INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI[®]

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

An Exemplar-Based Account of Contextual Effects

by
Mohsen Rais-Ghasem, B.Sc., M.Sc.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Ottawa-Carleton Institute for Computer Science

School of Computer Science

Carleton University
Ottawa, Ontario
June, 1998

© 1998, Mohsen Rais-Ghasem

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-37052-6

Canada

The undersigned hereby recommend to
The Faculty of Graduate Studies and Research
acceptance of the thesis,
An Exemplar-Based Account of Contextual Effects

submitted by

Mohsen Rais-Ghasem, B.Sc., M.Sc.

in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

External Examiner

Director, School of Computer Science

Thesis Supervisor

Carleton University
22 June 1998

Abstract

A great deal of psycholinguistic findings suggests that the meaning representation of a single concept changes in different contexts. Context modulates word senses by underlining or obscuring certain aspects of them. This modulation usually results in making some properties contextually more relevant than others.

In most computational models of lexicon, however, the effects of context are limited to favoring one reading over a small number of contrasting alternatives for a given word. Such alternatives are pre-defined and fixed. This strategy presents serious difficulties for any account of sense modulation, which requires flexibility in word senses.

This dissertation proposes a new model of lexicon that accounts for both sense selection and sense modulation. Most importantly, word senses in the proposed model are two-tiered, consisting of a sense-concept and a sense-view. While sense-concepts represent the meaning of words, sense-views describe how these meanings should be viewed in context.

The thesis also outlines an exemplar-based approach to 'learning' in which sense-views evolve as sense-concepts join them. A prototype implementation of the proposed model for sentential contexts is also discussed.

Acknowledgements

I would like to thank first my advisor, Jean-Pierre Corriveau, who in spite of his busy schedule, always made time to ensure that I would complete this research. I will always remain indebted to him for his guidance and support.

I am grateful to the members of my thesis committee: Graeme Hirst (the external), Stan Szpakowicz, Franz Oppacher, and John Logan. I also benefited from many comments that I received from Andy Brook and Rob Stainton in the earlier stages of this work.

My parents deserve special mention, for their endless encouragement, support and love. My sister and brother also have been most supportive. I am also grateful to my wife, Mitra, who in past few months patiently allowed me to spend so many hours working on the thesis.

In past few years, I have spent countless hours at Schools of Computer Science. By providing a friendly and lively environment, the school's capable staff allowed me to make the most of this time. In addition to its people and research programs, the school should also pride itself for its colorful cultural mosaic that offers equally exciting learning opportunities. Finally, financial support from the Natural Science and Engineering Research Council of Canada is gratefully acknowledged.

Table of Contents

<u>OVERVIEW</u>	<u>1</u>
1.1 INTRODUCTION	1
1.2 ET CETERA	3
1.3 TO VIEW CONCEPTS	5
1.3.1 CONCEPTS AND VIEWS	5
1.3.2 EVOLUTION OF SENSE-VIEWS	9
1.4 OVERVIEW OF THE IMPLEMENTED PROTOTYPE	11
1.5 GUIDE TO THE READER	12
<u>MOTIVATIONS</u>	<u>14</u>
2.1 INTRODUCTION	14
2.2 COMPUTATIONAL MODELS OF LEXICON	15
2.3 ON SENSE ENUMERATION	18
2.4 ENUMERATING ALL SENSES OF A WORD?	19
2.5 SEMANTIC FLEXIBILITY	21
2.6 AMBIGUITY VS. VAGUENESS	25
2.7 JUXTAPOSITION	28
2.8 IS MEANING DETERMINISTIC?	29
2.9 LEARNING NEW SENSES	30
2.10 IN NEED OF A NEW MODEL	32
<u>TOWARDS TWO-TIERED WORD SENSES</u>	<u>33</u>
3.1 INTRODUCTION	33
3.2 EPISTEMOLOGICAL ASSUMPTIONS	34
3.3 TWO-TIERED WORD SENSES	36
3.4 A LEXICON WITH EXEMPLARS	41
3.5 FROM WORDS TO SENSES	46
3.6 CONCEPT DEVELOPMENT	54
3.7 SPACE AND TIME CONSIDERATIONS	56
3.7.1 VIEW-SETS	57
3.7.2 DISTRIBUTED/CONCURRENT PROCESSING	58
3.7.3 SELECTIVE STORAGE OF EXEMPLARS	58
<u>A LEXICON FOR SENTENTIAL CONTEXTS</u>	<u>59</u>

4.1 INTRODUCTION	59
4.2 WORKING ASSUMPTIONS	60
4.2.1 CONTEXT REPRESENTATION	60
4.2.2 CASE-BASED CONTEXT RECOGNITION	62
4.3 DESIGN	64
4.3.1 SYSTEM ARCHITECTURE	64
4.3.2 PROCESSING AGENTS	66
4.3.3 IMPLEMENTATION	67
4.4 ONTOLOGY NETWORK	68
4.4.1 WHY AN ONTOLOGY?	68
4.4.2 NETWORK STRUCTURE	71
4.4.3 NETWORK BEHAVIOR	72
4.5 LEXICON NETWORK	76
4.5.1 NETWORK STRUCTURE	77
4.5.2 NETWORK BEHAVIOR	81
4.6 DEVELOPING VIEW-SETS AND SENSE-VIEWS	93

DEMOSNTRATIONS	96
-----------------------	-----------

5.1 INTRODUCTION	96
5.2 METHOD	97
5.3 SENSE-VIEW DEVELOPMENT (OBJECT-OF-MOVING)	100
5.4 SENSE-VIEW DEVELOPMENT (DESTINATION-OF-GOING-TO)	101
5.5 SENSE MODULATION (PIANO)	103
5.6 SENSE MODULATION (BOOK)	106
5.7 WORD SENSE SELECTION	107
5.8 POSSIBILITY OF MULTIPLE WORD SENSES	108
5.9 PROPERTY PROMOTION/DEMOTION (SHARK)	109
5.10 PROPERTY PROMOTION/DEMOTION (STUDENT)	110
5.11 FAILURE TO INTERPRET INPUT	110
5.12 UNKNOWN WORDS	113
5.13 IDIOSYNCRASY	113
5.14 DIACHRONY	114
5.15 POLYSEMIC/METONYMIC MAPPINGS	115

METAPHOR UNDERSTANDING	119
-------------------------------	------------

6.1 INTRODUCTION	119
6.2 THEORIES OF METAPHOR UNDERSTANDING	120
6.3 REVIEW OF COMPUTATIONAL MODELS	122
6.4 METAPHORS AS CLASS-INCLUSION ASSERTIONS	124
6.5 TESTS	128
6.5.1 TEST ONE	129
6.5.2 TEST TWO	134

DISCUSSION AND FUTURE WORK	137
-----------------------------------	------------

7.1 INTRODUCTION	137
-------------------------	------------

7.2 ACCOMPLISHMENTS	138
7.2.1 SENSE-VIEWS VS. SENSE-CONCEPTS	138
7.2.2 EVOLUTION OF SENSE-VIEWS	139
7.2.3 EXEMPLARS	140
7.2.4 AUGMENTING SENSE SELECTION WITH SENSE MODULATION	141
7.3 FUTURE WORK	142
7.3.1 BEYOND SIMPLE NOUNS/VERBS	142
7.3.2 FURTHER EVOLUTION OF SENSE-VIEWS	143
7.3.3 ENHANCING THE IMPLEMENTED PROTOTYPE	144
 REFERENCES	 146

Table of Figures

FIGURE 1-1: KNOWLEDGE STRUCTURE IN THE PROPOSED MODEL	8
FIGURE 2-1: AN EXAMPLE OF LEXICAL ENTRIES IN BORIS	16
FIGURE 2-2: ONE OF THE DEMONS REFERENCED IN FIGURE 2-1	16
FIGURE 2-3: SOME OF THE SENSES WORDNET LISTS FOR <i>BOOK</i>	19
FIGURE 3-1: TWO-TIERED WORD SENSES	37
FIGURE 3-2: HIGHLIGHTING/BACKGROUNDING WITH SENSE-VIEWS	38
FIGURE 3-3: PROPERTY PROMOTION/DEMOTION WITH SENSE-VIEWS	39
FIGURE 3-4: CONCEPTS ASSOCIATED WITH <i>BOOK</i> AND SOME OF THEIR READINGS	39
FIGURE 3-5: USING SENSE-VIEWS FOR POLYSEMIC MAPPINGS	40
FIGURE 3-6: JUXTAPOSITION	41
FIGURE 3-7: A SENSE-VIEW SHARED BY MANY SENSE-CONCEPTS	42
FIGURE 3-8: SCHEMATIC STRUCTURE OF THE LEXICON	44
FIGURE 3-9: LEXICAL ENTRIES IN THE PROPOSED LEXICON	45
FIGURE 3-10: PROCESSING STEPS FOR THE PROPOSED LEXICON	48
FIGURE 3-11: PARTS OF THE LEXICON BEFORE INPUTTING (1)	49
FIGURE 3-12: THE LEXICON AFTER PROCESSING (1)	51
FIGURE 3-13: EXAMPLE OF A VIEW-SET AND ITS EXEMPLARS	57
FIGURE 4-1: THE SYSTEM'S OVERALL ARCHITECTURE	65
FIGURE 4-2: AN EXAMPLE OF MIKROKOSMOS HIERARCHIES	70
FIGURE 4-3: EXAMPLES OF MIKROKOSMOS PROPERTIES	70
FIGURE 4-4: THE STRUCTURE OF THE ONTOLOGY NODES	72
FIGURE 4-5: THE STRUCTURE OF THE LEXICON NETWORK	77
FIGURE 4-6: THE STRUCTURE OF THE LEXEME NODES	78
FIGURE 4-7: THE STRUCTURE OF THE CONCEPT NODES	78
FIGURE 4-8: STRUCTURES OF EXEMPLAR NODES/SENSE-CONCEPTS	79
FIGURE 4-9: STRUCTURES OF VIEW-SET NODES/SENSE-VIEWS	80
FIGURE 4-10: TWO EXEMPLARS AND THEIR ASSOCIATED VIEW-SET	81
FIGURE 5-1: SYSTEM PARAMETERS AND THEIR GIVEN VALUES	98
FIGURE 6-1: <i>ENCYCLOPEDIA</i> AND ITS ASSOCIATED SENSE-VIEWS	125
FIGURE 6-2: <i>ENCYCLOPEDIA</i> AND ITS ASSOCIATED SENSE-VIEWS AFTER EXEMPLARS 8-16	126
FIGURE 6-3: <i>HUMAN</i> HIERARCHY IN MIKROKOSMOS	130

Table of Algorithms

ALGORITHM 4-1: THE BEHAVIOR OF THE ONTOLOGY NODE AGENTS (GETPROPERTIES)	74
ALGORITHM 4-2: THE BEHAVIOR OF THE ONTOLOGY NODE AGENTS (ACTIVATE)	75
ALGORITHM 4-3: THE OVERALL OF BEHAVIOR OF THE LEXICON NETWORK	82
ALGORITHM 4-4: MEASURING ALIKENESS	85
ALGORITHM 6-1: UNDERSTANDING METAPHORS AS CLASS INCLUSION ASSERTIONS	127

Chapter One

Overview

1.1 Introduction

The main focus of this dissertation is the variations that context causes in the meaning representation of a single concept, as it appears in different contexts. Consider the following sentences:

I read the book.

Many books were burnt in the fire.

The book broke the window.

I printed a copy of the book on my printer.

His arms were so weak that he could not hold a book.

Though the word *book* is hardly considered ambiguous, it is easy to see that in each of the above sentences certain aspects of its associated meaning are underlined. For example,

properties portraying *book* as a literary work seem more relevant in the first sentence than in the second one, which highlights the flammability of *book*.

This phenomenon, usually referred to as *semantic flexibility* (see Barclay et al., 1974; Greenspan, 1986) has been long studied by psycholinguists. It has been argued that context affects the encoding of concepts in memory, and that contextually relevant properties of a concept become more effective cues in recall than others (Barclay, et al., 1974), or that they are verified significantly faster (Barsalou, 1982). For example, Greenspan (1986) reports similar evidence for concrete nouns. He examined the effects of context on central and peripheral properties separately, and demonstrated that central properties were activated regardless of context, but peripheral properties were activated only if they were contextually relevant. He further demonstrated that contextually relevant central properties were more activated than contextually irrelevant central properties.

The role of context in most existing computational models is, however, limited to selecting a reading, out of a number of contrasting readings, for a given word. Therefore, in order to account for variations of the sort I suggested in the earlier examples, each such variation must become a separate reading. That is, each possible interpretation of the word *book* must ultimately be enumerated in the lexicon.

In this thesis, I reject this enumerative approach and instead, propose a generative lexicon-based strategy in which the selection of word senses is combined with another process that contextually modulates the selected word senses. Unlike generative approaches that involve complicated generation rules (e.g. Pustejovsky, 1995; Franks,

1995) , the proposed model adopts an exemplar-based approach to account for contextual effects, as explained subsequently.

1.2 Et Cetera

The phrase “*et cetera*” has played a significant role in inspiring the work presented in this dissertation, and perhaps, it can best summarize its gist. Consider the following entries from the *Oxford Student’s Dictionary*:

nuisance : a person, insect, thing, act, etc. that causes trouble or offense.

rule : to govern people, a country, etc.

shake something out : to remove lazy workers, bad methods, etc. from a business.

strip : a long, narrow piece of cloth, land, etc.

What is interesting in these definitions is the fact that in every one of them, objects that have very little in common *per se* are brought together. For instance, one can hardly find similarities between a *person*, an *insect* and an *act* under normal circumstances. Nevertheless, seen from the *perspective* of causing trouble, they seem alike.

Let us consider the way *etc.* is understood in each of the above definitions. More precisely, we want to characterize what *etc.* can instantiate to. For instance, can *noise* be a *nuisance*? We first need to build a description of things that *etc.* is used for. Intuitively, we take the definition presented in each case (e.g. *things that cause trouble or offense*) as a cue and attempt to recall other instances from our previous experiences that match this cue. Consequently, we obtain a category of objects, including those mentioned in the definition, whose members can be used in place of *etc.* in a given definition. It should be clear that members of such a category may have very little in common at first sight but, that from the viewpoint of the definition of their category, they seem alike.

From the characteristics of the members of a category we can extrapolate a generalization of the things that the category stands for. In other words, we have a description of things to which *etc.* can refer in each definition. Note that unlike in most situations where the definition of a category precedes its membership, here, we seem to extract a category's definition from its membership. This parallels the *causal* and *derivative* effects of similarity, as observed by Tversky:

Similarity of objects is modified by the manner in which they are classified. Thus similarity has two faces: causal and derivative. It serves as a basis for the classification of objects, but it is also influenced by the adopted classification (Tversky, 1977: 344).

Our intuitive account of the interpretation of *etc.* leads to a number of observations. First, it appears that the description of such a category is built on the fly each time *etc.* is used. In other words, there aren't any pre-defined and pre-stored characterizations for *nuisances*, things to be *ruled*, or things that can be *shaken out*. Second, those dynamically built descriptions may vary from person to person: these descriptions depend on earlier experiences, which clearly can vary from one individual to the next. For example, asked to name other *nuisances*, people can come up with different lists. For example, while one can remember earlier experiences in which *noise* and *humid weather* were nuisances, another may disagree on *humid weather*. Thus, descriptions can be personalized, based on past experiences.

