a finite amount of additional information to the discourse. The amount of additional information is dependent on the dictionary meaning of the chosen word for the anaphor. I also argued that in the process of interpreting the meaning of the anaphor word, one is also able to identify it. When the anaphor consists of multiple words, the complete meaning is also contributed by the relation between the composite words. These relations are described by the theories behind the formation of general multi-word phrases which was discussed in Section 3.3. I then used these theories on compound noun

generation to formulate a framework for resolving anaphora. The framework is designed as a complete solution to resolving a very wide range of anaphora¹ rather than focussing on any particular category as done by most of the works reviewed in chapter 2. The proposed framework uses most of the existing features used in previous works such as morphological compatibility, syntactic binding constraints and parallelism as discussed in section 2.1. However, our framework enhances the previous frameworks by integrating the constraints originating from predicate deletion and normalization as discussed in sections 3.4 and 3.5. These constraints, based on relational knowledge structure, are used to help correctly resolve all types of anaphora, particularly anaphora represented by compound nouns.

The chapter as a whole presented anaphora in a new light. This perspective puts more emphasis on its semantic aspects which gives us a better insight into the reasons for the wide variety of anaphora found in natural discourses. In addition, the accompanying anaphora resolution framework illustrates how the semantic aspects of anaphora can be used to identify the antecedent. The next chapter presents the details of how the framework was implemented as a general purpose anaphora resolution system.

¹Exclusions are *it*, *one/some*, *this/that* and *who/which* categories

Chapter 4

Anaphora Resolution Implementation

This chapter describes the implementation details of the anaphora resolution system which is based on the new framework described in Chapter 3. Before being able to use the framework, we need to perform several auxiliary tasks in order to transform the data existing as free text into a form which can be used by the resolution algorithm using the framework. Our system achieves this by implementing these tasks as separate support modules which are used by the main resolution algorithm. I describe these modules first in Section 4.1. After laying this groundwork I next describe the main anaphora resolution algorithm in Section 4.2. One of the novel aspects of our framework is that it uses embedded relational knowledge between nouns to help resolve anaphora. The techniques used to mine these relations are described in Section 4.3. This is followed by Sections 4.5 and 4.6 which describe the full set of rules that were used by the resolution algorithm. The rules are separated into “hard constraints” in Section 4.5 and preferences in Section 4.6. Finally, Section 4.7 gives a summary of this chapter.

4.1 Descriptions of Major Modules

This section gives the descriptions of the main modules that support the higher level anaphora resolution algorithm. These are divided into the following subsections:

- **Section 4.1.1** - describes input text and the pre-processing module.
- **Section 4.1.2** - describes all other minor preprocessing adjustments to the parsed text before being presented to the anaphora resolution algorithm.
- **Section 4.1.3** - describes how a complex NP which can consist of multiple determiners, names, adjectives and nouns was de-constructed in order to facilitate the extraction of the knowledge required for anaphora resolution.
- **Section 4.1.4** - describes how anaphora represented by synonymous NPs were handled in our implementation.
- **Section 4.1.5** - describes how our implementation handles some other forms of phrases which can be interpreted as an NP in the common $N + N$ form.
- **Section 4.1.6** - describes how our system treats word senses.

4.1.1 Parsing

The input to the anaphora resolution module is a shallow parsed clausal form of sentences parsed by a LISP based parser. The output from the parser is a text file containing a structured clausal representation of the sentences from a discourse. A sample output from the parser is shown in Figure 4.1 for a single sentence from one of the discourses analyzed. Each of the sentences as well as the clauses within the sentence are labeled with an ID number. The output from the parser is meant for general purpose NLP tasks, hence all the components of the parsed structure were not necessarily used for anaphora resolution. The parser parses the sentences into basic syntactical roles as

defined by the rules of the language. In addition, it also identifies the attached prepositions and their target nouns which are used to derive the relations to indirect objects by the anaphora resolution module.

An analysis contained in the parsed form that goes deeper than surface level is the creation of the “LINKS” slot which links a clause, or an object in it to another clause or object based on the semantics of cue phrases such as “IF”. An example of this is illustrated in the first clause of the sample in Figure 4.1. Although this is a useful semantic information that can potentially be used for anaphora resolution, it was out of scope for this study, however, is a planned future extension. The other fields from the LISP style parsed structure from Figure 4.1 is first processed by a java based document model builder which is then used by all the modules in the anaphor resolver for accessing the document.

4.1 Descriptions of Major Modules

```
<sentence>
<sent-number>1</sent-number>
<sent-string>If Ahmed Zaoui's family succeed in their application to join him in
New Zealand as refugees, the decision will annoy NZ First leader Winston Peters
but delight a boy who has not seen his father for nearly three years.</sent-string>
<sent-cr>(#S(CR :ID (0)
:SID 0
:TYPE (MAIN-CLAUSE)
:VERB (SUCCEED-TO-JOIN*
(:IN*
(APPLICATION* (:NOUN) (:MODIFIER (THEIR*))))))
:SUBJ (FAMILY* (:NOUN) (:MODIFIER (|AHMED-ZAOUI'S*|))))
:OBJ (HIM* (:PNOUN)
(:IN*
(NEW-ZEALAND* (:UNKNOWN) (:NAME)
(:AS* (REFUGEES* (:NOUN))))))
:IOBJ NIL
:TIME NIL
:LINKS ((:LINK (IF*) (1) (:F))))
#S(CR :ID (1)
:SID 0
:TYPE (MAIN-CLAUSE)
:VERB (WILL-ANNOY*)
:SUBJ (DECISION* (:NOUN) (:MODIFIER (THE*)))
:OBJ (WINSTON-PETERS*
(:X-WORDS (*NZ*-FIRST*) (LEADER*)
(WINSTON-PETERS*))
(:NAME))
:IOBJ NIL
:TIME NIL
:LINKS ((:LINK (BUT*) (2) (:F))
(:LINK (IF*) (0) (:B))))
#S(CR :ID (2)
:SID 0
:TYPE (MAIN-CLAUSE)
:VERB (DELIGHT*)
:SUBJ (DECISION* (:NOUN) (:MODIFIER (THE*)) (:FILLER))
:OBJ (BOY* (:NOUN) (:MODIFIER (A*))
(:WHO* (:LINK (:WHO*) (3) (:F))))
:IOBJ NIL
:TIME NIL
:LINKS ((:LINK (BUT*) (1) (:B))))
#S(CR :ID (3)
:SID 0
:TYPE (SUB-CLAUSE)
:VERB (HAS-NOT-SEEN*)
:SUBJ (BOY* (:NOUN) (:MODIFIER (A*)) (:FILLER))
:OBJ (FATHER* (:NOUN) (:MODIFIER (HIS*))
(:FOR*
(YEARS* (:NOUN)
(:MODIFIER (NEARLY*) (THREE*))))))
:IOBJ NIL
:TIME NIL
:LINKS ((:OBJ (:WHO*) (2) (:B))))</sent-cr>
</sentence>
```

Figure 4.1: Lisp Style Sample Clausal Output from the Parser.

4.1.2 Other Pre-Processing Adjustments

Apart from parsing of the sentences into clausal forms there were other preprocessing adjustments done to the input data in order for the data structures to be populated for critical components such as the agent for a clause. Some of these were done by the parser while others were done manually in the text file before processing by the anaphora resolution module. These were:

- **Ellipse substitution** – This process involved substitution of ellipses (clauses in which agents are omitted) in sentences in order to ensure that all clauses contain both an Agent and Object/I-Object. This substitution was done at the parser level and was differentiated by the label “:FILLER”. The sample output in Figure 4.1 shows the two clauses corresponding to the phrase “the decision will annoy NZ First leader Winston Peters but delight a boy who ...”, represented in clauses 1 and 2. The second clause has a filler for the noun “decision” for the Agent, which is the Agent entity of the previous clause. Such occurrences were not counted as an anaphoric hence were not included in the statistics.
- **Relative pronoun substitution** – Relative pronouns predominantly refer to the previous NP hence is trivial to resolve. They were substituted by the antecedent NP at the preprocessor level rather than passing them on the main anaphora resolution module. Hence, all occurrences of the relative pronouns (*which, who, whom, what, whose, where and that*) were substituted by the immediately previous NP. For instance, in the phrase “... were gutted by the inferno which erupted seconds after ...”, the interrogative pronoun *which* in the second clause was substituted by the most recent entity, “inferno”, in the example. Similarly, a phrase such as “The man who was wearing a hat shot the shopkeeper” was parsed into the two clauses, “The man shot the shopkeeper” and “who was wearing the hat”. The pronoun *who* in the second clause was substituted with “The man” in the second clause which by default was set to be the most recent entity before the

relative pronoun. Some manual substitutions were done for the incorrect cases in order to minimize the transmission of preprocessing errors.

- **Speaker substitution** – One of the characteristics of newspaper articles is that there is an abundance of quoted speeches, either in direct or reported form. Frequently, the speaker of a series of sentences is only stated once at the beginning while the subsequent sentences are assumed to be interpretable by the consumer to be either directly or indirectly reported from the same speaker. In our object based representation all utterance are assigned a producer¹, hence subsequent reported sentences without an explicit producer were assigned the most recent producer of a reported speech. For example in the excerpt below, the producer entities for the second and third sentences were substituted by “Hirschfeld”.

Hirschfeld said separation had a strong effect on son Youssef, who was 4 1/2 years old when he last saw his father.

”Ahmed phones regularly but I think the satellite link was so powerful because they hadn’t actually seen each other in the flesh for nearly three years.

”Seeing how much Youssef had grown was something Ahmed really responded to.”

4.1.3 Handling of Multi-Word Noun Phrases

A complex NP can be constructed by juxtaposing two nouns in the form $N + N$. However, we can construct relatively complex formations by combining determiners and using various morphemes of nouns such as names and adjectives, for example, “the senior Hamilton fire station officer”. In an example such as this we need to be able to extract “senior” to be the determiner, “Hamilton” to be the name and the rest as common nouns. Note that in the NP “Senior Hamilton fire officer”, the determiner

¹For unreported speech the producer is the writer

4.1 Descriptions of Major Modules

Determiner	Examples
Articles	a,an, the.
Demonstratives	this, that, these, those.
Possessives	my, its, their, his, her.
Titles	Mr, Ms, Dr, Miss etc.
Quantifiers	three, several, many, few, couple, etc.
Qualifiers	senior, medium, top, young etc.
Complements	other, another, second etc.

Table 4.1: List of determiner types with examples

“Senior” would be interpreted as a name rather than a determiner because of the use of capital “S”. In order to facilitate this we used a template of the form described in Figure 4.2 to populate a corresponding java object.

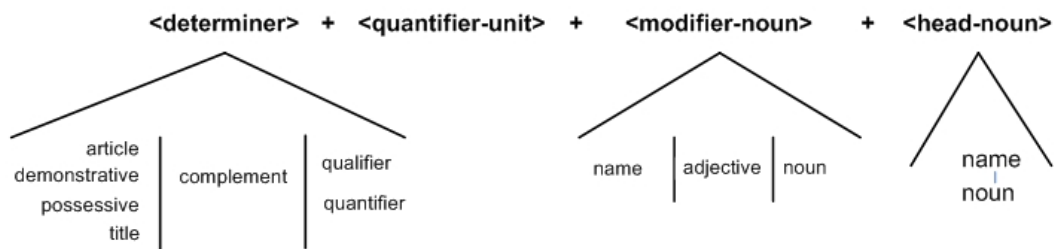


Figure 4.2: The Syntactical Order of Components for a Generic NP - there can be more than one occurrence of the modifier-noun component.

A majority of generic NPs¹ found in our data can be represented using a structure shown in Figure 4.2. The determiners consist of articles, demonstratives, possessives, quantifiers, qualifiers and complements. Table 4.1 shows the set of determiners with a corresponding list of examples, exhaustive in the case of articles, demonstratives, and possessives and a representative sample in the case of the others.

The structure in Figure 4.2 outlines the order of the components of an NP. The vertical lines indicate the categories of words that can be used for each of the components. For instance, a determiner can consist of the three components, one from each of the groups

¹an example of exception is “both the articles”.

4.1 Descriptions of Major Modules

separated by the vertical lines. The modifier-noun component can consist of names, adjectives and nouns which can reoccur multiple times, however, the order *name*, *adjective* and then *noun* seems to be maintained. There are no stringent grammatical rules specifying this order, however the examples used by English language teaching sites always seem to maintain this order, hence our use of the template. The articles used for this research did not show any non-conforming NPs. An example enforcing the order is the NP “Fonterra driver Ted Collins” which was processed as:

name (Fonterra), **noun** (driver) **name** (Ted Collins)

In a more complex NP such as “Senior Hamilton fire station officer Daryl Trim” the components would be:

name (Senior Hamilton), **noun** (fire), **noun** (station), **noun** (officer), **name** (Daryl Trim)

The right most noun or name was taken to be the head noun and two or more names occurring consecutively were taken to be a single name consisting of multiple words. Names consisting of multiple words were also stored as composite names and the composites were also made available as candidates since, for example, the name “James Smith” could be co-referred with “Mr Smith”.

An aspect of compound nouns that has been simplified in compound noun generation theories is its interpretation when there are more than two composite nouns. That is, in a compound noun of the form N1 N2 N3, does the compound noun generation relation for N1 extend to N2, N3 or both. In the example “Hamilton fire station officer Daryl Tim” the modifier “Hamilton” applies to “fire station” while in example “Auckland rugby player” the modifier “Auckland” applies to “player”. Compound nouns with multiple composite nouns can have both possibilities, that is, a modifier can apply to the head noun as well as other nouns which are part of the modifier. There is still no theory that completely describes this (Girju et al., 2005). For the purpose of anaphora resolution we were interested in only finding the relations between the head noun and a modifier. Hence, in processing a compound noun with multiple modifiers, we first

apply the template from Figure 4.2 to extract any determiners and quantifier units as these were interpreted differently compared to the rest of the compound noun. The leftover NP will now contain only nouns and adjectives. Next, any adjectives were reduced into their equivalent stem nouns, e.g. the noun “industry” for “industrial”. This reduces the leftover NP to a series of nouns with the right most noun acting as the head noun. We then attempt to find c-relations (compound noun generation relations discussed in Chapter 4j) between the head noun and each one of the modifier nouns among the previously mentioned nouns. Any relation found is then used by the anaphora resolution module.

4.1.4 Handling of Synonymous Nouns

The synonymous terms were found from Wordnet using two ways. The first one was using the hypernym, hyponym and synonym pointers to traverse one level in order to identify similar nouns such as matching “automobile” and “car”. We did not go any higher than single level because it starts becoming computationally expensive and hauls in too many unrelated synsets making the process of deciding on the antecedent difficult. To mitigate this we used a tool, also based on WordNet, that enabled us to consider candidates from higher than a single level in a more constrained way. The tool, described in Pedersen and Patwardhan (2004), can be used to compute the relative similarity of two synsets from WordNet. This is a freely available software package which uses WordNet to compute a similarity index between any two given synsets, such as, an automobile is more like a boat than tree. This is due to the fact that an automobile and boat share vehicle as an ancestor in the WordNet hierarchy. The measure of similarity uses information found in the *is-a* hierarchy of synsets and uses path lengths between a pair of synsets as the index of measure. The similarity index does not cross part of speech boundaries hence it compares nouns with nouns (e.g. boat and automobile) and verbs with verbs (e.g. walk and run), but not across them. A second measure called “Relatedness” is more general which goes beyond part of speech boundaries,

however this tended to deteriorate the performance and thus was not used. The similarity measure was effectively used to find relations between dissimilar nouns where a different noun was used for co-reference, such as “vehicle” and “car”. Some examples from input data that were successfully determined were fire/inferno, rig/truck/tanker and car/vehicle. Hence, for two NPs to be co-referential, they either had to have the same spelling, be a direct hypernym, hyponym, synonym or a similarity index of more than 0.5. To compute this we used the most common sense of the anaphoric noun (listed first in WordNet) and computed the index for all the senses of the second NP. If one of the senses scores more than 0.5, then we assumed the two NPs to be coreferential.

4.1.5 Handling of Other Forms of Noun Phrases

- **Complex Nouns** - these are nouns written as a single word, however they are made up of two simple nouns, e.g. “firefighter” made up of “fire” and “fighter” and “girlfriend” made up of “girl” and “friend”. Some of the complex nouns can have the composite nouns being referred to individually by an anaphor. For instance, the compound noun “girlfriend” can be co-referenced by the NP referents “the friend” or “the girl”. In order to resolve such anaphora we need to de-construct complex noun phrases and avail the composite nouns as candidates. A rigorous option for implementation would be to search the dictionary for all complex nouns and propose the composite nouns as candidates whenever the complex noun is proposed as a candidate. Due to time constraints we chose a simpler approach of searching for the anaphor string at the beginning and end of all candidate NPs. This enabled us to resolve anaphoric uses such as use of “girl” and “friend” in “girlfriend”. However this implementation will not enable one to resolve cases when alternative nouns such as “she” and “the woman” are used as the anaphor. Our data had only two such cases, hence our implementation currently does not contain a more robust implementation on complex nouns.

- **Hyphenated NPs** - Compounding of nouns is also done by hyphenating the constituent nouns, usually to avoid confusion with the scope of application of a modifier. For example the phrase “more important reasons” written as an un-compounded NP can have the qualifier *more* applicable to either the head-noun *reasons* or the modifier *important*. In the former case the interpretation would be “more reasons which are important”. In the latter case it would be “reasons which are more important”. This ambiguity is usually avoided by hyphenating nouns in order to constrain the scope. Hence *more-important reasons* and *more important-reasons* would constrain the hyphen connected nouns to be interpreted together removing the ambiguity. In terms of anaphora resolution this has critical consequences since the determiners are used to deduce the number of instances of an entity. In the examples “21-year-old girl” and “35-millimeter film” the use of the hyphens implies that there is only one instance of the head-noun instead of 21 and 35, which would lead to erroneous processing of the number of instances of the head noun. For non-numeric hyphen connected nouns, this means finding relations between the connected nouns instead of relation to the head noun. Hence our implementation treats a hyphen connected noun in the form N1-N2, as the meaning of N1 applying to only N2, where N1 can be a number a determiner or a common noun.

- **Appositive-constructs** These are forms where two NPs are either juxtaposed with or without a comma separation. Some examples are :

- Ted Collins, the president of RSA ...
- Ted Collins president of RSA ...
- Alexander the great ...

All of the appositive constructions have consistent semantics which can be reduced to equivalent NPs. For example the NPs above can be translated to “RSA president Ted Collins” and “great Alexander”. This is how we implemented the appositive constructions such as above in our system.

of-construction Type	Examples
Quantify	two of three/two of my friends.
Qualify	big amount of smoke/large number of cars.
function	owner of car/driver of car/sound of car/place of refuge.
Characteristic	sighting of car/description of car.
Part-of	edge of road.
Structure	plume of smoke/a bucket of water.
Relative	north of Auckland/brother of John.
Predicative	spoke of separation/knew of him.

Table 4.2: List of of-construction types with examples

- **Of-constructs** Table 4.2 shows the range of of-constructs found in the corpus, which have been categorized into groups based on their associated semantics.

Of-constructs have the form N_1 of N_2 where either N_1 , N_2 , or both could be anaphoric. Some of the of-constructs can be expressed as equivalent NP by translating the of-construction in the form NP_1 of NP_2 to NP_2 NP_1 with NP_2 representing the modifier and NP_1 representing the head-noun. Hence the of-construction, “owner of car” can be translated to the NP “car owner”. In order to resolve for the antecedent “car” we can search for a c-relation between “car” and “owner” as a source of knowledge for resolution. However, some of the of-construction categories such as Quantify and Quantify can not be translated into equivalent NP forms. Since it is not computationally trivial to determine the specific category of the of-constructs, we treated all the of-constructs as if they can be translated, and assumed that whichever ones were not syntactically correct would not exist in a c-relation in the discourse.

4.1.6 Treatment of Word Senses

An aspect of anaphora interpretation, pertinent to NP anaphora, is the disambiguation of the sense of the anaphoric NP. Any natural language typically has words which are polysemous and it is possible to have two different senses used in the same discourse.

Although modifiers are often used to distinguish between the entities to which an NP may refer to, a single word, without any modifiers, can sometimes be used to refer to different sense of a noun. The antecedent of such polysemous words is often derived from the context of usage. For example, the word “car” used multiple times in a discourse may refer to one or more instances of “automobile”, as well as, instances of “gondola” which are two very different senses of “car”. Disambiguation of word senses using various techniques is an active research area and studies such as Fragos et al. (2003) and Canas et al. (2003) have come up with complex learning algorithms that use WordNet as well as other corpora in order to disambiguate word senses with various levels of success. Since the primary focus of this study was anaphora resolution we implemented a simplified version based on these studies that relies only on WordNet. Our purpose for sense disambiguation is to determine if more than one sense is used, and not, to determine WHICH sense is used. For example the noun “car” has the following 5 senses in WordNet:

1. **car**, *auto, automobile, machine, motorcar* (a motor vehicle with four wheels; usually propelled by an internal combustion engine) ”he needs a car to get to work”
2. **car**, *railcar, railway car, railroad car* (a wheeled vehicle adapted to the rails of railroad) ”three cars had jumped the rails”
3. **car**, *gondola* (the compartment that is suspended from an airship and that carries personnel and the cargo and the power plant)
4. **car**, *elevator car* (where passengers ride up and down) ”the car was on the top floor”
5. **car**, *cable car* (a conveyance for passengers or freight on a cable railway) ”they took a cable car to the top of the mountain”

Multiple mentions of the noun *car* in a discourse can represent one of four possibilities:

- *Multiple occurrence* - represents the same entity participating in multiple events, hence mentioned multiple times.
- *Multiple instances* - represents multiple instances of the same entity. e.g. a red car and a white car.
- *Multiple senses* - where the different mentions represent different senses of the word. e.g. motorcar and a cable car.

For an accurate interpretation, an NP resolution system has to be able to disambiguate between the above three possibilities. Some NPs can be interpreted based on the use of modifiers while others require the use of encyclopedic knowledge and inferences.

For the purpose of word sense disambiguation we used WordNet to firstly determine if a mentioned word or head noun of a NP is polysemous or not. If WordNet does not state more than one sense for a noun then we assume that it is not polysemous, hence the task is reduced by one dimension to dealing with either multiple instances or occurrences. For resolution between instances and occurrences we use modifiers and determiners discussed in detail in Section 4.2.

NPs with polysemous head nouns add a third dimension in that we have to resolve between the different senses in addition to the instances and occurrences. For the purpose of anaphora interpretation, we do not need to disambiguate an NP to a particular sense, what we need to do is to **differentiate** between them, so that we can resolve them to the correct sense. To differentiate between the senses we used WordNet to firstly determine if more than one sense of the word is used by comparing the modifiers of the NP in the discourse to the modifiers of the different senses of the head noun in WordNet. However, there still remains the difficulty of disambiguation of a noun without a modifier or when it is used with a modifier that can not be identified as one of the senses. For instance, consider the excerpt below:

“We took the cable car to the top of the mountain. After spending 2 hours we took the car back.”

In this case, does the second mention of the noun “car” belong to the “cable car” sense or the “automotive sense”?

The senses of polysemous words are often organized so that there is a single dominant sense and one or more subordinate senses (Fragos et al., 2003). In the absence of “external forces” in the form of modifiers and/or contextual information, the dominant sense is the default sense that we apply when we encounter a word in a discourse. When we encounter a word such as car, house, ball, truck, we first apply the dominant sense, and then look for presence of “external forces” in case a change in interpretation is required. The subordinate sense is only maintained for as long as the “external forces” are present and reverts back to the dominant sense as soon as the “external force” disappears. In the cable car example, the usage of the noun “car” in the second sentence is somewhat ambiguous between the subordinate *cable car sense* and *automotive car sense*. The “external forces” in the form of the modifier “cable” required to maintain the subordinate sense is not present in the second mention, hence there is a shift towards the dominant sense. According to WordNet there are 3 other senses for car but we do not move towards any of these senses since they are subordinates.

We implement this observation of dominant and subordinate sense with the aid of WordNet. WordNet is organized so that the dominant sense is mentioned first and we use this as a benchmark to determine the default sense. We use a simplified technique of determining the sense of a word by using modifiers (e.g. cable car) and/or different words (e.g. gondola). More complex techniques, based on WordNet, can also be implemented to harvest word sense information from the gloss definitions, see Canas et al. (2003) for an example. For a discourse containing a dominant and subordinate senses we assume the dominant sense in the absence of a modifier corresponding to a subordinate sense. Hence in the previously mentioned excerpt “*We took the cable car to the top of the mountain. After spending 2 hours we took the car back.*” the first occurrence of the noun “car” was interpreted to be the subordinate sense (cable car) while the second occurrence was interpreted as the dominant (automotive) sense.

4.2 Noun Phrase Resolution Algorithm

4.2.1 Data Structures Used in the Implementation

The anaphora resolution system (aCAR) is written in Java hence the first task it does is to further parse the parsed input structure resembling lisp into a hierarchial object model shown in Figure 4.3. The upper level Document object is then used for all the processing by the resolution module. The Document object contains a list of sentences which in turn contains a list of clauses. All objects in the document class are identifiable using a unique identifier which is a number. In addition each of the clauses in a sentence are identified with a relative number starting from one, which is used to determine the order of the clauses within the sentence.

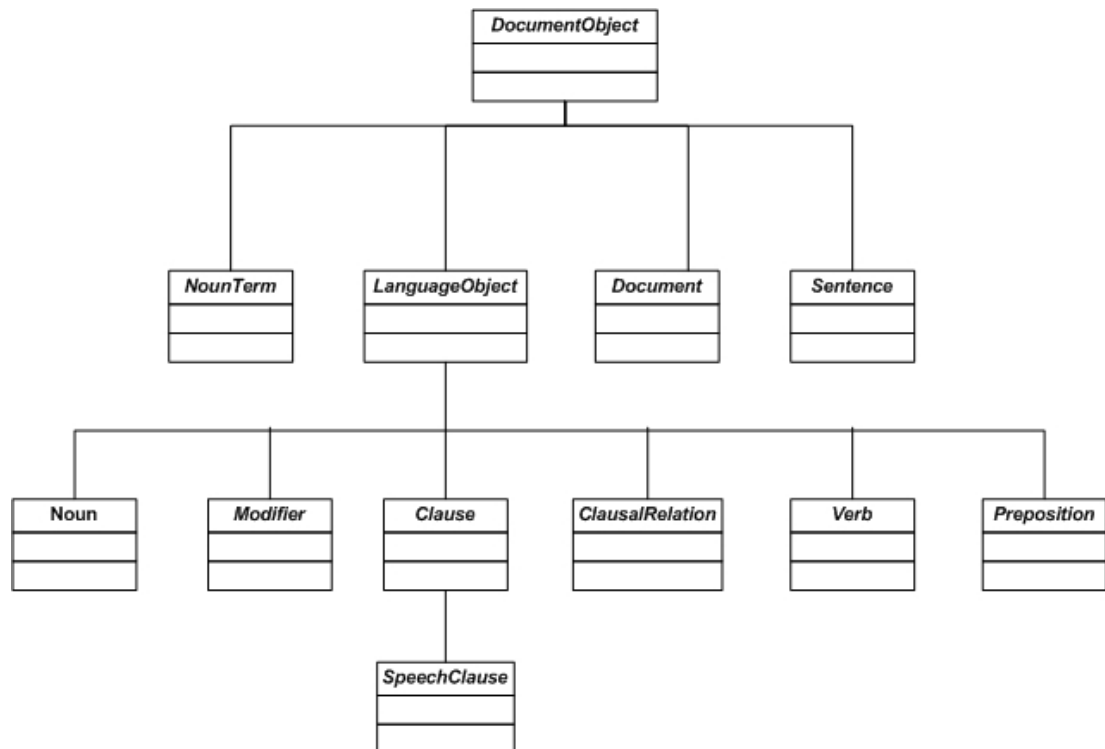


Figure 4.3: Document Class Diagram - Represents the hierarchal object model used by the anaphora resolution system for discourse processing

4.2 Noun Phrase Resolution Algorithm

The process of anaphora resolution in our implementation is a subsumption action in which the antecedent subsumes the anaphor. Hence, when an NP is identified to co-refer to another NP, all the attached information to each of the individual NPs are merged in order to create a richer information base corresponding to the entity referred to by the two NPs. To facilitate this we defined a conceptual data structure called a **Reference Unit (RU)**. This is a conceptualization of the real world entity co-referenced using different nouns in the discourse. A RU also encapsulates all the associated relations and property values of the entity. At the beginning of processing a discourse, a RU is formed for each noun giving us the same number of RUs as the number of nouns. As anaphora gets resolved, RUs and the associated knowledge get merged resulting in far fewer number of RUs compared to the initial number of NPs in the discourse. After the RUs get formed all processing transactions are done using RUs, and not individual nouns. Hence, a richer set of knowledge, originating from more than one NP is available to be used for the transaction. Recall that the scope of this study also includes normalized NPs, some of which co-refer to a clause as a whole, instead of only an NP. In order to include this, we formed RUs from clauses so that they can be made available as a target for co-reference relations. Hence at the start of processing, RUs get formed from all the NPs as well as the clauses in the document giving us a total equal to the sum of the number of clauses and the number of NPs. As the anaphora resolution progresses the number of RUs get continually reduced.

The components of a RU were defined to be:

- **ID** - A unique numerical identifier
- **Name** - A string name. The first mentioned NP was used as the name.
- **Member List** - the list of RUs co-referring to this entity, e.g. John, he, man and driver. The list can also contain a RU which encapsulates a whole clause.
- **Property List** - the list of properties and their values, e.g. gender and number.

- **Relation List** - the list of relations to other RUs. Relations were represented using tuples, e.g. HAVE(car, tyre) and ON(diesel, truck).
- **Preference List** - the list of potential resolutions to other RUs. Each preference candidate has an attached salience value.

When a RU (child RU) is subsumed as a member to another RU (parent RU), the child RU becomes invisible and all subsequent transactions are done with the parent RU. The subsumption process also merges the Property, Relation and Preference lists of the parent and child RUs. This subsumption process represents anaphora resolution.

4.2.2 Upper Level Algorithm

This section specifies the implementation of a range of constraints discussed as rules and preferences later in this chapter.