There are significant similarities between this intuitive account of the interpretation of *etc.* and the model that will be discussed in this dissertation. First, both emphasize the notion of *perspective* or *viewpoint*. The idea is that, as with *etc.*, context brings together concepts that taken individually may have little in common. For instance, I suggest that the sense of *book* in *The book broke the window* and the sense of *bat* in *The*

man smashed the windshield with a bat are **alike** from the perspective of being the instrument of breaking. Furthermore, I suggest that *book* from this perspective is more similar to *bat* than to *book* in *I read the book*.

Also, as in the case of *etc.*, I will contend that context leads to dynamic categories, that is, categories whose descriptions are built on the fly. For example, my claim is that there is no pre-defined category of objects as “breakers”, that is, of objects that can be used in breaking other objects. Instead, such categories are built on demand, and therefore are similar to *ad hoc* categories as described in Barsalou (1983).

1.3 To View Concepts

This thesis proposes a model of lexicon that takes a number of words as input and generates word senses as output. In this section first, I introduce my proposed structure for word senses, and then describe how sense-views evolve over time.

1.3.1 Concepts and Views

Word senses in the proposed model consist of two components: **sense-concept** and **sense-view**. Given a word sense, its sense-concept will determine the concept it represents, whereas, the sense-view describes how this sense-concept is to be viewed in its surrounding context. For example, the word sense generated for *book* in *The book broke the window* will consist of the sense-concept *book*, representing the concept BOOK (by means of its central properties), and a sense-view which portrays *book* as an instrument of breaking (by underlining the physical properties of *book*, relevant to the act of breaking).

This separation of senses into sense-concepts and sense-views is crucial in my work. My thesis is that a two-tiered word sense structure that separates representation of concepts from the way the latter are viewed in various contexts is useful in order to deal with semantic flexibility (Rais-Ghasem and Corriveau, 1996, 1998a, 1998b). Such a separation parallels the two different roles that Franks (1995) proposed for concepts when he distinguished between their *representational* and *classificatory* functions. While the former is used to discern instances of one concept from instances of others, the latter specifies how an instance of a concept should be classified. For example, Franks (*ibid.*) argues that, depending on context, *fake gun* could be classified along with a *gun*, a *toy*, a *replica*, and a *model*.

Borrowing from Frank's terminology, I can say that sense-views bring together sense-concepts that with respect to their classification are **alike**. For instance, earlier I suggested that *book* in *The book broke the window* (S1) is like *bat* in *The man smashed the windshield with a bat* (S2). That is to say that, *book* in S1 is best classified with *bat* in S2 and not with *book* in a context such as *I read the book*. This means that the above occurrences (i.e. sense-concepts) of *book* and *bat* in S1 and S2 share the same sense-view, which can be roughly termed breaking-instrument.

Recognizing that, to a large extent, the concepts associated with a word are conventionalized, I propose, in this thesis, a model in which only sense-views are gradually and automatically developed from the sense-concepts that will become associated with them over time.

To model contextual effects, I propose a lexicon that structures its knowledge into four levels (see Figure 1-1). Words appear at the bottom level. Concepts associated with

words appear at the second lowest level. It is assumed that concepts are represented using their central properties. Properties centrally associated with a concept express its meaning or some part of its meaning in the standard sense. How centrality is determined for concepts is not the focus of this dissertation.

All models of lexicon, in one way or another, maintain such word-concept associations. But unlike most conventional models of lexicon, the proposed model also maintains examples of how each concept is typically used. Such usage examples, also called **exemplars**, appear at the third level. Each exemplar consists of a number of sense-concepts, each representing the occurrence of a concept in the context given by the exemplar. Sense-views constitute the fourth level. Each sense-concept is connected to a sense-view. Together, a sense-view and a sense-concept form a word sense. Thus, occurrences of a single concept in different contexts will be connected to different sense-views. Similarly, occurrences of different concepts can be connected to a single sense-view.

Figure 1-1 illustrates an example of two words, *bat* and *book* (represented as rectangles) and of their associated concepts: MAMMAL-BAT and BASEBALL-BAT for *bat*, and BOOK for *book* (represented as rounded rectangles).

specified by exemplars (represented as double-lined rectangles) each representing one use case for its associated concepts.

For the sake of simplicity, sense-concepts constituting each exemplar are not displayed. Examples of sense-concepts are the occurrence of BASEBALL-BAT in *Jimmy got a bat for his birthday* and the occurrence of MAMMAL-BAT in *Several bats flew in*

the cave. Each sense-concept is connected to a sense-view. Two such sense-views are shown: breaking-instrument (shared by sense-concepts BASEBALL-BAT and BOOK) and gift (sense-concept BASEBALL-BAT).

Figure 1-1: Knowledge structure in the proposed model.

Note that both sense-views and sense-concepts are in fact represented by a set of properties, and that the names used in this figure and all other figures in this dissertation are only labels used for the sake of clarity and brevity. For example, here is the output generated by the prototype discussed in the next section for *book* in *The book broke the window*.

```
SENSE Generated for Input Word Book
[Sense-Concept(s)] BOOK
[Sense View]
Thematic Role: Instrument - No. of Exemplars: 4 - Marker(s): with-pp, p-subj,
STATE-OF-MATTER(0.73305)---->:SOLID,
MADE-OF(0.82805)---->: MATERIAL,
```

existing exemplar, or if it is possible to use them as in an existing exemplar. Obviously, it is this second case that deserves more attention.

Let me elaborate. Presented with a number of words as an input context, the model attempts to find a set of interactable senses for them. For each word, this involves the selection of a sense-concept and then its modulation. Owing to lexical ambiguity, several combinations of concepts can possibly be found for the same set of input words. To decide if a set of concepts can interact with each other, the model tries to find an exemplar in which these concepts can be used. To decide if a concept is usable in place of a sense-concept in an exemplar, the model needs to determine if this concept can be associated with the same sense-view that that sense-concept belongs to.

Let me explain this. For example, suppose I want to find proper word senses for *book*, *break* and *window* in *The book broke the window*, assuming that the following exemplars are already available in the system:

The rock smashed the window.

Jon broke the law.

For simplicity, suppose that only *break* is ambiguous (e.g. SEPARATE-INTO-PIECES and VIOLATING-LAW) and that other input words have only one associated concept. It follows that, in this case, there will be two possible combinations of concepts for these three words. However, concepts associated with *book* and *window* can be used in the first exemplar, but not in the second one. That is, it is possible to associate *book* with the sense-view breaking-instrument, but not with law-violator¹. Similarly, only one (i.e., SEPARATE-INTO-PIECES) of the concepts associated with the ambiguous word *break*

can be used in the first exemplar. Consequently, one combination is rejected and the other one is accepted. And thus, the model has determined which concepts should be selected for each word. All that is left to do is to modulate these concepts. This is achieved by associating each selected concept with the same sense-view that its corresponding sense-concept is associated with. Therefore, BOOK will be associated with breaking-instrument, SEPARATE-INTO-PIECES with breaking-action, and WINDOW with breaking-object. And thus, the corresponding sense-views will evolve.

Finally, user intervention would be required if the model fails to find a set of interactable senses for input words. Such interventions provide the source of training for the model.

1.4 Overview of the Implemented Prototype

A prototype of the proposed model has been implemented for simple sentential contexts. In this implementation, a case-based context recognition approach is used in which case-markers (see Delisle et al., 1993) are used to match input words with sense-concepts in exemplars. For example, if an input word is marked as positional subject, most likely it corresponds to instrument or agent in a breaking exemplar.

The prototype is implemented as a hybrid system. Architecturally, it consists of two structured networks: the ontology network and the lexicon network. The ontology network is primarily used as the system's knowledge base. It provides definitions for concepts by first, associating them with properties and, second, organizing concepts in a taxonomy. The ontology network is based on Mikrokosmos, a general ontological system

¹ Unless it is a case of metonymy (see section 5.15).

described in (Mahesh & Nirenburg, 1995). On the other hand, the lexicon network implements the 4-level lexicon structure described in 1.3.

Although based on two structured networks, much of the system's functionality is carried out through the execution of **agents** attached to various nodes in both networks. Agents are processing elements that are activated once their attached nodes receive activation. At that point in time, an agent starts executing a series of instructions, as will be discussed in section 3.5. Agents can work simultaneously.

The implemented prototype² has been tested extensively to verify the model's capabilities for the incremental development of sense-views, as well as in dealing with both sense selection and modulation. Some of the results are demonstrated in chapter 5.

1.5 Guide to the Reader

Chapter 2 motivates this work by reviewing computational models of lexicon and arguing that most existing models do not address issues such as semantic flexibility, juxtaposition, idiosyncrasy and diachrony. In chapter 3, I discuss two-tiered word senses and their advantages in detail. The structure of the proposed model of lexicon and the process of generating word senses are also discussed in that chapter. Chapter 4 is dedicated to the implemented prototype. It provides details about the structure and function of the ontology and lexicon networks. Some of the tests conducted on the implemented prototype are discussed in chapter 5. These examples are intended to provide evidence of the capabilities and usefulness of the proposed model. Using sense-

² The system is programmed in Java and runs on a PC. The size of the application's bytecodes totals 88KB.

views for metaphor understanding is introduced in chapter 6. The thesis concludes with chapter 7, in which I reflect on the proposed model and discuss future research.

Chapter Two

Motivations

2.1 Introduction

People use dictionaries to look up words and obtain information about them. Similarly, most NLP systems rely on a lexicon to obtain information about a given word (Allen, 1995). Like dictionaries, lexicons are primarily used to find the meaning or meanings of a word. These meanings are commonly known as *word senses*.

Traditionally, word senses associated with a word are presented as concepts to which the word is thought to refer. In other words, the lexicon maps the word forms onto a concept network (Cruse, 1995). Concepts are typically represented as a set of properties. These properties are generally used to *facilitate comparison* and *explain similarity* (Medin and Goldstone, 1990). In general, three distinct concept representation approaches are used in AI systems: symbolic representations, representations based on

semantic networks, and distributed connectionist representations. Thus, this chapter classifies and discusses computational lexicons based on their underlying conceptual representation schemas.

Section 2.2 provides a brief overview of existing models of lexicon. Then, in section 2.3, I argue that despite their differences, existing models generally come short of accounting for a number of important issues, most notably, the effects of context in making certain properties more relevant than others during the course of interpretation. These issues are individually discussed in sections 2.4 to 2.9. Section 2.10 summarizes the chapter.

2.2 Computational Models of Lexicon

Three types of computational lexicons, I repeat, can be distinguished: symbolic lexicons, lexicons based on semantic networks, and distributed connectionist lexicons. In symbolic lexicons, word meanings are generally represented by means of *ad hoc* and complicated structures, which may encapsulate both declarative and procedural knowledge. Typically, such structures are predefined, that is, hand-coded and fixed. Most early NLP models (e.g., SAM, Cullingford, 1978; PAM, Wilensky, 1978) fall under this category (for a complete review, see chapter 2 of Corriveau, 1995). For example, in BORIS (Dyer, 1983), “lexical entries point to sources of knowledge and demons” (*Ibid.*:183). *Demons* are attached procedures that “wait until their conditions are satisfied, at which point they fire and execute their actions” (*Ibid.*:165). Figure 2-1 displays the entry of BORIS for *wife*. References to demons are preceded by *. One of the associated demons is shown in Figure 2-2.

(in order to account for context) eventually leads to one of these senses being selected as the correct meaning for the word *star* in that context. For example, in *The astronomer married a star*, in spite of the strong correlation between *astronomer* and celestial bodies, the movie star sense will be selected for the word *star* (*Ibid.*).

Distributed connectionist (PDP) networks are at the basis of the last group of lexicons. In this family, the representation of a concept is distributed across a network of unlabeled and massively connected nodes. Unlike semantic network-based lexicons where every concept is assigned to a node, here, a single set of nodes is used to represent all concepts. That is why Rumelhart and Norman termed such models *superpositional* since “each storage element contains information from many different memories *superimposed* upon one another” (see Oden, 1987). Examples of such lexicons are reported by Veronis and Ide (1995), Kawamoto (1993), Miikkulainen and Dyer (1991), and McClelland and Kawamoto (1986).

Connectionist lexicons can be further divided into two groups. In the first group, concepts are not represented as a set of properties. Instead, concepts are encoded along a set of unnamed features. The lexicons reported by Veronis and Ide, and Miikkulainen and Dyer belong to this category. Given the focus of this dissertation on the effects of context on the properties of concepts, such models are taken to be too distant from this work, and therefore, will not be further discussed.

Models in the second group of connectionist lexicons rely on microfeatures to encode their concepts. For example, in McClelland and Kawamoto’s model, every word is associated with one or more word senses. Word senses are encoded along dimensions of semantic microfeatures (henceforth, simply features). Examples of dimensions are

softness (soft/hard), gender (male/female/neuter), and volume (small/medium/large). In case of 'unclear' cases³ (e.g. *chicken*, cooked or living), the input pattern is the average of the feature patterns for each of the two readings of the word. The network is then generally trained to associate every word with a contextually appropriate meaning. Implicit here is the necessity for all concepts to be defined using a universal set of features.

In spite of their differences, it appears that existing models of lexicon typically fail to account for a number of important issues, and in particular for the effects of context in underlining certain properties of a single word sense in different contexts. I explore these shortcomings in more detail in the next section.

2.3 On Sense Enumeration

What is common in all of the three types of lexicon (especially in symbolic and semantic network-based lexicons) is their commitment to a meaning theory called *sense enumeration*⁴ (Simpson & Burgess, 1985; Seidenberg et al., 1982). This theory suggests that, upon hearing a word, listeners access a mental dictionary containing an exhaustive list of potential senses for that word. Eventually, with the help of contextual clues, one of these senses is selected. WordNet (Miller, 1990) is most representative of such comprehensive lexicons. In WordNet, a great deal of effort has been made to list as many

³ Despite the text referring to such cases as ambiguous terms, they are in fact vague (see the vagueness versus ambiguity discussion, later in this chapter).

⁴ This commitment is made at the architectural level in most symbolic and localist models. In distributed connectionist models, however, this is mostly a matter of representation (see 2.5).

senses as possible for a fairly large set of English words. Figure 2-3 shows some of the senses listed for the noun *book* in WordNet.

Fundamental to the sense enumeration theory are i) the existence of such a comprehensive list of possible senses for each word, and ii) the existence of a selection process where one sense is chosen. However, a number of issues cast doubt on the validity and applicability of such an enumerative approach (see chapter 2, Corriveau, 1995). These issues will be discussed in the subsequent sections.

- A copy of a written work or composition that has been published.
- A book as a physical object: a number of pages bound together.
- The known facts regarding something or someone.
- A written version of a play or other dramatic composition.
- A record in which commercial accounts are recorded.
- A major division of a long written composition.
- An accounting book as a physical object.
- A number of sheets (ticket or stamps etc.) bound together on one edge.

Figure 2-3: Some of the senses WordNet lists for book.

2.4 Enumerating all Senses of a Word?

There are theoretical objections to this position. Taking a sharply contrasting stance, meaning minimalists attribute more importance to the pragmatic rules of interpretation as opposed to the literal meanings. Ruhl for example, in the introduction of his book on monosemy writes:

I claim that current linguistic theories accept too uncritically the conclusion of dictionaries that words in general have multiple meanings. I argue that words contributes much less to meaning than usually supposed; the apparent lexical

meaning of a word includes in large part a contribution of contextual factors, both linguistic and extralinguistic (Ruhl, 1989:vii).

Kilgarriff (1997) contends that standard dictionaries specify the range of meanings of a word in a list merely in response to “constraints imposed by tradition, the printed page, compactness, a single, simple method of access, and resolving disputes about what a word does and does not mean.” He, then, criticizes word sense disambiguation (WSD) approaches for committing themselves to such lists. He remarks:

Much WSD word proceeds on the basis of there being a computationally relevant, or useful, or interesting, set of word senses in the language, approximating to those stated in a dictionary... Meanwhile, the theoreticians provide various kinds of reason to believe there is no such set of senses (Kilgarriff, 1998:95).

Moreover, it is not always possible to enumerate all potential senses of a word. The use of language, even non-figuratively, by its very nature is creative. That is, “words can take on an infinite number of meanings in novel contexts” (Pustejovsky, 1995: 42). No matter how comprehensive a list can be, it is generally possible to find a new sense (i.e., one that does not appear in that list) for some words. For example, eponymous expressions, that is, expressions built around references to people, provide a good source of examples. Clark and Gerrig (1983) showed how understanding eponymous expressions such as *Please do a Napoleon for the camera* requires sense generation rather than sense selection. They concluded that we might use both mechanisms. Gerrig (1989) later studied the time course of sense generation and concluded that sense generation and sense selection operate simultaneously.

Cognitive parsimony also makes the existence of such a comprehensive list of word senses unlikely. Indeed, in the context of machine learning, Michalski (1989; 1990)

dismisses the possibility and rationality of storing all possible meanings for a term. He writes:

[T]he meaning of a concept is a dynamic structure built each time anew, in the course of an interaction between some initial base meaning and the interpreter's background knowledge in the given context of discourse (Michalski, 1989: 122).

2.5 Semantic Flexibility

Even more damaging to existing computational models of lexicon are a number of psycholinguistic findings, which have been mostly ignored by the NLP community. For long, psycholinguists have studied the effects of context on storage and retrieval of concepts by human subjects. One early experiment was reported by Barclay et al. (1974). These researchers were concerned with the relation between encoding and retrieval conditions in human memory. The “encoding specificity principle” postulates that the goodness of remembering an event or a fact depends on two critical factors: (1) the initial encoding operations and resultant memory trace, and (2) the subsequent retrieval conditions or retrieval cues (Tulving, 1990). The retrieval condition in question in that work was context. Barclay et al. (*ibid.*) demonstrated how interpretations of familiar, unambiguous words vary with context. For instance, the choice of attributes for *piano* is affected by verb selection in *The man (lifted) (tuned) (smashed) (sat on) (photographed) the piano*. In particular, these researchers showed how the prior acquisition of a sentence like *The man lifted the piano* (vs. *The man tuned the piano*) influences the effectiveness of cues like “something heavy” (vs. “something with a nice sound”) in recall.

In a property-verification experiment, Barsalou (1982) also showed how subjects verify some properties (e.g., “where cooking can occur”) significantly faster in relevant

contexts (e.g., *The hospital was quiet when the dinner was served*) than in irrelevant contexts (e.g., *The hospital was internationally famous for its progressive techniques*).

Witney et al. (1985) also report similar results. Their work is particularly interesting since they distinguished between two levels of property activation: 1) the functional level (useful in activities such as sentence comprehension and classification) and 2) the semantic access level (corresponding to the information that is actually accessed upon seeing a word). They used intervals of 0, 300 and 600 ms. and found that all properties of a word were initially activated (i.e., accessed), regardless of the context. However, the contextually irrelevant ('low-dominant' in their terminology) properties would die off rapidly, and hence play a negligible role in the overall sentence comprehension.

Greenspan (1986) studied the effect of sentential context on concrete nouns (e.g. *piano*, *boot*, and *stream*). He examined the presence of central and peripheral properties of a single noun in different contexts. Greenspan presented his subjects with a pair of sentences where, in each sentence, a different type of properties was emphasized. For example, consider the noun *basket*. Being a container is central for *basket*, whereas being made of straw is peripheral. Each of the following sentences focuses on one of these properties.

The butler placed the letter in the *basket*. (Container)

Sally took several days to weave the *basket*. (Straw)

Later he examined subjects' recall in various memory tasks and found that the central properties were activated regardless of the context, but peripheral properties were activated only if they were emphasized by the sentence. He further showed that the

emphasized central properties were more activated than unemphasized central properties. He concluded that the interpretation of a concrete noun is a function of both the sentential context and the noun's central properties (*ibid.*).

Anderson et al. (1976) demonstrated that general terms are contextually instantiated to more specific terms. For example, they used *fish* in contexts like the following sentences and hypothesized that it was likely to be instantiated respectively to *goldfish*, *trout* and *shark*.

The grocer stared at the fish in the bowl.

The grocer stared at the fish in the stream.

The fish attacked the swimmer.

They examined their hypothesis in a number of experiments and found that an expected instantiated term was a better cue for the recall of a sentence than the general term itself, even though the general term had appeared in the sentence and the particular term had not. For instance, they discovered that *shark* was a better cue in recalling *The fish attacked the swimmer* than *fish* itself. They argued that this could have not happened unless subjects had used *shark*, rather than *fish*, in encoding and storage of the sentence.

In my opinion, the above findings clearly show that, in many cases, selecting a word sense from a number of alternatives is simply not adequate. This is because the exact semantic characteristics of the chosen sense may vary from context to context. Thus, even if one attempts to treat all possible sense variations as separate senses, such a list of senses will be far too large, if not open-ended.

To the best of my knowledge, none of the existing computational models of lexicon have adequately addressed the above mentioned findings. Having subscribed to

the sense enumeration theory, the effects of context in these models are limited to the selection of one sense over others. Any divergence from a sense either has to come in the form of a new sense or is ignored.

Among the three discussed categories of lexicon, this criticism is more relevant for symbolic and semantic network-based models, for in these models, contextual effects are limited to sense selection.

Conversely, connectionist models appear to be capable of offering some form of sense modulation. For example, McClelland and Kawamoto (1986) reported an unintended yet interesting result. They had presented their model with *The ball broke the vase*. Despite the fact that throughout the training the *ball* was always associated with the microfeature *soft*, in the output generated for this sentence, it was the microfeature *hard* that was activated. They attributed this result to the fact that *breakers* in their experiment were all *hard*: the model had *shaded* the meaning of *ball* accordingly. Kawamoto more recently remarked

[T]he flexibility of a distributed representation allows a natural account of polysemy and homonymy, and provides a mechanism for new senses to be learned or generated on-line (1993: 510).

However, the connectionist approach is still limited in its treatment of context. For example, connectionist lexicons generally presuppose a large fixed set of universal microfeatures and demand that every possible sense be characterized using this set. This is far from practical for any real world application (Bookman, 1994). And furthermore, such an approach indirectly subscribes to the enumeration theory.

Also, it is generally quite difficult to reason about and interpret the underlying representation of distributed connectionist models (*ibid.*). It is particularly difficult to

separate patterns of contextual effects from meaning representations. In other words, contextual patterns are also *superimposed* on the network's nodes. For example, it is impossible to extract characteristics of *breakers* in McClelland and Kawamoto (1986). These characteristics are embedded in the representation of the instrument of breaking.

I believe that separating patterns of contextual effects from representations of word senses is important. First, as previously suggested, such patterns can be thought of as *ad hoc* categories: categories built by people to achieve goals (Barsalou, 1983). For instance, *breakers* can be instrumental in achieving the goal of "breaking a window". Second, with respect to learning, such patterns can be very useful. Indeed, in section 3.6, I will show how a concept can evolve (i.e., acquire new properties) from such patterns as it appears in various contexts. And finally, in chapter 6, I use such patterns to implement a metaphor understanding theory that interprets metaphors as class inclusion assertions (see Glucksberg & Keysar, 1990).

2.6 Ambiguity vs. Vagueness

Another difficulty facing existing models of lexicon concerns the treatment of ambiguity and vagueness. Consider these examples:

Nadia's *plane* taxied to the *terminal*.

Mary ate a *fig* for lunch

Mary watered the *figs* in the garden.

As discussed in Hirst (1987), both *plane* (aircraft, carpentry tool) and *terminal* (computer terminal, place of embarking) are ambiguous in the first sentence and must be disambiguated. In the case of lexical ambiguity, alternative senses are clearly distinct. This is not the case for vague terms, for which alternative senses are related. For

example, as Pustejovsky (1995) suggests, the two given senses of *fig* in the above sentences can be considered as food/plant facets of a common entity. Frawley separates vagueness from ambiguity as follows:

[An expression] is *ambiguous* if there are at least two distinct semantic specifications underlying a single overt form. An expression is *vague* if it is *unspecified* for particular meanings and takes them from context (1992:58).

Linguists have proposed several methods to determine whether an expression is vague or ambiguous (e.g., Cruse, 1995, 1986; Frawley, 1992,).

This issue is a point of contention among researchers. Some (e.g., Fodor & Lepore, 1996) suggest vague words are no different than ambiguous ones (strong homonymy stance). Others (e.g., Cruse, 1995; Pustejovsky, 1995; Lakoff, 1987) maintain that there exists a basic meaning shared by all vague expressions (weak homonymy). For example, Lakoff (1987), dismissing separate and independent senses, proposes a process in which basic metaphors are used to derive new senses from existing senses. Examples of formal approaches towards generative meaning are reported by Pustejovsky (1995), Franks (1995) and Pustejovsky and Boguraev (1993).

For Cruse (1995, 1986) the distinction is rather fuzzy. He further investigated polysemyous cases and identified cases in which alternative senses of a word could be fully accounted by what he calls “differential contextual enrichment of a single lexical meaning” (1995: 33). The word *teacher* (male teacher and female teacher, respectively) in the following sentences is an example of such cases:

The teacher stroked his beard.

Our math teacher is on maternity leave.

The point here is that *teacher* by its definition is gender neutral, that is, the gender is left unspecified, but can be specified contextually. He also found other cases that involve slightly different contextual enrichment. For example, in the following sentences, different aspects of *car* are highlighted, while others are dimmed.

John washed the car.

The mechanic lubricated the car.

The thief fell on a car.

The crane lifted the car.

The car crashed into a huge rock.

He concludes:

One of the main points to be made ... is that the traditional sharp dichotomy between polysemy and monosemy is misleading: there is, in fact, a continuum of degrees of distinctness ranging from not distinct at all (as with the two readings of *teacher*), to fully distinct (as with the two readings of *bank*) (*ibid.*: 36).

An interesting set of empirical findings on this issue is reported in Frazier and Rayner (1990). By recording eye movements, these researchers showed how ambiguity and vagueness are treated differently by human readers. They tested their subjects on sets of sentences containing both ambiguous and vague words. Each sentence also was accompanied by a disambiguation phrase that either preceded the problematic term (prior disambiguation) or followed it (late disambiguation).

Here are examples of their test materials:

Walking in late, her *date* surprised Jane. (date is ambiguous)(prior disambiguation)

Tasting so bitter, her *date* surprised Jane.

Yesterday her *date* surprised Jane, by walking in late. (late disambiguation)

Yesterday her *date* surprised Jane, by tasting so bitter.

Lying in the rain, the *newspaper* was destroyed. (*newspaper* is vague)(prior disambiguation)

Managing advertising so poorly, the *newspaper* was destroyed.

Unfortunately, the *newspaper* was destroyed, lying in the rain. (late disambiguation)

Unfortunately, the *newspaper* was destroyed, managing advertising so poorly.

They noticed that for vague words, the difference between reading times in the prior and late disambiguation forms was much smaller than for ambiguous words. They concluded that:

[T]his effect is not due to garden path [as it is for ambiguous words] but only [to] additional specification of the intended sense of the target word together with end of sentence effects (*ibid.*: 194).

Again, existing models of lexicon typically deal with vagueness the same way they deal with ambiguity: by allocating separate entries for each reading.

2.7 Juxtaposition

Another problematic issue for conventional lexicons arises in cases where different usages of a word are juxtaposed in the same context. Consider these sentences:

Boys will be boys.

A dime is a dime.

Even disciplining his son, a father is a father.

In these examples, the first occurrences of *boys*, *dime* and *father* each point to a contextually specific instance of these words, whereas the second occurrences refer to more typical readings of these words. Traditional lexicons are inherently incapable of accounting for such cases unless they maintain two separate readings for every possible sense of a word: one typical and one contextually specific. Even so, as discussed earlier, if not unlikely, it is impractical to foresee and list all possible contextual variations of a sense. What we need for every sense is a typical reading plus a way of deriving contextual variations from it (see next chapter).

2.8 Is Meaning Deterministic?

Corriveau (1995) has argued that language understanding is both idiosyncratic and diachronic. Idiosyncrasy tells us that comprehending an utterance is an interactive process that begins with the speaker/writer and ends in the listener/reader. Both parties are actively involved, therefore, what the listener/reader ultimately *grasps* from the message are her idiosyncratic interpretations of what the speaker/writer intended to communicate. Diachrony, on the other hand, is concerned with the changes that may occur over a period of time in one reader's comprehension of a single utterance. Such changes occur as the reader's knowledge of world and words changes. For our purpose, this means that the model of lexicon must be able to account for changes in the senses of a single word over time and from reader to reader. But of course, a great deal of overlap exists among readers' senses, otherwise no communication would be possible. In other words, it seems likely that a common nucleus of properties is shared across readers for a single concept.

In existing lexicons, senses, to a large extent, remain fixed over time. New information has little or no effect on existing senses stored in the lexicon. Consider a hypothetical case where a toddler is learning about a *crane* and objects it moves. In her first few experiences, she witnesses a *piano* and an *automobile* being moved. She may conclude that cranes are used to lift heavy, inanimate and big objects. Later, in another experience, she sees an *elephant* being moved by a crane. It is probable that this will lead her to revise her understanding of objects moved by a crane to include animates. This revision process could even eventually eliminate *size* as a characteristic of such objects (as both large and small objects can be heavy).

In the context of language processing, the point is that there must be a direct relation between what a language processing system learns about the world (by processing linguistic utterances) and its lexicon (its knowledge of words). More precisely, seeing a word in a novel context must have an impact on the system's knowledge of that word, and the impact may be retroactive! This approach not only leads to an idiosyncratic (i.e. personalized) lexicon, but also enables diachronic interpretation. For instance, in the above scenario, if the toddler again witnesses a piano being moved by a crane, she will likely realize that this is only because the *piano* is massive and not because it is big or inanimate.

2.9 Learning New Senses

The issue of learning is closely related to the previous topic. Here, however, I am not addressing 'learning' in the rich sense typically ascribed to this word by the machine learning community. Instead, following Bookman (1994), I limit here what I mean by the term 'learning' to the question of how to augment the lexicon from the processing of a context.

I suggest that learning in a lexicon can take two forms. First, a lexicon can learn new word senses. For example, consider a lexicon in which the word *palladium* is not defined. In a context like the following, the model is expected to assume that *palladium* is a place that people can go to (see 5.12 for more details).

John went to the palladium.

The evolution of existing senses through the acquisition of new properties constitutes another simple form of learning. For example, assume a lexicon in which *piano* is strictly defined as musical instrument and its physical characteristics (such as weight and size)

are not considered. Such properties must be added to the initial definition of *piano* as it appears in a context such as the following.