Before describing the algorithm, I will summarize below the pertinent aspects of our approach to resolving anaphora:

- We only consider co-reference as anaphora, hence resolve only co-reference anaphora.
- We use relations between entities as constraints for anaphora resolution.
- We do not use a separate task to identify anaphora before attempting to resolve them. All NPs are considered potentially anaphoric. Note that no other syntactical components such as verbs can be used as an anaphor, hence only NPs are considered.
- The antecedent search space for NPs is up to the beginning of the discourse for NPs and for pronouns, it is variable starting from the previous sentence.

4.2 Noun Phrase Resolution Algorithm

- NP and Pronominal anaphora are resolved in the order in which they appear in the discourse.
- If sufficient knowledge is not available to completely resolve an anaphor, it is left in a semi-resolved state. Its resolution is attempted multiple times as more knowledge is accumulated by resolution and merging of knowledge for other anaphora in the discourse.
- Any semi-resolved anaphora is resolved at the end using preferences.

The implemented system (named aCar) was designed as a set of independent modules to facilitate extensions at a later date. The overall architecture of the modules and their interactions are shown in Figure 4.4. The input to the system is a parsed clausal form shown in Figure 4.1. Due to lack of easy integration between the LISP based parser and the Java based resolver, the parsed form is obtained as a plain text file. Some manual adjustments were done to this file to correct any parsing errors. This file is then processed by the modules in aCAR to resolve the anaphora. The results is currently output as a text file, however can be easily delivered in XML format or as a java object for use by third party applications such as a document visualizer. The following are brief descriptions of the modules from Figure 4.4.

Document Object Builder Builds a java objects by reading lisp style clausal form from the parser. The whole document is constructed using lists of java objects corresponding to the clausal form from the parsed input. The anaphora resolution module uses this document structure for all traversals when searching for antecedents.

Morpho Miner This module extracts the gender, animacy and number information for nouns from the document. it uses syntactical structure and use of determiners such as “Mr” to derive morphological values where possible.

4.2 Noun Phrase Resolution Algorithm

Relations Miner This module extracts the relations between nouns using a combination of surface structures and knowledge sources. The techniques used are described in Section 4.3

Constraints Manager Applies a list of *hard* constraints on a candidate. A hard constraint can either completely resolve the anaphor or eliminate a candidate from contention. These constraints are described in Section 4.5

Preferences Manager Applies a list of *soft* constraints or preferences. The module uses a salience scores corresponding to a list of features to boost the likelihood of a candidate to be the antecedent. The list of preferences used are described in Section 4.6.

Statistics Manager This module updates and collects the statistics on the anaphora and the candidates as the anaphora resolution algorithm processes the NPs from the beginning of the document. It also formats and prints the summary statistics for the whole document at the end of processing of the whole document. aCar can also be configured to process multiple documents, in which case a summary of the series of documents can also be printed at the end of processing.

RU Manager Manages the merging of RUs when an anaphor gets resolved. This includes merging of relational knowledge as well as the other lists described earlier in Section 4.2.1.

Knowledge Sources The main knowledge source used was WordNet, however, some contextual knowledge was also mined from the document itself. In addition, local lists such as list of irregular verbs, morphological-values of father, mother, etc. were also used.

Anaphora Resolution Algorithm This is the main module that resolves the anaphora with the help of the other described modules. This module is described in more detail in Figure 4.5.

4.2 Noun Phrase Resolution Algorithm

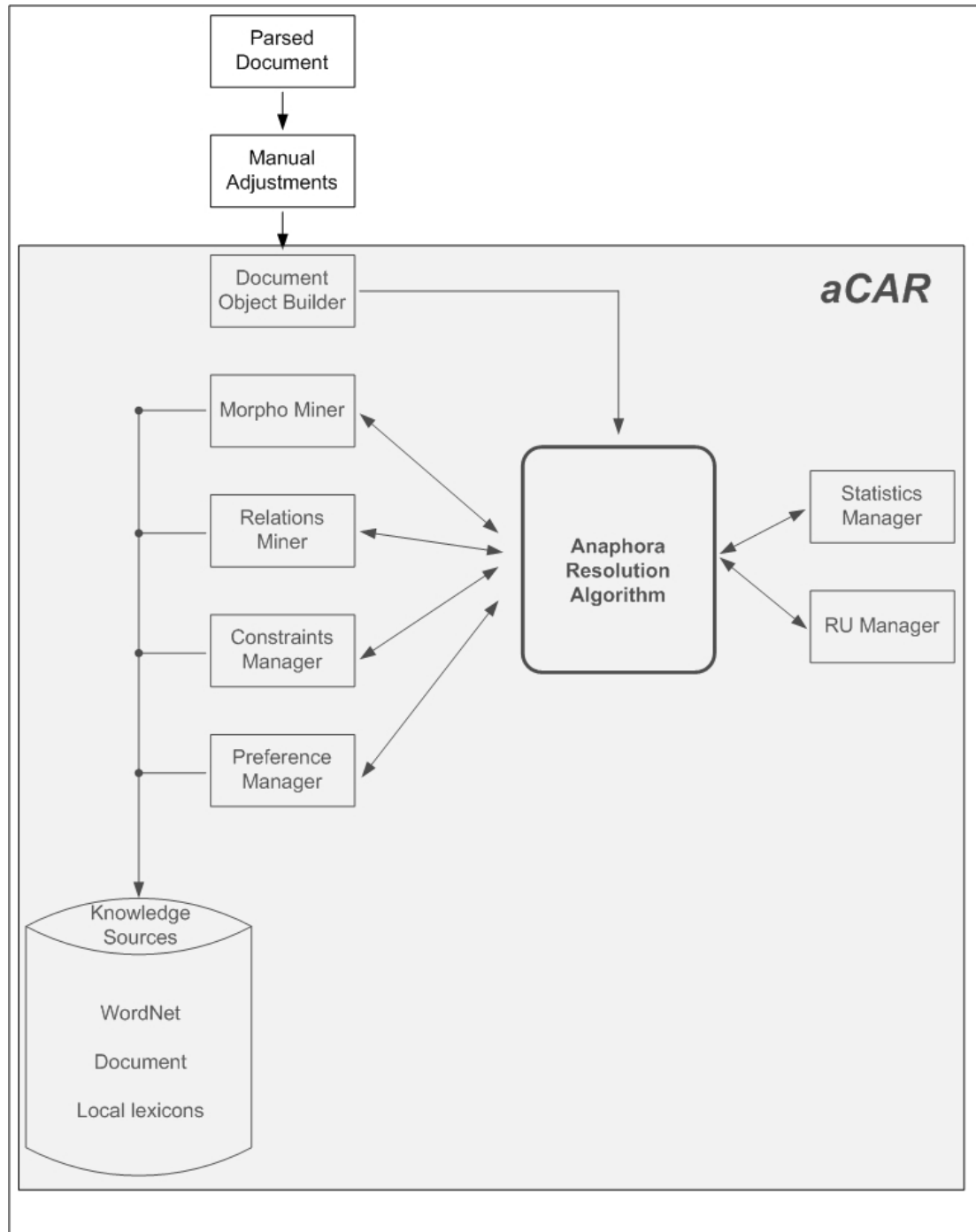


Figure 4.4: The Main Modules of aCAR - The diagram illustrating the architecture of the main modules of the anaphora resolver

4.2 Noun Phrase Resolution Algorithm

The algorithm in Figure 4.5 specifies the top most level steps as implemented in *aCAR*.

```
1: while (!End of document) do
2:   Get the next NP
3:   if (Pronoun) then
4:     Set window to previous sentence
5:     for all (RUs in window) do
6:       Apply Constraints
7:       if (No. of candidates == 0) then
8:         Extend window by 1 sentence
9:       end if
10:    end for
11:   else
12:     Set window to start of document
13:     for all (RUs in window) do
14:       Apply Constraints
15:     end for
16:   end if
17:   if (No. of candidates == 1) then
18:     while ( $\Delta$  Entropy != 0) do
19:       for all (Semi-resolved RUs) do
20:         Apply Constraints
21:       end for
22:     end while
23:   end if
24: end while
    {End of one complete pass. Apply preferences now}
25: for all (Semi-resolved RUs) do
26:   Apply preferences
27:   while ( $\Delta$  Entropy != 0) do
28:     for all (Semi-resolved RUs) do
29:       Apply Constraints
30:     end for
31:   end while
32: end for
```

Figure 4.5: Top Level Anaphora Resolution Algorithm.

4.2 Noun Phrase Resolution Algorithm

Processing of a document starts with the outermost *whileloop* from step 1 to 24. Since it was shown in Chapter 2 that pronouns are used more in a local context while NP anaphora is used in a more global context, we use different search spaces for pronominal and NP anaphora. Steps 3 to 10 resolve pronominal anaphora by initially starting with the previous sentence as the search window. This window is progressively increased, one sentence at a time, if the application of the constraints happens to eliminate all candidate RUs, in step 6. Application of the constraints in step 6 as well as the other places in the algorithm (steps 14, 20 and 29) can result in one of the following:

- Complete resolution.
- Partial resolution with more than one candidate remaining.
- No candidates left. In the case of pronouns the window is extended. In the case of NPs the NP is declared non-anaphoric.

For NP anaphora, steps 11 to 16, the corresponding search window was taken to be all of the previously encountered RUs up to the beginning of the document. Similar to the pronouns, the constraints applicable to the NP anaphora are applied in step 14. If the constraints alone are able to resolve the anaphor, then the number of remaining candidates remaining will be equal to one, represented in step 17. A complete resolution is interpreted as a gain in knowledge which is represented by the term *Entropy* in step 18. The term Entropy is used to represent “units of knowledge” extracted from the document. A unit of knowledge is defined to be a single piece of information about an entity. Some examples of a unit of information are :

- determining that an NP co-refers to the same entity referred to by another NP in the discourse, i.e. resolution.
- determining that the gender of an entity is male.
- determining a c-relation between two nouns.

4.2 Noun Phrase Resolution Algorithm

The value of Entropy was incremented by 1 for any information gain such as above. A change in Entropy, at any point in the process means there is a gain in new knowledge which can potentially be used to resolve any anaphora which is in a semi-resolved state. Hence, an Entropy change triggers an iterative pass through all the semi-resolved RUs in steps 18 to 21, applying the constraints again, this time with the new knowledge integrated.

Step 24 represents the end of a complete pass through the document. At this stage some of the anaphora will be fully resolved while the rest of them will be in a semi-resolved state. Steps 25 to 32 starts again from the beginning of the document and apply the soft-constraints as preferences in order to completely resolve the semi-resolved anaphora. Note that the preferences are not applied in a single pass, rather, after every resolution using preferences, a constraint application pass is again made through all the semi-resolved RUs (steps 28 to 30), in case a resolution can be made due to the latest resolution from the application of preferences. Hence, this algorithm gives a first chance to resolution using hard-constraints before using any of the soft-constraints. Hard constraints can be found anywhere in the document, including after the mention of the anaphor. The algorithm uses the fact that a discourse consists of a finite, and usually a small number of entities participating in events described in the discourse. Hence sometimes the knowledge required for resolution of an anaphor may be embedded after the mention of the anaphor. Humans may be able to resolve the same anaphor at the point of mention because of the better processing ability and/or availability of lexical knowledge. For a computational system, we can make up for this shortfall by mining for knowledge which might appear even after the anaphor.

The “Apply constraints” process in steps 14, 20 and 29 are described in more detail in Figure 4.6. The first step “Process Determiners” processes any existing determiners which are described in greater detail in Section 4.2.3. We then treat existence of names as a special case and attempt to resolve them using string matching of the entire name followed by matching of the last name. This can result in the name getting resolved, however, if it is not, then it is declared as non-anaphoric. If the head noun of the

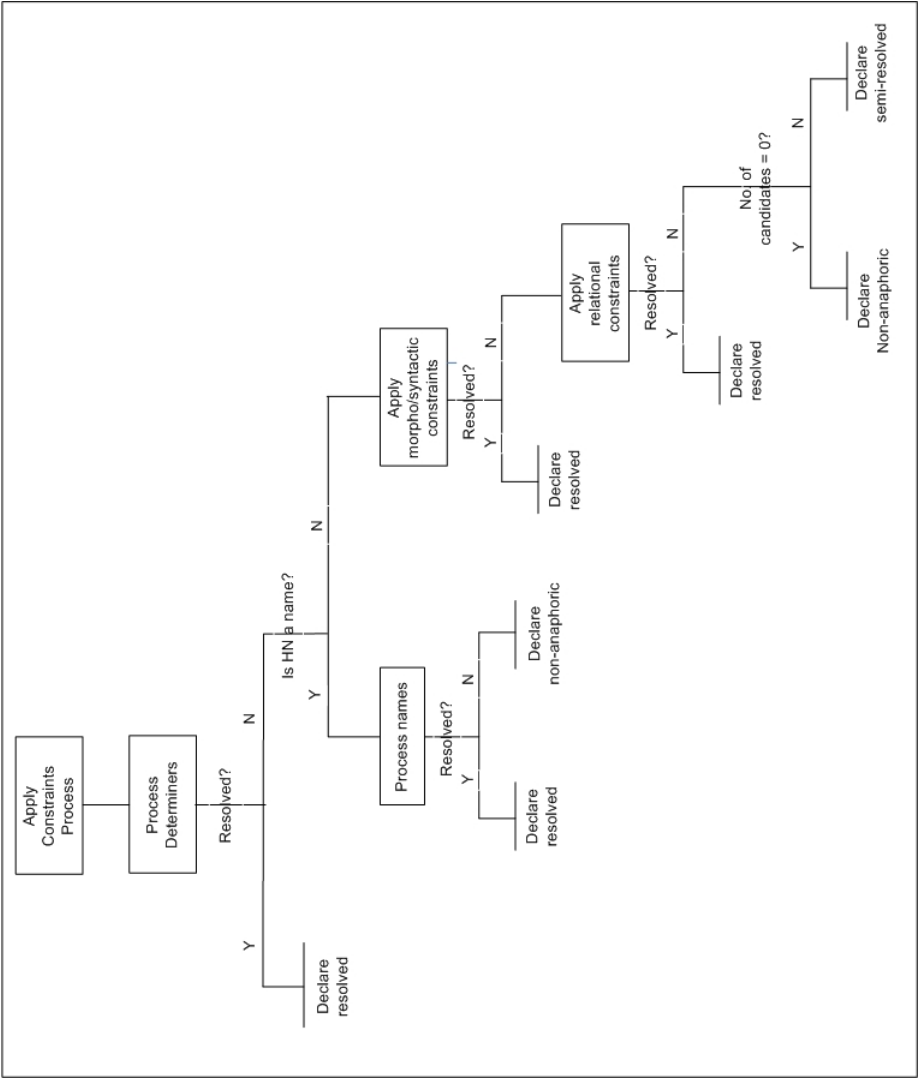


Figure 4.6: Diagram Showing the Inner Details of the Constraints Application Module

NP is not a name then we apply a list of morphological and syntactical constraints followed by relational constraints. After application of these constraints, if the NP has no candidates left then its declared non-anaphoric, otherwise it is left in an unresolved state to be attempted in a later pass. The next subsection describes the details of how determiners were used in the resolution process.

4.2.3 Treatment of Determiners

The composites of an NP (illustrated in Figure 4.2) can be represented in the form <determiner/s> <modifer/s> <HN>. An NP was processed starting from the determiners in order to first identify the presence of an *instance*. The scope of determiners was taken to apply to the head noun unless a hyphen is used to constrain the scope as discussed in Section 4.1.5. The value of the determiners were used to determine if an NP is:

- a co-reference to a single instance,
- a co-reference to one of the instances,
- an introduction of a new instance,
- a co-reference to more than one instances,

The following are the descriptions of the determiners and how they were used to identify instances of entities.

Article - The articles *a* and *an* were used only as an indicator of a singular noun. Apart from this use as an indication of the number, the articles were not used for any other purpose.

4.2 Noun Phrase Resolution Algorithm

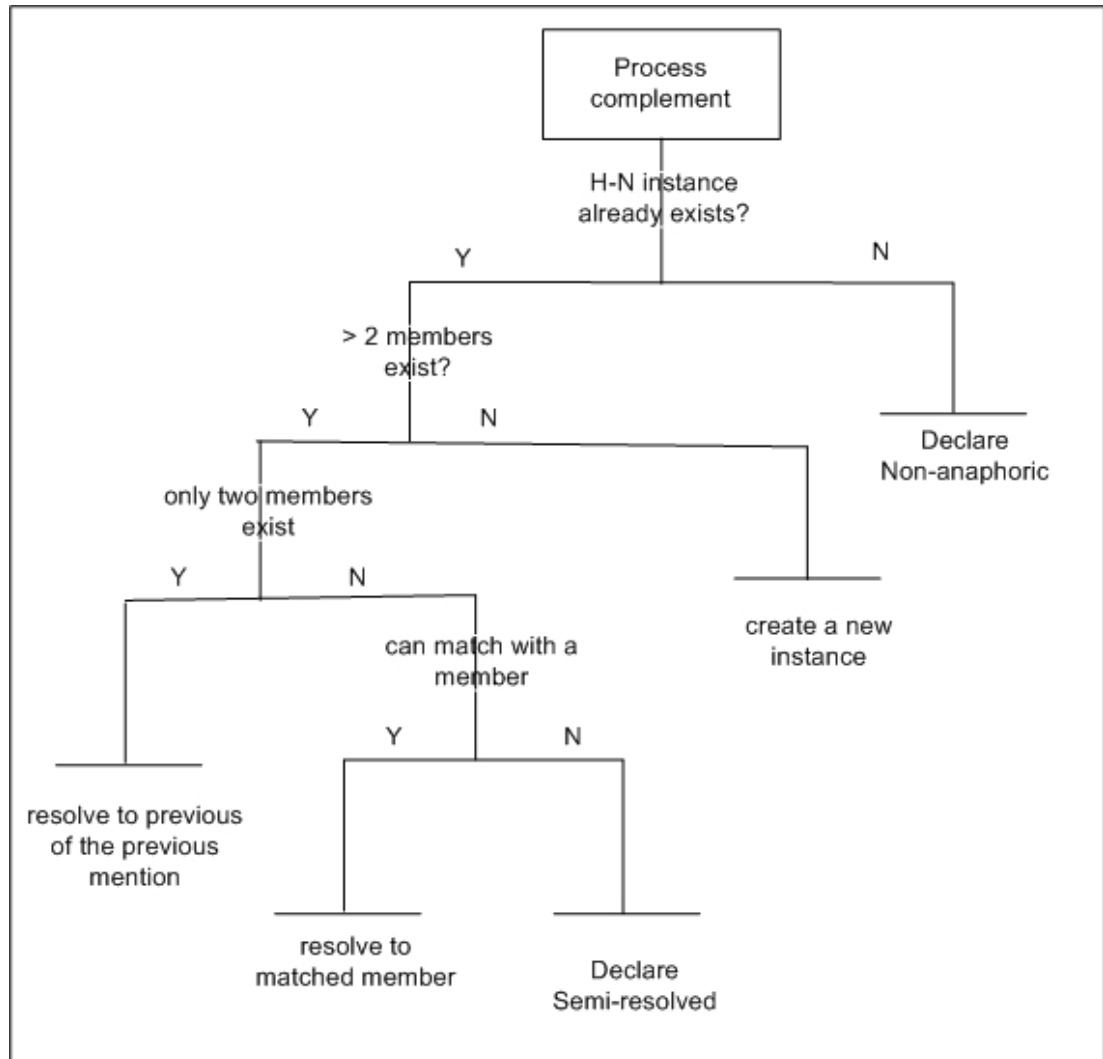


Figure 4.7: Diagram Illustrating the Complement Process

4.2 Noun Phrase Resolution Algorithm

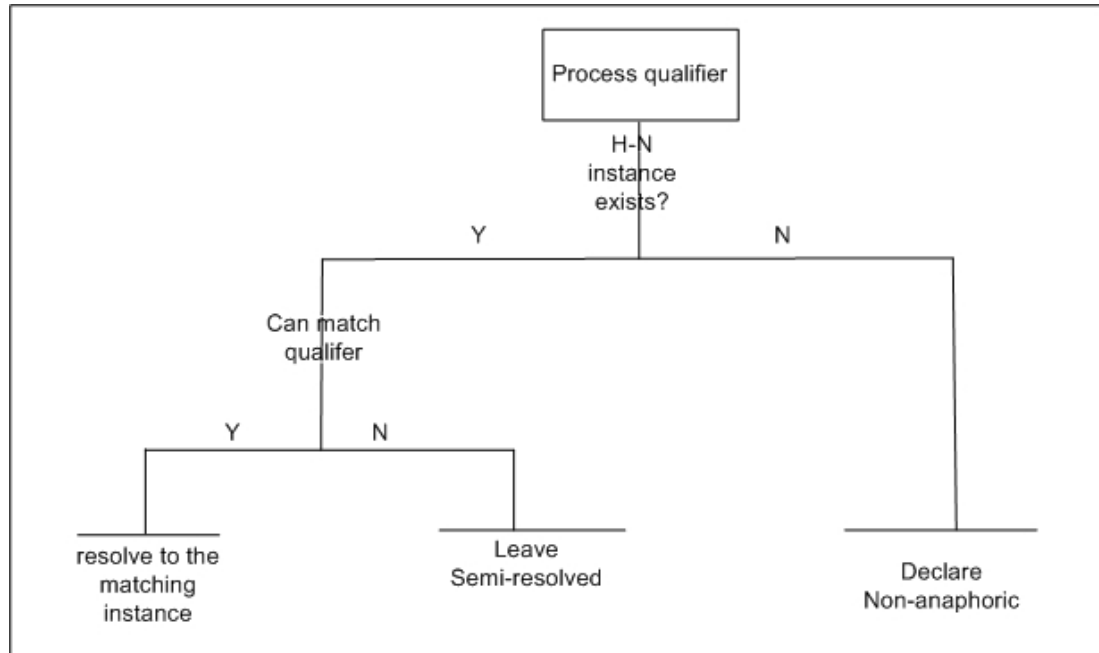


Figure 4.8: Diagram Illustrating the Qualifier Process

Demonstrative - The proximal demonstratives (*this, these*) were used to identify the most recently mentioned entity in the case of two instances while the distal demonstratives (*that, those* and *yonder*) were used to identify the previous to the most recently mentioned entity represented by the head noun.

Possessive - The possessives were used to determine multiple instances of “possessee” in the case of multiple instances of the “possessor” entity. For example, for the NP “his truck”, where “his” is equal to “driver”, in a context where it has already been determined that there are two instances of “truck”, the pronoun “his” was used to determine that there are two instances of “driver” as well.

Title - Titles are used with entities which are names. These were used to identify that two entities are co-referential when the subsequent one is used with a title and the last name.

Complements - These are frequently used to distinguish between two instances. The complement of an entity is another entity that makes it complete. A common

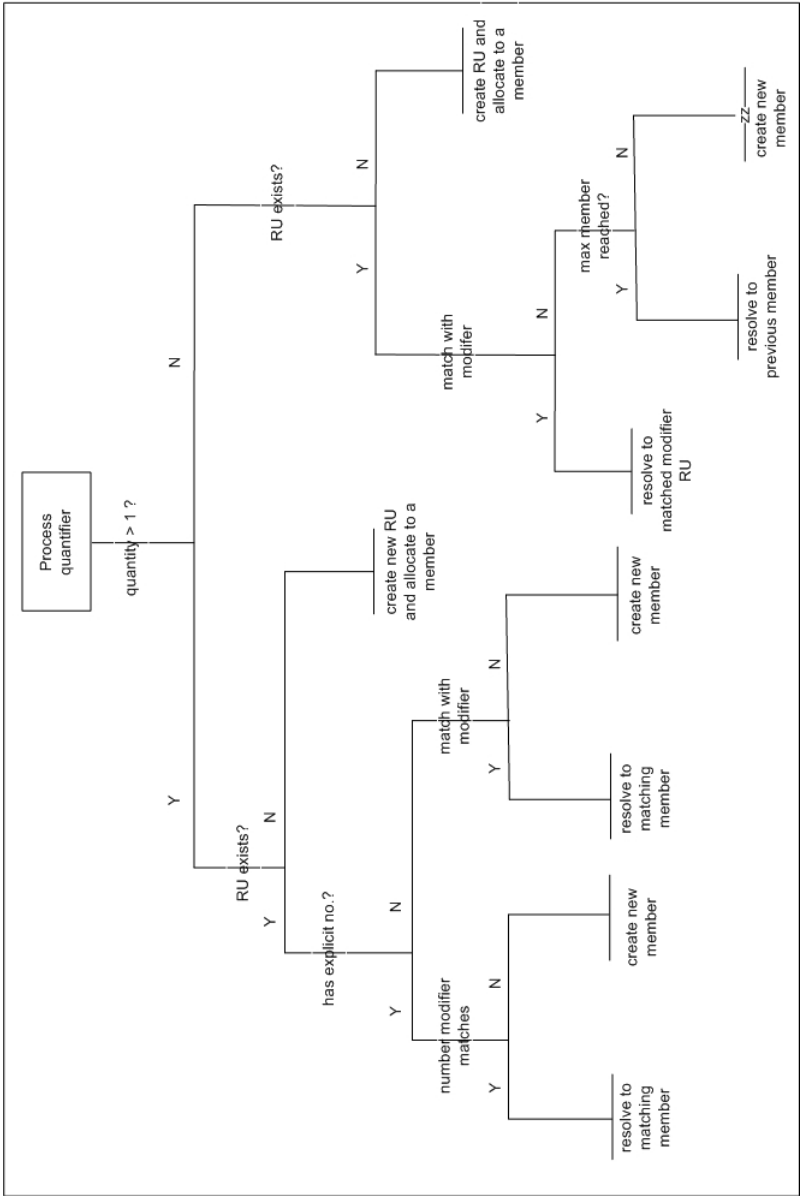


Figure 4.9: Diagram Illustrating the Quantifier Process

means of expressing complement is by the use of the word *other*, e.g. “other truck”. Currently our system contains only one complement word, “other”. The inner details of how this was used to identify instances is illustrated in Figure 4.7. When a complement is used it implies at most two instances of an entity. This means either both instances have already been identified or only one has been identified. In the former case we resolve it to the “previous to the previous” mention. For the latter case we create a new instance, as illustrated in Figure 4.7.

qualifier - While the complement has only two components to complete the whole, a qualifier could have more than two constituents, e.g. top/middle/bottom, and slow/medium/fast. These are more difficult to implement and can not be easily extracted from WordNet. Fortunately, use of qualifiers to distinguish between more than two entities are rare and were found only 6 times in the input corpora for this research, and in all the cases, use of additional modifiers were used. Some examples of qualifiers used were:

- *Colors* - red, blue, green etc.
- *Shapes* - round, square, circular etc.
- *Temperature* - cold, cool, hot, comfortable etc.
- *Size* - tiny, big, large, huge etc.
- *Extent* - slight, medium, extreme.
- *Direction* up, down, north, south, southbound etc.

Figure 4.8 illustrates how qualifiers were processed by the system. If the qualifier could not be matched, we left it in a semi-resolved state to be further processed by other constraints.

Quantifiers - These are determiners that can be used to express and identify two instances as well as any number more than 2. Non numeric quantifiers such as *most*, *some*, *rest* were treated as qualifiers hence the category of *quantifiers* were

only used to deal with the ones that represent cardinality. These include quantifiers such as *pair*, *couple*, *another*, *first*, *third*, *dozen* as well as the numeric representation of numbers. The interpretations for quantifiers fall into two groups; the first applies to quantifiers such as “couple”, where the word represents the actual number of entities involved. The second interpretation refers to one entity out of an identified number (e.g. *third* represents the third entity out of three) or out of an unidentified number (e.g. *another*). Figure 4.9 illustrates the overall process of how quantifiers were used in the resolution process. Note that when the quantifier can not be used to resolve an NP it is left in an unresolved state. An attempt is made later to resolve it using other constraints as shown in the higher level algorithm in Figure 4.6.

4.3 Extracting Relational Knowledge

This section describes the computational techniques that were used to identify the eight compound noun formation relations (c-relations) from Table 3.1. It should be noted that for the purpose of this study, it is not absolutely critical that we are able to identify the specific relation type between two entities, but to be able to identify two entities that are related by ANY of the relations in the list. To do this, we searched for the existence of any of the c-relations among the nouns in the document using a combination surface structures, WordNet and some local lexicons. The following sub-sections, correspond to each of the c-relations in the list CAUSE, HAVE, MAKE, USE, IN, FOR, FROM, and ABOUT. The sub-sections firstly give a brief background on the associated semantics of each of the relation types followed by a description of the strategies used to extract the corresponding relation from a document. The strategies used for identifying each of the relation types, although different, are repetitive. Hence, a summary of the sections from 4.3.1 to 4.3.8 is provided in Section 4.4 for readers who want to skip the derivation details.

4.3.1 The CAUSE Relation

In order to extract CAUSE relations we need an overview of what it means and how it is expressed in terms of surface structures. The semantic concept of *cause* is complex and has received considerable attention from language and semantic researchers (e.g. Kaplan and Berry-Rogghe, 1991; Khoo et al., 2000; Kim, 1971). The concept of *cause* has been described to have three components; *cause*, *effect* and optionally, a *consequence*. Consider the example “Peter was sick.” As a result, “Susan was late for work.” In this case the *cause* is represented by Peter being sick somehow causing Susan to be late for work. This is a case of **implicit** cause relation since we need to draw inferences as to how the two events are related, for instance, Peter might be Susan’s son, and she might have had to go the doctors before work. Kim (1971) describes causation to be often events, but can also be conditions, states, phenomena, processes and sometimes even facts. He says that “coherent causal talk is only possible within a coherent ontological framework of such states of affairs.” In the case of Susan getting late for work, it is only possible to draw the relevant inferences in order to determine the causative relation, if we are familiar with the context that Peter is Susan’s son which corresponds to Kim’s “ontological framework”. This type of implicit causal relations are based highly on semantics which is outside the scope of this study hence were not attempted to be extracted.