They needed a crane to move the piano to the second floor.

Sense evolution is important because it eliminates the need for extensive *a priori* coding and allows gradual and incremental learning of senses (see also 6.3 for another example).

Symbolic lexicons usually depend on hand-coding, although a few attempts have been made to learn word meanings (see Martin, 1991 for a review). For example, Berwick (1989) reports on a system that can deduce the meaning of a novel verb in a context from the surrounding causal network. In that work, the model is given two versions of the story of Macbeth. In one of them, “Macbeth *murders* Duncan” while in the other one “Macbeth *assassinates* Duncan”. The difference is that, unlike *murder*, *assassinate* is unknown to the system. Comparing the causal network for the two stories, the system then concludes that most likely *assassinate* plays the same role that *murder* does. But such systems still number few and are quite constrained.

Learning in semantic network models is notoriously difficult. Any form of learning typically requires some reconfiguration of the network. One exception, although not exactly a model of lexicon, is LeMICON (Bookman, 1994), which can dynamically connect highly active nodes, to facilitate activation of those nodes in any future recall.

Connectionist lexicons are inherently good at handling incomplete inputs (see Miikkulainen, 1993 chapter 13). But it appears that this capability has not yet been used for learning new word senses in a lexicon. On the other hand, as discussed in McClelland and Kawamoto (1986), new properties can be learned (as in the case of associating *ball*

with *hardness*). Unfortunately, this is limited only to those properties used to encode senses (i.e., dimensions of the model's semantic space).

2.10 In Need of a New Model

In this chapter, I reviewed computational lexicons and argued that, despite the differences in their concept representation approaches, they generally rely on sense enumeration. I also suggested that several issues, and most notably semantic flexibility, raise serious questions about the possibility and validity of sense enumeration.

In the next chapter, I will propose an alternative model that intends to address these issues. The proposed model has a hybrid nature: it attempts to offer

- a) flexibility of representation and
- b) training by example, as in connectionist models, while maintaining the advantages of symbolic models, and in particular, the ability to retrieve and interpret the underlying representation.

Chapter Three

Towards Two-Tiered Word Senses

3.1 Introduction

The previous chapter criticized existing computational models of lexicon for adopting a sense enumeration position towards word senses. I contended that such a position is incapable of accounting for several problems such as semantic flexibility, juxtaposition, idiosyncrasy, and diachrony. In this chapter, I propose a new model that attempts to address some of those problems.

I start with a discussion of my postulates. Then, in section 3.3, I will argue for a two-tiered approach to word sense and show how this approach allows us to account for semantic flexibility and similar issues. Section 3.4 describes the basic structure of the proposed lexicon. Section 3.5 explains the process of finding, from input words, word senses in the proposed lexicon. I will claim that, due to its reliance on previously processed contexts, the proposed model is capable of dealing with diachrony and

idiosyncrasy. Section 3.6 discusses concept evolution in the proposed model and shows how new properties can be associated with concepts. The last section in this chapter, section 3.7, considers space and time requirements of the proposed model.

3.2 Epistemological Assumptions

This section discusses the epistemological assumptions behind the proposed model. The first assumption concerns the nature of flexibility in word senses. For example, the question is whether different readings of *piano* in *The man played the piano* and *The man moved the piano* reflect two separate underlying concepts?

I subscribe to the position advanced by Barsalou (1993). He argues that flexibility “arises not from differences in knowledge but from differences in the *retrieval* of this knowledge”. He continues:

These varying subsets of features are what I am defining as concepts. Rather than being stable structures stored in long term memory and retrieved as needed, concepts are temporary constructions in working memory (*ibid.*:34)⁵.

The next assumption concerns the existence of semantic features (i.e., properties). Such semantic features are the subject of heated debates. Le Ny remarks:

It is well known that [the existence of semantic features] is a controversial issue, and there is no general agreement among theories on the existence of this type of cognitive entity. Nevertheless, models of this type are in fairly good agreement with the facts, and can be extended (Le Ny, 1995:62).

Though not committed to any particular epistemological theory, nor to any specific set of properties, I postulate the existence of such semantic features. The following citation from Le Ny best describes this position:

My line of reasoning will this be: to postulate the existence of semantic features in word representations, namely in their meaning part... My present postulate ... states merely that lexical units, in particular their meaning parts, can be analyzed into smaller pieces of meaning, sub-units which can be (in set conditions) separately processed by the speaker (*ibid.*).

Throughout this dissertation, I will be using the term *property* or *concept property* to refer to these “sub-units”. I reserve the term *semantic features* to refer to these “sub-units” in general. It should be constantly borne in mind that this thesis neither makes any claim about some particular property set, nor depends on any.

As the third postulate, I emphasize that the proposed model should not be considered as a complete NLP system. Indeed, there are many syntactic and semantic issues that are not being addressed here (some possible extensions will be discussed in chapter 7). The focus of this work is the lexicon: the goal is to build a two-tier model of lexicon capable of accounting for various contextual effects. In their introduction to the application of lexical semantics in natural language processing, Saint-Dizier and Viegas (1995) write:

One of the main problems is to identify exactly the lexical semantic resources that one needs to solve a particular problem. Another main difficulty is to know how best to organize this knowledge in order to keep the system reasonably efficient and maintainable (*ibid.*:349).

This dissertation must be considered as an exercise in this direction, that is, as the development of a model of lexicon to be used eventually in a more complete NLP system.

⁵ In this dissertation, what Barsalou terms a *concept* is called a *word sense*. He uses the term *knowledge* to refer to the stable body of information in long term memory, which is what we call a *concept*. (See the discussion on the representational function of concepts in 3.3.)

Moreover, although motivated by psycholinguistic findings, my stance is resolutely a computational one and I am not making any claim about the cognitive plausibility of the proposed model.

The next issue concerns the generative nature of the proposed lexicon. Instead of using complicated generation rules (e.g., Pustejovsky, 1995; Franks, 1995), word-senses are taken to evolve from a number of exemplars. And, although presented as a lexicon, this model focuses solely on semantic aspects of words. In other words, morphological and phonological issues are not considered. Of course, this does not mean that they are of lesser significance or play no role in the meaning. This decision is primarily due to the necessity to scope the thesis to a workable problem.

And lastly, although contextual effects are not limited to nouns, this work is mainly concerned with simple noun phrases. In fact, the ideas put forward in this work are not readily applicable to compound nouns or noun phrases. Such cases require a compositional theory that can specify the effects of each component on the head noun. Such a theory lies beyond the scope of this thesis.

3.3 Two-Tiered Word Senses

The claim behind this dissertation is that a simple two-tiered word sense structure is capable of accounting for many of the issues motivated in the previous chapter. In the proposed model, a word sense is structured into two tiers: a **sense-concept** and a **sense-view** (see Figure 3-1). While the sense-concept tier specifies the core meaning of a word in any given context, the sense-view tier describes how this meaning is to be viewed in that context. Analogous to masks, sense-views typically hide or dim certain aspects of

their underlying sense-concepts. They can also highlight properties of their sense-concepts; or even introduce new features. Thus, sense-views are made of properties. And it is through these properties that the associated sense-concepts are to be viewed. (Thus, please do recall that the names used for sense-views in this section and throughout this dissertation are merely mnemonic labels, and that sense-views are in fact specified only using the properties they possess.)

The next section provides more details about sense-views. For now, I intend to show how the use of two tiers allows us to address various types of contextual effects.

Figure 3-1: Two-tiered word sense.

Cruse (1986) specifies two ways in which context affects the semantic contribution of a word: *sense selection* and *sense modulation*. Sense selection happens in cases of lexical ambiguity where one sense is chosen among a number of distinct senses associated with a word. Sense selection is followed by sense modulation in which the semantic characteristics of selected senses are modulated according to their surrounding contexts. It is in this type of contextual effects that I am particularly interested in this work.

Cruse distinguishes two types of sense modulation: a) feature highlighting or backgrounding, and b) promotion or demotion of features. In the former, context

underlines certain aspects of the selected meaning for a word while dimming others, as it happens for *car* in the following examples:

John washed the car.

The mechanic lubricated the car.

The crane lifted the car.

Context can also promote, or demote, certain aspects of a word's basic meaning. As discussed earlier, by definition *teacher* is gender neutral. However, in a context like the following, the feature gender is promoted for *teacher*.

The teacher stroked his beard.

Similarly, certain aspects can be demoted, as is *solidity*, a feature usually associated with *butter*, in the following context:

Mary poured the butter into the jar.

Both forms of sense modulation can be supported by the proposed two-tiered structure. In either case, the sense-concept tier points to the core meaning of the selected sense word. Sense modulation happens at the sense-view tier. More specifically, it is through sense-views that some aspects of their underlying sense-concepts become highlighted or dimmed. Figure 3-2 shows how a single sense-concept for *car* is viewed differently in the above sentences.

Figure 3-2: Highlighting/Backgrounding with sense-views.

Sense-views also can introduce new properties to their underlying sense-concepts. Figure 3-3 illustrates such a situation for *butter* in *Mary poured the butter into the jar*. In this example, *butter* is portrayed as an object of pouring.

Figure 3-3: Property Promotion/Demotion with sense-views.

As another example of using sense-views, consider the word *book* (which is vague) in referring to *book* as a physical object or a literary work, as in the following examples:

The book broke the window.

John read the book.

Unlike traditional lexicons which treat these two readings of *book* as separate senses, the proposed sense-views allow us to have a shared underlying meaning that can be projected differently according to the context. Figure 3-4 illustrates the word *book* along with two of its associated concepts and different readings for one of these meanings (sense-views).

*Figure 3-4: Concepts associated with *book* and some of their readings.*

In addition to establishing unspecified aspects of a meaning, sense-views can also be used to implement polysemic mappings. Let's revisit some of the previous examples:

Mary ate a *fig* for lunch

Mary watered the *figs* in the garden.

Lying in the rain, the *newspaper* was destroyed.

Managing advertising so poorly, the *newspaper* was destroyed.

As discussed in Pustejovsky (1995), the two occurrences of *fig* are related through a plant/food mapping. Similarly, a product/producer mapping (*ibid.*; Lakoff, 1987) can account for the occurrences of *newspaper* (see Pustejovsky, 1995:31 for examples of other mappings). In each case, sense-views can be used to map the underlying sense-concept into the target meaning. Figure 3-5 displays the mapping for the plant/food example. In this picture, the given sense-view is associated with a predefined mapping that projects a food type concept (*fig* in this case) to its related plant type.

Figure 3-5: Using sense-views for polysemic mappings.

Another advantage of the proposed two-tiered word senses is their ability to account for juxtaposition examples. As suggested in 2.7, to deal with sentences such as *Even disciplining his son, a father is a father*, we should be able to separate a typical reading of a word, in this case a *fatherly father*, from its more specific reading, here a *punishing father*.

As discussed earlier, a sense-concept represents the core meaning of a word, therefore, it can naturally be used to refer to the typical meaning of a word, as opposed to a sense-view, which reflects a specific view of its associated sense-concept. Figure 3-6 outlines this approach for the above example.

Figure 3-6: Juxtaposition.

3.4 A Lexicon with Exemplars

The goal of this section is to introduce a model of lexicon in which sense-views are developed in an incremental fashion. That is, over time, sense-views are acquired by and evolve in this model. The key idea in achieving this is to use *exemplars*. In essence, exemplars represent typical use cases, therefore, in this sense they resemble *semantic relations* in Cruse (1986). However, the proposed model has no way of judging the typicality of use cases in advance. Therefore, exemplars represent examples of usage. In this sense, it resembles exemplars in exemplar-based learning approaches (*cf.* Bareiss, 1989).

Consider a dictionary analogy. Imagine a dictionary (such as the *Oxford Student's Dictionary*) that not only lists different meanings for a word, but also further illustrates each meaning with one or more examples, each exemplifying one typical use of the given meaning. Similarly, lexical entries in the proposed model not only represent various word

senses, but also maintain a number of exemplars for each word sense⁶. In this section, I will show how sense-views can be developed automatically and incrementally using exemplars.

The key idea is that as one sense-concept can project many sense-views, one sense-view too can be shared by many sense-concepts. For example, in Figure 3-2, we saw that a *car* can be viewed as a massive object, a mechanical apparatus, and a washable surface. Let's only consider the massive object sense-view. This is a sense-view that can be shared by many other sense-concepts such as a *piano*, a piece of *rock*, or an *elephant*. In other words, despite the differences in what each represents, from the perspective of being weighty, a *car*, a *piano*, a piece of *rock* and an *elephant* are alike (Figure 3-7). From this viewpoint, these representationally different sense-concepts can be classified together. This reflects the distinction Franks (1995) made between two basic functions of concepts, namely *representational* and *classificatory*. He writes:

The *representation* function portrays concepts as stable representations in long-term memory that act as the contents for thinking and inference. The *classification* function postulates some method of feature-detection and comparison for determining whether an entity "falls under" a concept (*ibid.*:442).

Figure 3-7: A sense-view shared by many sense-concepts.

⁶ Similarly, DiMarco and Hirst (1993) propose to exploit *usage notes* to tackle the problem of near-synonymy for lexical choice in text generation.

A similar idea had been advanced by Landman (1986) in the form of *pegs* and *alecs*. Pegs are informational objects, and like conventional pegs are used to hang things on. Things we hang on pegs are properties. When a peg takes a specific role in a domain and becomes an indiscernible approach to the role, it turns into an *alec*⁷.

The idea behind the model of lexicon that I am proposing is that every sense-view brings together a number of sense-concepts that, from that point of view, can be classified together. This means that a *piano*, a piece of *rock*, a *car* and an *elephant* all can be classified together as massive objects. This is to say that a sense-view consists of classificatorily alike, but representationally different elements. Therefore, a sense-view can be considered as a generalization of its associated sense-concepts. The proposed approach to the automatic development of sense-views proceeds from this observation.

To find classificatorily alike sense-concepts in exemplar sentences, I introduce the following relationship:

Definition 1: Sense-concepts SC1 and SC2 are **alike** if they appear in similar contexts, and they share the same thematic role⁸.

For example, in the following contexts *book* and *baseball bat* are alike (both are instruments of breaking):

The book broke the window.

The angry mob smashed the shop-windows with baseball bats.

⁷ After Alec Guinness who plays seven different characters in *Kind Hearts and Coronets*.

⁸ I use the phrase “thematic roles” as commonly understood in linguistics (Fillmore, 1968).

This definition allows us to organize sense-concepts of exemplars into a number of sense-views according to their classificatory functions. In fact, this is exactly how the proposed lexicon is structured. Figure 3-8 gives a schematic view of this structure.

The lexicon is structured in four layers. Lexemes or canonical word forms (displayed with rectangles) appear at the bottom layer. Each lexeme is mapped onto one concept (displayed as a rounded rectangle) or more in the case of lexical ambiguity. Also, more than one lexeme can point to the same concept.

The nature of word-concept associations is rather controversial. For instance, as pointed out by Miller (in Lenat et al., 1995), the concept-word ratio in the EDR dictionary (Yokoi, 1995) is about 4 to 3. However the same ratio in WordNet (Miller, 1995) is 3 to 4. But this issue does not directly concern this dissertation and will not be discussed any further.

Figure 3-8: The schematic structure of the lexicon.