The other form of causation relation is **explicit**, which can be determined by the associated verb, for instance “earthquake causes Tsunamis” and “smoking leads to lung cancer”. These are expressed in the form <NP1> <CAUSE VP> <NP2>. A large proportion of causative relations can be determined by examining such verbal constructions. A verbal construction can be decomposed into a taxonomy of causative verbs in terms of whether they define a causal link between two NPs or additionally define other components of a causal relation. In the phrase “Peter killed his wife with poison”, what is the cause of “kill”, Peter or the poison? In this case, the causative relation’s *cause* component has two sub-components; an *agent* and an *instrument*. Hence we can define two types of cause relations:

Simple causatives (*cause, generate, bring about, make, leads to, force etc.*) In this case the linking verb expresses the causal link between the NPs, which is synonymous with the verb *cause*. E.g. “Smoking *leads to* lung cancer”.

Compound causatives (*kill, melt, provoke, poison, dry, etc*) The verb refers to the causal link which has either an explicit or an implicit instrument that is the composite cause of the effect. For example in “Peter killed his wife” there is an implicit instrument and in “Peter killed his wife with poison”, poison is the explicit instrument. In the second case even human interpretation of the cause of death is arguable between “Peter” and “poison”. In our implementation, if both are specified, we treat them both as causative agents and form two CAUSE relations, *CAUSE(Peter, wife)* and *CAUSE(knife, wife)*. If only one is specified then we form only one relation with whichever, out of cause and instrument, is specified.

It would be inaccurate to base the determination of CAUSE relations solely on the categorization of verbs since a verb’s “causativeness” can also depend on the subject and object. For instance the verb “produce” in “The tree produces fruit” is not causative contra to its use in “This procedure produces a curious effect”. This gives us three classes of verbs characterizing causativeness (Kim, 1971):

- Ones that are always causative. These are the synonyms of Cause such as *cause, generate, make, etc.*
- Ones that are ambiguous and depends on the connecting subject and the object such as *produce, provoke, trigger off, etc.*
- Ones that are not causative in any usage such as *write, talk, eat, etc.*

To extract the causal relations the use of a verb was first checked against a list of verbs which were classified to be always causative. If, the verb was not in the list then it was classified to be potentially causative depending on the subject or the object. To determine the set of NPs that can give rise to ambiguous verbs to be causative, we used the technique outlined in Girju (2003) in which the author used a set of semantic features

4.3 Extracting Relational Knowledge

to automatically train a system to determine causative use of verbs using WordNet. In her study she claimed a precision of 73.91% and a recall of 88.69%. Her strategy involves classifying all NPs in terms of the nine noun hierarchies used by WordNet, they are; *entity*, *psychological feature*, *abstraction*, *state*, *event*, *act*, *group*, *possession*, and *phenomenon*. This can be found by traversing via the direct hypernym pointer to the farthest synset from the noun. A clause with a verb that can be ambiguous is dependent on the semantic features of the connecting NPs as shown in Figure 4.10.

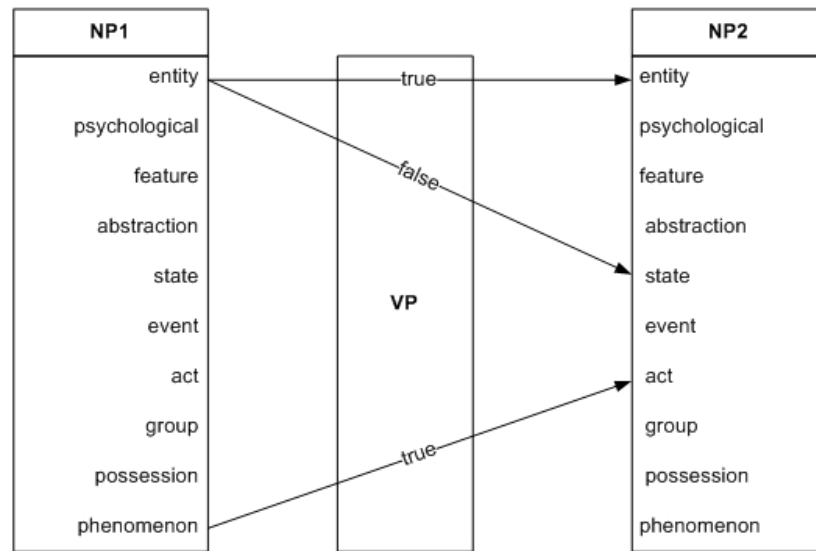


Figure 4.10: Illustration of a Causative Relation between Two NPs with a Connecting Verb - The nine semantic features for the NPs correspond to the features used in WordNet. The two true-arrows indicate a causative use of a verb while the false-arrow represents a non-causative usage.

The example in Figure 4.10 shows that for an ambiguous VP, if NP1 and NP2 are both entities, then the use of the verb is causative. The VP is also causative if NP1 is a phenomenon and NP2 is an act but it is not causative if NP1 is an entity and NP2 is a state. The full set of rules used were adopted from Girju (2003), shown in Table 4.3.

To use this result, we checked the semantic feature of the connected subject and object of a verb in order to classify whether the verb expresses a causative relation. All other uses of verbs were classified to be non-causative.

4.3 Extracting Relational Knowledge

NP1 class	VP	NP2 class	causative?
*	unambiguous causative verbs (e.g. cause, make, force etc)	*	yes
*	*	phenomenon	yes
! entity	associated-with or related-to	! abstraction and ! group and ! possession	yes
! entity	*	event	yes
! abstraction	*	event or act	yes
*	lead-to	! entity and ! group	yes
*	induce	entity or abstraction	no
*	*	! state and ! event and !act and group	no
entity	*	!state and ! event and !phenomenon	no

Table 4.3: Rules used to determine if a verb is classified as causative (adopted from Girju (2003)). An * means anything.

4.3.2 The HAVE Relation

The HAVE relation has a very wide scope and can include several forms of possession relations between two entities. The following list outlines the sub-categories with illustrative examples.

PART-OF When an entity is physically attached or conceptually comprises another entity. For example, a window attached to a house and Calculus comprises Maths form PART-OF relation.

OWN When an entity legally or otherwise owns another entity. For example, John's car and Mary's cat.

ANIMATE-RELATION Animals, including humans are related to one another by several types of relations. These relations do not have a strictly possessor/possessee relation since either can be possessed by the other, signified by *renaming* of the relation. For example, a woman can be a man's mother, where the man possesses the woman with the relation "mother". The same relation can be renamed

as “son” in which case the woman possesses the man. For some relations even the renaming of the relation is not necessary, e.g. cousin relation, where two humans are cousins of each other. Some other examples of such bi-directional relations are brother-in-law, partner and friend. In this case the two entities in the HAVE relation are interchangeable.

BELONG When an entity belongs to another entity without the legal or other implications of ownership. This represents a weaker and transient form of ownership such as a person belonging to a country or a politician belonging to a party.

PARTICIPATION An entity can also form a HAVE relation with the eventuality resulting from the entity’s participation in an action. For instance a person surviving an air crash results in the “eventuality of surviving” belonging to the person who participated in the action “surviving” as in the phrase “His survival ...”

The types of HAVE relations discussed above can be expressed using a number of surface structures. The following are the list of structures that were used to extract the HAVE relations:

- **Personal pronouns** - E.g. “his car” represents a HAVE relation between the antecedent of “his” with “car”
- **Apostrophe** - E.g. “John’s car” represents a HAVE relation between “John” and “car”.
- **HAVE verbs** - Some verbs such as *have*, *own*, *contain*, *belong* can be used to express the HAVE relation. A local list containing examples of HAVE verbs was used to extract the HAVE relation expresses using verbs.

4.3.3 The MAKE Relation

This category consists of relations between two entities where one entity is physically or conceptually used to construct the second entity. Hence the MAKE relation includes

all the cases where one entity *resides* in another entity. A range of different type of examples from the MAKE category are:

- window/house
- wood/house
- poison/snake
- sugar/cake
- flower/wreath
- calculus/Maths

The MAKE relation was established using WordNet supplemented by manually created lexicons corresponding to the relations encountered in the input data. WordNet(1.7) has a number of synsets connected by the MAKE relation. These were found by traversing the following pointers in WordNet:

- *part-holonym* e.g. door/house
- *part-meronym* e.g. house/door
- *substance-holonym* e.g. wood/lumber
- *substance-meronym* e.g. lumber/wood

4.3.4 The USE Relation

This relation is meant to capture the relation between two entities where the first entity is used as a consumable by the second entity to achieve the second entity's principle functionality. This includes examples such as an electric drill uses electricity, a pen

uses ink, a car uses petrol and a painter uses paint. These consumables may be expressed explicitly as in the case of “electric drill” or assumed to be a default value as part of encyclopedic knowledge. The default value for a car for example, is petrol, however default values can be overridden by another, contextually relevant entity such as electricity for an electric car.

The USE relation is not directly encoded in WordNet, however some relations can be mined from the gloss definitions using the verb forms *used to*, *used with*, *used in*, *used by* and *used as*. In this case the noun that precedes the verb form is the consumable, and the one that follows it is the user, for example, “fuel used by internal combustion engines”.

Other syntactic patterns used to mine the Use relation were:

- USE verbs - local list of verbs such as *use*, *utilize*, *employ* and *apply* were used to declare the connecting nouns in a USE relation.
- ING-form - The syntactic form <NP1> <VP> <NP2> <VP-USE_{ing}> <NP3> was used to declare a USE relation between NP1 and NP3.
- WITH-form - The syntactic form <NP1> <VP> <NP2> <with> <NP3> was declared to have a USE relation between NP3 and NP1.

4.3.5 The IN Relation

The IN relation captures entities that are related by their proximity either in terms of time or location. It can be easily determined by the use of prepositions *in*, *on* and *at* in one of the following syntactical templates where <NP1> is related by the IN relation to <NP3>:

1. <NP1> <VP> <NP2> in/on <NP3> e.g. John went to town in his car.
2. <NP1> <VP> in/on <NP3> <NP2> e.g. John lives in Auckland.

3. <NP1> in/on <NP3> <VP> <NP2> e.g. The person on the footpath was John.
4. In/On <NP3> <NP1> <VP> <NP2> e.g. At 2 o'clock, I have an exam.

4.3.6 The FOR Relation

The FOR relation is captured predominantly by use of the preposition *for*, however the interpretations are slightly different as summarized below:

- **Purpose** of an action. E.g. run for fitness.
- **Recipient** of an action. E.g. raised money for the homeless.
- **Equivalence** for an action. E.g. paid lot of money for a car.
- **Amount** indicator. E.g. walked for 2 hours.

We used all of the above interpretations of the preposition *for* to extract the FOR relation. The object of the clause and the target of the for relation were taken as the arguments of the relation. For instance, the phrase “John practiced Karate for self defence” would result in FOR(Karate, self defence).

In addition a local list containing a list of “FOR-verbs”, for example, “meant to”, “purpose of” and “used as” were also used to extract the FOR-relations expressed using verbs.

4.3.7 The FROM Relation

This relation has the following variations in semantics:

- **Movement** This can be either in terms of time/location. E.g. from 2’oclock to 3 o’clock, from home to school.

4.4 A Summary of Derivational Aspects of Relations

- **Origin** Used to indicate source. E.g. oil from olive, book from shelf.
- **Separation** Separate an entity from another. E.g. liberation from bondage.
- **Differentiation** Distinction of one entity from another. E.g. good from bad.
- **Causal** Overlaps with the CAUSE relation. E.g. die from heart attack.

All of the above forms of FROM relation were extracted using the “from” preposition in the form NP_1 from NP_2 . The pattern was searched both in the discourse as well as WordNet gloss definitions.

4.3.8 The ABOUT Relation

This is a relation where two entities are related by the virtue of an entity being a topic of another, hence cannot be determined by any surface structure. Because this relation is inherently challenging to extract and there were two instances in the corpus used for this study, we did not implement any automatic extraction of this relation. We instead manually encoded the relation and used it for anaphora resolution.

4.4 A Summary of Derivational Aspects of Relations

Table 4.4 summarizes the essential derivational aspects of the relations discussed in Sections 4.3.1 to 4.3.8.

4.4 A Summary of Derivational Aspects of Relations

Table 4.4: Table summarizing the derivational aspects of the relations discussed in Sections 4.3.1 to 4.3.8

Relation	Surface indicators	WordNet used	Hand coded lexicon used
CAUSE	Always causative verbs. e.g. <i>cause</i> , <i>generate</i> . Ambiguous verbs.e.g. <i>produce</i> . Use connecting subject and object WordNet hierarchies. All else treated as non-causative.	YES	YES
HAVE	Personal pronouns. Use of apostrophe Verb morphemes for PARTICIPATION relations Handcoded lexicon for ANIMATE relations	YES	YES
MAKE	Used lexicons to determine all relations. Search for verbs such as <i>used in/as</i> , <i>made from/of</i> etc.	YES	YES
USE	Search WordNet for “use patterns”, e.g. <i>used to</i> , <i>used in</i> etc. Search discourse for “use verbs” e.g. <i>employ</i> , <i>utilize</i> etc. Other syntactic patterns used were: <NP1> <VP> <NP2> <VP-USEing> <NP3> <NP1> <VP> <NP2> <with> <NP3>	YES	YES
Continued on next page			

Table 4.4 – continued from previous page

Relation	Surface indicators	WordNet	Hand coded
IN	<p>Clauses containing the prepositions <i>in</i>, <i>on</i> and <i>at</i> in the following syntactic patterns:</p> <p><NP1> <VP> <NP2> in/on <NP3> e.g. John went to town in his car.</p> <p><NP1> <VP> in/on <NP3> <NP2> e.g. John lives in Auckland.</p> <p><NP1> in/on <NP3> <VP> <NP2> e.g. The person on the footpath was John.</p> <p><NP3> <NP1> <VP> <NP2> e.g. At 2 o'clock, I have an exam.</p>	NO	NO
FOR	<p>Preposition <i>for</i> used in any pattern</p> <p>Other <i>for</i> verbs, e.g. <i>purpose of</i>, <i>meant to</i>, <i>used as</i>.</p>	NO	NO
FROM	<p>Preposition <i>from</i> used in NP₁ from NP₂ pattern.</p>	YES	YES
ABOUT	Not implemented.	NO	YES

4.5 Hard Constraints Guiding Anaphora Usage

This section describes the non-defeasible constraints guiding the use of anaphora by a producer. These constraints were used in the algorithm in reverse for the process of resolving anaphora.

Let us define some terms with specific meanings that will be used in the following

discussion. We will use the term *rule* to refer to hard constraints which is always valid and use the term *preference* to refer to soft constraints which can be violated in a specific context. An *entity* is a real life object that can be represented by an NP. We define a relation between two entities to be *orthogonal* if they **cannot** be coreferential under any circumstance. If an entity **can be** coreferential to another, it will be referred to as *parallel* and if an entity is determined to be **definitely** coreferential to another, then we will refer to them as *equal*. Thus the process of anaphora resolution amounts to finding ways to enforce *orthogonality*, *parallelism*, or *equality* of NPs in a discourse.

4.5.1 Relational Constraints

4.5.1.1 Predicate Deletion

This constraint applies to anaphoric NPs which are compound nouns consisting of one or more modifier and a head noun. In this case the primitive nouns corresponding to the modifier and the head noun have to be related by a relation from the set CAUSE, HAVE, MAKE, USE, IN, FOR, FROM AND ABOUT. Using this constraint in reverse for anaphora resolution, involves extracting all possible relations in the set from the discourse in the form $REL(N_1, N_2)$. Once these relations are extracted, we can assume that any compound noun of the form $N_1 + N_2/N_2 + N_1$ for the bi-directional relations (CAUSE, HAVE AND MAKE), and $N_1 + N_2$ for unidirectional relations (USE, IN, FOR, FROM, ABOUT), originated from the source clause of the relation. If N_1 and N_2 are exactly the same noun in both the relation and the compound noun, then resolution is trivially a matter of matching the strings. However, if either N_1 , N_2 or both are synonymous, then we used the WordNet and *Similarity* tool (described in Section 4.1.4 for matching with the nearest synonymous noun. For a compound noun of the form $N_1 + N_2$, if a string match was not found then an indirect attempt was made with a list of synonymous nouns for both N_1 and N_2 . **Both** the primitive nouns had to have some degree of match (similarity value more than 0.5) before declaring a co-reference

relation. This threshold was enforced to prevent matches such as “Freightlines truck” to “Freightlines driver”. For NPs with more than one noun-based modifiers, the relation between each of the modifiers with the head noun were tested against the set of relations found in the discourse. That is, an NP of the form $NP_n + NP_2 + NP_1 + HN$ was tested for existence of any of the relations between NP_1 and HN, NP_2 and HN, $\dots NP_n$ and HN.

4.5.1.2 Normalization

This constraint determines the types of morphemes corresponding to a verb or adverb which is available to a producer to produce an anaphoric NP by the process of normalization. As discussed in Section 3.5, an anaphoric NP produced by this process can be anaphoric to the agent or the object of the source clause, but in addition, can also be anaphoric to the clause as a whole. In order to determine the co-referential properties of normalized NPs, we firstly identified the normalized NPs in the base corpora used for this study consisting of 120 newspaper articles and classified their antecedents into AGENT, OBJECT or the CLAUSE. In addition we also used The Corpus of Contemporary American English (Davies, 2010) for morpheme types which were either sparse or non-existent in the base corpus. This freely available corpora consisting of some 410 million words from a variety of genre has an online web interface which can be used to do fairly complex searches for words and phrases hence forms an excellent resource for manual content analysis for NLP tasks. Table 4.5 outlines the range of morpheme types and the corresponding antecedents. The second column in Table 4.5 gives the common part-of-speech usage for the morpheme types however note that some of the adjectives can also be used as nouns. The last column gives the component of the clause to which the normalized NP co-referenced in the majority usage in our content analysis. The information from the table was used as constraints by using the suffix of NPs to search for previous occurrence of the stem verb and if one was found then the respective component from the third column of Table 4.5 was used as the antecedent.

4.5 Hard Constraints Guiding Anaphora Usage

Suffix	POS	Examples	Antecedent
-er, -or, -ie	noun	teacher, director, truckie	AGENT
-ant, -ent	noun	applicant, president	AGENT
-ess (not ness)	noun	waitress	AGENT
-ist	noun	plagiarist, abortionist	AGENT
-ian	noun	politician, beautician	AGENT
-ee	noun	employee, trainee	OBJECT
-uct	noun	product, conduct	OBJECT
-ness	noun	abusiveness, sharpness	CLAUSE
-ity	noun	security, stability	CLAUSE
No suffix	noun	run, talk	CLAUSE
-ment	noun	achievement, astonishment	CLAUSE
-ance, -ence	noun	violence, attendance	CLAUSE
-tion	noun	information, expression	CLAUSE
-ism	noun	criticism, terrorism	CLAUSE
-ship	noun	leadership, dictatorship	CLAUSE
-able, -ible	adj	reliable, valuable	INDETERMINATE
-al	adj	arrival, theoretical	INDETERMINATE
-ful	adj	wonderful, helpful	INDETERMINATE
-ish	adj	greenish, girlish	INDETERMINATE
-ive	adj	creative, captive	INDETERMINATE
-ous, -ious	adj	glamorous, flirtatious	INDETERMINATE
-ic	adj	hypnotic, theoretic	INDETERMINATE
-ed, -en	adj	played, written	INDETERMINATE
-ing	adj	running, thinking	INDETERMINATE
No suffix	adj	wet, silent	INDETERMINATE

Table 4.5: Table showing the possible suffixes inflicted by morphemes of finite verbs that can function either as a head-noun or a modifier of an anaphoric NP.

For the cases where the normalization resulted in an adjective (bottom 10 cases in Table 4.5), the antecedent was found to be indeterminate. In these cases, the adjective modifier could not be used to determine the antecedent, as it was dictated entirely by the head noun of the NP.

4.5 Hard Constraints Guiding Anaphora Usage

Property	Possible property values
<i>Gender</i>	male,female,neutral.
<i>Animacity</i>	true,false,neutral.
<i>Number</i>	singular,plural.

Table 4.6: Table morphological properties and their possible values.

4.5.2 Morphological Constraints

The most commonly used criteria by anaphora resolution systems (e.g. Baldwin, 1997; Jerry, 1978; Lappin and Leass, 1994; Mitkov, 1998; Nasukawa, 1994; Prasad and Strube, 2000) is the compatibility of morphological properties, *gender*, *number* and *animacity*. Morphological compatibility is the starting point for any pronominal anaphora resolution system. Table 4.6 gives a list of morphological properties and their respective values.

The morphological constraints enforce that the corresponding property values for the anaphor and the antecedent must be the same. In order to be able to use this basic constraint we need to know the property values of the anaphor and the antecedent, which is not readily available in any lexicon including WordNet. We used two sources for the morphological property values. The first was a manually encoded list containing all the pronouns as well as commonly used nouns and names such as man, woman, father, Mary, John and driver. In addition, we also use simple extraction techniques to extract the values that are indirectly expressed in the discourse, either in the NP itself or in the syntax. For example, “Mr Lenzie” implies male, singular and animate while “spokeswoman” implies female, singular and animate. Similarly some of the property values can also be expressed in the syntactical structure of the language, for example, “The Smiths have...” implies that the number value for “Smiths” is plural. At the beginning of processing a discourse the property values were populated for all nouns for which values were known either from the lexicons used or from mining the discourse. Not all values get to be known for all the nouns in a particular discourse. The property values which were not known were populated with the string *unknown*

4.5 Hard Constraints Guiding Anaphora Usage

to encode that this value is currently unknown. When applying the morphological constraints a candidate was discarded only if a **known** property value got violated. Thus we defined a rule based on the morphological constraint as :

(Rule 1) **Morphological Orthogonal Rule** Two entities are orthogonal if any one known property value is incompatible.

If one or both of the values for a property are unknown, then it remains in the candidate list until preferences get to be applied after a complete pass of the document. Some of the unknown values can become known in the process of anaphora resolution by the subsumption process in which the anaphor and the antecedent inherit the knowledge from each other. This can lead to rejection or resolution of a candidate to an anaphor at any point in the first pass of the document.

4.5.3 Syntactical Constraints

This section discusses the clause level syntactical constraints that guide the use of NPs for constructing a clause. These are described in the following sub-sections.

4.5.3.1 Orthogonality

We defined a constraint called *Orthogonal Rule*, which is based on the binding theory (Chomsky, 1969), formalized as *c-command* by Reinhart (1983). This theory explains how in the sentence, “Mary’s brother helped her”, the pronoun “her” can be co-referential to either “Mary” or someone else while in “Mary helped her”, the pronoun “her” can only be coreferential to someone else. A simplified version of this theory was implemented as the following orthogonal rule:

(Rule 2) **Orthogonal Rule** All participating entities in a single event are orthogonal except entities represented by a reflexive pronoun.

We defined an event to be represented by a clause as parsed by the parser used (described in Chapter 5) and participating entities were defined as the head nouns of all the entities. This excludes modifier nouns hence excludes “Mary” in the sentence “Mary’s brother helped her” and enforces that “brother” and “her” are orthogonal.

The *Orthogonal Rule* can be utilized to either eliminate a highly competitive candidate or in some cases resolve to the correct antecedent in a case in which there were two candidates to start with. As an example, consider two clauses represented as the following two events:

- (4.1) (a) Event₁ {subject₁, object₁}.
(b) Event₂ {subject₂, object₂}.

Suppose both the entities, subject₂ and object₂ in (4.1b) are pronouns. The candidate list for both pronouns would contain the entities subject₁ and object₁. If for instance, subject₂ can be resolved to object₁ using some other rule, then when we come to resolve object₂, we can straightaway eliminate object₁ from the candidate list using *Orthogonal Rule* resulting in resolution of object₂ to subject₁. The orthogonal rule was implemented by maintaining an *orthogonal list* for each RU and before admitting an entity in the candidate list it was checked against the orthogonal list. As a consequence of the orthogonal rule we can also define the following for reflexive pronouns:

- (Rule 3) **Reflexive Entity equal Rule** Reflexive pronouns in an event are equal to the subject entity.

The rule correctly describes the use of reflexive pronouns in object and indirect object positions in examples (4.2a) and (4.2b) as well as the illicit use in (4.2c).

- (4.2) (a) John hurt himself with knife.
(b) John compared Peter with himself.
(c) * John hurt himself with himself.

4.5.3.2 Pronouns

English language has a list of 33 pronouns listed in the first row of Table 4.7. Each of the pronouns have specific property values from the list gender, number and animacy. In addition to this, the pronouns also have attached syntactical roles. For example, the sentences “Him went to town” and “John gave he a pencil” are syntactically incorrect. The syntactical rules constrain a producer to use “he” in the former and “him” in the latter. Similarly, the pronoun “she” has to be used to co-refer to an entity whose gender is female. We can use these syntactical and morphological properties of pronouns as constraints to help resolve pronominal anaphora. The list of the properties and the particular values that were used in the implementation are specified in the rows of Table 4.7. The list of pronouns were divided into *N-Person*, *Reflexive* and *Possessive* categories. The first row in the same table provides a complete list of pronouns.

First person pronouns *I*, *me*, *myself*, *my*, *mine*, *we*, *us*, *our*, *ours*, *ourselves* and *ourselves* are used by the writer/speaker to refer to oneself, either individually, or a group to which the speaker/writer belongs. The pronouns *I*, *me*, *myself*, *my*, *mine* can only refer to the person speaking/writing the utterance. Resolution of these singular, first person pronouns is trivial in that, it can only refer to the speaker. When it comes to the plural first person pronouns (*we*, *us*, *our*, *ours*, *ourselves* and *ourselves*), they refer to the speaker belonging to a group which could have two or more undefined members. For instance in the clause “John said “I picked up Peter and we went to town””, the pronoun “we” refers to ‘John’ and ‘Peter’ which are mentioned in the clause itself. However in the clause “We shouldn’t have a system in New Zealand where you get one refugee and end up with several others”, the pronoun “We” refers to the group “New Zealanders” to which the utterer of the clause belongs to, but was never mentioned in the discourse. The antecedent of such pronouns is heavily based on semantics and context of the discourse, hence is a challenge for computational anaphora resolution systems. Our algorithm uses various some simple heuristics to resolve such anaphora. For example, looking for recently mentioned plural and animate nouns and existence of more than one animate entities in the same sentence as the mention of the first person plural

4.5 Hard Constraints Guiding Anaphora Usage

Category	List
<i>Complete list</i>	I,me,myself,my,mine,we,us,our,ours,ourselves,ourself, you,your,yours,yourself,yourselves. he,him,his,himself,she,her,hers,herself, they,them,their,theirs,themselves,it,its,itself,oneself.
<i>First-person</i>	I,me,myself,my,mine,we,us,our,ours,ourself,ourselves.
<i>Second-person</i>	you,your,yours,yourself,yourselves.
<i>Third-person</i>	he,him,his,himself,she,her,hers,herself, they,them,their,theirs,themselves,it,its,itself,oneself.
<i>Gender-male</i>	he,him,his,himself
<i>Gender-neutral</i>	I,me,myself,my,mine,we,us,our,ours.ourself,ourselves,you, your,yours,yourself,yourselves,it,its,itself, they,them,their,theirs,themselves
<i>Non-animate</i>	it,its,itself,oneself.
<i>Neutral-animate</i>	they,them,their,theirs,themselves,
<i>Singular</i>	I,me,myself,my,mine,you,your,yours,yourself,he him,his,himself,she,her,hers,herself,it,its,itself,oneself.
<i>Reflexive</i>	myself,ourself,ourselves,yourself,yourselves,himself, herself,itself,oneself,themselves,
<i>Possessive</i>	my,mine,our,ours,your,his,her,hers,its,their,theirs.

Table 4.7: Table showing the complete list of pronouns in the first row. The other rows show pronoun lists categorized according to commonly used property values for pronominal anaphora resolution.

4.5 Hard Constraints Guiding Anaphora Usage

noun. If no such entities can be found we leave the antecedent as *producer-member group*. Formulating this constraint as a rule gives us the *First Person pronoun equal Rule* (4) and *Third Person pronoun orthogonal Rule* (5) which enforces the associated constraint for third person pronouns.

(Rule 4) **First Person pronoun equal Rule** First person pronouns are equal to either the producer of the utterance or a group of which the producer belongs.

(Rule 5) **Third Person pronoun orthogonal Rule** Third person pronouns in a directly reported event are orthogonal to the producer.

The second person pronouns *you, your, yours, yourself* and *yourselves* can only be co-referential to the consumer, or consumers in the case of the plural pronoun “yourselves”. For a written discourse, there is usually no explicit consumer, hence in a clause such as “If you pay peanuts you get monkeys” the pronoun “you” refers to the person reading the sentence. It is also possible for the producer to explicitly specify the addressee of the clause such as in the clause “John said “Peter, you should go home now” or “John told Peter “you should go home now”. The addressee entity is captured in our representation at the preprocessing level and the entity is attached to the producer entity via an addressee relation. Rule (6) below enforces this constraint.

(Rule 6) **Second Person pronoun equal Rule** Second person pronouns are equal to addressee/s of the clause or the consumer of the discourse.

The largest and the most challenging category of pronouns in terms of anaphora resolution are the third person pronouns. Syntactical rules constrain specific pronouns into certain roles. The list of third person pronouns and their categories by role as used in our implementation is shown in Table 4.8.

Syntactical rules of English constrain us from using clauses like “John called he” and “him called John” because the pronoun is “he” can only be correctly used as a subject

4.5 Hard Constraints Guiding Anaphora Usage

Subject	object,i-object prep-object	possessive	reflexive
he	him	his	himself
she	her	her,hers	herself
they	them	their,theirs	themselves
it	it	its	itself,oneself

Table 4.8: Table showing the various syntactical roles that third person pronouns can be used for.

and similarly the pronoun “him” can only be used as an object, i-object or as a target object of a preposition. Different pronouns are provided for male, female, plural, and non-animate properties for the possible syntactical roles in a clause and only the appropriate one can be used for a grammatically correct clause. The only exception is that the pronoun “her” can be used both as a possessive and as an object/i-object/prep-object pronoun (hereafter referred to as object pronouns). Out of the list of third person pronouns we can eliminate the reflexive pronouns as we have covered their usage with the *Reflexive Entity equal Rule* (3).

The subject pronouns (*he, she, they, it*) can only function as the subject in an event hence can be parallel to entities only in one or more of the previous events. Hence for pronouns from this category, only the entities from the previous clauses were made available as candidates compared to the other pronouns for which entities from the same event were also eligible. This exclusion can be very useful for eliminating highly competitive candidates which increases chances for correct antecedent selection. The object and the possessive pronouns are the most challenging categories which were resolved using morphological compatibility and preference factors described in Section 4.6.

4.6 Soft Constraints Guiding Anaphora Usage

The rules discussed in Section 4.5 can be used to either trim a candidate list, in the case of the *orthogonal* rules, or resolve an anaphor in the case of *equal* rules. In this section we discuss the soft-constraints that were used as preferences. This point in the algorithm is only reached after a single complete pass through the document. At this point all NPs are either resolved, have been declared non-anaphoric or are in semi-resolved state. The preference application algorithm starts again from the beginning of the document with the first semi-resolved NP. Preferences are applied to the first semi-resolved NP which changes the Entropy because of this resolution. This triggers a forward pass up to the end of the document applying the constraints again to the remaining semi-resolved NPs. Then the preferences are applied to the next semi-resolved NP.

Previous research on the use of preferences have used surface factors such as recency, syntactical role and frequency of occurrences to implement higher level concepts such as focuss and coherence. In our implementation, we used these factors as well as additional ones originating from relation in compound nouns. We implemented preferences using scores based on anaphora factors, referred to as *salience factors* when applied using preferences. For each salience factor the candidate entity scores a certain number of points called *salience value*, which are totaled at the end to determine the candidate with the highest salience score as the antecedent. The absolute value of salience value used for each of the factors is immaterial. We are only interested in the relative values. The relative value gives us the relative suitability of candidates for a particular anaphor. Hence the maximum value for each of the factors was arbitrarily decided to be 10 and the salience value corresponding to each salience factor was normalized to be between 0 and a maximum value value of 10. Table 4.9 outlines the list of salience factors that were used in our implementation.

4.6 Soft Constraints Guiding Anaphora Usage

Salience factor	Classification	Description
Subject role	Discourse structure	A RU with any of its members in a subject role has proportionately higher salience score.
Topic score	Discourse structure	A RU representing entities in the first two paragraphs scores a higher salience score.
Recency	Discourse structure	A RU with any of its members being more recent has a higher salience score. Recency was computed in terms of the order of mention in the discourse.
Morphological compatibility	Morphological	A RU with known property compatibility has a higher salience score than one where it is unknown.
Name entity	Semantic-Similarity	A RU with a named entity as a member has a higher salience score.
Entity Frequency	Discourse structure	A RU which occurs more frequently has higher salience score. (Does not apply to members which are pronouns.)
Participating entity	Semantic-Relational	A RU which occurs as a head noun has a higher salience score than a RU represented by a modifier.
Related entity	Semantic-Relational	A RU corresponding to a modifier related by one of the eight c-relations has higher salience score.
HN Similarity score	Semantic-Similarity (HN)	The salience score is proportional to the SIMILARITY quotient. That is, a matching spelling RU has the highest score and it decreases proportionally with the quotient.
Modifier Similarity score	Semantic-Similarity (Modifier)	The salience score is proportional to the SIMILARITY quotient of one or more modifiers. That is, a matching spelling modifier has the highest score and it decreases proportionally with the quotient.

Table 4.9: Table showing the salience factors and their descriptions.

4.6.1 A Summary of the List of Constraints Used

Type	Classification	Description
=	Syntactical	A reflexive pronoun is equal to the subject RU.
=	Syntactical	A first person pronoun is equal to either the producer of the utterance or producer-member group.
=	Syntactical	A second person pronoun is equal to the either reader or addressee of the clause.
=	Semantic-Relational	A RU represented by a normalized NP is equal to the clause component as described in Table 4.5 .
⊥	Syntactical	A third person pronoun is orthogonal to the producer.
⊥	Syntactical	All participating RUs in an event are orthogonal excepting RUs represented by a reflexive .
⊥	Morphological	Any two RUs are orthogonal if any ONE of KNOWN gender, number and animacy are incompatible.
⊥	Semantic-Determiners	Any two NP RUs are orthogonal as determined by the semantics of the determiner/s.
	Semantic-Similarity	Any two NP RUs are parallel if they have the same spelling.
	Semantic-Similarity	Any two NP RUs are equal if they have a SIMILARITY quotient of more than 0.5.
	Semantic-Relational	A RU represented by a compound noun is parallel to a RU related to its modifier-RU (same spelling or SIMILARITY quotient > 0.5). The relation has to be one of the eight discussed relations.

Table 4.10: Table showing the summary of hard constraints used in the implementation.

As discussed in Section 4.5, we categorized the constraints in the following three categories:

- Equal (=) - two RUS are definitely co-referential.
- Orthogonal (⊥) - two RUs are definitely NOT co-referential.
- Parallel (||) - two RUs could be co-referential.

The rules were applied in the order, Equal, Orthogonal and then Parallel. The Equal rule was applied first because it is mostly applicable between entities in close proximity

and since the rules were applied in the order, most recent to the least recent, this rule manages to catch the applicable antecedent very early. Hence we do not have to expend any more effort applying the rest of the constraints. Similarly, the Orthogonal rules applied next eliminated a candidate without application of the numerous Parallel rules. The parallel rules apply a set of constraints to rule in a RU as a possible co-referent for the anaphor in question. Some of the rules have already been discussed from Rule (2) to (6) but are restated in Table 4.10 again for completeness.

4.7 Chapter Summary

This chapter described the implementation details of the anaphora resolution system implemented as part of this study. The described system makes use of a lot of the existing strategies for resolution, however also integrates the novel use of relations between nouns for resolving multi-word anaphora. Sections 4.1 and 4.2 described the upper level algorithm as well as the supporting modules. In Section 4.3 I described how the relations between nouns were mined to be used by the resolution algorithm. Finally Sections 4.5 and 4.6 described the implementation details of the sum of rules used by the algorithm, thus providing enough details for one to be able to replicate the algorithm.

The chapter as a whole, describes the implementation of an algorithm which is based on the new theoretical framework proposed in this thesis. It provides an advancement on the previous strategies by providing a technique for integrating semantics for resolving anaphora. The next chapter describes the evaluation of the described algorithm.

Chapter 5

Input Data and Experimental Results

This chapter reports the results from the implementation of the algorithm specified in Chapter 4. In Section 5.1, I describe the test data and in Section 5.2 I present the overall results. Section 5.2.1 presents a discussion of the precautions in interpreting statistics from direct comparison of anaphora resolution systems. I then do a comparison of the relevant statistics from our results with a list of major reported anaphora resolution systems. In Section 5.2.2 I present a breakdown of the results and analyze the proportional contribution of the types of rules used in the resolution. In particular, I present an analysis of the proportion of anaphora that were resolved using rules originating from the integration of semantics proposed in this study. In the last section of this chapter I examine a selection of case studies from the corpora in order to illustrate the strengths of the proposed algorithm as well as some of its limitations. To do this, I firstly use a complete article to illustrate how the proposed algorithm resolved multiple NPs consisting of pronouns, simple nouns and compound nouns to refer to the correct instance of the entity. This example is also used to illustrate how an anaphor can be correctly resolved at a local level but may not necessarily be resolved correctly at a discourse level. This is one of the factors which complicates the comparison of results across

Statistic	Train Data	Test Data
No. of articles	20	15
No. of Sentences	352	273
No. of paragraphs	246	204
No. of clauses	1302	860
No. of nouns	2323	1895
No. of pronouns	598	414
No. of anaphoric NPs	915	723

Table 5.1: Table showing the raw data statistics for the articles that were used for this research.

different studies. In the latter part of the last section I examine a selection of short cases from the corpora which would have been difficult to resolve using the existing systems, but were correctly resolved using the framework proposed in this thesis.

5.1 Input Data

The input data used for content analysis and for all aspects of NP usage consisted of 20 articles (of mixed genre) from *The New Zealand Herald*, *The Dominion Post* and *The Press* which are three major online newspapers from three different cities in New Zealand. We extracted a further 15 articles to be used purely as test data. The choice of the articles was not completely random. The corpora was developed using articles which were not too short (had more than 15 sentences), exhibited the use of a variety of anaphoric uses (including pronominal anaphora) and had been written by different writers. The statistical details of the whole corpora is shown in Table 5.1.

5.2 Results

Category	Training data						Test data					
	P	R	F	P*	R*	F*	P	R	F	P*	R*	F*
Pronouns	0.88	0.84	0.86	0.82	0.77	0.79	0.86	0.83	0.84	0.81	0.78	0.79
NP anaphora	0.81	0.72	0.76	0.78	0.63	0.70	0.79	0.69	0.74	0.75	0.63	0.68
Overall	0.85	0.78	0.81	0.80	0.70	0.75	0.83	0.76	0.79	0.78	0.71	0.74

Table 5.2: Table showing the overall performance of the implemented system on training and test data. P = Precision, R = Recall and F = F-measure

5.2 Results

Table 5.2 shows the overall results obtained from the anaphora resolver. Each result set consists of precision, recall and F-measure (Makhoul et al., 1999) columns where the F-measure was calculated using the following equation:

$$\text{F-measure} = \frac{2 \cdot P \cdot R}{P + R}$$

The precision value gives the resolution rate for targeted anaphora while the recall value gives the rate relative to all identified anaphora in the discourse. The F-value combines the precision and recall values into a single figure, giving us a single tool to compare systems for both Precision and Recall. We will generally use F-measure to make observations instead of individual recall and precision values.

The columns with asterisks (*) show results for *discourse wide* resolution, compared to *first level* or *local* resolution in the other column. First level resolution refers to an anaphor being correctly resolved to the correct entity in the first instance, at a local level. However with our multi-pass approach, it is possible that the resolved anaphor can later get resolved to another, incorrect antecedent in a subsequent pass of the document. For example, a pronoun “he” was correctly resolved to the antecedent “driver”, however in a later pass, “driver was incorrectly resolved to “Freightlines driver” instead of “Southbound driver”. This was counted as a single correct resolution represented by the column without asterisks. If the whole chain is correctly resolved then it was

counted as discourse wide correct resolution represented by the column with asterisks. The F^* -values are slightly lower than the F -values in all the cases indicating that there were cases where anaphora was resolved correctly at a local level but they were resolved incorrectly at the discourse level. Accurate interpretation of a discourse requires resolution of anaphora both at local as well as discourse level, which is indicated by the F^* value columns in Table 5.2. The values for pronominal anaphora is higher than that of NP anaphora as expected. In addition there is a much smaller difference (0.02 and 0.01*) in the F -values between training and test data for pronouns indicating that the rules are better generalizable for pronouns compared to NP anaphora. For NP anaphora, the difference in the training and test F -values is much bigger (0.04 and 0.05*) indicating a lower generalization ability. Some of this was directly attributed to gaps in basic knowledge such as morphological values, while others were due to either extraction of incorrect relations or inability to extract required relational knowledge. Some illustrative examples are discussed later in this chapter. The individual results for pronominal and NP anaphora compare well with other reported results which are discussed in the next subsection.

5.2.1 Some Comparisons

Most studies on anaphora focus on specific areas or types of anaphora which makes it difficult to compare results across studies. This difficulty is exacerbated by the fact that different systems use varying levels of preprocessing hence the availability of preprocessed data becomes a prerequisite adding another error factor. Additionally, the size of data used for testing also varies, hence we can not reliably compare the results of a small data size with a large one. In spite of these factors we can still get a ballpark idea about the performance of a system by looking at results from other similar systems. To do this we need to look at the results for pronouns and NP anaphora separately since there are few reported results which resolve both types of anaphora. Table 5.3 shows a summary of pertinent factors as well as the reported results for pronominal

anaphora. Putting aside all the other factors, the reported precision rates range from 75% for Kennedy and Boguraev to 90% for Baldwin. This compares with the precision rate of 81% on the test data and 88% on the training data obtained by our system. We cannot compare the asterisks value (69%) since none of the systems evaluated indicate any form of discourse wide resolution. Our system was also tested on a larger sample size of 414 pronouns (except Kennedy and Boguraev (1996)) compared to most of the studies in Table 5.3.

A study that resolves both pronominal and NP anaphora is reported in Palomar et al. (2001). This paper reports a precision value of 81.3% for pronominal anaphora and an impressive 81.5% (compared to 76%) for adjectival (NP) anaphora. These results are based on dialogue based discourses in Spanish. The system requires manual annotation of the dialogue structure, nonetheless, the results for NP anaphora are impressive compared to a precision value of 81% for training and 79% for testing data from our implementation. NP anaphora rates reported in most studies lie between a lower value of 40% to an upper bound of 80% (e.g. Markert and Nissim, 2003; Poesio et al., 2004a; University et al., 2003; Vieira and Poesio, 2000). An implementation for NP anaphora that is also based on rules, uses WordNet and solved a range of anaphoric types is reported in Poesio et al. (2000). This study (reviewed in detail in Chapter 2) reports an overall precision rate of 76% however the recall rate is relatively low at 53% resulting in a F-value of 62%. Although the precision for our system is only slightly higher at 81%, the recall is relatively high at 72% (compared to 53%) giving us a F-value of 76% which is significantly higher than that in Vieira and Poesio (2000).

The systems that participated in the MUC-6 (van Deemter and Kibble, 1995) co-reference task competition achieved recall values ranging from 35.69% to 62.78% and precision values ranging from 44.23% to 71.88%. This gives us F-values ranging from 40% to 67% (compared to 74% for our system). The resolutions rates achieved by the MUC-6 systems is a better comparison with our results because the co-reference task has several similarities with our implementation.

Factor	Lappin and Leass (1994) - RAP	Kennedy and Boguraev (1996) - CogNIAC	Baldwin (1997)	Mitkov (1998) - MARS	Ge et al. (1998)
Genre of strategy	Mixed	Non-learning	Non-learning	Non-learning	Learning
Recall rate	–	–	60%	–	–
Precision rate	85%	75%	90%	89.7%	82.9%
Genre of data	Technical manual	Newspaper	Narratives and Newspaper	Technical manual	Journal
Input corpus	Computer Manuals	Randomly from newspapers	Narrative text about 2 persons	Manuals for VCR	Wall Street Journal
Sample size	5 Manuals, 82,000 words	27 texts	15 articles, 200 pronouns	140 pages, 71 pro-nouns, 48 consid-ered	3975 sentences, 2477 pronouns, 1371 considered
Amount of knowledge used	Morphological info supplied	Morphological info from LING-SOFT	Morphological info supplied	Morphological info supplied	Morphological info supplied and mined
Pre-processing	McCords's Grammar parser	Slot tagger (LING-SOFT)	Part-Of-Speech tagger	Part-Of-Speech tagger	None

Table 5.3: Table showing the pronominal anaphora resolution rates and some pertinent factors used in five implemented systems.

Heuristic type	Hard-constraint	Soft -Constraint	Total	Proportion
Morphological	586	372	958	0.24
Syntactical	689	0	689	0.17
Semantic Determiner	365	0	365	0.09
Semantic Similarity	459	197	656	0.16
Semantic Relational	589	372	961	0.24
Discourse structure	0	358	358	0.09
Totals	2688	1299	3987	1.0

Table 5.4: Table showing the breakdown of heuristics used in the resolution process.

Firstly the systems are required to resolve all, and only, identity anaphora. This can be in the form of pronouns and NPs including non-definite NPs. Secondly, the systems are required to consider the co-reference chain as whole rather than only one correct antecedent. This is similar to our representation of the RU data structure. Hence, this compares better with our results in the asterisks column. Our implementation achieves an overall precision value of 78% and a F-value of 74%. These values are slightly better compared to the MUC- systems (71% and 67%), however, it is important to note that the systems competing in the MUC competitions are required to parse the text before processing them. This introduces errors including some sentences that cannot be parsed. All such errors were manually corrected for the results reported in this thesis.

The next section gives various breakdown of results in order to present some insights into the overall results obtained.

5.2.2 Result Breakdowns

Table 5.4 shows the proportional breakdown of the various types of constraints discussed in Chapter 4. The categories correspond to the hard and soft constraints detailed in Tables 4.10 and 4.9 respectively. The data shows how often the rules of a particular category were used for any purpose. This includes both correct/incorrect resolutions as well as inclusion/exclusion of an entity from the candidate list. There were no

5.2 Results

	Training data			Test data			Totals	Pro- propotion
	Correct	Correct*	Not Correct	Correct	Correct*	Not Correct		
Normalization (Tiebreaker)	38	31	16	28	27	12	152	0.19
Normalization (Assisted)	21	20	6	19	17	6	89	0.11
Relations (Tiebreaker)	64	53	22	48	43	14	244	0.30
Relations (Assisted)	79	72	29	62	55	21	318	0.40
Totals	202	176	73	157	142	53	803	–
Proportion	0.25	0.22	0.09	0.20	0.18	0.07	–	–

Table 5.5: Table showing the breakdown of resolutions that used any form of relational knowledge.

soft-constraints of types Syntactical and Semantic-Determiner and no hard constraints based on the Discourse structure, hence the corresponding zero values. The last column in Table 5.4 shows the proportions for each of the heuristic types. The highest proportion of heuristic types were Morphological and Semantic-Relational at 24% followed by Syntactical at 17% and Semantic-Similarity at 16%. Use of Discourse Structure and Semantic Determiner were below 10%. The overall results firstly indicate that anaphora resolution requires an integrated approach using a variety of constraints from different aspects of language usage. Secondly, the results significantly affirm the use of relational knowledge in anaphora usage and hence its resolution. Out of all the types of constraints, semantic constraints (of types Determiner, Similarity, and Relational) were used 49% of the time and approximately half of these were based on relational knowledge. Since use of this relational knowledge is the core part of this thesis I next look at a breakdown of the results that contributed to its use in the 961 or 24% of the cases.

Table 5.5 shows the breakdown of resolution results according to the role played by relational knowledge. Again the Correct and Correct* respectively represents local and discourse wide correct resolutions. The Tiebreaker and Assisted rows respectively

represents cases for which relational knowledge was critical in selection of the antecedent and cases in which it assisted in the resolution. Combining training and test resolutions, relational knowledge was considered to be the tiebreaker for correct resolutions in 58 (31 + 27) cases for normalization relations and 96 (53 + 43) for the other relations. This makes 154 out of 803 (19.2%) relation based resolutions which were made critically based on relational knowledge. A tiebreaker resolution was classified to be one in which a candidate was considered to be the antecedent either due to a hard constraint because of a relation or when the preference score from a relation was the difference between two highest scoring candidates, resulting in the choice of an antecedent. Apart from these cases relational knowledge was also used to either rule an entity into a candidate list or rule it out, which was later resolved by some other non-relation based constraint/s. These cases were classified as “Assisted” shown in rows 2 and 4 in Table 5.5. A total of 92 (20 + 72) were found in the training data and 72 (17 + 55) were found in the test data making a total of 164 out of 803 (20.4%) correct resolutions that used some form of assistance from relational knowledge.

Before discussing case examples that were correctly resolved using relational knowledge I will demonstrate the resolution of a set of entities in a detailed step-by-step manner which will illustrate the overall workings of the algorithm in addition to the use of relational knowledge. Consider the article in Figure 5.1. The various mentions of the entity “truck” are highlighted. In the following discussion I will only mention the significant entities for clarity. After processing sentence 2, the following RUs were formed:

ID#:	RU1
Members:	Both trucks _[1,1]
Relations:	

5.2 Results

- [1] **Both trucks**_[1,1] involved in yesterday's crash were gutted by the inferno which erupted seconds after Ted Collins saved one of the drivers.
- [2] A dairy worker pulled an injured truck driver from the wreckage of his cab moments before **a second big rig laden**_[2,1] with 33,000 litres of diesel erupted in flames.
- [3] Explosions rocked the surrounding area shortly after the fire broke out, and a plume of thick black smoke drifted over Hamilton.
- [4] Fonterra driver Ted Collins was following **the northbound Freightlines truck**_[4,1] on State Highway 1, and saw **it**_[4,2] collided with **the Shell diesel tanker**_[4,3] at Horotiu, north of Hamilton.
- [5] **"The fuel tanker**_[5,1] heading south got into trouble and jack-knifed across the road and on to **its**_[5,2] side. The guy in front of me didn't have much time to do anything and didn't quite stop in time."
- [6] **The Freightlines truck**_[6,1], driven by a 51-year-old contractor, ploughed into **the diesel tanker**_[6,2], Mr Collins said.
- [7] "With 40 tonnes you've got a bit of momentum, and he had almost stopped when he hit **it**_[7,1]."
- [8] Mr Collins leaped from his cab and ran to the aid of his fellow truckies.
- [9] "I did what I could to save somebody, it's a natural instinct."
- [10] Senior Hamilton fire station officer Daryl Trim said Mr Collins' actions were heroic.
- [11] "It's fantastic. If he hadn't done that, it may have had some dire consequences. It's pretty heroic really."
- [12] Mr Collins said: "I didn't really think about it, I just knew one of my truck friends could be dead or dying. **It**_[12,1] was **a fuel tanker**_[12,2] and I thought I've just got to get there quickly."
- [13] Mr Collins approached the Freightlines truck's cab as the dazed man was getting out of his seat, struggling to get away.
- [14] "The other truck driver I could not get to - **his truck**_[14,1] was ablaze and **it**_[14,2] was blocking the road, there was no way I could get to him.
- [15] "That was the thing that was worrying me. I thought the other guy had gone, it was a hell of a blaze."
- [16] Mr Collins helped the injured man to the side of the road, away from the fire.
- [17] "He was slightly out of it, a bit hurt. We sat down for a few minutes then some big bangs hit. I could see fuel going into the drains at the edge of the road and thought it [the fire] might get to us, so we had to move along another 50 yards or so."
- [18] The 28-year-old tanker driver escaped from his cab on the other side of the blaze with minor injuries.
- [19] Both drivers were taken to Waikato Hospital. The tanker driver was discharged, but the Freightlines driver was admitted with moderate injuries.
- [20] Those battling the blaze faced temperatures of up to 1000C. The fire gutted **both trucks**_[20,1].
- [21] Gavin Crook, who lives metres north of where the crash occurred, said he heard a loud bang, followed by the sound of cars slowing down.
- [22] "It was like a bomb, and a big amount of smoke and flames were coming out. You could feel the ground shaking."
- [23] The accident closed SH1 between Hamilton and Ngaruawahia.
- [24] Transit spokesman Chris Allen said the highway surface was badly damaged, and a 70m stretch had to be rebuilt and resealed during the day at a cost of about \$20,000.

Figure 5.1: Trucks Article.

ID#:	RU2
Members:	a second big rig _[2,1]
Relations:	INSTANCE(Both trucks _[1,1]) ON(diesel)

The INSTANCE relation in the second RU was formed from the synonymous relation between “rig” and “truck”. The ON relation was formed from the surface form “laden with”. When the processing reaches the NP “the northbound Freightlines truck_[4,1]” it would have been resolved to RU2 (rig = truck) if the relation INSTANCE(Both trucks) was not present in RU2. This relation tells us that the NP “the northbound Freightlines truck_[4,1]” could be a member of RU2, however at this point we are not sure. Hence instead of resolution, this results in a third RU being formed. This RU together with the RU formed for the pronoun it_[4,2] with a search space up to the previous sentence is shown below: (note PARALLEL relation is equivalent to a candidate)

ID#:	RU3
Members:	the northbound Freightlines truck _[4,1]
Relations:	INSTANCE(Both trucks _[1,1]) PARALLEL(RU2)

ID#:	RU4
Members:	it _[4,2]
Relations:	PARALLEL(RU3, Hamilton, smoke, fire, area, Explosions ORTHOG(the Shell diesel tanker _[4,3])

The next anaphor under consideration, “the Shell diesel tanker_[4,3]”, has RU4 and RU2 as candidates on the basis of synonymous relation between “tanker” to both “truck” and “rig”. However, the modifier “diesel” in the “Shell diesel tanker_[4,3]” and the existence of ON relation to the entity “diesel” in RU2 is used as the relational knowledge to resolve it to RU2 instead of RU3. Note that at this point RU3 is still in a semi-resolved

state with RU2 as the potential candidate. Moving further, the anaphor “fuel tanker_[5,1]” gets resolved to RU2 (and not RU3) because of the ON(diesel) relation, where fuel has a synonymous relation to diesel extractable from WordNet.

Another crucial use of relational knowledge is illustrated by the orthogonal relation to “Shell diesel tanker_[4,3]” in RU4. In a subsequent pass of the document RU4 gets resolved to RU3 using preferences. This results in the orthogonal relation ORTHOG(Shell diesel tanker_[4,3]) merged into RU3. Merging of orthogonal relations was treated as a special case compared to merging of other relations. In this merge, a check is made to remove any entities from parallel list of the target RU that also exists in the orthogonal list of the source RU. This resulted in RU2 being removed from the parallel list of RU3 establishing that all members of RU3 are orthogonal to RU2. Further, the NPs “the diesel tanker_[6,2]” and “a fuel tanker_[12,2]” get resolved to RU2 and “The Freight-lines truck_[6,1]” to RU3 using synonymous relations. The NP “it_[7,1]” was incorrectly resolved to “momentum” in the same sentence. NP “it_[12,1]” was correctly resolved to RU2 because copula constructions were taken to represent co-reference between the subject and the object.

The NP “his truck_[14,1]” illustrates local and discourse wide resolution corresponding to Correct and Correct* columns in Table 5.5. The pronoun “his” in “his truck_[14,1]” was correctly resolved to “other truck driver” in the same sentence, however “other truck driver” was incorrectly resolved to “one of my truck friends” in sentence 12. This resolution would increment the count in Correct column but not in Correct* column in Table 5.5. The NP “his truck_[14,1]” as a whole was incorrectly resolved to RU3 because of the relation HAVE between RU3 and “other truck driver”. Similarly “it_[14,2]” was correctly resolved locally to “his truck_[14,1]”, however it was incorrectly resolved to RU3 at discourse level. Finally “both trucks_[20,1]” was correctly resolved to RU1, resulting the following set of RUs for the truck entities.

ID#:	RU1
Members:	Both trucks _[1,1] , both trucks _[20,1]
Relations:	

ID#:	RU2
Members:	a second big rig _[2,1] , the Shell diesel tanker _[4,3] , The fuel tanker _[5,1] , its _[5,2] , the diesel tanker _[6,2] a fuel tanker _[12,2] , It _[12,1]
Relations:	INSTANCE(Both trucks _[1,1]) ON(diesel)

ID#:	RU3
Members:	the northbound Freightlines truck _[4,1] , it _[4,2] , The Freightlines truck _[6,1] , <u>his truck</u> _[14,1] , <u>it</u> _[14,2]
Relations:	INSTANCE(Both trucks _[1,1]) ORTHOG(the Shell diesel tanker _[4,3] , HAVE(his)

Use of relational knowledge between entities becomes challenging when we start dealing with cardinalities represented by the INSTANCE relation. It is possible for more than one type of relation from the set CAUSE, HAVE, MAKE USE, IN, FOR, FROM ABOUT AND INSTANCE to be present between a pair of entities. For example, it is possible to have, USE, IN, FROM, HAVE and FOR relations between the entities “truck” and “diesel”. Hence, if more than one relation can be established in a single discourse, than we also need to determine if the relations are with the same instance of the entity. Consider the following example to illustrate this, as well as, the use of normalization for resolution.

The owner of a 4WD vehicle seized in connection with the Birgit Brauer murder case says it was stolen by his employee six weeks ago. The

*Toyota Hilux is registered to Palmerston North man Brent Cleverley, who confirmed to the Herald yesterday that **his vehicle** had been taken by a **man** who worked for **him**. The Herald understands **the worker**, **who**[the worker] used **the vehicle** in his job cutting firewood, failed to turn up to work one day last month.*

In the absence of gender information for the name “Birgit Brauer” the pronoun “his” in the first sentence was unable to be fully resolved, with “owner” and “Birgit Brauer” remaining as candidates. However, the pronoun “his” in the NP “his vehicle” (second sentence) was able to be resolved to “Palmerston North man Brent Cleverley”. In addition, the relation HAVE(owner, 4WD vehicle) was established from the verb “own” as an indicator of a HAVE relation. For the purpose of processing cardinalates we used the policy, “consider the same instance unless there are determiners suggesting otherwise”. This resulted in “his vehicle” in the second sentence resolved to “4WD vehicle” in the first sentence. Merging of the relational knowledge due to this resolution results in establishing that the gender property of “The owner” in the first sentence is male. Hence, the pronoun “his” in “his employee” in the first sentence gets resolved to “The owner” instead of “Birgit Brauer”. When the third sentence gets processed, the NP “the worker” is immediately resolved to “the man” from the normalization process. The phrase “who used the vehicle” in the third sentence, (where “who” was substituted with “worker” at the preprocessing level) results in the USE(worker, vehicle). Since there are no determiners such as “other”, “second” etc, the noun “vehicle” in the third sentence was resolved to the same instance as the first and second sentences. This results in the relations HAVE(owner, vehicle) and USE(worker, vehicle) where the entity “vehicle” is same and “owner” and “worker” are different. It is also possible for “vehicle” in the two relations to represent two different instances in which case relational knowledge falls short, hence additional knowledge becomes necessary to distinguish between the instances.

Use of the ORTHOGONAL relation also played a significant role in the implemented algorithm. In the excerpt just discussed, the phrase “a man who [man] who walked

for him”, establishes the *ORTHOGONAL*(man,him) which rules out the entity “man” from the candidate list of the anaphor “him”. This resulted in “him” getting correctly resolved to “Palmerston North man Brent Cleverley” rather than the possibility of getting incorrectly resolved to “man”. Since the orthogonal relation is more deterministic from the surface structure compared to the other relation types, the constraints arising from orthogonality is “harder” than those from other relation types. This helped eliminate false positives in some instances, as in the following example:

A player from Varsity Legends, part of Auckland University Rugby Football Club, ...

A Varsity Legends official said the player, who did not want to be named, woke up to ...

From the first sentence the relation *FROM*(Varsity Legends, A player) was established which licences a producer to form the NP “Varsity Legends player”. The NP “A Varsity Legends official” in the second sentence has a high probability to be resolved to “A player” based on the relational knowledge. However, the orthogonal relation between “A Varsity Legends official” and “the player” in the second sentence correctly eliminates “A player” from the candidate list, preventing an incorrect resolution.