Each concept expresses one meaning of its associated word(s). Following the traditional approach in AI, concepts in this model are represented by a number of *central* properties. These properties are aspects of things that *facilitate comparison* and *explain similarity* (Medin and Goldstone, 1990). These representations fulfill the representational function

of concepts, as discussed by Franks (1995). Properties centrally associated with a concept partially define the concept's meaning. A concept is also defined through the connections it has with other concepts (see the discussion on the ontology network in section 4.3).

Each concept is exemplified by some exemplars (double-lined rectangles). An exemplar is made of a number of sense-concepts (not shown). A sense-concept is the occurrence of a concept in an exemplar. Each sense-concept has a related sense-view and together they describe the sense of the word in question in the given context. Each sense-view (displayed as a wave) is generally shared by a number of representationally different but classificatorily similar sense-concepts. Figure 3-9 gives an example.

Figure 3-9: Lexical entries in the proposed model.

For the sake of simplicity, sense-concepts (i.e. occurrences of concepts in exemplars) are not shown separately. Examples of sense-concepts are MAMMAL-BAT in *Several bats flew in the cave*, and BASEBALL-BAT in *The bat broke the window*⁹.

3.5 From Words to Senses

Being a lexicon, the primary goal of the model is to find proper senses for input words. That is to say that given an input word, the model is expected to first, select a proper sense, and then modulate this sense according to the context surrounding the word. As suggested earlier, this is accomplished by finding an appropriate pair of sense-concept and sense-view for a given word.

The other equally important goal of the proposed model is to enable the incremental and automatic development of sense-views. This section overviews these two tasks. Chapter 4 will detail the implementation of these tasks for sentential contexts.

The fundamental assumption here is that input words are presented to the model as parts of a context. Since both sense selection and sense modulation are strictly contextual processes, it is important to ensure that the words presented to the model constitute a context. Notice this assumption does not entail that the deep structure of the input context or word senses are known. It merely states that the input words are making up a meaningful context.

Sense selection in the proposed model is governed by one criterion: selected senses must allow interaction among themselves; I will say they must be **interactable**.

⁹ These mappings are a simplification we do for clarity. That is, throughout the dissertation, we will assume that first sentence suggests MAMMAL-BAT and that the second does not.

For example, consider choosing between the financial institute and riverbank for the word *bank* in following sentences:

Yesterday I opened a checking account in the bank.

Yesterday I spent some time at the bank.

It is clear that in the first sentence, financial institute is the only sense that can interact with other senses of the sentence¹⁰. In the second sentence, however, both senses can interact with other senses of the sentence. Implicit in the interactability of senses is the postulate that the meaning of a word is being dynamically made up, at least in part, of the meaning of other words it occurs with in a context (for more on this, see the discussion of *semantic traits* in Cruse, 1986)¹¹.

To account for interactability two questions must be answered: a) how can we represent and verify valid interactions? And b) how can we check all possible interactions to find a set of interactable senses for a number of input words. The answer to the first question lies in the use of exemplars: every exemplar represents one valid interaction among its participating sense-concepts. In other words, every exemplar exemplifies one use case of the word senses appearing in the exemplar. That is, each exemplar sets a precedent of one valid interaction between the word senses appearing in that exemplar.

The process described in Figure 3-10 intends to address the second question. The input to this process is a context made out of a number of words. This process is detailed

¹⁰ Given that the other words are not ambiguous.

¹¹ WordExpert (Small, 1987) is an exercise in utilizing this idea for parsing natural language.

in chapter 4. Here, instead, the goal is to overview the model with the help of a walkthrough example.

1. All exemplars in which at least one of the input words appears are located.
2. Search begins to find an exemplar that is *adaptable* to the input context.
3. If an *adaptable* exemplar exists then the input context is patterned after it and output word senses are generated.
4. If the input context is not *adaptable* to any of the exemplars, the trainer intervenes and selects proper word senses.
5. Newly generated word senses modify their associated sense-views.
6. The output generated for the input context is stored along with the existing exemplars.

Figure 3-10: Processing steps in the proposed lexicon.

Suppose (1) is the input to the model:

1. The book broke the window.

Also assume that the following exemplars are already present in the model (Figure 3-11 illustrates parts of the lexicon before inputting (1)).

2. Mary read the book.
3. I bought this book recently.
4. The power outage broke several computers.
5. The thief broke the windshield with a bat.
6. John broke the traffic law.
7. Mary opened the window.

The process begins when input words are looked up in the lexicon and their corresponding lexemes are found. Concepts connected to these lexemes are then reached.

From concepts, all exemplars in which input words appear are located. This set of exemplars represents the model's knowledge, up to that point, of various ways that the input words can interact with other words. In our example, exemplars (2), (3) are located for *book*, exemplars (4), (5) and (6) for *broke*, and exemplar (7) for *window*.

Having located all such possible interactions (i.e. the set of exemplars in which the word appears) for every input word, the challenge becomes to find a set of *interactable* senses for the input. Remember that I assume a number of word senses are interactable only if together they appear in a context. Therefore, to find a set of interactable senses for input words, one can attempt to find an **adaptable** exemplar for the input context.

Figure 3-11: Parts of the lexicon before inputting (1).

An input context is adaptable to an exemplar if it is possible to find a set of word senses for the context's words that are compatible with their corresponding sense-views in that exemplar context. The following definition explains this:

Moreover, these sense-concepts will be connected to their corresponding sense-views in the exemplar. These pairs of sense-concept and sense-view will constitute the output word sense. Figure 3-12 shows the lexicon displayed in 3-11 after processing (1).

The output senses generated for (1) consist of sense-concepts BOOK, SEPARATE - INTO-PIECES and WINDOW-GLASS (as opposed to WINDOW-FRAME), and their associated sense-views (shared with exemplar (5)), namely Breaking-Instrument, Breaking-Action, and Breaking-Object.

Figure 3-12: The lexicon after processing (1).

If no adaptable exemplar for the input context is found, the trainer should intervene. Two distinct situations may occur. First, an adaptable exemplar exists but the system has failed to locate it. This happens specially when the sense-views are not sufficiently developed. For example, consider a Reading-Material sense-view that is instantiated with two sense-concepts, *book* and *magazine* (both printed-media) as in these exemplars:

John read the book.

The student read the magazine.

The model would fail to identify the above exemplars adaptable to an input like: /

The boy read the story.

mostly because *story* is not compatible with the Reading-Material sense-view in which being a printed medium has a strong presence (see 5.11 for more detail). In such cases, the trainer will point to the adaptable exemplar and have the system use it to interpret the input context.

The second situation occurs when the input context is unprecedented, that is, when the input is not adaptable to any of the available exemplars. Consider a situation where after presenting the model with an input like:

John played the piano.

only the following exemplar is located.

John played the tape.

This exemplar is not adaptable to the input (piano and tape have a low compatibility), therefore, the model fails to interpret the input.

In this second case, the input context should become an exemplar itself. To become so, the trainer needs to disambiguate the input words, that is, to choose their sense-concepts. Instantiated with only one exemplar, sense-views for the new exemplar will be created automatically based on its sense-concepts. In the above playing-piano input, the trainer must provide proper sense-concepts for *John*, *play* and *piano*. From them, then, sense-views, Music-Player, Playing-Music and Musical-Instrument will be automatically created (see section 4.6). Being exemplified with only one sense-concept, each sense-view will reflect features of its sole sense-concept; e.g., Musical-Instrument

will be represented only with *piano*'s sense-concept. As more sense-concepts (e.g., various musical instruments) join in, this sense-view will become more representative of musical instruments in general.

In the processing algorithm given in Figure 3.10, following steps 3 (when an adaptable exemplar is found) or 4 (when user intervention was necessary), every input word is disambiguated to a sense-concept and has an associated sense-view. In step 5, input words' sense-concepts will be added (i.e., connected) to their corresponding sense-views. This may lead to some modifications in the sense-views. Such modifications are plausible because being adaptable to an exemplar does not entail that all sense-concepts selected for input words are fully compatible with their corresponding sense-views. In our example, prior to adding *book*, the sense-view Breaking-Instrument contains properties such as made of wood and being a sport gear. Such properties are present because *baseball bat* is the only member of this sense-view. Clearly, these properties will fade away as new members are included in this sense-view.

As another example, recall from section 2.8 how the understanding of our toddler of *objects lifted by a crane* was evolving over time as she encountered new examples. Section 4.4 provides more details on the implementation of this idea. Also, sections 5.2, 5.3 and 5.10 report some empirical results.

And finally, the output generated for the input context will be stored along other exemplars maintained by the lexicon¹³. It is this last step that enables the model to

¹³ In reality, not all outputs are stored in the model. See 4.5.2 on how redundancy is avoided.

address idiosyncrasy, diachrony, and ‘learning’. I will very briefly overview idiosyncrasy and diachrony in this section, and concept development in the next.

Let me begin with idiosyncrasy. Clearly, the process overviewed above is highly dependent on the set of exemplars maintained by the model. These exemplars constitute the model’s knowledge of how words (and concepts they denote) can interact with each other, and hence, provide a repository of use-cases. Therefore, differences in exemplar sets can potentially lead to differences in the model’s interpretations, thus, to idiosyncrasy. As an example, imagine a reader for whom being or not being *kosher* has become part of the sense-view food. This clearly may not be the case for another reader.

The proposed process also supports diachrony. As discussed earlier, output senses generated by the model consist of pairs of sense-views and sense-concepts. As we have seen, sense-views are subject to constant revision, that is, they evolve as new sense-concepts are added. Consequently, because its sense-view(s) can change over time, any word sense can also change over time. For instance, in the example discussed earlier in this section, we saw that adding *book* to the sense-view Breaking-Instrument led to some changes in this sense-view. This also indirectly affects the sense for *bat* in (5) since the sense-concept *Baseball Bat* is now viewed differently through Breaking-Instrument.

3.6 Concept Development

Recall that Barsalou (1982) classifies properties associated with a concept into two categories: 1) context-independent (CI) properties, 2) and context dependent (CD) or *ad hoc* properties. On context-independent properties he writes:

[T]wo types of properties are likely to be frequently active during the processing of a word. First properties having high diagnosticity may often be active, since

they are useful for distinguishing instances of a concept from instances of other concepts...The second type of property likely to be frequently active during the processing of a word includes properties relevant to how people typically interact with instances of the respective concepts. “Edible” becomes CI for “apple” because it is central to how people typically interact with them (*ibid.*:82).

On context-dependent properties he remarks:

In contrast, other properties in a concept rarely seem to come to mind, and when they do, it is only in relevant contexts. For example, “basketball” rarely makes people think of “floats”. However, the sentence frame “Chris used X as a life preserver when the boat sank” would probably bring “floats” to mind for “basketball” when “X” is “basketball” (*ibid.*:82)

The definition of concepts in the proposed model clearly reflects this partitioning. Properties associated with a concept at the concept layer in Figure 3-8 can be thought of as diagnostic properties: properties used to distinguish instances of one concept from another. Each exemplar context constitutes a use-case for participating concepts. On the other hand, sense-views account for the effects that a context “inflicts” on the properties of its constituting concepts.

I claim that the separation of sense-concepts and sense-views offers a neat way for ‘learning’ new concepts, as well as for expanding already known concepts. And by enabling the incremental specification of concepts, we eliminate the need for full specification of concepts in advance. This is extremely useful simply because, generally, there are too many properties that could be associated with a concept. Consider the concept *human* for instance. Several properties can be associated with *human*, such as cognitive/perceptual abilities, all different facts about *human* body, skills specific to various occupations taken up by *humans*, and many more. Rather than expecting a comprehensive set that includes every conceivable property, we can start with a minimum set and allow the system to expand this definition as occurrences of *human* appear in various contexts. Note that here I am not concerned with whether such

properties can be inferred from relations a concept has with other concepts (e.g., a *book* can be a *breaker* because it is a physical object, or one can float on a basketball because it holds pressurized air). My concern only addresses knowledge bootstrapping.

The concept development approach adopted here stems from the fact that a sense-view is more than a simple generalization for a number of concepts with overlapping properties. As discussed earlier, a sense-view may introduce new properties to some of its associated sense-concepts (see sections 5.8 and 5.9 for detailed examples). Therefore, a comprehensive and up-to-the minute set of properties for a concept can be obtained from both 1) the concept's base definition, and 2) all sense-views associated with the concept. A concept is initially defined by its representational properties, present at the concept layer of the lexicon. This base definition is completed with the properties of all sense-views to which the concept connects. This is achieved during the second step of the process displayed in Figure 3.10, when all occurrences of the concept are activated (see the discussion in section 6.5.1 for an example).

Similarly, as previously suggested, some properties of an unknown (i.e., new) word can be established, if the system succeeds in determining the overall context surrounding that word. This is again accomplished using the sense-view assigned to the word in question (see section 5.11 for a detailed example).

3.7 Space and Time Considerations

The process described in Figure 3.10 is a search process aiming to find an adaptable exemplar for a given input. The set of exemplars maintained by the model constitutes the search space. Considering the fact that the model stores generated outputs

as new exemplars, the cardinality of this set could potentially grow very large. This raises serious concerns about the storage space and search time. This section describes measures to deal with these issues.

3.7.1 View-Sets

Recall that *adaptability* is essentially defined with respect to sense-views. Also, a set of sense-views can be shared by many exemplars. Therefore, computing adaptability will be more efficient if we group such a set of exemplars together and compare their shared sense-views once per group, rather than once for every exemplar. Such a group of sense-views, which are shared by a number of exemplars, is called a **view-set**.

Figure 3-13 illustrates a view-set and its exemplars. As displayed, the three exemplars are essentially similar, meaning their sense-concepts share similar sense-views. In this case, rather than examining these exemplars individually, we can instead measure the adaptability of their view-set to the input.

Figure 3-13: Example of a view-set and its exemplars.

A view-set is initially defined by one exemplar. It develops as new exemplars are included in it. Section 4.5 discusses the development of view-sets in more detail. For

now, the point is that through the use of view-sets, we can considerably reduce our search space, by focusing on groups of similar exemplars, instead of individual exemplars.

3.7.2 Distributed/Concurrent Processing

To increase performance (with respect to processing time), I propose a distributed and concurrent process for locating adaptable exemplars. This strategy is close to the memory-based reasoning paradigm (Stanfill & Waltz, 1986) in which concurrency is exploited to address the efficient processing of large amounts of data.

As described in Figure 3-10, following the presentation of an input to the model, all exemplars in which at least one of the input words appear will be located. But in fact, it is a relatively small number of view-sets that will be identified. And, as I will explain later, each view-set can independently and simultaneously measure its adaptability to the input. Should a view-set find itself adaptable to the input, it can proceed with the generation of word senses. (Details about the implementation of this process will be discussed for the prototype described in Chapter 4.)

3.7.3 Selective Storage of Exemplars

Finally, not all generated outputs need to be stored by the model. Only those outputs that cause any modification in sense-views, and therefore, bear new information for the model, will be stored.

Chapter Four

A Lexicon for Sentential Contexts

4.1 Introduction

In the preceding chapter, I sketched out a model to account for effects of context in highlighting or dimming certain features of word senses during interpretation. Mechanisms for selecting and modulating word senses, as well as for the incremental development of patterns of contextual effects were also overviewed.

What was presupposed throughout chapter 3 was the ability to recognize context. In other words, given a number of words as an input context, I assumed that the model would know how to assess the similarity of the input context with the current stored exemplars. Since it is only after locating a similar context that the model can resolve sense selection and sense modulation, context recognition is obviously a crucial facet of the proposed model.