Table 5.6 shows the corresponding results for incorrect resolutions split by the type of role played by relational knowledge for both training and test data. The rows in Table 5.6 give the number of cases that were incorrectly resolved either because of rule exception or missing knowledge. Rule exception includes cases in which a relation was correctly derived but the antecedent identified due to the relation was incorrect. The second case of missing knowledge includes false positive cases as well as the inability to extract relations which were outside the scope of the surface structure based mining techniques used. The table firstly shows that a much smaller percentage (approximately 30%) of incorrect resolutions were due to normalization, most of these

5.2 Results

	Training data		Test data	
	No. Incorrect	Proportion Incorrect	No. Incorrect	Proportion Incorrect
Normalization rule exception	4	0.05	3	0.06
Normalization Knowledge missing	18	0.25	15	0.28
Relations rule exception	22	0.30	13	0.25
Relations knowledge missing	29	0.40	22	0.42
Totals	73	1	53	1

Table 5.6: Table showing the breakdown of incorrect resolutions due to relational knowledge.

were due to the inability to extract the relation. These were cases in which a verb entailment was used to derive a normalized verb, for example, the verb “produce” was normalized to “grower” from the equivalent verb “grow”. The small number of rule exceptions for normalization represent cases in which the suffix based matching was incorrect, for example, the normalization rule resolves words ending with the suffix “ent” to the Agent, however the normalized noun *assessment* from the phrase “The fire department assessed the damage . . .” co-refers to the Clause.

A much larger number of incorrect resolutions were due to non-normalization relational knowledge, however these relations were more applicable, hence apply to more NPs resulting in a larger proportion of both correct and incorrect resolutions. A total of 51 (from both training and test data) were due to the system’s inability to extract relations from the surface structure while 35 cases were due to exceptions in the application of the rule based on a derived relation, or false negative. That is, derivation of a non-existent relation. Some of these examples are discussed below. An example that was resolved incorrectly even though the corresponding relational knowledge was able to be derived from the surface structure is illustrated in the following excerpt.

*Even his parents, **who [parents] live in Hong Kong**, thought it was legitimate and agreed to lend him the money.*

⋮

...and his mother told him that a fortnight ago **Hong Kong officers** arrested a few people thought to be linked to the scam.

The first sentence establishes the relation IN(parents,Hong Kong) which licences a producer to form the NP “Hong Kong Parents”. When processing the NP “Hong Kong officers”, the noun “parents” was pulled into the candidate list due to the IN relation and a SIMILARITY score of more than 0.3 between “parents” and “officers”. The final selection as the antecedent was done by the use of salience score which was boosted because of the number and animacy compatibility. A higher SIMILARITY threshold was able to distinguish between “officer” and “parent”, however it deteriorated the overall performance and was not able to resolve legitimate anaphoric uses such as driver/truckie and tanker/rig. This is an example that illustrates a correctly derived relation giving rise to an incorrect resolution. Such cases were counted as rule exceptions in the results in Table 5.6. Another example of rule exception is when anaphoric compound nouns were formed from relations which were outside the set proposed in this thesis. As an illustration consider the following excerpt:

...when she was confronted by a **second person brandishing a machete**

•

⋮

...the **machete man** took off when ...

In the excerpt above the NP “machete man” was formed on the basis of a relation formed between “man” and “machete” by the verb “brandishing”. The verb “brandishing” was interpreted to be not equivalent to any of the relations from CAUSE, HAVE, MAKE, USE, IN, FOR, FROM, ABOUT and INSTANCE, hence was determined to be outside the scope for forming a compound noun. This led to “second person” not being identified as the antecedent. In another example the phrase “*man wearing a hooded*

sweatshirt ...” followed by *... camera showed a hooded person ...*” was found in the test data. The verb “wearing” can be interpreted to represent an ON relation since *A wear B* implies *B ON A*. Although the verb “wear” is now added to the templates for mining the ON relation, the test data results reported in this thesis did not include patterns arising from analysis of the test data.

The excerpt below illustrates an example that could be resolved with the proposed relational framework, however, due to the knowledge gap in the lexicon, relational knowledge was not able to be used.

*Police searching for **a missing 5-year-old girl** are investigating whether the woman who reported her gone was at home at the time the girl vanished, an officer working on the case said Wednesday.*

:

Croslin, 17, is the girlfriend of Haleigh’s father, Ronald Cummings.

*The police officer rebuffed a question about whether Croslin had changed her story about where **she** was when **the girl disappeared**, though he called the tip “a very important one.”*

The anaphor represented by the NP “the girl” in the second paragraph co-refers to “a missing 5-year-old girl” in the first paragraph. There is an occurrence of another female in the intervening paragraphs, hence the presence of an equally competing candidate. The clue to resolving the anaphor to the correct antecedent lies in interpreting that the missing girl has to be the same as the girl who disappeared. With our framework, the process of normalization would interpret “the girl” as the agent of the verb “disappear” which is an alternative to the verb “missing” used as the modifier of the antecedent. The bridging knowledge between the verbs “miss” and “disappear” is not encoded in WordNet, hence this knowledge was not used in the resolution. However, the anaphor was correctly resolved due to the orthogonal relation between “she” and “the girl” in the last paragraph. This highlights the effectiveness of a multi-strategy approach to approach to anaphora resolution.

The relational knowledge used for resolving anaphora were derived from the relations CAUSE, HAVE, MAKE, USE, IN, FOR, FROM ABOUT AND INSTANCE. Out of these, the first eight are relations used in the generation of compound nouns while the last one, the INSTANCE relation, was used to represent the group-membership relation. The BE or *is-a* relation, which is described in Levi (1978) as a relation which can also be used to form compound nouns, is a special relation which itself represents an anaphoric relation. That is, instead of identifying and using the relation to help resolve anaphora, the process of identifying the relation itself is resolution of anaphora. A simple example is when the noun “car” is co-referenced with the noun “vehicle”. In this case the process of identifying the BE relation between “car” and “vehicle” is effectively resolving anaphora. When the relation is used for generating a compound noun, the resulting compound noun will have the modifier and the head noun in a BE relation. Hence, we can have the modifier (or its morpheme) used as a co-reference to the head noun. Two examples from the corpus exhibiting this characteristic are:

- **4WD vehicle** co-referred with the anaphor **4WD**.
- **16-month-old baby girl** co-referred with the anaphor **baby**.

In both the examples there is a BE relation between the modifier and the head noun which can be respectively expressed as “4WD” *is-a* “vehicle” and “baby” *is-a* “girl”. This special relation enables one to use just the modifier on its own as a co-reference to refer to the head noun. Note that the head noun is a noun with a completely different dictionary meaning. Any of the other compound noun generation relations, other than the BE relation, can not be used in a similar fashion. For example, in the discourse in Figure 5.1, it has the NPs “Freightlines cab” and “Freightlines truck” which have a HAVE relation between the nouns “Freightlines” and “cab”/“truck” respectively. A subsequent use of the anaphor “Freightlines” co-refers to the actual modifier noun “Freightlines” rather than the head noun “truck” or “cab” as in the case of the previous BE relations. Hence, even though the BE relation functions the same as the other eight

relations in generation of compound nouns, it has to be treated differently when applied to anaphora usage and its resolution.

5.3 Some Identified Shortfalls of this Research

Collection of “search data” for the knowledge mined from WordNet would have given us better insights on its suitability for anaphora resolution. Knowledge was mined from various synsets such as *direct-hypernym* and *inherited-hypernym* by traversing different distances into each of the synset paths. For this study we did not record the synset path or the path distance which was used for mining the required knowledge. Recording and analysis of this data, although time intensive, would have given us invaluable information which could have been used to fine-tune the mining process as well as shed light on the organization of information in a lexicon such as WordNet.

Approximately 70% of the incorrect resolutions in the test data were identified to be due to knowledge gaps and this could have been potentially filled in by using a supplementary knowledge base. Mining the web for additional knowledge was an option however due to time constraints it was not explored. However this is a planned extension and will be explored next.

Lastly, we used very basic word sense disambiguation for simple NPs or NPs with same modifiers. Our implementation of word sense disambiguation was, either the one matching one of the gloss definitions in WordNet, or the most common one. The most common one is defined as the first synset in WordNet. In any NP interpretation framework word sense disambiguation is one of the cornerstones hence needs a more detailed implementation than the extent adopted for this study. This is also planned to be implemented as a future improvement of this project.

5.4 Chapter Summary

The first part of this chapter reported the results obtained from the implementation of the rules derived from relational knowledge as well as other established rules. I made evaluations with several works, however noted that direct comparisons can not be done due to several variances in individual studies. The MUC-6 co-reference task was noted to be the closest to the work in this thesis and our precision value of 78% was a good comparison with the highest value of 71% for MUC-6. In the latter part of the chapter I provided a breakdown of the results showing data for correct and incorrect resolutions for the different categories of rules used in the system followed by some case studies to illustrate the application of the algorithm as well as some examples of correct and incorrect resolutions due to specific rules. The chapter concluded by outlining some of the shortfalls of this research which were either identified hindsight or were taken as acceptable due to resource/time constraints.

Chapter 6

Conclusions and Future Work

6.1 The Contributions of this Research

The work presented in this thesis contributes towards the advancement of research in the NLP discipline on two fronts.

Firstly, it enhances our theoretical understanding of a particular aspect of natural language usage, namely, the use of anaphora. We observed that a producer always attempts to communicate the goal of a discourse in the most concise way possible, hence uses whatever shortcuts are possible in order to reduce the number of words. Using multiple nouns together as a compound noun has already been known to be a shortcut way in which one omits the relation between the words in order to make it shorter. We proposed that the use of anaphora is another similar phenomenon in which the anaphor-NP is used in a shortcut way to elaborate on the antecedent entity, in addition to identifying it. This theory on anaphora thus unifies the phenomena of anaphora usage and generation of compound nouns. An immediate consequence of this is that the two processes should be guided by the same underlying relational framework. We

tested this by using a framework that uses compound noun generation relations in an anaphora resolution algorithm. The results obtained were very encouraging and compared well when evaluated against other similar systems. However, the system has scope for an even better resolution rate with integration of additional knowledge sources and better relation extraction techniques.

In terms of implementation, we have demonstrated an anaphora resolution algorithm which integrates a wide range of existing resolution strategies with the use of a well defined knowledge structure which can be extracted from domain independent lexicons and corpora. The knowledge structure is a step forward for anaphora researchers for two reasons. Firstly, it provides a framework for integrating knowledge in anaphora resolution systems. Secondly, since the same knowledge structure also underpins compound noun generation, advances in knowledge extraction techniques can be “borrowed” by researchers in either of the fields.

Apart from resolved anaphora, a consequential benefit of using our knowledge framework for resolving anaphora is that the algorithm results in a rich knowledge structure which can be useful for higher level NLP applications such as document summarization and visualization. As part of the anaphora resolution process, the relational knowledge between entities are merged and stored in Reference Units (RUs), which effectively contain knowledge about the pertinent relations of an entity with all other entities in the discourse. This can be used to compute the relative significance of entities in a discourse for the purpose of higher level discourse structures such as discourse focus or for direct applications such as document visualization.

6.2 Future Works

In this thesis we have identified a specific form of knowledge, namely relational knowledge, that can be used for resolving anaphora. We have also demonstrated ways of

extracting this knowledge from an existing lexicon, WordNet. However a large percentage (70%) of the incorrectly resolved anaphora were identified to be due to knowledge gaps or errors in the extracted knowledge. Hence, an immediate future work is to improve this by extending the search for knowledge beyond WordNet. An attractive option for this is to use The Web. Deployment of *Google Web Toolkit* is an example of momentum towards mining knowledge from The Web. Ways of extracting the required relational knowledge using this toolkit can be investigated in order to further improve the performance of the proposed framework.

The second aspect that requires further investigation is the validity of the proposed framework for discourses in other genres and even across different languages. Some of the basic syntactical factors such as morphological compatibility have been shown to be across genres, however, the extent to which the other factors hold true across genres is still largely unknown. The genre specific factors in Mitkov (1998) used on technical manuals produced an impressive success rate of 90%, however, some of these factors had low validity in the corpus of newspaper articles used in this study. It would be interesting to investigate the extent to which the relational framework from this thesis is valid across other genres such as dialogues and longer discourses, and if it is also valid for other languages. A planned extension to this study is to firstly test the system on data from other genres in English and later to investigate if the relational knowledge can be used to resolve anaphora in other languages.

The last future work identified from this research deals with normalized NPs. In this study we have shown that the anaphoric properties of a normalized anaphoric NP are dependent on the suffix used in the formation of the morpheme, however this has opened up the research question; *what property in a verb allows us to form a morpheme with a certain suffix?*. For instance the language allows us to form only *runner* from the verb *run* while *government*, *governance* and *governor* can be formed from the verb *govern*, each with different anaphoric properties. The factors that constrain a verb has to lay either in the action it represents or in the arguments (subject and object/s). Further investigation is needed to answer the following two questions:

1. Are there any underlying factors that constrain generation of anaphoric morphemes from verbs?
2. Are the anaphoric properties of verb morphemes dependent on these factors or the suffix of the morpheme?

The analysis on normalization was done on the suffix of the normalized NPs versus their anaphoric properties and this showed a relatively high conformance rate of 86% with the hypothesis. We took this result to be a statistical property of natural discourses, however frequently, statistical results are good indicators of some underlying phenomenon which could well be true in the case of normalization. As a reader of this thesis, you are welcome to pursue this open topic.

6.3 Concluding Remarks

Anaphora has been an active research area for more than two decades and there has been substantial progress in terms of understanding it from a linguistic and psychological perspective. There has also been an impressive amount of progress in terms of computationally resolving it. Some studies have reported up to 99% success rates for specific types of anaphora while more generic systems have achieved success rates in the region of 50% to 80%. However an area that needs more research is the integration of knowledge into computational anaphora resolution systems. While a majority of anaphora in naturally occurring discourses can be resolved using syntactical and morphological constraints, the leftover ones require use of knowledge. Although use of knowledge has been reported in some works, most of these have been used on so called “bridging” anaphora. We have argued that considering bridging relations as anaphora is not within the scope of the original definition of anaphora and doing so has been a distraction for anaphora researchers. This has created the perception, that

co-reference anaphora is largely resolved, hence attention and resources has been diverted to bridging anaphora.

In this thesis we have presented a case to re-focuss our efforts back to co-reference anaphora as it still needs development, especially in the area of knowledge integration. To this effect, we presented a fundamental theory on anaphora which also underpins the separate phenomenon of compound noun generation. We then described a knowledge framework that guides both, the use of anaphora as well as generation of compound nouns. The knowledge framework was then used as a source of constraints to resolve anaphora. We used these constraints in an implemented system to resolve anaphora in discourses from the news media genre. The results obtained were encouraging, even with basic knowledge extraction techniques and only WordNet as the source of knowledge. However, the structure of the knowledge used enables one to easily extend the extraction techniques as well as the source of knowledge to any corpora such as the Web.

In summary, in this thesis we have:

- presented an enhanced definition of anaphora;
- proposed a new theoretical framework supporting the new definition of anaphora;
- presented an algorithm that is based on the proposed framework;
- presented the results and evaluation of the implemented algorithm.

Appendix A

Appendix

A.1 Local Lexicons used in the Algorithm

This section outlines the list of lexicons used by the resolution algorithm. Some of the lists are complete while others are semi-complete. The whole set of lexicons are contained in an XML file which can be updated by the user for processing new texts.

List of Object Types

NOUN	ADJECTIVE	VERB
BE_VERB	ADVERB	ARTICLE
PREPOSITION	CONJUNCTIVE_ADVERB	COMMA
COORD_CONJUNCT	QUANTIFIER	QUALIFIER
TITLE	DEMONSTRATIVE	IRREGULAR_PLURAL
ORAL_VERB	PRONOUN	SUBORD_CONJUNCT
MARKER	RELATION	POS

Table A.1: Table showing the list of object types used by the resolution algorithm.

A.1 Local Lexicons used in the Algorithm

Pronoun	Gender	Animacity	Nperson	Number	Reflexive	Possessive
I	neutral	true	1	singular	false	false
me	neutral	true	1	singular	false	false
myself	neutral	true	1	singular	true	false
my	neutral	true	1	singular	false	false
mine	neutral	true	1	singular	false	true
we	neutral	true	1	plural	false	false
us	neutral	true	1	plural	false	false
our	neutral	true	1	plural	false	true
ours	neutral	true	1	plural	false	true
ourselves	neutral	true	1	plural	false	false
you	neutral	true	2	singular	false	false
your	neutral	true	2	singular	false	true
yours	neutral	true	2	singular	false	true
yourself	neutral	true	2	singular	true	false
he	male	true	3	singular	false	false
him	male	true	3	singular	false	false
his	male	true	3	singular	false	true
himself	male	true	3	singular	true	false
she	female	true	3	singular	false	false
her	female	true	3	singular	false	true
hers	female	true	3	singular	false	true
herself	female	true	3	singular	true	false
it	neutral	false	2	singular	false	false
its	neutral	false	2	singular	false	true
itself	neutral	false	2	singular	true	false
they	neutral	unknown	3	plural	false	false
them	neutral	unknown	3	plural	false	false
their	neutral	unknown	3	plural	false	true
theirs	neutral	unknown	3	plural	false	true
themselves	neutral	unknown	3	plural	true	false

Table A.2: Table showing the list of pronouns and their property values.

A.1 Local Lexicons used in the Algorithm

Single word prepositions

aboard	astride	circa	like	per	under
about	at	concerning	mid	plus	underneath
above	athwart	despite	minus	pro	unlike
absent	atop	down	near	qua	until
across	barring	during	next	regarding	up
after	before	except	notwithstanding	round	upon
against	behind	excluding	of	save	versus
along	below	failing	off	since	via
alongside	beneath	following	on	than	vice
amid	beside	for	onto	through	with
amidst	besides	from	opposite	throughout	within
among	between	given	out	till	without
amongst	betwixt	in	outside	times	worth
around	beyond	including	over	to	
as	but	inside	pace	toward	
aside	by	into	past	towards	

Table A.3: Table showing the alphabetical list of prepositions in the English language (source : wikipedia).

A.1 Local Lexicons used in the Algorithm

Be_Verbs List	
be	is
am	are
was	were
been	could
would	might
might	will
can	

Table A.4: Table showing the list of Be-Verbs used in by the algorithm.

Value	Meaning	Number
this	near	singular
that	far	singular
these	near	plural
those	far	plural
who	animate	unknown
whose	possessive	singular
which	non-animate	singular

Table A.5: Table showing the list of Qualifiers used by the resolution algorithm.

Time	Place
at	at
on	on
in	in
for	
since	

Table A.6: Table showing the prepositions that were used to identify time and place locations.

A.1 Local Lexicons used in the Algorithm

Value	Meaning
senior	one of 2
medium	one of 3
top	one of 3
bottom	one of 2
young	one of 2
older	one of 2
16-year-old	one of many
dark colored	one of 2
thick	one of 2
black	one of many

Table A.7: Table showing the list of qualifiers used in the resolution algorithm

Value	Meaning
one	1
two	2
both	2
couple	2
three	3
four	4
five	5
several	plural
few	plural
many	plural
group	plural

Table A.8: Table showing the list of quantifiers used in the resolution algorithm

Oral Verb List		
said	told	confirmed

Table A.9: Table showing the list of oral verbs used in the resolution algorithm

Article List		
An	A	The

Table A.10: Table showing the list of articles used in the resolution algorithm

A.1 Local Lexicons used in the Algorithm

Value	Gender
Dr	male
Miss	female
Mr	male
Mrs	female
Mrs	female
Ms	female
Sir	male
Madam	female
Master	male

Table A.11: Table showing the list of titles used in the resolution algorithm

Singular	Plural
man	men
woman	women
child	children
mouse	mice
tooth	teeth
foot	feet
cactus	cacti
person	people
phenomenon	phenomena
thesis	theses
datum	data
family	family

Table A.12: Table showing the list of irregular plurals used in the resolution algorithm

Coordinating Conjunction List		
And	Or	But
So	For	Yet
Nor		

Table A.13: Table showing the list of coordinating conjunctions used in the resolution algorithm

Subordinating Conjunction List

After	Although	As
As if	As long as	Because
Before	Even if	Even though
If	Once	Provided
Since	So that	That
Though	Till	Unless
Until	What	When
Whenever	Wherever	Whether
While	Accordingly	Also
Anyway	Besides	Consequently
Finally	For example	For instance
Further	Furthermore	Hence
However	Incidentally	Indeed
In fact	Instead	Likewise
Meanwhile	Moreover	Namely
Now	Of Course	On the contrary
On the other hand	Otherwise	Nevertheless
Next	Nonetheless	Similarly
So far	Until now	Still
Then	Therefore	Thus

Table A.14: Table showing the list of subordinating conjunctions used in the resolution algorithm

A.2 Sample Annotator Forms Used

A.2.1 Normalization Annotator Form Sample

Example of a form used by human annotators to annotate co-reference targets for NPs formed by normalization. The example clause/NP pairs are were derived from both base corpus as well as the extended corpus.

A.2 Sample Annotator Forms Used

Source Clause	Anaphor	Target Types
The 4WD driven by Jim Baker...	the driver	<input type="checkbox"/> AGENT <input type="checkbox"/> OBJECT <input type="checkbox"/> ACTION <input type="checkbox"/> CLAUSE <input type="checkbox"/> PREDICATE
The industry produces bacteria...	the industrial production	<input type="checkbox"/> AGENT <input type="checkbox"/> OBJECT <input type="checkbox"/> ACTION <input type="checkbox"/> CLAUSE <input type="checkbox"/> PREDICATE
The industry produces bacteria...	the bacterial production	<input type="checkbox"/> AGENT <input type="checkbox"/> OBJECT <input type="checkbox"/> ACTION <input type="checkbox"/> CLAUSE <input type="checkbox"/> PREDICATE
Kevin M. was admitted to the board...	the admissibility	<input type="checkbox"/> AGENT <input type="checkbox"/> OBJECT <input type="checkbox"/> ACTION <input type="checkbox"/> CLAUSE <input type="checkbox"/> PREDICATE
Over 500 000 people ran on the day...	the run	<input type="checkbox"/> AGENT <input type="checkbox"/> OBJECT <input type="checkbox"/> ACTION <input type="checkbox"/> CLAUSE <input type="checkbox"/> PREDICATE
The plastic protects the meat...	the protective layer	<input type="checkbox"/> AGENT <input type="checkbox"/> OBJECT <input type="checkbox"/> ACTION <input type="checkbox"/> CLAUSE <input type="checkbox"/> PREDICATE
The council worker worked on it ...	the workman-ship	<input type="checkbox"/> AGENT <input type="checkbox"/> OBJECT <input type="checkbox"/> ACTION <input type="checkbox"/> CLAUSE <input type="checkbox"/> PREDICATE
The chemical industry produces it to...	the product	<input type="checkbox"/> AGENT <input type="checkbox"/> OBJECT <input type="checkbox"/> ACTION <input type="checkbox"/> CLAUSE <input type="checkbox"/> PREDICATE

A.3 Samples of Input Text

This section contains a sample of full articles from input data with a selection of anaphoric NPs highlighted to illustrate some “interesting” cases. Excerpts from some of these articles have been used in the thesis while some others are included as further illustration.

A.3.1 The Full “Trucks” Article

[1] Both trucks involved in yesterday’s crash were gutted by the inferno which erupted seconds after Ted Collins saved one of the drivers.

[2] A dairy worker pulled an injured truck driver from the wreckage of his cab moments before a second big rig laden with 33,000 litres of diesel erupted in flames.

[3] Explosions rocked the surrounding area shortly after the fire broke out, and a plume of thick black smoke drifted over Hamilton.

[4] Fonterra driver Ted Collins was following the northbound Freightlines truck on State Highway 1, and saw it collide with the Shell diesel tanker at Horotiu, north of Hamilton.

[5] “The fuel tanker heading south got into trouble and jack-knifed across the road and on to its side. The guy in front of me didn’t have much time to do anything and didn’t quite stop in time.”

[6] The Freightlines truck, driven by a 51-year-old contractor, ploughed into the diesel tanker, Mr Collins said.

[7] "With 40 tonnes you've got a bit of momentum, and **he** had almost stopped when **he** hit it."

[8] **Mr Collins** leaped from **his cab** and ran to the aid of **his fellow truckies**.

[9] "I did what **I** could to save somebody, it's a natural instinct."

[10] Senior Hamilton fire station officer Daryl Trim said **Mr Collins'** actions were heroic.

[11] "It's fantastic. If **he** hadn't done that, it may have had some dire consequences. It's pretty heroic really."

[12] **Mr Collins** said: "I didn't really think about it, **I** just knew one of my truck friends could be dead or dying. **It** was **a fuel tanker** and **I** thought **I**'ve just got to get there quickly."

[13] **Mr Collins** approached the **Freightlines truck's cab** as **the dazed man** was getting out of **his seat**, struggling to get away.

[14] "The **other truck driver** **I** could not get to - **his truck** was ablaze and **it** was blocking the road, there was no way **I** could get to **him**."

[15] "That was the thing that was worrying **me**. **I** thought **the other guy** had gone, it was a hell of a blaze."

[16] **Mr Collins** helped **the injured man** to the side of the road, away from the fire.

[17] "**He** was slightly out of it, a bit hurt. **We** sat down for a few minutes then some big bangs hit. **I** could see fuel going into the drains at the edge of the road and thought

it [the fire] might get to us, so we had to move along another 50 yards or so.”

*[18] **The 28-year-old tanker driver** escaped from **his cab** on the other side of the blaze with minor injuries.*

*[19] **Both drivers** were taken to Waikato Hospital. **The tanker driver** was discharged, but the **Freightlines driver** was admitted with moderate injuries.*

[20] Those battling the blaze faced temperatures of up to 1000C. The fire gutted both trucks.

[21] Gavin Crook, who lives metres north of where the crash occurred, said he heard a loud bang, followed by the sound of cars slowing down.

[22] ”It was like a bomb, and a big amount of smoke and flames were coming out. You could feel the ground shaking.”

[23] The accident closed SH1 between Hamilton and Ngaruawahia.

[24] Transit spokesman Chris Allen said the highway surface was badly damaged, and a 70m stretch had to be rebuilt and resealed during the day at a cost of about \$20,000.

A.3.2 The Full “Taupo” Article

*[1] A **man** and **his son** are dead, while **another son** survived after a night boating trip turned to tragedy on Lake Taupo.*

[2] *The **three** were heading back to Tokaanu at the southern end of Lake Taupo when their small aluminium dinghy sank about midnight.*

[3] *The **father** and **one son** drowned while the survivor was found on the shore about 9.30am today - nine hours after the capsized.*

[4] *Senior Sergeant Tony Jeurissen said police from Taupo and Turangi were told that three people in an aluminium dingy had got into difficulties at the southern end of the lake.*

[5] *The coastguard and Lion Foundation Rescue helicopter were called out.*

[6] *Two **bodies** were found floating in the Tongariro Delta and **a man** was found and treated by ambulance on the lake shore this morning.*

[7] *The matter was now under investigation by Turangi police, who were making inquiries on behalf of the Bay of Plenty coroner.*

[8] *Dan Harcourt, pilot of the Taupo-based Lion Foundation Rescue helicopter, said **only three people** were thought to be in the dinghy.*

[9] *"Two **bodies** have been recovered. **One survivor** has been checked by ambulance," he said.*

[10] *He said there were no difficulties recovering **the bodies** or **the survivor**.*

[11] *"We have got Turangi Coastguard on the scene as well recovering **the bodies** and we are just assisting in the visual search of the area," he said.*

A.3.3 The Full “Levinman” Article

*[1] The owner of a 4WD vehicle seized in connection with the **Birgit Brauer murder case** says **it** was stolen by **his employee** six weeks ago.*

*[2] The **Toyota Hilux** is registered to **Palmerston North man Brent Cleverley**, who confirmed to the Herald yesterday that **his vehicle** had been taken by **a man** who worked for **him**.*

*[3] The Herald understands **the worker**, who used **the vehicle** in **his job** cutting firewood, failed to turn up to work one day last month.*

*[4] **He** has not been seen since.*

*[5] **Mr Cleverley**, whose **4WD** was taken for forensic tests, said **he** had been asked by police not to talk to the media.*

*[6] **His worker** is understood to have grown up in the Himatangi/Foxton area but had recently spent time in the South Island before moving back to Levin.*

*[7] **He** is known by various names.*

*[8] **He** is in his mid-30s and lived in Levin, about 4km from the Ohau River where the vehicle was found yesterday.*

[9] After it was discovered dozens of police officers, including forensic experts and a dive squad, arrived at the river bed to secure and examine the scene.

*[10] Last night Detective Senior Sergeant Grant Coward would not call **the man** a suspect in the killing of **the German backpacker**, but has confirmed police know his name and want to find **him**.*

*[11] Last Friday police made inquiries about the ownership of **the specific 4WD vehicle** as part of routine inquiries to account for all **dark grey or black Toyota Hiluxes**, like the one **Ms Brauer** travelled in before her death last week.*

*[12] **Her body** was found at Lucys Gully, near New Plymouth, on September 20.*

*[13] **The 28-year-old** had been stabbed and had head injuries.*

[14] Mr Coward, the head of the investigation, has repeatedly said the key to catching **the killer** lay in finding **the mid-1980s Hilux Ms Brauer** was seen getting into.

[15] **The one** airlifted out of the Ohau River yesterday clearly matched the description.

[16] Retired schoolteacher Jenny Burnell, who lives on the farm which backs on to the riverbed, said **the vehicle** was dumped there on Monday night.

[17] Her nephews saw **its** lights along a track on Miss Burnell's farm about 9.30pm.

[18] **They** investigated but did not find anything.

[19] **The vehicle** was seen on the edge of the riverbed yesterday morning and reported to police, who used a helicopter to remove it from the water.

[20] Miss Burnell said it was an unusual place for the vehicle to be dumped as few people would know how to get there.

[21] Taranaki helicopter company owner Alan Beck said police called for one of his heavy-lift Super Huey Iroquois helicopters.

[22] "**Our** job was to lift **it** out because there were fears if the water suddenly rose they might lose valuable evidence," Mr Beck said.

[23] Mr Coward said there was no evidence that **the person** who dumped **the vehicle** in the river was **the murder suspect**, but police would still like to talk to **him**.

[24] "Obviously we want **the person** who stole **that vehicle** to come forward. That would be most helpful."

[25] "**Someone** who's committed this crime would, in my view, be acting differently than **they** would beforehand.

[26] **They** have done something horrendous and **they** know it.

[27] "I can assure you if **someone** knows who **the killer** is and **they** want to be treated confidentially, **they** will be treated with confidence."

[28] Mr Coward said police had received many reports of vehicles matching the description of the one **Ms Brauer** was last seen in.

[29] *Several* had been recovered, including **three stolen ones**.

[30] Land Transport NZ figures show that almost 23,000 **Toyota Hiluxes** were registered in New Zealand between 1985 and 1990.

[31] *Of those*, **673** were in the Taranaki area.

[32] Mr Coward said police had tracked sightings of the vehicle from Waitotara, where **Ms Brauer** was seen getting into a **Toyota Hilux**, to Lucys Gully, where **her body** was found. [33] **The vehicle** had then been seen at Cardiff, near Stratford.

A.3.4 The Full “Zaoui” Article

[1] If **Ahmed Zaoui’s family** succeed in **their** application to join **him** in New Zealand as refugees, the decision will annoy **NZ First leader Winston Peters** but delight **a boy** who has not seen **his father** for nearly three years.

[2] **Mr Zaoui’s wife**, Leila, and four sons - **Youssef**, 7, **Adbel**, 14, **Soheib**, 17, and **Hicham**, 19 - have applied through **Mr Zaoui’s lawyer, Deborah Manning**, to move to New Zealand as refugees.

[3] **Ms Manning** told the Sunday Star-Times that supporters had offered **the family** financial assistance, but **she** could not preclude the need for **the family** to apply for a benefit, to which refugees are entitled.

[4] The Government has refused to consider the applications until **Mr Zaoui’s case** is resolved (he is on bail awaiting a review of his SIS security risk certificate, which could mean **his** deportation).

[5] The applications have been criticised by the National and New Zealand First parties. While the politicians bicker over issues of cost, security risk and human rights, for **Mr Zaoui’s youngest son** all it means is growing up without a father.

[6] The family spoke of the separation to TV3's Campbell Live producer Carol Hirschfeld, who travelled to interview them at a Southeast Asian location for a segment to screen tonight.

[7] They also got a taste of what they were missing - a satellite link let Mr Zaoui and the family talk face to face for the first time since he left for New Zealand in December 2002.

[8] Hirschfeld said separation had a strong effect on son Youssef, who was 4 1/2 years old when he last saw his father. "Ahmed phones regularly but I think the satellite link was so powerful because they hadn't actually seen each other in the flesh for nearly three years.

[9] "Seeing how much Youssef had grown was something Ahmed really responded to."

[10] Hirschfeld said the family had perceived NZ as "a peaceful land which offered a place of refuge with a good record for democracy".

[11] National's immigration spokesman, Tony Ryall, said calls for the family to be allowed in before Mr Zaoui's case was settled were ludicrous.

[12] "We shouldn't have a system in New Zealand where you get one refugee and end up with several others. It puts a lot of pressure on the system."

[13] If it was decided Mr Zaoui should go but his family had already been let in, then Mr Zaoui would have another argument for coming back, said Mr Ryall.

[14] Mr Peters said at his campaign launch in Takapuna yesterday that the Zaoui case had already cost the country \$2 million.

[15] "Mr Zaoui, you can see your family tomorrow if you would just get on a plane and go and see them."

[16] But the Green Party said Mr Zaoui should be shown some "Kiwi compassion" because his separation from his family was the fault of deficiencies in New Zealand's own systems. He had also spent two years in jail "unnecessarily".

[17] *The party said allowing the family to get together could not represent a security danger to New Zealand.*

A.3.5 The Full “Rugby” Article

[1] *The Welsh rugby fan who killed a young Waikato woman when his campervan smashed into her oncoming car has been ordered to pay almost \$10,000 in fines and reparation.*

[2] *James Berry, 23, had previously pleaded guilty to careless driving causing injury and careless driving causing the death of 18-year-old Cambridge woman Liz Neels.*

[3] *This morning in the Hamilton District Court he stood looking scared and vulnerable as he was handed down a \$500 fine, plus costs, for the injury charge.*

[4] *Judge Anne McAloon said she took into account that Berry had been held in custody for three days after the June 23 crash and the volunteer work he had been doing since his initial court appearance.*

[5] *The accounting graduate from Swansea had come to New Zealand to follow the Lions tour but had not seen any of the games.*

[6] *The judge ordered him to pay \$9000 in reparation to the Neels family on the causing death charge.*

[7] *Berry’s mother was in the court for the sentencing, and cried throughout the proceedings.*

[8] *Liz Neels’ father Michael told the court he thought more statistics needed to be kept on the number of accidents caused by tired drivers.*

[9] *Mr Neels and his wife, Lori, said before the sentencing that they did not want to see Berry jailed or “financially crippled with some huge fine”.*

[10] But if **Berry** was to be released with no more than a "bad luck, my **boy** - hope you learned something by all this", then the wrong message would be going out, said **Mr Neels**.

[12] The clear message was that it was not acceptable to drive when tired, **he** said.

[13] **Berry** had arrived from Britain with two friends mid-morning on June 23.

[14] After a break at **his** sister's house in Auckland, **they** left about 4pm, intending to stop for the night at 10pm and catch the interisland ferry from Wellington the next day to get to Christchurch for the first Lions versus All Blacks test.

[15] About 7.40pm, the campervan crossed the centre line on State Highway 1, near Karapiro, slamming into a car driven by **Liz Neels**, a student chef.

[16] After initial reluctance to meet **Berry** at a restorative justice conference, a deeply grieving **Mr and Mrs Neels** came to think of **him** as a victim too, "just in a completely different tragedy, connected only by **Liz**".

[17] **They** thought **they** should take the opportunity before **Berry** returned to Wales.

[18] "We're glad **we** did."

[19] **Mr Neels** doubted **they** could have faced **him** had he been a "boy-racer or a drunk driver" instead of an intelligent, sensitive young **man** with a promising future.

[20] A nephew and **Liz**'s two best friends accompanied the couple to the voluntary meeting, run by facilitators.

[21] **Berry** came alone.

[22] **Berry** expressed a wish to meet the **Neels**, appeared remorseful and even sent a card.

[23] "It seemed that, in falling asleep [at the wheel], he had made a very human mistake with devastating consequences," **Mr Neels** said.

[24] **The Neels** said they wanted **Berry** to live a good life, both for **himself** and for **Liz**, and not carry **her** in **his** memory as a burden, rather "letting our **Lizzie** rest lightly on **his** soul".

[25] *No further communication was planned, said **Mr Neels**.*

[26] *"But if **he** wants to contact us in five years' time and let **us** know how **he** is doing, what **his wife** is like, I daresay **we** will show interest."*

A.3.6 The Full "Aaron" Article

[1] *A **man** watched in horror as emergency services tried in vain to save **his brother-in-law**'s life after **he** was hit by a speeding car and left hanging through the windscreen without an arm.*

[2] *Aaron Chan, 21, of Manukau, was with **his brother-in-law** unloading goods from **their** delivery van when a car ploughed into **him** on West Coast Rd, Glen Eden, about 6.30pm on Monday.*

[3] *Police last night arrested **two men** and charged **them** with reckless driving causing death.*

[4] *Dion Wells, 35, of New Lynn, appeared in Waitakere District Court this morning on a charge of operating a motor vehicle recklessly and thereby caused the death of Aaron Chan.*

[5] *Wells did not make a plea and was remanded on bail until December 17.*

[6] *As Wells entered the dock, having being held in custody overnight, **the unemployed father** appeared to have one arm in a sling beneath **his** jumper.*

[7] *The Police prosecutor asked for a curfew to put in place, which was granted by the judge. Wells has been ordered not to drive or associate with the **other driver** alleged to be involved in the accident. [8] **The other driver**, 28 from Whangaparoa, will appear in Waitakere District Court this afternoon.*

[9] *Mr Chan's friends have expressed anger at what appears to have been a needless death caused by what witnesses described as boy-racing.*

[10] *"I hope the [**drivers**] live the rest of **their** life knowing **they** took the life of **one of the nicest people I know**," said **Ryan Leong**.*

[11] *"**His** life has been robbed from **him** by **some selfish idiots**. **I** don't even want to think about how **his** parents will cope, losing **their youngest son**."*

[12] *Witnesses say the accident happened when a blue Nissan Skyline - that appeared to be racing a white Subaru WRX - hit the back of **Mr Chan**'s truck.*

[13] ***He** was thrown on to the bonnet of the car as it ran along the footpath, over the top of a small tree and into a brick wall. When it came to rest **Mr Chan** was lying across the bonnet with the upper part of **his** body through the windscreen. **His** arm was severed and most bones appeared to be badly broken.*

[14] *The Subaru lost control and crashed on the other side of the road.*

[15] *One of the first people on the scene, local woman Nicole King, ran outside to find **a young woman** and **an injured man** - who are both believed to have been in the Nissan - looking dazed.*

[16] *"**They** were coming towards **me** and **he** was all bloody so **I** got **him** to sit down."*

[17] ***Mrs King** said the **young woman**, who was holding her stomach and complaining of pain, thought **she** was pregnant.*

[18] *It wasn't until **Mrs King** stood up again from treating **the bleeding man** that **she** saw legs hanging over the front of the car and realised someone had been hit. **She** went to **Mr Chan** but was unable to help.*

[19] *"**He** had a pulse. **I** just lifted his head but **I** heard a bit of gurgling. **I** couldn't pull **him** out through the window. There was nothing **I** could do."*

[20] *Station Officer Bryan Marsden of the Glen Eden Fire Service said firefighters pulled **Mr Chan** from the window and tried to administer first aid.*

[21] *"At that stage **he** was still alive but **he** slipped away while **he** was there. **He** didn't last long."*

[22] *The accident happened outside National Party MP Paula Bennett's office. She said last night that something needed to be done to discourage boyracers. [23] "Obviously what we are doing now isn't working and I think that's pretty clear from what happened."*

[24] *She said she would write to the Minister for Road Safety to see what other measures could be investigated, such as destroying boyracers' cars after an offence.*

[25] *West Coast Rd residents say a group of people - believed to be friends or family - stood at the scene of the crash about 3am yesterday and burned incense and paper money in what looked like a blessing ceremony.*

[26] *Police say they are still investigating witness reports of a fight occurring at traffic lights near the crash scene and that further charges could be laid.*

A.3.7 The Full "Poisoning" Article

[1] *Two **rugby players** were treated in hospital for blood poisoning and several others developed infections after playing on an **Auckland club field** contaminated with what was believed to be sewage.*

[2] *The board of **College Rifles** met yesterday to discuss the incident three weekends ago, and decided that if the drain overflowed sewage again all play on **the field** would be stopped.*

[3] *The **Remuera field** was inspected yesterday by Public Health but no contamination was found.*

[4] *A **player** from **Varsity Legends**, part of **Auckland University Rugby Football Club**, grazed **his** knee while playing the **College Rifles Raiders** on June 25 was admitted to hospital with blood poisoning.*

[5] *He had continued to play the following week but **his** knee swelled up and became increasingly painful.*

[6] *Varsity Legends coach Mitch Canning said **the player**, who did not want to be named, woke up in "a hell of a lot of pain" the next Sunday morning and went to hospital.*

[7] *"**He** was in pretty bad shape and ended up spending a couple of days in hospital on an IV drip."*

[8] *Mr Canning believed **the player** had suffered a blood infection from playing on the ground. "It's the first time **I**'ve seen **people** getting poisoned on **the football field**.*

[9] *"A **guy** from the Raiders went to hospital straight after the game. **He** had a cut on his face and has had a couple of trips to the hospital. **I** don't believe **he**'s back playing yet."*

[10] *The Varsity fullback was released from hospital on Tuesday but was not back at work until last Thursday, Mr Canning said.*

[11] *Mr Canning said the state of the ground was not the **club**'s fault.*

[12] *"It was more a reflection of the weather. In the four years or so **I**'ve never seen torrential rain like that.*

[13] *"**You** couldn't really see the sewage but **you** could see quite a lot of flooding coming off the park and water cascading on the road.*

[14] *" ... The jerseys and gear smelled pretty bad afterwards."*

[15] *The drain being blamed is actually a swift-flowing creek on the northern side of the **fields** just a few metres from the goalposts. A large pipe drains into it and the water eventually ends up in Orakei Basin.*

[16] *Rifles operations manager Derek Rope said it had overflowed before but there had been no problems.*

[17] *"We can't say it was sewage but it was definitely something that wasn't good for you," he said.*

[18] *The overflow problem happened about once a year. On June 25 it coincided with a match day when four games were played.*

A.4 A Sample Evaluation File from aCAR

[19] *"It went over the top for a very brief period of time. Open sores got infected," Mr Rope said.*

[20] *However, **players** - aged in their 20s - were treated with antiseptic or antibiotics by **their** doctors and recovered quickly.*

[21] *"It was one of those freak things." said Mr Rope. "There has been nothing since or prior."*

[22] *"The open drain, I believe, is our responsibility. We would assume it is designed for stormwater but evidently contamination got in."*

[23] *Auckland medical officer of public health Dr Virginia Hope inspected the grounds yesterday but could find no evidence of sewage.*

[24] *She was told by **staff** at the **rugby club** they didn't know who **the players** were who reported getting ill or how many of **them** there were.*

[25] *She said illness caused by contaminated ground wasn't strictly a notifiable disease but GPs would likely have told public health if they thought **people** had caught gastro-enteritis from contamination.*

[26] *Board member John McKeaney said it was yet to decide how it was going to tackle the situation. He said the drain couldn't be screened off as "water goes through fences".*

[27] *Auckland City Council development manager John Duthie said the council was trying to contact the **club** last night to see if there was an ongoing problem.*

A.4 A Sample Evaluation File from aCAR

For the purpose of automated evaluation of correct resolutions aCAR currently reads a text file with the correct antecedents. This was done by initially processing an article which generates a uniquely identified file of NPs which is then modified to include the

A.4 A Sample Evaluation File from aCAR

antecedents for all anaphoric NPs. The following is the sample file for the article in Section A.3.1 (Sent - sentence number, Index - the order in which the nouns appear in the document, Uid - unique language object identifier, Gen, Hum and Typ: represent morphology values). The field's (*Ant:*) value is added manually by a huma with the Uid of the antecedent. This is then used by aCar for automated evaluation of the resolutions. The first 6 nouns with negative Uids are additional entities that can be referred to by NPs.

Sent: -1 Cl: -1 Index: 0 Uid: -1 PLEONASTIC Gen: U Num: U Hum: U Typ: U Ant:

Sent: -1 Cl: -1 Index: 1 Uid: -2 PREVIOUS CLAUSE NOUN-PHRASE Gen: U Num: U
Hum: U Typ: U Ant:

Sent: -1 Cl: -1 Index: 2 Uid: -3 READER Gen: U Num: U Hum: U Typ: U Ant:

Sent: -1 Cl: -1 Index: 3 Uid: -4 UNKNOWN NOUN-PHRASE Gen: U Num: U Hum: U Typ:
U Ant:

Sent: -1 Cl: -1 Index: 4 Uid: -5 IGNORE Gen: U Num: U Hum: U Typ: U Ant:

Sent: -1 Cl: -1 Index: 5 Uid: -10 PRODUCER-GROUP Gen: U Num: U Hum: U Typ: U Ant:

Sent: 1 Cl: 1 Index: 6 Uid: 4 trucks Gen: N Num: P Hum: U Typ: U Ant:

Sent: 1 Cl: 1 Index: 7 Uid: 6 crash Gen: U Num: U Hum: U Typ: U Ant:

Sent: 1 Cl: 2 Index: 8 Uid: 9 trucks Gen: N Num: P Hum: U Typ: U Ant:

Sent: 1 Cl: 2 Index: 9 Uid: 11 inferno Gen: U Num: U Hum: U Typ: U Ant:

Sent: 1 Cl: 2 Index: 10 Uid: 12 which Gen: U Num: U Hum: U Typ: U Ant:

A.4 A Sample Evaluation File from aCAR

Sent: 1 Cl: 3 Index: 11 Uid: 14 which Gen: U Num: U Hum: U Typ: U Ant:

Sent: 1 Cl: 4 Index: 12 Uid: 17 Ted-Collins Gen: U Num: S Hum: U Typ: U Ant:

Sent: 1 Cl: 4 Index: 13 Uid: 19 drivers Gen: N Num: P Hum: U Typ: U Ant:

Sent: 2 Cl: 1 Index: 14 Uid: 22 worker Gen: U Num: U Hum: U Typ: U Ant: 17

Sent: 2 Cl: 1 Index: 15 Uid: 24 driver Gen: U Num: U Hum: U Typ: U Ant: 19

Sent: 2 Cl: 1 Index: 16 Uid: 26 wreckage Gen: U Num: U Hum: U Typ: U Ant:

Sent: 2 Cl: 1 Index: 17 Uid: 28 cab Gen: U Num: U Hum: U Typ: U Ant:

Sent: 2 Cl: 1 Index: 18 Uid: 29 his Gen: M Num: S Hum: T Typ: U Ant: 24

Sent: 2 Cl: 2 Index: 19 Uid: 31 rig Gen: U Num: S Hum: U Typ: U Ant:

Sent: 2 Cl: 2 Index: 20 Uid: 33 33000 litres Gen: U Num: U Hum: U Typ: U Ant:

Sent: 2 Cl: 2 Index: 21 Uid: 35 diesel Gen: U Num: S Hum: U Typ: U Ant:

Sent: 3 Cl: 1 Index: 22 Uid: 39 explosions Gen: N Num: P Hum: U Typ: U Ant:

Sent: 3 Cl: 1 Index: 23 Uid: 41 area Gen: U Num: U Hum: U Typ: U Ant:

Sent: 3 Cl: 2 Index: 24 Uid: 43 fire Gen: U Num: U Hum: U Typ: U Ant: 11

Sent: 3 Cl: 3 Index: 25 Uid: 46 plume Gen: U Num: U Hum: U Typ: U Ant:

Sent: 3 Cl: 3 Index: 26 Uid: 48 smoke Gen: U Num: U Hum: U Typ: U Ant:

Sent: 3 Cl: 3 Index: 27 Uid: 50 Hamilton Gen: U Num: S Hum: U Typ: U Ant:

Sent: 4 Cl: 1 Index: 28 Uid: 53 Ted-Collins Gen: U Num: S Hum: U Typ: U Ant: 17

Sent: 4 Cl: 1 Index: 29 Uid: 55 Truck Gen: U Num: S Hum: U Typ: U Ant:

A.4 A Sample Evaluation File from aCAR

Sent: 4 Cl: 1 Index: 30 Uid: 57 1 Gen: U Num: U Hum: U Typ: U Ant:

Sent: 4 Cl: 2 Index: 31 Uid: 59 Ted-Collins Gen: U Num: S Hum: U Typ: U Ant: 17

Sent: 4 Cl: 2 Index: 32 Uid: 61 it Gen: N Num: S Hum: F Typ: U Ant: 55

Sent: 4 Cl: 3 Index: 33 Uid: 63 it Gen: N Num: S Hum: F Typ: U Ant: 55

Sent: 4 Cl: 3 Index: 34 Uid: 65 Tanker Gen: U Num: S Hum: U Typ: U Ant: 31

Sent: 4 Cl: 3 Index: 35 Uid: 67 Horotiu Gen: U Num: S Hum: U Typ: U Ant:

Sent: 4 Cl: 3 Index: 36 Uid: 68 north Gen: U Num: U Hum: U Typ: U Ant:

Sent: 4 Cl: 3 Index: 37 Uid: 70 Hamilton Gen: U Num: S Hum: U Typ: U Ant: 50

Sent: 5 Cl: 0 Index: 38 Uid: 73 Ted-Collins Gen: U Num: S Hum: T Typ: U Ant: 17

Sent: 5 Cl: 1 Index: 39 Uid: 77 tanker Gen: U Num: U Hum: U Typ: U Ant: 31

Sent: 5 Cl: 2 Index: 40 Uid: 80 tanker Gen: U Num: U Hum: U Typ: U Ant: 31

Sent: 5 Cl: 2 Index: 41 Uid: 82 trouble Gen: U Num: U Hum: U Typ: U Ant:

Sent: 5 Cl: 3 Index: 42 Uid: 84 tanker Gen: U Num: S Hum: U Typ: U Ant: 31

Sent: 5 Cl: 3 Index: 43 Uid: 86 road Gen: U Num: S Hum: U Typ: U Ant:

Sent: 5 Cl: 3 Index: 44 Uid: 88 side Gen: U Num: S Hum: U Typ: U Ant:

Sent: 5 Cl: 3 Index: 45 Uid: 89 its Gen: N Num: S Hum: F Typ: U Ant: 31

Sent: 5 Cl: 4 Index: 46 Uid: 91 guy Gen: M Num: S Hum: T Typ: U Ant:

Sent: 5 Cl: 4 Index: 47 Uid: 93 front Gen: U Num: U Hum: U Typ: U Ant:

Sent: 5 Cl: 4 Index: 48 Uid: 95 me Gen: N Num: S Hum: T Typ: U Ant: 17

A.4 A Sample Evaluation File from aCAR

Sent: 5 Cl: 4 Index: 49 Uid: 98 time Gen: U Num: U Hum: U Typ: U Ant:

Sent: 6 Cl: 0 Index: 50 Uid: 101 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 6 Cl: 1 Index: 51 Uid: 105 Truck Gen: U Num: S Hum: U Typ: U Ant: 55

Sent: 6 Cl: 1 Index: 52 Uid: 107 contractor Gen: U Num: U Hum: U Typ: U Ant: 24

Sent: 6 Cl: 2 Index: 53 Uid: 109 Truck Gen: U Num: S Hum: U Typ: U Ant: 17

Sent: 6 Cl: 2 Index: 54 Uid: 111 tanker Gen: U Num: U Hum: U Typ: U Ant: 31

Sent: 7 Cl: 0 Index: 55 Uid: 114 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 7 Cl: 1 Index: 56 Uid: 118 you Gen: N Num: S Hum: T Typ: U Ant: -3

Sent: 7 Cl: 1 Index: 57 Uid: 120 40 tonnes Gen: U Num: U Hum: U Typ: U Ant:

Sent: 7 Cl: 1 Index: 58 Uid: 122 bit Gen: U Num: U Hum: U Typ: U Ant:

Sent: 7 Cl: 1 Index: 59 Uid: 124 momentum Gen: U Num: U Hum: U Typ: U Ant:

Sent: 7 Cl: 2 Index: 60 Uid: 128 he Gen: M Num: S Hum: T Typ: U Ant: 24

Sent: 7 Cl: 3 Index: 61 Uid: 135 he Gen: M Num: S Hum: T Typ: U Ant: 24

Sent: 7 Cl: 3 Index: 62 Uid: 137 it Gen: N Num: S Hum: F Typ: U Ant: 31

Sent: 8 Cl: 1 Index: 63 Uid: 142 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 8 Cl: 1 Index: 64 Uid: 144 cab Gen: U Num: U Hum: U Typ: U Ant:

Sent: 8 Cl: 1 Index: 65 Uid: 145 his Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 8 Cl: 2 Index: 66 Uid: 149 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 8 Cl: 2 Index: 67 Uid: 151 aid Gen: U Num: U Hum: U Typ: U Ant:

A.4 A Sample Evaluation File from aCAR

Sent: 8 Cl: 2 Index: 68 Uid: 153 truckies Gen: N Num: P Hum: U Typ: U Ant:

Sent: 8 Cl: 2 Index: 69 Uid: 154 his Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 9 Cl: 0 Index: 70 Uid: 159 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 9 Cl: 1 Index: 71 Uid: 163 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 9 Cl: 1 Index: 72 Uid: 165 what Gen: U Num: U Hum: U Typ: U Ant:

Sent: 9 Cl: 2 Index: 73 Uid: 167 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 9 Cl: 2 Index: 74 Uid: 169 somebody Gen: U Num: U Hum: U Typ: U Ant:

Sent: 9 Cl: 3 Index: 75 Uid: 171 it Gen: N Num: S Hum: F Typ: U Ant: -1

Sent: 9 Cl: 3 Index: 76 Uid: 173 instinct Gen: U Num: U Hum: U Typ: U Ant:

Sent: 10 Cl: 0 Index: 77 Uid: 176 Daryl-Trim Gen: U Num: S Hum: T Typ: U Ant:

Sent: 10 Cl: 1 Index: 78 Uid: 180 actions Gen: N Num: P Hum: U Typ: U Ant:

Sent: 10 Cl: 1 Index: 79 Uid: 181 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 10 Cl: 1 Index: 80 Uid: 183 heroic Gen: U Num: U Hum: U Typ: U Ant:

Sent: 11 Cl: 0 Index: 81 Uid: 186 Daryl-Trim Gen: U Num: S Hum: T Typ: U Ant:

Sent: 11 Cl: 1 Index: 82 Uid: 190 it Gen: N Num: S Hum: F Typ: U Ant: -4

Sent: 11 Cl: 1 Index: 83 Uid: 192 fantastic Gen: U Num: S Hum: U Typ: U Ant:

Sent: 11 Cl: 2 Index: 84 Uid: 194 he Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 11 Cl: 2 Index: 85 Uid: 196 that Gen: U Num: U Hum: U Typ: U Ant:

Sent: 11 Cl: 3 Index: 86 Uid: 200 it Gen: N Num: S Hum: F Typ: U Ant: -4

A.4 A Sample Evaluation File from aCAR

Sent: 11 Cl: 3 Index: 87 Uid: 202 consequences Gen: N Num: P Hum: U Typ: U Ant:

Sent: 11 Cl: 4 Index: 88 Uid: 206 it Gen: N Num: S Hum: F Typ: U Ant: -4

Sent: 11 Cl: 4 Index: 89 Uid: 208 heroic Gen: U Num: S Hum: U Typ: U Ant: 183

Sent: 12 Cl: 0 Index: 90 Uid: 211 Collins Gen: M Num: S Hum: T Typ: U Ant:

Sent: 12 Cl: 1 Index: 91 Uid: 215 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 12 Cl: 1 Index: 92 Uid: 217 it Gen: N Num: S Hum: F Typ: U Ant: -4

Sent: 12 Cl: 2 Index: 93 Uid: 219 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 12 Cl: 3 Index: 94 Uid: 224 friends Gen: N Num: P Hum: U Typ: U Ant:

Sent: 12 Cl: 3 Index: 95 Uid: 225 my Gen: N Num: S Hum: T Typ: U Ant:

Sent: 12 Cl: 4 Index: 96 Uid: 230 friends Gen: N Num: P Hum: U Typ: U Ant: 224

Sent: 12 Cl: 4 Index: 97 Uid: 231 my Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 12 Cl: 5 Index: 98 Uid: 236 it Gen: N Num: S Hum: F Typ: U Ant: 31

Sent: 12 Cl: 5 Index: 99 Uid: 238 tanker Gen: U Num: S Hum: U Typ: U Ant: 31

Sent: 12 Cl: 6 Index: 100 Uid: 242 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 12 Cl: 7 Index: 101 Uid: 247 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 13 Cl: 1 Index: 102 Uid: 254 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 13 Cl: 1 Index: 103 Uid: 256 cab Gen: U Num: U Hum: U Typ: U Ant: 28

Sent: 13 Cl: 1 Index: 104 Uid: 257 Truck Gen: U Num: S Hum: U Typ: U Ant: 55

Sent: 13 Cl: 2 Index: 105 Uid: 259 man Gen: M Num: S Hum: T Typ: U Ant: 24

A.4 A Sample Evaluation File from aCAR

Sent: 13 Cl: 2 Index: 106 Uid: 261 seat Gen: U Num: U Hum: U Typ: U Ant:

Sent: 13 Cl: 2 Index: 107 Uid: 262 his Gen: M Num: S Hum: T Typ: U Ant: 24

Sent: 13 Cl: 3 Index: 108 Uid: 265 man Gen: M Num: S Hum: T Typ: U Ant:

Sent: 14 Cl: 0 Index: 109 Uid: 269 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 14 Cl: 1 Index: 110 Uid: 273 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 14 Cl: 1 Index: 111 Uid: 275 driver Gen: U Num: U Hum: U Typ: U Ant:

Sent: 14 Cl: 2 Index: 112 Uid: 277 truck Gen: U Num: S Hum: U Typ: U Ant: 31

Sent: 14 Cl: 2 Index: 113 Uid: 278 his Gen: M Num: S Hum: T Typ: U Ant: 275

Sent: 14 Cl: 2 Index: 114 Uid: 280 ablaze Gen: U Num: S Hum: U Typ: U Ant:

Sent: 14 Cl: 3 Index: 115 Uid: 282 it Gen: N Num: S Hum: F Typ: U Ant: 31

Sent: 14 Cl: 3 Index: 116 Uid: 285 road Gen: U Num: S Hum: U Typ: U Ant:

Sent: 14 Cl: 4 Index: 117 Uid: 287 there Gen: U Num: U Hum: U Typ: U Ant:

Sent: 14 Cl: 5 Index: 118 Uid: 290 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 14 Cl: 5 Index: 119 Uid: 292 him Gen: M Num: S Hum: T Typ: U Ant: 275

Sent: 15 Cl: 0 Index: 120 Uid: 295 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 15 Cl: 1 Index: 121 Uid: 299 that Gen: U Num: S Hum: U Typ: U Ant:

Sent: 15 Cl: 1 Index: 122 Uid: 301 thing Gen: U Num: S Hum: U Typ: U Ant:

Sent: 15 Cl: 2 Index: 123 Uid: 303 that Gen: U Num: U Hum: U Typ: U Ant:

Sent: 15 Cl: 2 Index: 124 Uid: 305 me Gen: N Num: S Hum: T Typ: U Ant: 17

A.4 A Sample Evaluation File from aCAR

Sent: 15 Cl: 3 Index: 125 Uid: 309 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 15 Cl: 4 Index: 126 Uid: 312 guy Gen: M Num: S Hum: T Typ: U Ant: 275