This chapter starts by addressing this issue and how it is dealt with in the implemented system. Then, in section 4.3, I will give an overall view of the implemented system. The system is structured in two separate yet connected structured networks: the lexicon network and the ontology network. These two networks are detailed in the two subsequent sections. Finally, in the last section, section 4.6, I will discuss ‘learning’ in the implemented model.

4.2 Working Assumptions

The first assumption concerns the adopted context recognition approach. Context and context recognition are possibly among the most complex issues facing any language understanding efforts. In this section, I first overview context and context representation approaches and then outline the context recognition approach I adopt.

4.2.1 Context Representation

The term context has different connotations in the literature, varying from words uttered before and after a word in question (linguistic context) to a more general form that includes linguistic context as well as all the non-linguistic information that a speaker knows about a particular situation (Miller & Charles, 1991). As for the nature of context recognition, contradicting ideas have been proposed. In Schank’s conceptual dependency theory (1975), an active context was represented through the expectations it caused. This top-down approach was further pursued with more sophisticated symbolic structures of various sorts such as goals, plans, and themes, that attempt to capture contextual information at different levels of abstraction (see Dyer, 1983; Schank, 1980; Schank & Abelson, 1977). Adopting a contrasting bottom-up view, Kintsch et al. (1990), argued for

three levels in discourse representation, 1) surface (lexical) level, 2) semantic level, and 3) situational level (see also Kintsch, 1988). At the lexical level, a text is specified by the exact words and phrases used. At the second level, the semantic content of the text is represented. What is characterized at the situation level is “the situation described by the text, detached from the text structure proper and embedded in pre-established fields of knowledge” (Kintsch et al., 1990: 135). In other words, text at this level is assimilated in a knowledge schema (e.g., an appropriate script or frame). Speelman and Kirsner (1990) have provided evidence regarding the availability of both text-based and situation-based information in memory. This diversity among context recognition approaches continues in representational approaches. It includes logical propositions (e.g., Kintsch et. al, 1990), distribution over micro-features (Waltz & Pollack, 1985), structured semantic networks (e.g., Alterman & Bookman, 1992), distribution over unnamed nodes (e.g., Miiikkulainen, 1993), and lexical chains (Hirst & St-Onge, 1995), to name a few. The advantages and disadvantages of these methods are beyond the scope of this dissertation (for a comprehensive review, see Corriveau, 1995, chapter 2).

In order to avoid such controversies and to have a workable scope, the present implementation was restricted to simple sentential contexts. That is, instead of working with full fledged contexts, both input contexts and exemplar contexts are limited to simple single-claused sentences (see next chapter for examples). The other rationale in favor of sentential context is the availability of considerable experimental psycholinguistic data on semantic flexibility for sentential contexts (e.g., Greenspan, 1986; Witney et. al, 1985; Barsalou, 1982; Anderson et al., 1976; Barclay et al., 1974).

Consequently, in the current implementation, a context is represented by means of its concepts and their thematic roles. For example, the context representation for a sentence like *The book broke the window* would include the representations of concepts BOOK (concept corresponding to the word *book*), SEPARATE-INTO-PIECES (for *break*), and WINDOW-GLASS (for *window*). The system will also know that BOOK and WINDOW-GLASS are respectively Instrument and Object of Action SEPARATE-INTO-PIECES.

4.2.2 Case-Based Context Recognition

Having restricted the system to simple sentential contexts, I must now describe the system's context recognition approach. As discussed earlier in 3.5, given an input context as a number of words, the model succeeds to select proper senses and to modulate them only when at least one of the stored exemplars is found to be *adaptable* to the input context. This adaptable exemplar will then be used as a pattern to process the input context. In other words, the model generally processes a novel case based on an already processed case. This strategy clearly favors a case-based context recognition approach (see Bareiss, 1989 for a comprehensive review of case-based reasoning applications and techniques; also Jones, 1996 for a NLP application). Indeed, this is how context recognition is handled in the current implementation.

Fundamental to any case-based system is its similarity measurement mechanism. In my case, this corresponds to how the similarity of an input context to an exemplar context is measured. In definition 2 (section 3.5) I introduced the notion of *adaptability* between an input context and an exemplar context. Now I must describe how this

measure can be implemented. More precisely, I must provide answers to the following questions:

- 1) How does every word participate in the making of the context it appears in (i.e., what is the composition function f)?
- 2) How are thematic roles recognized for an input context? And
- 3) How is the compatibility between a sense-view and a sense-concept determined?

Measuring Adaptability: Having restricted context to a sentence, the adaptability between two contexts could be approximated by a linear function f . More precisely, in order to measure the overall likeness between two contexts, one can measure the individual compatibility between corresponding elements and compose the partial results into a single measure (see Algorithm 4-5 for more details).

Determining Thematic Roles: Another assumption addresses the second question: how to determine thematic roles for an input context? It is assumed that case markers (Delisle et al., 1996; 1993) provide good indications about the thematic roles for the noun phrases they qualify. Examples of case markers are p-subj (positional subject) and with-pp (prepositional phrase *with*). Consider the role Instrument for example. We can assume that p-subj and with-pp, among other roles, signal an instrument in a context, as in the following examples:

The man broke the window with a bat.
The rock broke the window.

Sense-Concepts/Sense-Views Compatibility: The compatibility between a sense-concept and a sense-view is determined based on their property sets. This measure is calculated based on the number and weights of shared properties versus those different. Details of this measurement can be found in section 4.5.2.

4.3 Design

Architecturally, the system exhibits a hybrid nature: Although it is built on top of two localist networks, its behavior is not completely accomplished through a spreading activation process (see below). Instead, its functionality largely depends on the execution of a number of processing elements, called *agents*, attached to various nodes of the two networks. This section provides an overall picture of the architecture of the implemented system.

4.3.1 System Architecture

Localist models, also known as structured connectionist networks, are simple forms of semantic networks. Unlike semantic networks, localist networks do not have labeled arcs (Miikkulainen, 1993). They are also recognizable by their use of spreading activation as a general constraint satisfaction method. In spreading activation, subsequent to an initial activation stage, continuous values are propagated through weighted links between nodes and eventually settle into a stable pattern. The set of the most highly activated nodes represents the interpretation of the input (*ibid.*). Examples of such networks are those of Corriveau (1995), Bookman (1994), Lange and Dyer (1989), and Waltz and Pollack (1985).

Since they are based on relaxing the network activity over time, localist models are particularly successful in accounting for semantic priming effects and timing in cognitive tasks (Miikkulainen, 1993). They also offer knowledge-level parallelism: more than one inference path can be pursued simultaneously (*ibid.*; Oden, 1987). However, network structure and connection weights must usually be hand-coded. Learning is also typically a problem for such models. But more important, they hardly support the declaration and application of rules, and sequential control; features very well supported by symbolic models. These observations justify why the proposed model is a hybrid one (see Miikkulainen, 1993; Bookman, 1994).

Figure 4-1: The system's overall architecture.

Figure 4-1 gives an overview of the system. As illustrated, the system is structured into two separate but connected networks. The bottom network, the ontology network, represents the system's knowledge base, which allows the system to make sense of concepts in the lexicon network. The lexicon network, consisting of lexemes, concepts, exemplars and sense-views (see Figure 3-8), is responsible for receiving an input context

and generating proper senses in the output. These two networks are discussed in sections 4.4 and 4.5.

4.3.2 Processing Agents

Though grounded in two localist networks, the system achieves much of its functionality through symbolic processing. To explain this, we should take a closer look at the underlying networks. Unlike conventional localist networks, nodes here are not simple processing units that merely spread activation. Instead, attached to every node is an *agent*: an autonomous processing element capable of performing a sequence of instructions.

Agents work independently and simultaneously. They exchange *messages*, (i.e., typed activation signals) with other agents. When an agent receives a message, it becomes active and starts executing its instructions. Generally, associated with each message is a numerical value, representing some activation strength. This value ranges between 0 (minimum) and 1 (maximum activation). All messages are subject to decay. This decay occurs when a message spreads from one node to another. As discussed in Corriveau (1995, chapter 2), this approach is used by many systems (e.g., Hendler 1989, 1987) to address the convergence and promiscuity problems of spreading activation models.

This message decaying is governed by two factors. The first factor indicates the percentage of the activation strength exhausted as the result of decay. The second factor is a threshold representing a minimum strength for a message to be exchanged between nodes. In other words, only activation messages greater than this threshold are

propagated. The value of these factors varies for each message type (see 4.4.3 and 4.5.2 for actual values used for each type).

One agent type exists for every node type in the ontology and lexicon networks. Some of these agent types have a simple behavior. They receive a message and spread it along their outgoing connections. Others, however, have more complicated behavior. Sections 4.4.3 and 4.5.2 will respectively describe the exact behavior of agents in the system.

4.3.3 Implementation

The system is implemented using Java (Lea, 1997; Niemeyer & Peck, 1996). Java was selected because of its rich and robust concurrency support mechanisms. Adopting object-oriented technology, the implementation consists of a number of object classes and their operations. In total, more than 20 classes are implemented including: *Lexicon*, *Lexeme*, *Concept*, *Exemplar*, *ExemplarNode*, *ViewSet*, *Sense*, and *Ontology*. Examples of these classes and of their methods will be given during the discussion of the ontology and lexicon networks in the following sections.

Agents are implemented as procedures operating on instances of these classes. As for concurrency, although in theory all agents are activated and executed simultaneously, in the current implementation only agents attached to view-sets run concurrently. This is because only view-set agents are complicated enough to justify the use of concurrency. Agents attached to view-sets, implemented through the class *ViewSetThread*, are implemented as Java threads (see 4.5.2).

4.4 Ontology Network

The ontology network plays two important roles in the system. First and foremost, input words are interpreted in terms of their mappings onto the ontology network. More specifically, it is through the ontology network that properties associated with a concept and their centrality are determined. The second role of the ontology network is to provide a measure for **semantic distance** (cf. Hirst & St-Onge, 1995) by linking concepts using a taxonomic relation. This section discusses the implemented ontology network along these two roles. It also describes the structure and behavior of the ontology network.

4.4.1 Why An Ontology?

As pointed out in the preceding chapter, a word in this model is associated with one or more concept, each concept representing one distinct meaning for that word. Each concept in turn is represented by a number of properties that essentially distinguish it from other concepts. Properties associated with a concept are not of equal importance, or *centrality*, in identifying the concept. For instance, being solid is not very distinctive, though necessary, for *book*: there are too many other solid objects. In other words, *book* shares this property with many other objects. Viewing *book* from a taxonomic perspective, we can say that solidity is an *inherited* property, because book is a physical object, for which solidity is a central property.

It is *via* the ontology network that the system determines a) the set of properties associated with a concept, and b) the centrality of each property in that set.

The second application of the ontology network, namely the ability to locate *close* concepts, is more relevant to the adopted exemplar-based approach. One frequent

problem with exemplar-based models is *coverage* (Jones, 1996). Coverage considers the adequacy of exemplars maintained by a model in providing reasonable coverage of a problem domain. Here, locating and eventually activating concepts close to a given concept is how the system extends its coverage. The idea is that knowing something about concept *A* must facilitate the processing of concept *B* if they are conceptually close. For example, consider a case in which *book* appears in a printing context. Suppose this is an unprecedented usage of *book*. Therefore, the model would fail to correctly process *book* unless it can exploit a close concept, like *magazine*, in a similar context, e.g., publication.

Generally, finding an ontology is rather difficult. Delisle et al. (1996) writes:

[I]t is difficult to find an “off-the-shelf” ontology of the right grain, not so general as to be useless while not so specific as to be tied to the domain(*ibid.*: 41).

Presently, ontological knowledge is anything but standard (*ibid.*). I used the Mikrokosmos ontology (Mahesh & Nirenburg, 1995)¹⁴. Mikrokosmos offers a comprehensive and language independent ontology. Indeed, Mikrokosmos has been used in a large-scale machine translation application (Mahesh & Nirenburg, 1996). Figure 4-2 shows an example of taxonomic structures in Mikrokosmos. Some of the properties for *book* are listed in Figure 4-3. Notice properties appearing below the line are properties inherited from nodes that directly or indirectly subsume *book*.

Except for a handful of cases, sample runs reported in this thesis are entirely based on initial objects and properties of Mikrokosmos. More specifically, I added a few

¹⁴ Available for downloading at <http://crl.nmsu.edu/Research/Projects/mikro>.

concepts, such as *shark*, to Mikrokosmos. And, in some cases, I introduced new properties. This was necessary mainly because Mikrokosmos is more concerned with taxonomic relations between concepts than with specifying properties for concepts.

Figure 4-2: An example of Mikrokosmos hierarchies.

Figure 4-3: Examples of Mikrokosmos properties.

4.4.2 Network Structure

Every node in the ontology network consists of a set of properties representing a concept, pointers to super/sub classes, and the node's associated concept(s) in the lexicon network (see Figure 4-4.a). The property set is a frame consisting of a number of weighted slots (see Figure 4-4.b). Each slot represents a property and one or more values (slot fillers). The weight represents the property's centrality in the definition of the concept.

Here is an example of how an ontology node is constructed (given that N8 and N23 are already defined):

```
OntoNode N13= NewNode("BOOK", Book) ;
IsKindOf(N13, N8) ;      // book-document
IsKindOf(N13, N23) ;     // printed-media
```

The constructor takes two arguments: a title for the node (e.g., BOOK) and a property frame (e.g., Book).

Here is the definition for Book:

```
Frame Book = new Frame() ;
Book.AddSlot("HAS-PARTS", 1) ;
i = Book.SlotFiller("HAS-PARTS", "PAGE") ;
Book.AddSlot("MADE-OF", 1) ;
i = Book.SlotFiller("MADE-OF", "INK") ;
i = Book.SlotFiller("MADE-OF", "PAPER") ;
i = Book.slotFiller("MADE-OF", "PARCHMENT") ;
Book.AddSlot("CONTAINS", 1) ;
i = Book.SlotFiller("CONTAINS", "INFORMATION") ;
```

Note how weight and possible values are specified for each property (slot). For an example of connecting ontology nodes with their corresponding concepts in the lexicon network, see section 4.5.

Figure 4-4: (a) An ontology node, (b) A property frame.

4.4.3 Network Behavior

The behavior of the network is best described based on the activation messages sent to and from nodes. Corresponding to the two roles of the ontology network, two types of messages are sent from/to the network nodes: GetProperties and Activate. GetProperties is initially sent to a node in the ontology network by its corresponding concept node in the lexicon network. This happens when a concept node receives an activation message from one of its connected lexeme nodes (see next section).

Once received by an ontology node, the agent attached to the node becomes activated and carries out a few steps. It first makes sure that the message has not decayed below a given threshold. Assuming so, it then propagates the message to all its superclasses to get the inherited properties. The property sets returned by superclasses in

response are merged with the node's local property set. This is what will be sent back to the original sender of the message. Algorithm 4-1 summarizes this process.

In addition to controlling the extent of propagation, activation strength in GetProperties is also used to determine the centrality of properties. It is assumed that properties specific to a node directly activated by a concept node are all central (i.e., weight is 1) in that concept. Properties inherited from nodes activated as the result of the propagation of GetProperties to superclass nodes are less central, depending on their distance from the original node. This process is governed by two parameters, NonLocalityFactor and LocalityThreshold (set to 0.9 and 0.6 respectively). For instance, the centrality of a property inherited from a category two layers above would be 0.81. Also, choosing 0.6 entails that propagation will not go beyond the fourth level up (since $(0.9)^5$ is less than 0.6). That is to say that concepts inherit properties from concepts up to 4 levels above. Here are properties returned by the ontology node *book* in response to the GetProperties message. Note the number enclosed in brackets represents the property weights, and possible values for a property follow the arrow.

```
CONVEYS(0.6561) ---->: INFORMATION,
IsKindOf-REPRESENTATIONAL-OBJECT(0.6561) ---->:
IsKindOf-LANGUAGE-RELATED-OBJECT(0.729) ---->:
PRODUCED-BY(0.81) ---->: HUMAN,
IsKindOf-DOCUMENT(0.81) ---->:
IsKindOf-BOOK-DOCUMENT(0.9) ---->:
OPERATED-BY(0.6561) ---->: HUMAN,
AGE(0.6561) ---->:
COLOR(0.6561) ---->:
STATE-OF-MATTER(0.6561) ---->: SOLID,
IsKindOf-ARTIFACT(0.6561) ---->:
REPRESENTS(0.729) ---->: LANGUAGE, EVENT, OBJECT,
IsKindOf-MEDIA-ARTIFACT(0.729) ---->:
IsKindOf-VISUAL-MEDIA-ARTIFACT(0.81) ---->:
LOCATION(0.9) ---->: ACADEMIC-BUILDING,
IsKindOf-PRINTED-MEDIA(0.9) ---->:
```

HAS-PARTS(1) ---->:PAGE,
MADE-OF(1) ---->:INK, PAPER, PARCHMENT,
CONTAINS(1) ---->:INFORMATION,

Receive GetProperties (p)
If p has not decayed
 Get inherited properties (broadcast GetProperties
 to all superclasses).
Await the result.
Merge local properties with inherited ones.
Return the result set.

*Algorithm 4-1:*The behavior of the ontology node agent (GetProperties).

Many of these properties are of the type IsKindOf. These properties are generated by the ontology node agents whenever the GetProperties message reaches a class which subsumes the sender (i.e., the sender's superclass). For example, as shown in Figure 4-2, *book* is a subclass of *Printed-Media*, which in turn is subsumed by *Visual-Media-Artifact*. This explains the two properties *IsKindOf-PRINTED-MEDIA(0.9)* and *IsKindOf-VISUAL-MEDIA-ARTIFACT(0.81)*. Generating such properties is important since many concepts in *Mikrokosmos* do not have any properties, due to that fact that, I repeat, taxonomic relation between concepts in *Mikrokosmos* is more important than specifying properties for concepts.

The second type of messages exchanged between the ontology network nodes is Activate. Like the previous message, it is also originally sent by a concept in the lexicon network to its corresponding node in the ontology network. As discussed earlier, it is essential to muster the system's knowledge by activating all concepts (and ultimately their associated exemplars) close to an input concept. This is accomplished by spreading

the activation to all sub/superclass nodes. In addition to sub/superclass nodes, activation is also sent to the node's associated concepts (see Figure 4-4(a)). Algorithm 4-2 overviews this functionality¹⁵.

The two factors controlling the activation decay for Activate messages are ActivationLossFactor and ActivationThreshold. As explained before (section 4.2.3), the first parameter indicates the percentage of activation strength that a message loses due to decay when propagated from one node to another. The second factor, ActivationThreshold, specifies a minimum strength required for a message to be exchanged between nodes¹⁶.

Receive Activate α
If α has not decayed
 Propagate α to all subclasses.
 Propagate α to all superclasses.
 Activate all connected concept(s).

Algorithm 4-2: The behavior of the ontology node agent(Activate).

The following excerpts are from the activation log generated for inputs S1 and S2 for the test reported in section 5.2. First, the word *move* is presented to the system. Consequently, the lexeme node Move, and its associated concept MOVE-AWAY are activated (see 4.5 for more details). The activation is then spread to the ontology node corresponding to MOVE-AWAY concept. The activation is then recursively sent to sub/super classes, which eventually leads to the activation of the ontology node CROSS

¹⁵ Details on how the arc that the activation message was received from is excluded from the broadcast are not shown.

¹⁶ See Chapter 5 for the values of these parameters.

and its associated concept GO. This is an example of how a *close* concept can be reached using the ontology network. As shown, as the result of activating GO, two extra exemplars are activated for this input:

```
-----processing Move
>>>> activating lexeme Move
>>>> activating concept MOVE-AWAY
////////Activating OntoNode MOVE-AWAY
////////Activating OntoNode MOTION-EVENT
////////Activating OntoNode CHANGE-LOCATION
////////Activating OntoNode CROSS
>>>> activating concept GO
>>>> activating (Action) from Exemplar Ex11
>>>> activating (Action) from Exemplar Ex5
```

The next example also reports similar activity for the input word *John*, which ultimately activates concepts HUMAN-FEMALE and BOY:

```
-----processing John
>>>> activating lexeme John
>>>> activating concept HUMAN-MALE
////////Activating OntoNode MALE
////////Activating OntoNode HUMAN
////////Activating OntoNode PRIMATE
////////Activating OntoNode SOCIAL-ROLE
////////Activating OntoNode FEMALE
>>>> activating concept HUMAN-FEMALE
>>>> activating (Agent) from Exemplar Ex11
>>>> activating (Agent) from Exemplar Ex8
////////Activating OntoNode BOY
>>>> activating concept BOY
>>>> activating (Object) from Exemplar Ex12
>>>> activating (Agent) from Exemplar Ex10
>>>> activating (Agent) from Exemplar Ex5
>>>> activating (Agent) from Exemplar Ex4
>>>> activating (Agent) from Exemplar Ex3
```

4.5 Lexicon Network

The lexicon network implements the model of lexicon described in section 3.3 (see Figure 3-8). Thus, the network is structured into four layers (see Figure 4-5). These

layers parallel the layers of the proposed model of lexicon. The only difference is that, here, sense-views related to one particular context are grouped together into a view-set (see 3.7).

This section, first explains the structure of different nodes in the lexicon network, and continues with a detailed description of the tasks carried out by individual node agents.

View-sets

Exemplars

Concepts

Lexemes

Figure 4-5: The structure of the lexicon network.

4.5.1 Network Structure

Lexeme nodes appear in the bottom layer. Each node consists of a word form and its syntactic category (Figure 4-6). The model presently, I repeat, is not concerned with morphological processing. Therefore, lexemes are represented as word forms. For example, this is how *bank* is defined in the implemented system:

```
Lexeme l_Bank = NewLexeme("Bank", "noun") ;
```

Figure 4-6: The structure of lexeme nodes.

Every lexeme node also points to one or more (in case of lexical ambiguity) concept nodes in the next layer. Each such concept represents one *contrastive*¹⁷ meaning for the word represented by the given lexeme. The following statements specify the concepts associated with *bank* (see below for the definition of *c_Bank1* and *c_Bank2*):

```
l_Bank.Denotes(c_Bank1) ;      // river bank
l_Bank.Denotes(c_Bank2) ;      // financial institute
```

Concept nodes constitute the second layer of the network. Figure 4-7 displays the structure of a concept node. As shown, concept nodes do not store properties for concepts. In fact, only a label is maintained by a concept node and concept properties are obtained through connections with the ontology network (see GetProperties message in the previous section).

Figure 4-7: The structure of concept nodes.

¹⁷ As used by Pustejovsky (1995). Also see section 2.6.

Here are examples of concept definitions in the implemented system. The method `Concept2OntoNode(aConcept, anOntoNodeTitle)` connects a concept node with an ontology node (represented via its label):

```
Concept c_Bank1 = new Concept("RIVER-BANK") ;
anOntology.Concept2OntoNode(c_Bank1, "RIVER-BANK") ;
Concept c_Bank2 = new Concept("BANK-BRANCH") ;
anOntology.Concept2OntoNode(c_Bank2, "BANK-BRANCH") ;
```

A concept node also points to its occurrences, i.e. sense-concepts (SC), in a number of exemplar contexts. An exemplar node, representing an exemplar context, comprises a number of sense-concepts (Figure 4-8). As discussed earlier, a sense-concept is the occurrence of a concept in a context. Associated with every sense-concept are 1) the thematic role it takes in the context, and 2) the case-marker it was specified with in the sentence (see the example in Figure 4-10). Since by definition the role is the same for all sense-concepts that share a sense-view, sense-concept nodes do not have to hold thematic roles. Instead, the role of a sense-concept in an exemplar can be obtained from the sense-view to which this sense-concept points. This is an example of how exemplars are defined in the system:

```
Exemplar Ex8 = new Exemplar("Ex8") ;
```

Figure 4-8: Structures of a) exemplar nodes, b) sense-concepts.

Given the view-set `vs_ReadMaterial` (see below) and concepts `FemaleHuman`, `c_Read`, and `c_Article`, the following statements show how sense-concepts are added to this exemplar:

```
// Mary read the article.
Ex8.Exemplifies(c_FemaleHuman, "p-subj", "Agent", vs_ReadMaterial)
;
Ex8.Exemplifies(c_Read, "verb", "Action", vs_ReadMaterial) ;
Ex8.Exemplifies(c_Article, "p-obj", "Theme", vs_ReadMaterial) ;
```

A number of sense-views, representing various roles associated with a specific context, are grouped together into a view-set (Figure 4-9). A sense-view node holds information about: 1) the thematic role it represents, 2) syntactic case-markers signaling the role, 3) properties associated with the sense-view, 4) number of connected sense-concepts, and 5) a threshold value representing the minimum number of sense-concepts the sense-view needs to become *stable* (as explained later).

Figure 4-9: Structures of view-set nodes.

Figure 4-10 shows two exemplars and the view-set to which both are connected. The following code snippet is an example of view-set definition. As shown, besides the label, a stability threshold value is also specified for each view-set. This value will be used for all sense-views included in this view-set.

```
ViewSet vs_ReadMaterial = NewViewSet("ReadMaterial", 5) ;
```

Figure 4-10: Two exemplars and their associated view-set.

4.5.2 Network Behavior

Much of the functionality of the system comes from the lexicon network. Input contexts are processed by the lexicon network. The following is an example of how an input context (as a number of word forms with their case markers) is presented to the system (see section 5.6):

```
// The book broke the window.
TestLex1.NewInput("S9") ;
TestLex1.AddInputWord("Book", "p-subj") ;
TestLex1.AddInputWord("Break", "verb") ;
TestLex1.AddInputWord("Window", "p-obj") ;
TestLex1.Process(LogFile) ;
```

Algorithm 4-3 gives an overall view of the behavior of the network. This algorithm implements the processing steps of the proposed model, as discussed in section 3.5 (see

Figure 3-10). The algorithm is carried out by agents attached to the network nodes. Using the above example, different steps of the algorithm along with agents performing them will be discussed throughout the rest of this section.

1. A lexeme node is activated for an input word.
2. Activation is spread to concept nodes connected to the lexeme node.
3. Each activated concept node activates i) connected sense-concepts (exemplar nodes), and ii) the connected ontology node (to activate *close* concepts).
4. Activation is passed to sense-views connected to the activated sense-concepts, which eventually leads to the activation of view-set nodes.
5. Each activated view-set node measures its alikeness with the input context (determine if they are *adaptable*).
6. If an adaptable view-set for the input is found, output word senses are generated accordingly; otherwise a user intervention is necessary.
7. Newly generated sense-concepts are connected to their corresponding sense-views (which may result in the modification of the sense-views properties).
8. If not redundant, the generated output will be stored as an exemplar.

*Algorithm 4-3:*The overall behavior of the lexicon network.

Step 1: The process starts when input words are looked up in the lexicon and consequently their corresponding lexeme nodes are activated. Here is the activation log:

```
-----processing Window  
>>>>> activating lexeme Window
```

```

-----processing Break
>>>> activating lexeme Break

-----processing Book
>>>> activating lexeme Book

```

Step 2: Once activated, the attached lexeme agents spread the activation to their connected concepts. This is what happens in our example:

```

-----processing Window
>>>> activating lexeme Window
>>>> activating concept WINDOW

-----processing Break
>>>> activating lexeme Break
>>>> activating concept BREAK-VIOLATE
>>>> activating concept BREAK-SEPARATE INTO PIECES

-----processing Book
>>>> activating lexeme Book
>>>> activating concept BOOK

```

Step 3: Agents attached to the activated concepts, in turn, activate 1) their corresponding nodes in the ontology network, and 2) their occurrences (i.e., sense-concepts) in exemplar contexts. Recall that the activation of a node in the ontology network leads to the activation of close concepts in the lexicon networks. To continue with our example, below is the complete activation trace up to this point (see Section 5.1 for the list of exemplars used in this example). Notice how the exemplar Ex10 is activated because of *newspaper*, a close concept of *book*. Also notice how *break* activates, along with other exemplars, the output generated for one of the previously processed input (input S8). For more on the storage of generated output, see the discussion of **Step 8**.

```

-----processing Window
>>>> activating lexeme Window
>>>> activating concept WINDOW
////////Activating OntoNode WINDOW
////////Activating OntoNode BUILDING-PART
////////Activating OntoNode BUILDING-ARTIFACT

```

```

>>>> activating (Object) from Exemplar Ex2

-----processing Break
>>>> activating lexeme Break
>>>> activating concept BREAK-VIOLATE
////////Activating OntoNode VIOLATE-LAW
////////Activating OntoNode CRIMINAL-ACTIVITY
////////Activating OntoNode SOCIAL-EVENT
>>>> activating (Action) from Exemplar Ex3
>>>> activating concept BREAK-SEPARATE INTO PIECES
////////Activating OntoNode BREAK
////////Activating OntoNode CHANGE-IN-QUALITY
////////Activating OntoNode CHANGE-STATE
>>>> activating (Action) from Exemplar Added by Output(S8, Breaking)
>>>> activating (Action) from Exemplar Ex4
>>>> activating (Action) from Exemplar Ex1

-----processing Book
>>>> activating lexeme Book
>>>> activating concept BOOK
////////Activating OntoNode BOOK
////////Activating OntoNode PRINTED-MEDIA
////////Activating OntoNode VISUAL-MEDIA-ARTIFACT
////////Activating OntoNode PERIODICAL
////////Activating OntoNode PAGE
////////Activating OntoNode NEWSPAPER
>>>> activating concept NEWSPAPER
>>>> activating (Theme) from Exemplar Ex10
////////Activating OntoNode FORM
////////Activating OntoNode TICKET
////////Activating OntoNode ADVERTISING-INSERT
////////Activating OntoNode BOOK-DOCUMENT
////////Activating OntoNode DOCUMENT
////////Activating OntoNode TEXT-MANUAL
////////Activating OntoNode JOURNAL
////////Activating OntoNode DIRECTORY
////////Activating OntoNode MAIL-ORDER-CATALOG
////////Activating OntoNode ENCYCLOPEDIA

```

Step 4: Activated sense-concepts pass the activation to their connected sense-views. A view-set becomes activated once any of its sense-views receives activation. In this example, three view-sets become activated. Note that a total of five exemplars originally received activation.

```

***** Entering ViewSet Breaking
***** Entering ViewSet ReadMaterial

```

***** Entering ViewSet LawViolation

Step 5: Once a view-set receives activation, its attached agent awakes and starts measuring the adaptability of its view-set to the input context. Since view-set agents are implemented as Java threads, the adaptability between the input and all the activated view-sets is measured concurrently. Algorithm 4-4 describes how view-set agents determine whether an input is adaptable to their view-set, using a metric named *aliqueness*. This metric results from the individual compatibility measures between sense-views on one side and input words on the other side.