Sent: 15 Cl: 5 Index: 127 Uid: 315 it Gen: N Num: S Hum: F Typ: U Ant: -1

Sent: 15 Cl: 5 Index: 128 Uid: 317 hell Gen: U Num: S Hum: U Typ: U Ant:

Sent: 15 Cl: 5 Index: 129 Uid: 319 blaze Gen: U Num: S Hum: U Typ: U Ant:

Sent: 16 Cl: 1 Index: 130 Uid: 322 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 16 Cl: 1 Index: 131 Uid: 324 man Gen: M Num: S Hum: T Typ: U Ant: 24

Sent: 16 Cl: 1 Index: 132 Uid: 326 side Gen: U Num: U Hum: U Typ: U Ant:

Sent: 16 Cl: 1 Index: 133 Uid: 328 road Gen: U Num: U Hum: U Typ: U Ant: 285

Sent: 16 Cl: 1 Index: 134 Uid: 330 fire Gen: U Num: U Hum: U Typ: U Ant: 11

Sent: 17 Cl: 0 Index: 135 Uid: 333 Collins Gen: M Num: S Hum: T Typ: U Ant: 17

Sent: 17 Cl: 1 Index: 136 Uid: 337 he Gen: M Num: S Hum: T Typ: U Ant: 24

Sent: 17 Cl: 1 Index: 137 Uid: 339 it Gen: N Num: S Hum: F Typ: U Ant: 11

Sent: 17 Cl: 1 Index: 138 Uid: 340 hurt Gen: U Num: U Hum: U Typ: U Ant:

Sent: 17 Cl: 2 Index: 139 Uid: 342 we Gen: N Num: P Hum: T Typ: U Ant: -10

Sent: 17 Cl: 2 Index: 140 Uid: 344 minutes Gen: N Num: P Hum: U Typ: U Ant:

Sent: 17 Cl: 3 Index: 141 Uid: 347 bangs Gen: N Num: P Hum: U Typ: U Ant:

Sent: 17 Cl: 4 Index: 142 Uid: 350 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 17 Cl: 4 Index: 143 Uid: 352 fuel Gen: U Num: U Hum: U Typ: U Ant: 35

A.4 A Sample Evaluation File from aCAR

Sent: 17 Cl: 5 Index: 144 Uid: 354 fuel Gen: U Num: U Hum: U Typ: U Ant: 35

Sent: 17 Cl: 5 Index: 145 Uid: 356 drains Gen: N Num: P Hum: U Typ: U Ant:

Sent: 17 Cl: 5 Index: 146 Uid: 358 edge Gen: U Num: U Hum: U Typ: U Ant:

Sent: 17 Cl: 5 Index: 147 Uid: 360 road Gen: U Num: U Hum: U Typ: U Ant: 285

Sent: 17 Cl: 6 Index: 148 Uid: 362 I Gen: N Num: S Hum: T Typ: U Ant: 17

Sent: 17 Cl: 7 Index: 149 Uid: 365 us Gen: N Num: P Hum: T Typ: U Ant: -10

Sent: 17 Cl: 7 Index: 150 Uid: 367 fuel Gen: U Num: U Hum: U Typ: U Ant: 35

Sent: 17 Cl: 8 Index: 151 Uid: 369 we Gen: N Num: P Hum: T Typ: U Ant: -10

Sent: 17 Cl: 8 Index: 152 Uid: 372 yards Gen: N Num: P Hum: U Typ: U Ant:

Sent: 17 Cl: 8 Index: 153 Uid: 374 so Gen: U Num: U Hum: U Typ: U Ant:

Sent: 18 Cl: 1 Index: 154 Uid: 377 driver Gen: U Num: U Hum: U Typ: U Ant: 275

Sent: 18 Cl: 1 Index: 155 Uid: 379 cab Gen: U Num: U Hum: U Typ: U Ant:

Sent: 18 Cl: 1 Index: 156 Uid: 380 his Gen: M Num: S Hum: T Typ: U Ant: 275

Sent: 18 Cl: 1 Index: 157 Uid: 382 side Gen: U Num: U Hum: U Typ: U Ant:

Sent: 18 Cl: 1 Index: 158 Uid: 384 blaze Gen: U Num: U Hum: U Typ: U Ant: 319

Sent: 18 Cl: 1 Index: 159 Uid: 386 injuries Gen: N Num: P Hum: U Typ: U Ant:

Sent: 19 Cl: 1 Index: 160 Uid: 389 those Gen: U Num: U Hum: U Typ: U Ant:

Sent: 19 Cl: 1 Index: 161 Uid: 391 Hospital Gen: U Num: S Hum: U Typ: U Ant:

Sent: 19 Cl: 2 Index: 162 Uid: 393 driver Gen: U Num: U Hum: U Typ: U Ant: 24

A.4 A Sample Evaluation File from aCAR

Sent: 19 Cl: 2 Index: 163 Uid: 396 driver Gen: U Num: U Hum: U Typ: U Ant: 275

Sent: 19 Cl: 2 Index: 164 Uid: 398 injuries Gen: N Num: P Hum: U Typ: U Ant: 386

Sent: 20 Cl: 1 Index: 165 Uid: 401 those Gen: U Num: U Hum: U Typ: U Ant:

Sent: 20 Cl: 1 Index: 166 Uid: 403 blaze Gen: U Num: U Hum: U Typ: U Ant: 319 Sent: 20

Cl: 2 Index: 167 Uid: 405 those Gen: U Num: U Hum: U Typ: U Ant:

Sent: 20 Cl: 2 Index: 168 Uid: 407 temperatures Gen: N Num: P Hum: U Typ: U Ant:

Sent: 20 Cl: 2 Index: 169 Uid: 409 1000c Gen: U Num: U Hum: U Typ: U Ant:

Sent: 20 Cl: 3 Index: 170 Uid: 411 fire Gen: U Num: U Hum: U Typ: U Ant: 11

Sent: 20 Cl: 3 Index: 171 Uid: 413 trucks Gen: N Num: P Hum: U Typ: U Ant: 9

Sent: 21 Cl: 0 Index: 172 Uid: 416 Gavin-Crook Gen: U Num: S Hum: T Typ: U Ant:

Sent: 21 Cl: 0 Index: 173 Uid: 417 who Gen: U Num: U Hum: U Typ: U Ant:

Sent: 21 Cl: 1 Index: 174 Uid: 421 who Gen: U Num: U Hum: U Typ: U Ant:

Sent: 21 Cl: 1 Index: 175 Uid: 423 where Gen: U Num: U Hum: U Typ: U Ant:

Sent: 21 Cl: 1 Index: 176 Uid: 424 crash Gen: U Num: U Hum: U Typ: U Ant: 6

Sent: 21 Cl: 2 Index: 177 Uid: 428 crash Gen: U Num: U Hum: U Typ: U Ant: 6

Sent: 21 Cl: 3 Index: 178 Uid: 433 he Gen: M Num: S Hum: T Typ: U Ant:

Sent: 21 Cl: 3 Index: 179 Uid: 435 bang Gen: U Num: U Hum: U Typ: U Ant:

Sent: 21 Cl: 4 Index: 180 Uid: 439 sound Gen: U Num: U Hum: U Typ: U Ant:

Sent: 21 Cl: 4 Index: 181 Uid: 441 cars Gen: N Num: P Hum: U Typ: U Ant:

A.4 A Sample Evaluation File from aCAR

Sent: 22 Cl: 0 Index: 182 Uid: 447 Gavin-Crook Gen: U Num: S Hum: T Typ: U Ant:

Sent: 22 Cl: 1 Index: 183 Uid: 451 it Gen: N Num: S Hum: F Typ: U Ant: -1

Sent: 22 Cl: 1 Index: 184 Uid: 454 bomb Gen: U Num: S Hum: U Typ: U Ant:

Sent: 22 Cl: 2 Index: 185 Uid: 458 amount Gen: U Num: U Hum: U Typ: U Ant:

Sent: 22 Cl: 2 Index: 186 Uid: 460 smoke Gen: U Num: U Hum: U Typ: U Ant: 8

Sent: 22 Cl: 2 Index: 187 Uid: 462 flames Gen: N Num: P Hum: U Typ: U Ant:

Sent: 22 Cl: 3 Index: 188 Uid: 467 you Gen: N Num: S Hum: T Typ: U Ant: -3

Sent: 22 Cl: 3 Index: 189 Uid: 469 shaking Gen: U Num: U Hum: U Typ: U Ant:

Sent: 23 Cl: 1 Index: 190 Uid: 472 accident Gen: U Num: U Hum: U Typ: U Ant: 6

Sent: 23 Cl: 1 Index: 191 Uid: 474 SH1 Gen: U Num: U Hum: U Typ: U Ant:

Sent: 23 Cl: 1 Index: 192 Uid: 476 Hamilton Gen: U Num: S Hum: U Typ: U Ant: 50

Sent: 23 Cl: 1 Index: 193 Uid: 478 Ngaruawahia Gen: U Num: S Hum: U Typ: U Ant:

Sent: 24 Cl: 0 Index: 194 Uid: 481 Chris-Allen Gen: U Num: S Hum: T Typ: U Ant:

Sent: 24 Cl: 1 Index: 195 Uid: 485 surface Gen: U Num: U Hum: U Typ: U Ant:

Sent: 24 Cl: 2 Index: 196 Uid: 489 stretch Gen: U Num: U Hum: U Typ: U Ant:

Sent: 24 Cl: 3 Index: 197 Uid: 494 stretch Gen: U Num: U Hum: U Typ: U Ant: 489

Sent: 24 Cl: 3 Index: 198 Uid: 496 day Gen: U Num: U Hum: U Typ: U Ant:

Sent: 24 Cl: 3 Index: 199 Uid: 498 cost Gen: U Num: U Hum: U Typ: U Ant:

Sent: 24 Cl: 3 Index: 200 Uid: 500 \$20000 Gen: U Num: U Hum: U Typ: U Ant:

A.5 Sample Text Output from aCAR

The output from the aCAR is a document level Java object with the anaphoric NPs replaced by the antecedent entity. This can be easily integrated into any Java based document processing application. The resolution data was also output into a text file for debugging purposes as well as direct human analysis of resolutions. The following is a sample output of the summary part of the text file, corresponding to the “Trucks” article in Section A.3.1

```
EntityManager : 114 DocEntites

*****Merging exact Names Begin*****

Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: Ted-Collins(42)
Merged:DocEntity : Entity16 (16) ParentNoun : Hamilton(39) Source: Hamilton(62)
Merged:DocEntity : Entity16 (16) ParentNoun : Hamilton(39) Source: Hamilton(341)
Merged:DocEntity : Entity19 (19) ParentNoun : Freightlines Truck(45)
Source: Freightlines Truck(79)
Merged:DocEntity : Entity19 (19) ParentNoun : Freightlines Truck(45)
Source: Freightlines Truck(88)
Merged:DocEntity : Entity28 (28) ParentNoun : Collins(76) Source: Collins(120)
Merged:DocEntity : Entity28 (28) ParentNoun : Collins(76) Source: Collins(127)
Merged:DocEntity : Entity28 (28) ParentNoun : Collins(76) Source: Collins(137)
Merged:DocEntity : Entity28 (28) ParentNoun : Collins(76) Source: Collins(188)
Merged:DocEntity : Entity28 (28) ParentNoun : Collins(76) Source: Collins(243)
Merged:DocEntity : Entity28 (28) ParentNoun : Collins(76) Source: Collins(309)
Merged:DocEntity : Entity53 (53) ParentNoun : Daryl-Trim(154) Source: Daryl-Trim(163)
Merged:DocEntity : Entity54 (54) ParentNoun : Senior-Hamilton Fire Station Officer(155)
Source: Senior-Hamilton Fire Station

Officer(164)
Merged:DocEntity : Entity112 (112) ParentNoun : Ngaruawahia(343)
Source: Ngaruawahia(345)
```

A.5 Sample Text Output from aCAR

Exact Names merge: 14

*****Merging exact Names End*****

*****Merging part-names Begin*****

Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: Collins(76)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: Collins(120)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: Collins(127)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: Collins(137)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: Collins(188)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: Collins(243)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: Collins(309)

Part Names merge: 1

*****Merging part-names End*****

*****Merging exact Nouns Begin*****

Merged:DocEntity : Entity1 (1) ParentNoun : trucks(4) Source: trucks(287)
Merged:DocEntity : Entity3 (3) ParentNoun : inferno(8) Source: inferno(10)
Merged:DocEntity : Entity15 (15) ParentNoun : smoke(37) Source: smoke(319)
Merged:DocEntity : Entity15 (15) ParentNoun : smoke(37) Source: smoke(323)
Merged:DocEntity : Entity21 (21) ParentNoun : saw(49) Source: saw(52)
Merged:DocEntity : Entity30 (30) ParentNoun : 51-year-old contractor(83)
Source: 51-year-old contractor(93)
Merged:DocEntity : Entity42 (42) ParentNoun : cab(122) Source: cab(269)
Merged:DocEntity : Entity61 (61) ParentNoun : that(173) Source: that(247)
Merged:DocEntity : Entity82 (82) ParentNoun : blaze(264) Source: blaze(274)
Merged:DocEntity : Entity82 (82) ParentNoun : blaze(264) Source: blaze(279)
Merged:DocEntity : Entity102 (102) ParentNoun : amount(321)

A.5 Sample Text Output from aCAR

Source: amount(317)

Exact Nouns merge: 11

*****Merging exact Nouns End*****

*****Merging Pnouns Begin*****

Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: his(25)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: he(107)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: he(113)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: his(123)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: his(132)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: I(141)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: I(145)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: I(192)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: I(199)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: my(212)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: I(229)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: I(235)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: me(256)
Merged:DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13) Source: his(270)
Merged:DocEntity : Entity20 (20) ParentNoun : State-Highway 1(47) Source: it(55)
Merged:DocEntity : Entity36 (36) ParentNoun : bit(101) Source: it(115)
Merged:DocEntity : Entity53 (53) ParentNoun : Daryl-Trim(154) Source: he(171)
Merged:DocEntity : Entity69 (69) ParentNoun : one(209) Source: it(223)
Merged:DocEntity : Entity97 (97) ParentNoun : sound(301) Source: it(313)

PNouns merge: 11

*****Merging PNouns End*****

A.5 Sample Text Output from aCAR

```
*****Making Collection Entities List Begin*****

Parent Entity : Entity1 (1) ParentNoun : trucks(4)
Parent Entity : Entity6 (6) ParentNoun : drivers(15)
Parent Entity : Entity10 (10) ParentNoun : cab moments(24)
Parent Entity : Entity12 (12) ParentNoun : explosions(28)
Parent Entity : Entity35 (35) ParentNoun : 40 tonnes(99)
Parent Entity : Entity55 (55) ParentNoun : actions(158)
Parent Entity : Entity63 (63) ParentNoun : consequences(180)
Parent Entity : Entity70 (70) ParentNoun : truck friends(211)
Parent Entity : Entity90 (90) ParentNoun : temperatures(281)
Parent Entity : Entity92 (92) ParentNoun : trucks(287)
Parent Entity : Entity98 (98) ParentNoun : cars(304)
Parent Entity : Entity106 (106) ParentNoun : flames(327)

EntityManager : 114 CollectionEntites

*****End Collection Entities List*****


DocEntity : Entity1 (1) ParentNoun : trucks(4) ChildDnouns : trucks(287)
DocEntity : Entity2 (2) ParentNoun : crash(6)
DocEntity : Entity3 (3) ParentNoun : inferno(8) ChildDnouns : inferno(10)
DocEntity : Entity5 (5) ParentNoun : Ted-Collins(13)
ChildDnouns : Ted-Collins(42) Collins(76) Collins(120) Collins(127) Collins

(137) Collins(188) Collins(243) Collins(309) his(25) he(107) he(113)
his(123) his(132) I(141) I(145) I(192) I(199) my(212) I(229) I

(235) me(256) his(270)

DocEntity : Entity6 (6) ParentNoun : drivers(15)
DocEntity : Entity7 (7) ParentNoun : dairy worker(18)
DocEntity : Entity8 (8) ParentNoun : truck driver(20)
DocEntity : Entity9 (9) ParentNoun : wreckage(22)
```

A.5 Sample Text Output from aCAR

DocEntity : Entity10 (10) ParentNoun : cab moments(24)

DocEntity : Entity12 (12) ParentNoun : explosions(28)

DocEntity : Entity13 (13) ParentNoun : plume(35)

DocEntity : Entity14 (14) ParentNoun : area(30)

DocEntity : Entity15 (15) ParentNoun : smoke(37)

ChildDnouns : smoke(319) smoke(323)

DocEntity : Entity16 (16) ParentNoun : Hamilton(39)

ChildDnouns : Hamilton(62) Hamilton(341)

DocEntity : Entity18 (18) ParentNoun : fonterra driver(43)

DocEntity : Entity19 (19) ParentNoun : Freightlines Truck(45)

ChildDnouns : Freightlines Truck(79) Freightlines Truck(88)

DocEntity : Entity20 (20) ParentNoun : State-Highway 1(47)

ChildDnouns : it(55)

DocEntity : Entity21 (21) ParentNoun : saw(49) ChildDnouns : saw(52)

DocEntity : Entity24 (24) ParentNoun : Shell Diesel Tanker(57)

DocEntity : Entity25 (25) ParentNoun : Horotiu(59)

DocEntity : Entity26 (26) ParentNoun : north(60)

DocEntity : Entity30 (30) ParentNoun : 51-year-old contractor(83)

ChildDnouns : 51-year-old contractor(93)

DocEntity : Entity31 (31) ParentNoun : diesel tanker(85)

DocEntity : Entity34 (34) ParentNoun : you(97)

DocEntity : Entity35 (35) ParentNoun : 40 tonnes(99)

DocEntity : Entity36 (36) ParentNoun : bit(101) ChildDnouns : it(115)

DocEntity : Entity37 (37) ParentNoun : momentum(103)

DocEntity : Entity42 (42) ParentNoun : cab(122) ChildDnouns : cab(269)

DocEntity : Entity45 (45) ParentNoun : aid(129)

DocEntity : Entity49 (49) ParentNoun : what(143)

DocEntity : Entity51 (51) ParentNoun : somebody(147)

DocEntity : Entity52 (52) ParentNoun : it(149)

DocEntity : Entity53 (53) ParentNoun : Daryl-Trim(154)

ChildDnouns : Daryl-Trim(163) he(171)

A.5 Sample Text Output from aCAR

DocEntity : Entity54 (54) ParentNoun : Senior-Hamilton Fire Station Officer(155)
ChildDnouns : Senior-Hamilton Fire Station

Officer(164)

DocEntity : Entity55 (55) ParentNoun : actions(158)
DocEntity : Entity56 (56) ParentNoun : heroic(160)
DocEntity : Entity59 (59) ParentNoun : it(168)
DocEntity : Entity61 (61) ParentNoun : that(173) ChildDnouns : that(247)
DocEntity : Entity62 (62) ParentNoun : it(177)
DocEntity : Entity63 (63) ParentNoun : consequences(180)
DocEntity : Entity64 (64) ParentNoun : it(184)
DocEntity : Entity67 (67) ParentNoun : it(198)
DocEntity : Entity69 (69) ParentNoun : one(209) ChildDnouns : it(223)
DocEntity : Entity70 (70) ParentNoun : truck friends(211)
DocEntity : Entity73 (73) ParentNoun : fuel tanker(225)
DocEntity : Entity78 (78) ParentNoun : thing(249)
DocEntity : Entity80 (80) ParentNoun : it(260)
DocEntity : Entity81 (81) ParentNoun : hell(262)
DocEntity : Entity82 (82) ParentNoun : blaze(264) ChildDnouns : blaze(274) blaze(279)
DocEntity : Entity83 (83) ParentNoun : 28-year-old tanker driver(267)
DocEntity : Entity86 (86) ParentNoun : side(272)
DocEntity : Entity88 (88) ParentNoun : those(277)
DocEntity : Entity90 (90) ParentNoun : temperatures(281)
DocEntity : Entity91 (91) ParentNoun : fire(285)
DocEntity : Entity93 (93) ParentNoun : Gavin-Crook(290)
DocEntity : Entity94 (94) ParentNoun : he(293)
DocEntity : Entity95 (95) ParentNoun : bang(295)
DocEntity : Entity96 (96) ParentNoun : he(299)
DocEntity : Entity97 (97) ParentNoun : sound(301) ChildDnouns : it(313)
DocEntity : Entity98 (98) ParentNoun : cars(304)
DocEntity : Entity101 (101) ParentNoun : bomb(315)

A.5 Sample Text Output from aCAR

DocEntity : Entity102 (102) ParentNoun : amount(321) ChildDnouns : amount(317)

DocEntity : Entity106 (106) ParentNoun : flames(327)

DocEntity : Entity107 (107) ParentNoun : you(332)

DocEntity : Entity108 (108) ParentNoun : shaking(334)

DocEntity : Entity109 (109) ParentNoun : accident(337)

DocEntity : Entity110 (110) ParentNoun : SH1(339)

DocEntity : Entity112 (112) ParentNoun : Ngaruawahia(343)

ChildDnouns : Ngaruawahia(345)

DocEntity : Entity114 (114) ParentNoun : 70m stretch(351)

Total no. of Nouns: 114

Exact Names merge: 14

Part Names merge: 1

Exact Nouns merge: 11

ProNouns merge: 19

Total merged Nouns: 69

Percentage Nouns merged: 40

-----Summary of All Pronominal Anaphora in Document-----

Sent: 2	Cl: 1	Index: 10	Uid: 25	Anap: his	Ant: Ted-Collins	CorrectAnt: truck driver
Sent: 4	Cl: 2	Index: 22	Uid: 55	Anap: it	Ant: State-Highway 1	CorrectAnt: Freightlines Truck
Sent: 7	Cl: 2	Index: 33	Uid: 97	Anap: you	Ant:	CorrectAnt: READER
Sent: 7	Cl: 4	Index: 37	Uid: 107	Anap: he	Ant: 51-year-old contractor	CorrectAnt: 51-year-old
Sent: 7	Cl: 3	Index: 38	Uid: 113	Anap: he	Ant: he	CorrectAnt: he
Sent: 7	Cl: 3	Index: 39	Uid: 115	Anap: it	Ant: bit	CorrectAnt: diesel tanker
Sent: 8	Cl: 2	Index: 42	Uid: 123	Anap: his	Ant: Collins	CorrectAnt: Collins
Sent: 8	Cl: 1	Index: 45	Uid: 132	Anap: his	Ant: Collins	CorrectAnt: Collins
Sent: 9	Cl: 1	Index: 47	Uid: 141	Anap: I	Ant: Collins	CorrectAnt: Collins

A.5 Sample Text Output from aCAR

Sent: 9	Cl: 2	Index: 49	Uid: 145	Anap: I	Ant: I	CorrectAnt: I
Sent: 9	Cl: 3	Index: 51	Uid: 149	Anap: it	Ant:	CorrectAnt:
Sent: 11	Cl: 1	Index: 58	Uid: 168	Anap: it	Ant:	CorrectAnt: PLEONASTIC
Sent: 11	Cl: 1	Index: 59	Uid: 171	Anap: he	Ant: Daryl-Trim	CorrectAnt: I
Sent: 11	Cl: 2	Index: 61	Uid: 177	Anap: it	Ant:	CorrectAnt: NOUN-PHRASE
Sent: 11	Cl: 1	Index: 63	Uid: 184	Anap: it	Ant:	CorrectAnt: PLEONASTIC
Sent: 12	Cl: 1	Index: 65	Uid: 192	Anap: I	Ant: Collins	CorrectAnt: Collins
Sent: 12	Cl: 2	Index: 66	Uid: 198	Anap: it	Ant:	CorrectAnt: NOUN-PHRASE
Sent: 12	Cl: 2	Index: 67	Uid: 199	Anap: I	Ant: I	CorrectAnt: Collins
Sent: 12	Cl: 4	Index: 70	Uid: 212	Anap: my	Ant: I	CorrectAnt: I
Sent: 12	Cl: 2	Index: 71	Uid: 223	Anap: it	Ant: one	CorrectAnt: diesel tanker
Sent: 12	Cl: 1	Index: 73	Uid: 229	Anap: I	Ant: my	CorrectAnt: my
Sent: 12	Cl: 3	Index: 74	Uid: 235	Anap: I	Ant: I	CorrectAnt: I
Sent: 13	Cl: 2	Index: 78	Uid: 256	Anap: me	Ant: Collins	CorrectAnt: Collins
Sent: 13	Cl: 1	Index: 79	Uid: 260	Anap: it	Ant:	CorrectAnt: NOUN-PHRASE
Sent: 14	Cl: 1	Index: 84	Uid: 270	Anap: his	Ant: Collins	CorrectAnt: 28-year-old tanker drive
Sent: 17	Cl: 1	Index: 93	Uid: 293	Anap: he	Ant:	CorrectAnt: Gavin-Crook
Sent: 17	Cl: 2	Index: 95	Uid: 299	Anap: he	Ant: he	CorrectAnt: he
Sent: 18	Cl: 2	Index: 99	Uid: 313	Anap: it	Ant: sound	CorrectAnt: bang
Sent: 18	Cl: 1	Index: 106	Uid: 332	Anap: you	Ant:	CorrectAnt: READER

-----Summary of attempted Pronominal anaphora in Document-----

Sent: 2	Cl: 1	Index: 10	Uid: 25	Anap: his	Ant: truck driver(20)	truck driver(20)
CorrectAnt: truck driver(20)---CORRECT						
Sent: 4	Cl: 2	Index: 22	Uid: 55	Anap: it	Ant: State-Highway 1 (47)	State-Highway 1(47) Correct
)---INCORRECT						

A.5 Sample Text Output from aCAR

Sent: 7 Cl: 4 Index: 37 Uid: 107 Anap: he Ant: 51-year-old contractor(93)
51-year-old contractor(93) CorrectAnt: 51-year-

old contractor(93)---CORRECT

Sent: 7 Cl: 3 Index: 38 Uid: 113 Anap: he Ant: he (107)Collins(76)
CorrectAnt: he(107)---CORRECT

Sent: 7 Cl: 3 Index: 39 Uid: 115 Anap: it Ant: diesel tanker(85) diesel tanker(85)
CorrectAnt: diesel tanker(85)---

CORRECT

Sent: 8 Cl: 2 Index: 42 Uid: 123 Anap: his Ant: Collins (120)Collins(120)
CorrectAnt: Collins(120)---CORRECT

Sent: 8 Cl: 1 Index: 45 Uid: 132 Anap: his Ant: Collins (127)Collins(127)
CorrectAnt: Collins(127)---CORRECT

Sent: 9 Cl: 1 Index: 47 Uid: 141 Anap: I Ant: Collins (137)Collins(137)
CorrectAnt: Collins(137)---CORRECT

Sent: 9 Cl: 2 Index: 49 Uid: 145 Anap: I Ant: I (141)Collins(137)
CorrectAnt: I(141)---CORRECT

Sent: 11 Cl: 1 Index: 59 Uid: 171 Anap: he Ant: I(145)
CorrectAnt: I(145)---CORRECT

Sent: 12 Cl: 1 Index: 65 Uid: 192 Anap: I Ant: Collins (188)Collins(188)
CorrectAnt: Collins(188)---CORRECT

Sent: 12 Cl: 2 Index: 67 Uid: 199 Anap: I Ant: I (192)Collins(188)
CorrectAnt: Collins(188)---CORRECT

Sent: 12 Cl: 4 Index: 70 Uid: 212 Anap: my Ant: I (199)Collins(188)
CorrectAnt: I(199)---CORRECT

Sent: 12 Cl: 2 Index: 71 Uid: 223 Anap: it Ant: one (209)one(209)
CorrectAnt: diesel tanker(85)---INCORRECT

Sent: 12 Cl: 1 Index: 73 Uid: 229 Anap: I Ant: my (212)Collins(188)
CorrectAnt: my(212)---CORRECT

Sent: 12 Cl: 3 Index: 74 Uid: 235 Anap: I Ant: I (229)Collins(188)

A.5 Sample Text Output from aCAR

CorrectAnt: I(229)---CORRECT

Sent: 13 Cl: 2 Index: 78 Uid: 256 Anap: me Ant: Collins (243)Collins(243)

CorrectAnt: Collins(243)---CORRECT

Sent: 14 Cl: 1 Index: 84 Uid: 270 Anap: his Ant: 28-year-old tanker driver(267)

28-year-old tanker driver(267) CorrectAnt:

28-year-old tanker driver(267)---CORRECT

Sent: 17 Cl: 1 Index: 93 Uid: 293 Anap: he Ant: Gavin-Crook(290)

Gavin-Crook(290)

CorrectAnt: Gavin-Crook(290)---CORRECT

Sent: 17 Cl: 2 Index: 95 Uid: 299 Anap: he Ant: he (293)he(293)

CorrectAnt: he(293)---CORRECT

Sent: 18 Cl: 2 Index: 99 Uid: 313 Anap: it Ant: sound(301) sound(301)

CorrectAnt: bang(295)---INCORRECT

Total Number of Sentences : 20

Total Number of Clauses : 54

Total Number of Nouns : 114

		RECALL	PRECISION
Anaphors	:	29	21
Correctly Resolved	:	19(66%)	19(90%)
Incorrectly Resolved	:	10(34%)	2(10%)

-----Summary of Modifiers and Nouns -----

Sent: 1 Cl: 1 Index: 2 Uid: 8 the inferno

Sent: 1 Cl: 2 Index: 3 Uid: 10 the inferno

Sent: 1 Cl: 3 Index: 5 Uid: 15 one of the drivers

Sent: 2 Cl: 1 Index: 6 Uid: 18 a dairy worker

Sent: 2 Cl: 1 Index: 7 Uid: 20 an injured truck driver

Sent: 2 Cl: 1 Index: 8 Uid: 22 the wreckage

A.5 Sample Text Output from aCAR

Sent: 3 Cl: 1 Index: 12 Uid: 35 a plume
Sent: 3 Cl: 1 Index: 13 Uid: 30 the surrounding area
Sent: 3 Cl: 1 Index: 14 Uid: 37 thick black smoke
Sent: 6 Cl: 1 Index: 29 Uid: 83 a 51-year-old contractor
Sent: 6 Cl: 1 Index: 30 Uid: 85 the diesel tanker
Sent: 7 Cl: 0 Index: 32 Uid: 93 a 51-year-old contractor
Sent: 7 Cl: 2 Index: 35 Uid: 101 a bit
Sent: 8 Cl: 1 Index: 44 Uid: 129 the aid
Sent: 12 Cl: 2 Index: 72 Uid: 225 a fuel tanker
Sent: 13 Cl: 1 Index: 77 Uid: 249 the thing
Sent: 13 Cl: 1 Index: 80 Uid: 262 a hell
Sent: 13 Cl: 1 Index: 81 Uid: 264 a blaze
Sent: 14 Cl: 1 Index: 82 Uid: 267 the 28-year-old tanker driver
Sent: 14 Cl: 1 Index: 85 Uid: 272 the other side
Sent: 14 Cl: 1 Index: 86 Uid: 274 the blaze
Sent: 15 Cl: 1 Index: 88 Uid: 279 the blaze
Sent: 16 Cl: 1 Index: 90 Uid: 285 the fire
Sent: 17 Cl: 1 Index: 94 Uid: 295 a loud bang
Sent: 17 Cl: 2 Index: 96 Uid: 301 the sound
Sent: 18 Cl: 2 Index: 100 Uid: 315 a bomb
Sent: 18 Cl: 2 Index: 101 Uid: 321 a big amount
Sent: 18 Cl: 2 Index: 102 Uid: 317 a big amount
Sent: 18 Cl: 1 Index: 107 Uid: 334 the ground shaking
Sent: 19 Cl: 1 Index: 108 Uid: 337 the accident
Sent: 20 Cl: 2 Index: 113 Uid: 351 a 70m stretch
\\end{verbaitim}

-----Summary of attempted anaphors in Document-----

Sent: 2 Cl: 1 Index: 10 Uid: 25 Anap: his Ant: Ted-Collins (13)Ted-Collins(13)
CorrectAnt: truck driver(20)---INCORRECT

A.5 Sample Text Output from aCAR

Sent: 4 Cl: 2 Index: 22 Uid: 55 Anap: it Ant: State-Highway 1 (47)
State-Highway 1(47) CorrectAnt: Freightlines Truck(45)---INCORRECT

Sent: 7 Cl: 4 Index: 37 Uid: 107 Anap: he Ant: Collins (76)Collins(76)
CorrectAnt: 51-year-old contractor(93)---INCORRECT

Sent: 7 Cl: 3 Index: 38 Uid: 113 Anap: he Ant: he (107)Collins(76)
CorrectAnt: he(107)---CORRECT

Sent: 7 Cl: 3 Index: 39 Uid: 115 Anap: it Ant: bit (101)bit(101)
CorrectAnt: diesel tanker(85)---INCORRECT

Sent: 8 Cl: 2 Index: 42 Uid: 123 Anap: his Ant: Collins (120)Collins(120)
CorrectAnt: Collins(120)---CORRECT

Sent: 8 Cl: 1 Index: 45 Uid: 132 Anap: his Ant: Collins (127)Collins(127)
CorrectAnt: Collins(127)---CORRECT

Sent: 9 Cl: 1 Index: 47 Uid: 141 Anap: I Ant: Collins (137)Collins(137)
CorrectAnt: Collins(137)---CORRECT

Sent: 9 Cl: 2 Index: 49 Uid: 145 Anap: I Ant: I (141)Collins(137)
CorrectAnt: I(141)---CORRECT

Sent: 11 Cl: 1 Index: 59 Uid: 171 Anap: he Ant: Daryl-Trim (163)Daryl-Trim(163)
CorrectAnt: I(145)---INCORRECT

Sent: 12 Cl: 1 Index: 65 Uid: 192 Anap: I Ant: Collins (188)Collins(188)
CorrectAnt: Collins(188)---CORRECT

Sent: 12 Cl: 2 Index: 67 Uid: 199 Anap: I Ant: I (192)Collins(188)
CorrectAnt: Collins(188)---INCORRECT

Sent: 12 Cl: 4 Index: 70 Uid: 212 Anap: my Ant: I (199)Collins(188)
CorrectAnt: I(199)---CORRECT

Sent: 12 Cl: 2 Index: 71 Uid: 223 Anap: it Ant: one (209)one(209)
CorrectAnt: diesel tanker(85)---INCORRECT

Sent: 12 Cl: 1 Index: 73 Uid: 229 Anap: I Ant: my (212)Collins(188)
CorrectAnt: my(212)---CORRECT

Sent: 12 Cl: 3 Index: 74 Uid: 235 Anap: I Ant: I (229)Collins(188)
CorrectAnt: I(229)---CORRECT

Sent: 13 Cl: 2 Index: 78 Uid: 256 Anap: me Ant: Collins (243)Collins(243)

A.5 Sample Text Output from aCAR

CorrectAnt: Collins(243)---CORRECT
Sent: 14 Cl: 1 Index: 84 Uid: 270 Anap: his Ant: Collins (243)Collins(243)
CorrectAnt: 28-year-old tanker driver(267)---INCORRECT
Sent: 17 Cl: 1 Index: 93 Uid: 293 Anap: he Ant:
CorrectAnt: Gavin-Crook(290)---INCORRECT
Sent: 17 Cl: 2 Index: 95 Uid: 299 Anap: he Ant: he (293)he(293)
CorrectAnt: he(293)---CORRECT
Sent: 18 Cl: 2 Index: 99 Uid: 313 Anap: it Ant: sound (301)sound(301)
CorrectAnt: bang(295)---INCORRECT

Total Number of Sentences : 20

Total Number of Clauses : 54

Total Number of Nouns : 114

	RECALL	PRECISION	
Anaphors	: 29	21	(ignore 8)
Resolved	: 20(68\%)	20(95\%)	(unresolved 1)
Correctly Resolved	: 11(37\%)	11(52\%)	
Incorrectly Resolved	: 9(31\%)	9(42\%)	

-----Summary of Modifiers and Nouns -----

Sent: 1 Cl: 1 Index: 2 Uid: 8 the inferno
Sent: 1 Cl: 2 Index: 3 Uid: 10 the inferno
Sent: 1 Cl: 3 Index: 5 Uid: 15 one of the drivers
Sent: 2 Cl: 1 Index: 6 Uid: 18 a dairy worker
Sent: 2 Cl: 1 Index: 7 Uid: 20 an injured truck driver
Sent: 2 Cl: 1 Index: 8 Uid: 22 the wreckage
Sent: 3 Cl: 1 Index: 12 Uid: 35 a plume
Sent: 3 Cl: 1 Index: 13 Uid: 30 the surrounding area
Sent: 3 Cl: 1 Index: 14 Uid: 37 thick black smoke
Sent: 6 Cl: 1 Index: 29 Uid: 83 a 51-year-old contractor

A.5 Sample Text Output from aCAR

Sent: 6 Cl: 1 Index: 30 Uid: 85 the diesel tanker
Sent: 7 Cl: 0 Index: 32 Uid: 93 a 51-year-old contractor
Sent: 7 Cl: 2 Index: 35 Uid: 101 a bit
Sent: 8 Cl: 1 Index: 44 Uid: 129 the aid
Sent: 12 Cl: 2 Index: 72 Uid: 225 a fuel tanker
Sent: 13 Cl: 1 Index: 77 Uid: 249 the thing
Sent: 13 Cl: 1 Index: 80 Uid: 262 a hell
Sent: 13 Cl: 1 Index: 81 Uid: 264 a blaze
Sent: 14 Cl: 1 Index: 82 Uid: 267 the 28-year-old tanker driver
Sent: 14 Cl: 1 Index: 85 Uid: 272 the other side
Sent: 14 Cl: 1 Index: 86 Uid: 274 the blaze
Sent: 15 Cl: 1 Index: 88 Uid: 279 the blaze
Sent: 16 Cl: 1 Index: 90 Uid: 285 the fire
Sent: 17 Cl: 1 Index: 94 Uid: 295 a loud bang
Sent: 17 Cl: 2 Index: 96 Uid: 301 the sound
Sent: 18 Cl: 2 Index: 100 Uid: 315 a bomb
Sent: 18 Cl: 2 Index: 101 Uid: 321 a big amount
Sent: 18 Cl: 2 Index: 102 Uid: 317 a big amount
Sent: 18 Cl: 1 Index: 107 Uid: 334 the ground shaking
Sent: 19 Cl: 1 Index: 108 Uid: 337 the accident
Sent: 20 Cl: 2 Index: 113 Uid: 351 a 70m stretch

Bibliography

- Abad, A., Bentivogli, L., Dagan, I., Giampiccolo, D., Mirkin, S., Pianta, E., and Stern, A. (2010). A resource for investigating the impact of anaphora and coreference on inference. 5, 8, 61
- Abbott, B. (2005). Definite and indefinite. *Encyclopedia of language and linguistics*, 2nd ed. New York: Elsevier. 15
- Akman, V. and Ersan, E. (1994). Focusing for pronoun resolution in discourse: An implementation. Technical report, Bilkent University, Bilkent, Ankara 06533, Turkey. 20
- Asher, N. (1993). *Reference to Abstract Objects in Discourse*. Kluwer Academic Publisher. 7
- Asher, N. and Lascarides, A. (1998). Bridging. *Journal of Semantics*, 15(1):83–113. 6
- Baldwin, B. (1997). Cogniac: high precision coreference with limited knowledge and linguistic resources. In *Proceedings of the ACL*, volume 97, pages 38–45. 32, 116
- Barbu, C. and Mitkov, R. (2001). Evaluation tool for rule-based anaphora resolution methods. In *Annual Meeting–Association for Computational Linguistics*, volume 39, pages 34–41. 26
- Barker, K. and Szpakowicz, S. (1998). Semi-automatic recognition of noun modifier relationships. 59, 60

- Brennan, S. E., Friedman, M. W., and Pollard, C. J. (1987). A centering approach to pronouns. In *Proceedings of the 25th annual meeting on Association for Computational Linguistics*, ACL '87, pages 155–162, Stroudsburg, PA, USA. Association for Computational Linguistics. 2, 5, 36
- Bunescu, R. (2003). Associative anaphora resolution: A web-based approach. In *In Proceedings of the EACL2003 Workshop on the Computational Treatment of Anaphora*, pages 47–52. 4
- Butnariu, C. and Veale, T. (2008). A concept-centered approach to noun-compound interpretation. In *Proceedings of the 22nd International Conference on Computational Linguistics - Volume 1*, COLING '08, pages 81–88, Stroudsburg, PA, USA. Association for Computational Linguistics. 8, 59
- Butterworth, B. (1974). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research*, 4:75–81. 2
- Canas, A., Valerio, A., Lalinde-Pulido, J., Carvalho, M., and Arguedas, M. (2003). Using wordnet for word sense disambiguation to support concept map construction. In *String Processing and Information Retrieval*, pages 350–359. Springer. 83, 85
- Carbonell, J. G. and Brown, R. D. (1988). Anaphora resolution: a multi-strategy approach. In *Proceedings of the 12th conference on Computational linguistics - Volume 1*, COLING '88, pages 96–101, Stroudsburg, PA, USA. Association for Computational Linguistics. 16, 25
- Carter, D. (1987). *Interpreting anaphors in natural language texts*. Halsted Press, New York, NY, USA. 24
- Chan, S. W. K. and T'sou, B. K. (1999). Semantic inference for anaphora resolution: Toward a framework in machine translation. *Machine Translation*, 14:163–190. 10.1023/A:1011138323599. 4
- Chomsky, N. (1969). *Aspects of the Theory of Syntax*. The MIT Press. 18, 57, 117

BIBLIOGRAPHY

- Clark, H. H. (1977). *Thinking: Readings in cognitive science*. Cambridge University Press, Cambridge. 7
- Clark, H. H. and Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory and Cognition*, 1(7):35–41. 2
- Dahl, D. (1984). *The structure and function of One-Anaphora in English*. University of Minnesota. 15
- Dahl, D. A. (1986). Focusing and reference resolution in pundit. In *Proceedings of the workshop on Strategic computing natural language*, HLT '86, pages 114–126, Morristown, NJ, USA. Association for Computational Linguistics. 15, 35
- Dahl, D. A., Palmer, M. S., and Passonneau, R. J. (1987). Nominalizations in pundit. In *Proceedings of the 25th annual meeting on Association for Computational Linguistics*, pages 131–139, Morristown, NJ, USA. Association for Computational Linguistics. 16, 35, 66
- Davies, M. (2010). Corpus of contemporary american english. <http://www.americancorpus.org>. 114
- Downing, P. (1977). On the creation and use of english compound nouns. *Language*, 53(4):810–842. xiii, 57, 58, 59
- Fabb, N. (1988). English suffixation is constrained only by selectional restrictions. *Natural Language & Linguistic Theory*, 6:527–539. 68
- Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. Bradford Books. 4
- Fox, B. (1993). *Discourse structure and anaphora: Written and conversational English*. Cambridge Univ Pr. 21, 38
- Fragos, K., Maistros, Y., and Skourlas, C. (2003). Word sense disambiguation using wordnet relations. In *First Balkan Conference in Informatics, Thessaloniki*. Citeseer. 83, 85

BIBLIOGRAPHY

- Fraurud, K. (1990). Definiteness and the processing of noun phrases in natural discourse. *Journal of Semantics*, 7(4):395–433. 37, 38
- Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Psychology Pr. 22
- Girju, R. (2003). Automatic detection of causal relations for question answering. In *Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering-Volume 12*, page 83. Association for Computational Linguistics. 103, 104, 105
- Girju, R. (2007). Improving the interpretation of noun phrases with crosslinguistic information. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 568–575. 60
- Girju, R., Moldovan, D., Tatu, M., and Antohe, D. (2005). On the semantics of noun compounds. *Computer Speech and Language*, 19(4):479–496. 59, 60, 78
- Girju, R., Nakov, P., Nastase, V., Szpakowicz, S., Turney, P., and Yuret, D. (2009). Classification of semantic relations between nominals. *Language Resources and Evaluation*, 43:105–121. 10.1007/s10579-009-9083-2. 8
- Gordon, P. C., Grosz, B. J., and Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17:311–348. 2, 52
- Grosz, B. J., Joshi, A. K., and Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, 21(2):202–225. 19
- Grosz, B. J. and Sidner, C. L. (1986). Attention, intention, and the structure of discourse. *Computational Linguistics*, 12(3):175–204. 5, 14, 19, 20, 21, 35, 38, 44
- Halliday, M. A. K. and Hasan, R. (1976). *Cohesion in English*. London: Longman. 5, 22

BIBLIOGRAPHY

- Hawkins, J. (1978). *Definiteness and indefiniteness: a study in reference and grammaticality prediction*. Taylor & Francis. 37
- Heim, I. (1982). The semantics of definite and indefinite noun phrases. *Electronic Doctoral Dissertations for UMass Amherst*. 15
- Hendrickx, I., Kim, S. N., Kozareva, Z., Nakov, P., Séaghdha, D. O., Padó, S., Pennacchiotti, M., Romano, L., and Szpakowicz, S. (2010). Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. In *Proceedings of the 5th International Workshop on Semantic Evaluation, SemEval '10*, pages 33–38, Stroudsburg, PA, USA. Association for Computational Linguistics. 8, 61
- Hirst, G. (1981). *Anaphora in natural language understanding*. 14, 24
- Hobbs, J. R. (1979). Coherence and coreference. *Cognitive Science*, 67:67–90. 2, 22
- Ide, N. and Cristea, D. (2000). A hierarchical account of referential accessibility. In *ACL '00: Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*, pages 416–424, Morristown, NJ, USA. Association for Computational Linguistics. 20
- Ingria, R. J. and Stallard, D. (1989). A computational mechanism for pronominal reference. 18
- Jerry, H. (1978). Resolving pronoun references. *Lingua*, 44:311–338. 23, 116
- Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Harvard University Press, Cambridge, MA, USA. 2, 22
- Kameyama, M. (1986). A property-sharing constraint in centering. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pages 200–206. 2

- Kaplan, R. and Berry-Rogghe, G. (1991). Knowledge-based acquisition of causal relationships in text. *Knowledge Acquisition*, 3(3):317–337. 102
- Kasper, R. T., Davis, P. C., and Roberts, C. (1999). An integrated approach to reference and presupposition resolution. In *In The Relation of Discourse/Dialogue Structure and Reference: Workshop at the 37th Annual Meeting of the Association for Computational Linguistics (ACL-99)*, pages 1–10. 5, 36
- Kaufmann, M., editor (1995). *MUC-6*, San Francisco, CA. Proceedings of the 6th conference on Message understanding (MUC-6). 43
- Kaufmann, M., editor (1998). *MUC-7*, San Francisco, CA. Proceedings of the 7th conference on Message understanding (MUC-7). 43
- Kehler, A. (1993). The effect of establishing coherence in ellipsis and anaphora resolution. In *Proceedings of the 31st Conference of the Association for Computational Linguistics (ACL-93)*., pages 62–69. 2
- Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. CSLI publications Stanford, Calif. 22
- Kennedy, C. and Boguraev, B. (1996). Anaphora for everyone: Pronominal anaphora resolution without a parser. In *Proceedings of the 16th International Conference on Computational Linguistics (COLING'96)*, pages 113–118. 31, 131
- Khoo, C., Chan, S., and Niu, Y. (2000). Extracting causal knowledge from a medical database using graphical patterns. In *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*, page 343. Association for Computational Linguistics. 102
- Kim, J. (1971). Causes and events: Mackie on causation. *The Journal of Philosophy*, 68(14):426–441. 102, 103

- Kim, S. N. and Baldwin, T. (2005). Automatic interpretation of noun compounds using wordnet similarity. In *In Proceedings of the 2nd International Joint Conference on Natural Language Processing, Jeju Island, South Korea, 1113*, pages 945–956. 59
- Kim, S. N. and Baldwin, T. (2006). Interpreting semantic relations in noun compounds via verb semantics. In *Proceedings of the COLING/ACL on Main conference poster sessions, COLING-ACL '06*, pages 491–498, Stroudsburg, PA, USA. Association for Computational Linguistics. 8
- Kim, S. N. and Nakov, P. (2011). Large-scale noun compound interpretation using bootstrapping and the web as a corpus. In *The Proceeding of Conference on Empirical Methods in Natural Language Processing, EMNLP*, Edinburgh, UK. 8, 59, 61
- Lappin, S. (2003). A sequenced model of anaphora and ellipsis resolution. 4
- Lappin, S. and Leass, H. (1994). An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20(4):535–561. 10, 29, 31, 116
- Lees, R. B. (1966). *On a Transformational Analysis of Compounds: A Reply to Hans Marchand* Lees, Robert B. Indogermanische Forschungen, Berlin, New York (Walter de Gruyter). 55, 57, 61
- Levi, J. (1973). Where do all those other adjectives come from. In *Chicago Linguistic Society*, volume 9, pages 332–354. 48, 57, 61
- Levi, J. N. (1978). *The syntax and semantics of complex nominals*. Academic Press, New York :. 55, 59, 60, 61, 62, 64, 66, 145
- Li, C. N. (1971). *Semantics and the Structure of Compounds in Chinese*. PhD thesis, University of Carlifornia dissertation. 57
- Li, Y., Musilek, P., Reformat, M., and Wyard-Scott, L. (2009). Identification of pleonastic it using the web. *Journal of Artificial Intelligence Research.*, 34(1):339–389. 15

- Livia Polanyi, M. v. d. B. (1999). Logical structure and discourse anaphora resolution. In *Proceedings of ACL99 Workshop on the Relation of Discourse/Dialogue Structure and Reference.*, pages 110–117, College Park, USA. 4
- Makhoul, J., Kubala, F., Schwartz, R., and Weischedel, R. (1999). Performance measures for information extraction. In *Proceedings of DARPA Broadcast News Workshop*, pages 33–38. 129
- Mann, W. C. and Thompson, S. A. (1988). Rhetorical structure theory: Towards a functional theory of text organization. *Text*, 8(3):243–281. 2
- Marie, M. A. (2006). The resolution of bridging anaphora in ot. Technical report. 6
- Markert, K. and Nissim, M. (2003). Using the web for nominal anaphora resolution. In *EACL Workshop on the Computational Treatment of Anaphora*, pages 39–46. 131
- Markert, K. and Nissim, M. (2005). Comparing knowledge sources for nominal anaphora resolution. *Computational Linguistics*, 31:367–401. 5
- Marslen-Wilson, W., Levy, E., and Komisarjevsky Tyler, L. (1982). Producing interpretable discourse: The establishment and maintenance of reference. *Language*, pages 339–378. 2
- McCarthy, J. and Lehnert, W. (1995). Using decision trees for coreference resolution. *Arxiv preprint cmp-lg/9505043*. 43
- Mccord, M. (1990). Anaphora resolution in slot grammar. *Computational Linguistics*, 16:197–212. 2, 30
- Mitkov, R. (1993). A knowledge-based and sublanguage-oriented model for abstract anaphora resolution. 5, 26, 36
- Mitkov, R. (1994). An integrated model for anaphora resolution. 26, 28, 31

- Mitkov, R. (1995). An uncertainty reasoning approach for anaphora resolution. In *Proceedings of the Natural Language Processing Pacific Rim Symposium (NLPRS'95)*, pages 149–154, Seoul, Korea. 16, 26
- Mitkov, R. (1996). Anaphora resolution: a combination of linguistic and statistical approaches. In *Proceedings of the discourse Anaphora and Anaphor Resolution, DAARC'96, Lancaster, UK*. 28
- Mitkov, R. (1998). Robust pronoun resolution with limited knowledge. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 2*, pages 869–875. Association for Computational Linguistics Morristown, NJ, USA. 16, 26, 28, 116, 150
- Mitkov, R. (1999). Anaphora resolution: The state of the art. Technical report. 18, 19, 26
- Mitkov, R. (2002). Anaphora resolution. *Computational Linguistics*, 29(4). 28
- Mitkov, R., Evans, R., and Orasan, C. (2002). A new, fully automatic version of mitkov's knowledge-poor pronoun resolution method. In *Proceedings of the Third International Conference on Computational Linguistics and Intelligent Text Processing, CICLing '02*, pages 168–186, London, UK. Springer-Verlag. 4, 10, 26
- Nakov, P. (2008). Noun compound interpretation using paraphrasing verbs: Feasibility study. 8, 59
- Nand, P. and Yeap, W. (2007). Challenges in anaphora resolution in the news media genre. In *The Proceedings of the fifth International Conference on Information Technology in Asia*, pages 156–160, Sarawak, Malaysia.
- Nand, P. and Yeap, W. (2008). On the use of salience weights in anaphora resolution. In *The Proceedings of the New Zealand Computer Science Research Student Conference, NZCSRSC*, Christchurch, New Zealand.

- Nand, P. and Yeap, W. (2010a). Generation of anaphoric noun phrases and their coreferential properties. In *The Proceedings of The Annual International Conference on Advanced Topics in Artificial Intelligence (ATAI)*, pages A75–A80, Phuket, Thailand.
- Nand, P. and Yeap, W. (2010b). Towards a discourse level coherence structure. In *The Proceedings of the 2nd International Conference on Agents and Artificial Intelligence*, pages 11–19, Valencia, Spain.
- Nand, P. and Yeap, W. (2011). A relational framework for interpreting anaphora. In *The Proceedings of the Annual International Conference on Advanced Topics in Artificial Intelligence (ATAI)*, pages 44–49, Canning, Singapore.
- Nand, P. and Yeap, W. (2012). Using relations to interpret anaphora. In *The Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, volume 1, pages 92–99, Vilamoura, Algarve, Portugal.
- Nastase, V. and Szpakowicz, S. (2003). Exploring noun-modifier semantic relations. In *Proceedings of the 5th International Workshop on Computational Semantics*. 59, 60
- Nastase, V. S.-S., Sokolova, J., and M. Szpakowicz, S. (2006). Learning noun-modifier semantic relations with corpus-based and wordnet-based features. *Proceedings of the National Conference on Artificial Intelligence.*, 21(Part 1):781–787. 8
- Nasukawa, T. (1994). Robust method of pronoun resolution using full-text information. In *Proceedings of the 15th conference on Computational linguistics*, pages 1157–1163, Morristown, NJ, USA. Association for Computational Linguistics. 116
- Nedoluzhko, A., Mrovsk, J., Ocelk, R., and Pergler, J. (2009). Extended coreferential relations and bridging anaphora in the prague dependency treebank. In *DAARC 2009, 7th Discourse Anaphora and Anaphor Resolution Colloquium, Goa, India*. 5

- Ng, V. and Cardie, C. (2002). Improving machine learning approaches to coreference resolution. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 104–111. 42, 43, 44
- Palomar, M., Moreno, L., Peral, J., Munoz, R., Ferrandez, A., Martinez-Barco, P., and Saiz-noeda, M. (2001). An algorithm for anaphora resolution in spanish texts. *Computational Linguistics*, 27:567. 5, 35, 131
- Passonneau, R. J. (1998). *Interaction of discourse structure with explicitness of discourse anaphoric noun phrases.*, pages 327–356. Published by Oxford University Press. 2
- Pedersen, T. and Patwardhan, S. (2004). Wordnet::similarity - measuring the relatedness of concepts. pages 1024–1025. 79
- Poesio, M., Ishikawa, T., im Walde, S. S., and Vieira, R. (2002). Acquiring lexical knowledge for anaphora resolution. In *In Proceedings of the 3rd Conference on Language Resources and Evaluation (LREC)*, pages 1220–1224. 37
- Poesio, M., Mehta, R., Maroudas, A., and Hitzeman, J. (2004a). Learning to resolve bridging references. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, ACL '04, Morristown, NJ, USA. Association for Computational Linguistics. 4, 7, 10, 37, 44, 131
- Poesio, M., Stevenson, R., Eugenio, B. D., and Hitzeman, J. (2000). Specifying the parameters of centering theory: a corpus-based evaluation using text from application-oriented domains. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*, volume 38. 37, 131
- Poesio, M., Stevenson, R., Eugenio, B. D., and Hitzeman, J. (2004b). Centering: a parametric theory and its instantiations. *Computational Linguistics*, 30(3):309–363. 20

BIBLIOGRAPHY

- Poesio, M., Vieira, R., and Teufel, S. (1997). Resolving bridging references in unrestricted text. In *Proceedings of a Workshop on Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts*, ANARESOLUTION '97, pages 1–6, Morristown, NJ, USA. Association for Computational Linguistics. 37
- Prasad, R. and Strube, M. (2000). Discourse salience and pronoun resolution in hindi. In *Penn Working Papers in Linguistics*, volume 6, pages 189–208. 116
- Prince, E. (1992). The zpg letter: subjects, definiteness, and information status. *Discourse description: diverse analyses of a fund raising text*, pages 295–325. 37
- Quinlan, J. R. (1993). *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. 40
- Reichman, R. (1985). *Getting computers to talk like you and me*. MIT Press, Cambridge, MA, USA. 21, 38
- Reichman-Adar, R. (1984). Extended person-machine interface. *Artif. Intell.*, 22(1):157–218. 2
- Reinhart, T. (1983). *Anaphora and semantic interpretation*. Chicago University Press, Chicago. 18, 117
- Sidner, C. L. (1979). *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*. Ph.d. thesis, Massachusetts Institute of Technology, Cambridge, MA. 2, 19, 20, 24, 35
- Soon, W., Ng, H., and Lim, D. (2001). A machine learning approach to coreference resolution of noun phrases. *Computational Linguistics*, 27(4):521–544. 40, 41, 43, 44
- Strube, M. and Hahn, U. (1999). Functional centering-grounding referential coherence in information structure. *Computational Linguistics*, 25(3):309–344. 4

BIBLIOGRAPHY

- Tasmowski, L. and Verluyten, S. (1985). Control mechanisms of anaphora. *Journal of Semantics*, 4(4):341–370. 2
- Terry, W. (1972). Understanding natural language. *Cognitive Psychology*, 3(1):1 – 191. 22
- Tratz, S. and Hovy, E. (2010). A taxonomy, dataset, and classifier for automatic noun compound interpretation. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, ACL '10*, pages 678–687, Stroudsburg, PA, USA. Association for Computational Linguistics. xiii, 8, 59, 60, 61
- University, N. M., Modjeska, N. N., Markert, K., and Nissim, M. (2003). Using the web in machine learning for other-anaphora resolution. In *Proc of the 2003 Conference on Empirical Methods in Natural Language Processing*, pages 176–183. 131
- van Deemter, K. and Kibble, R. (1995). On coreferring: Coreference in muc and related annotation schemes. *Computational Linguistics*, 26:629–637. 11, 131
- Vanderwende, L. (1994). Algorithm for automatic interpretation of noun sequences. In *Proceedings of the 15th conference on Computational linguistics - Volume 2, COLING '94*, pages 782–788, Stroudsburg, PA, USA. Association for Computational Linguistics. 60
- Vieira, R., Bick, E., Coelho, J., Muller, V., Collovini, S., Souza, J., and Rino, L. (2006). Semantic tagging for resolution of indirect anaphora. In *Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue, SigDIAL '06*, pages 76–79, Stroudsburg, PA, USA. Association for Computational Linguistics. 5
- Vieira, R. and Poesio, M. (2000). An empirically based system for processing definite descriptions. *Computational Linguistics*, 26(4):539–593. 4, 20, 21, 37, 38, 39, 66, 131
- Voutilainen, A., Heikkilä, J., Fries, P. I. U., Tottie, G., and Schneider, P. (1993). An english constraint grammar (engcg) a surface-syntactic parser of english. 31

BIBLIOGRAPHY

- Walker, M. (1998). Centering, anaphora resolution, and discourse structure. *Centering theory in discourse*, pages 401–435. 36
- Walker, M. A. (1997). Centering, anaphora resolution, and discourse structure. In Walker, M. A., Joshi, A. K., and Prince, E., editors, *Centering Theory in Discourse*. Oxford University Press, Oxford. 2
- Warren, B. (1978). *Semantic patterns of noun-noun compounds*. Gothenburg studies in English. Acta Universitatis thoburgensis. 57, 59, 60
- Wolf, F. and Gibson, E. (2004). Discourse coherence and pronoun resolution. *Language and Cognitive Processes*, 19(6):665–675. 23
- Zimmer, K. E. (1971). Some general observations about nominal compounds. *Working Papers on Language Universals*, pages C1–21. 57