1. For every input word *W*
 - Using its case marker, find *W*'s corresponding sense-view(s).
 - Measure the compatibility of each sense-view with each of the *W*'s associated concepts.
 - If there are more than one corresponding sense-views or associated concepts, use the pair with maximum compatibility.
2. Calculate the overall *aliqueness*.

Algorithm 4-4: Measuring aliqueness.

Before going any further let us look at an example. One of the activated view-sets during Step 4 was Breaking. To determine if the input is adaptable to this view-set, its agents start measuring the compatibility of each input word with its corresponding (with respect to case markers) sense-view. Consider the input word *window* for instance. Being marked with p-obj, it will match with the sense-view OBJECT in the Breaking view-set. To decide on the compatibility of these two, *window*'s associated concept(s) is found and its property set will be compared with the properties present in the Object sense-view. Here is the result:

Compatibility of Sense-View (Breaking, Object) with WINDOW is: 1

Therefore, the compatibility of a sense-view with an input word is measured in terms of the compatibility of the sense-view's property set with the property set of the word's associated concept(s). In case of lexical ambiguity, where more than one concept is associated with a word, the greatest value will be considered. For example, consider the word *break* when it is compared with the sense-view Action of the same view-set. Being ambiguous, *break* is connected to two concepts: separate into pieces and violating law. Clearly, the former is a perfect match for this sense-view (because their respective properties overlap completely), and therefore, this reading of *break* is determined more compatible with this sense-view than the other.

Compatibility of Sense-View (Breaking, Action) with BREAK-SEPARATE INTO PIECES is: 1

The compatibility between properties of a sense-view and properties of a concept is defined as the sum of those properties of the concept that are shared by the sense-view divided by the total weight of the sense-view properties.

The following formula describes the compatibility between a sense-view and a concept:

$$\sum \min(VP_i(w) , CP_i (w)) / \sum VP_i (w) \quad \text{for every property of the sense-view}$$

where $VP_i(w)$ is the weight of the i -th property of the sense-view and $CP_i (w)$ is the weight of the same property (if any) in the concept's property set. For those properties of sense-view that do not appear in the concept's property set, this value will be 0. For example, here are the property sets for the sense-view Object in view-set BREAKING and *window*.

```

ViewSet:Breaking
[Sense View]
Thematic Role: Object - No. of Exemplars: 3 - Marker(s): p-obj,
STATE-OF-MATTER(0.513)---->:SOLID,
FRAGILITY(0.666667)---->:FRAGILE,
MADE-OF(0.903333)---->:METAL,CERAMIC,PLASTIC,GLASS,

[WINDOW]
IsKindOf-INANIMATE(0.6561)---->:
OPERATED-BY(0.729)---->:HUMAN,
AGE(0.729)---->:
COLOR(0.729)---->:
STATE-OF-MATTER(0.729)---->:SOLID,
IsKindOf-ARTIFACT(0.729)---->:
IsKindOf-BUILDING-ARTIFACT(0.81)---->:
IsKindOf-BUILDING-PART(0.9)---->:
MADE-OF(1)---->:GLASS,
FRAGILITY(1)---->:FRAGILE,

```

Having described the calculation of individual compatibility measures, I should now specify how the overall likeness is obtained. In fact, this overall likeness constitutes the similarity measurement mechanism adopted in this implementation. As discussed in 4.1.2, the similarity measurement in an analogy-based system specifies how one of previously processed cases is selected to guide the processing of a novel case.

Considering the intricacies of contexts, even of simple sentential contexts, integrating individual compatibility measures into a single measure indicating the overall likeness of an input context to one of the stored exemplars is not a trivial task. For this work, I only consider two intuitive situations, each leading to a different formula.

The first case happens when a view-set becomes activated by the verb in the input context. For instance, suppose the following exemplars are available in the system repository:

The judge read the verdict.

The man read the book.

John reads physics in the university.

For an input like *The teacher read the novel*, all three exemplars will receive activation due to the appearance of the verb *read* in the input. Clearly in such cases, noun phrases play a crucial role in disambiguation of the verb.

The second case happens when a view-set is activated differently, that is, *via* a noun phrase. For example, suppose the input context *The judge instructed the jury* with the same set of exemplars. In this case, the first exemplar is activated again but this time due to *judge*. Unlike previous cases, the verb here plays a more crucial role in the proper interpretation of the input.

The overall likeness, i.e., adaptability between an input context and a view-set, is measured by the following two formulas, one for each of the above cases:

$$[(1 + (1 - \lambda) / (N - 1)) * C_n + \lambda * C_v] / N \quad (\text{view-set activated by a verb})$$

$$[(1 + (1 - \lambda) * (N - 1)) * C_v + \lambda * C_n] / N \quad (\text{view-set is not activated by a verb})$$

where λ is a constant, N is the number of input words, C_n is the total compatibility for nouns, and C_v is the compatibility of verbs. These two formulas ensure that in measuring the likeness between an input context and a view-set, individual compatibility values between sense-views and input words are properly weighted according to one of the above mentioned cases. Finally, here is the activation log generated for this step:

```
***** Entering ViewSet Breaking
Compatibility of Sense-View (Breaking, Object) with WINDOW is: 1
Compatibility of Sense-View (Breaking, Action) with BREAK-SEPARATE INTO PIECES is: 1
***** Entering ViewSet ReadMaterial
Compatibility of Sense-View (ReadMaterial, Theme) with WINDOW is: 0
Compatibility of Sense-View (Breaking, Instrument) with BOOK is: 0.427695
Compatibility of Sense-View (ReadMaterial, Action) with BREAK-VIOLATE is: 0.160216
Compatibility of Sense-View (Breaking, Agent) with BOOK is: 0
####Aliveness:0.742463 (with Breaking)
***** Entering ViewSet LawViolation
```



```
Compatibility of Sense-View (LawViolation, Theme) with WINDOW is: 0
Compatibility of Sense-View (ReadMaterial, Agent) with BOOK is: 0
####Alikeness:0.128173 (with ReadMaterial)
Compatibility of Sense-View (LawViolation, Action) with BREAK-VIOLATE is: 1
Compatibility of Sense-View (LawViolation, Agent) with BOOK is: 0
####Alikeness:0.1 (with LawViolation)
```

Step 6: Once an adaptable view-set has been found for the input context (based on a minimum likeness factor¹⁸), the view-set agent proceeds to interpret the input context according to its associated view-set. In other words, based on their corresponding sense-views, output senses (i.e., pairs of sense-concept and sense-view) will be generated for each input word. More specifically, finding an adaptable view-set not only entails finding appropriate sense-views for the input words, but it also entails the selection of proper sense-concepts for the input words. Recall **Step 5**, during which, in order to measure the compatibility between a sense-view and an input word, all concepts associated with the word were compared with the sense-view, and the highest compatibility value was taken. The same idea is used to disambiguate an input word, that is, to select the most compatible concept, among all associated with the word, as the sense-concept for the given word.

For instance, in our example, finding that the input context is adaptable to the view-set BREAKING entails that among the two concepts associated with the input word *broke* (in fact, its word form *break*), the one most compatible with the sense-view ACTION must be selected (see below for the output).

In some cases, ambiguity must be preserved, that is, the context does not favor any of the potential readings. For instance, consider this sentence:

¹⁸ Assumed 0.65 in this implementation (see Section 5.1 for more details).

John went to the bank.

In this case, both readings of the word *bank*, namely the river-bank and financial institute, are highly compatible with the sense-view Destination (see Section 5.8 for details).

In theory, more than one concept can be compatible with a sense-view in the proposed model. To account for such cases, the algorithm not only looks for the concept with the highest compatibility, but also considers other concepts with high compatibility. This is controlled by another parameter, CompatibilityThreshold, which sets a minimum for a concept to be highly compatible with a sense-view. Currently, a sense-view can have up to two compatible concepts.

Finding an adaptable view-set also allows the system to establish some properties for unknown words (see Section 5.12 for an example).

Step 7: The next step is to connect the sense-concepts selected for the input words to their corresponding sense-views. To do so, properties of the sense-concepts and their corresponding sense-views must be merged. If a property is already present in a sense-view, its weight will be added to the weight of the existing one. Otherwise, the property must be inserted into the sense-view's property set. Including a new member will also increase the value of the view's number of exemplars.

The centrality of a property in a sense-view is defined as the property's weight divided by the number of exemplars. Here is the sense-view Instrument in the view-set BREAKING before including the sense-concept *book*.

```
ViewSet:Breaking
Thematic Role: Instrument - No. of Exemplars: 3 - Marker(s): with-pp,p-subj,
OPACITY(0.666667)---->:OPAQUE,
IsKindOf-EARTH-MATERIAL(0.6)---->:
IsKindOf-SOLID(0.54)---->:
```

```
STATE-OF-MATTER(0.7587) ---->:SOLID,  
MADE-OF(0.770733) ---->:SILVER, LEAD, STEEL, PLASTIC, METAL, MATERIAL,
```

Notice how this sense-view evolves once the sense-concept book is added:

```
[Sense View]  
Thematic Role: Instrument - No. of Exemplars: 4 - Marker(s): with-pp,p-subj,  
STATE-OF-MATTER(0.73305) ---->:SOLID,  
MADE-OF(0.82805) ---->:MATERIAL,
```

This example explains how sense-view properties are gradually and incrementally developed. The more a property appears in the sense-view's related sense-concepts the stronger it becomes. Similarly, a property with strong original weight will fade away if not supported by newly added sense-concepts. (Also see sample runs reported in the next chapter, particularly 5.3 and 5.4.)

Step 8: To complete my discussion of the implemented system, let us look at the way the output is added to the collection of stored exemplars. Since the generated output itself constitutes an exemplar, storing it along with other exemplars is potentially useful.

Redundancy of stored exemplars is the major concern: how could the system avoid storing redundant exemplars? The system applies two criteria to decide if a generated output, or any exemplar in general, is worth storing. The output generated for an input context will be stored in the repository only if 1) the output helps stabilize any of the view-set's sense-views, and/or 2) the output causes modifications in any of the view-set's sense-views.

The stability of a view-set, specified by StabilityThreshold parameter, simply controls the maximum number of sense-concepts that can be included in a sense-view. Once this limit is reached, no new sense-concepts can be included in the view and

therefore view's characteristics will not change. In reality, this could be a limitation of the implementation, since "better" exemplars could come later.

As for the second criterion, it simply states that only those output that cause any modifications in the sense-views must be stored and those that do not bear any new information for the system, (i.e., do not exemplify any changes in sense-views) are redundant and should not be retained.

To conclude, here is the complete output generated for the example discussed throughout this section. The first line below indicates the exemplar and view-set the output is patterned after and the likeness of the input to them. This value can serve as a certainty measure for the generated output.

The output displays generated word senses for three input words, *book*, *break* and *window*, as present in *The book broke the window*. Each word sense consists of a sense-concept (e.g., BOOK for *book*), and a sense-view, represented by a) a thematic role (e.g., Instrument), b) a number of exemplars at this point in time, c) syntactic markers, and d) a number of properties, each appearing on a separate line. Specified for each property is the name of the property, its centrality (weight), and some possible values.

```
-----Output #1--Output(S9, Breaking)(0.742463)
SENSE Generated for Input Word Book
[Sense-Concept(s)] BOOK
[Sense View]
Thematic Role: Instrument - No. of Exemplars: 4 - Marker(s): with-pp,p-subj,
STATE-OF-MATTER(0.73305)---->:SOLID,
MADE-OF(0.82805)---->: MATERIAL,

SENSE Generated for Input Word Break
[Sense-Concept(s)] BREAK-SEPARATE INTO PIECES
[Sense View]
Thematic Role: Action - No. of Exemplars: 5 - Marker(s): verb,
SYNTHESIS-DECOMPOSITION-ATTRIBUTE(1)---->:DECOMPOSITION,
```

```

PRECONDITION(1)---->:HIT,
IsKindOf-CHANGE-IN-QUALITY(0.9)---->:
TYPE-OF-CHANGE(0.9)---->:PHYSICAL,
IsKindOf-CHANGE-STATE(0.81)---->:
IsKindOf-PHYSICAL-EVENT(0.729)---->:
IsKindOf-EVENT(0.6561)---->:

SENSE Generated for Input Word Window
[Sense-Concept(s)] WINDOW
[Sense View]
Thematic Role: Object - No. of Exemplars: 5 - Marker(s): p-obj,
IsKindOf-ARTIFACT(0.5832)---->:
STATE-OF-MATTER(0.5994)---->:SOLID,
COLOR(0.5832)---->:
AGE(0.5832)---->:
OPERATED-BY(0.5832)---->:HUMAN,
IsKindOf-INANIMATE(0.52488)---->:
FRAGILITY(0.6)---->:FRAGILE,
MADE-OF(0.742)---->:METAL,CERAMIC,PLASTIC,GLASS,

```

4.6 Developing View-Sets and Sense-Views

The system develops view-sets in a semi-automatic fashion. A view-set is created by an exemplar. Presenting the system with an unprecedented exemplar, the trainer must specify the proper concept and thematic role for each input word (except for prepositions and articles). For example, given a novel exemplar such as:

The rock broke the glass.

the trainer should first specify to which concepts the words *rock*, *broke*, and *glass* are referring to, and then what role each plays in this context. To specify the concepts, the trainer has to connect exemplar nodes corresponding to each word to their appropriate concept nodes. Notice this assumes the connections between these concepts and the input words' corresponding lexeme nodes already exist in the lexicon network. If this is not the case, the trainer should first specify appropriate concept(s) associated with an unknown word.

The next step is to assign a view-set to the new exemplar. Being an unprecedented exemplar, a view-set node is automatically created. Since this view-set is exemplified only with one exemplar, its sense-views will have the same characteristics as the sense-concepts of that exemplar. In other words, sense-views in the view-set will have the same properties as their corresponding sense-concepts in the exemplar. To complete the specification of a new sense-view, the trainer should also specify its thematic role¹⁹.

Once created, a sense-view can evolve as new sense-concepts are connected to it. I have already described how adding new sense-concepts to a sense-view can cause modifications in the sense-views properties. Adding a sense-concept can also introduce a new case marker to a sense-view. Consider the view-set *Breaking* with only two exemplars as follows:

The rock broke the glass.

The bullet smashed the window.

In light of these exemplars, the sense-view Instrument would recognize only one syntactic marker, namely p-subj, for this role. Adding an exemplar such as the following can add another marker to this view-set (see Figure 4-9 for the structure of sense-views):

The man broke the computer with a rock.

Here is the content of this sense-view after connecting *rock* in the above exemplar to it.

Notice the value of the case marker field, which now contains with-pp as well.

```
[Sense View]
Thematic Role: Instrument - No. of Exemplars: 3 - Marker(s): with-pp,p-subj,
OPACITY(0.666667)---->:OPAQUE,
IsKindOf-EARTH-MATERIAL(0.6)---->:
IsKindOf-SOLID(0.54)---->:
```

¹⁹ Of course the trainer should not have to handle all the details. A friendly user interface can hide unnecessary details. Such an interface is not yet implemented.

STATE-OF-MATTER(0.7587)---->:SOLID,
MADE-OF(0.770733)---->:SILVER,LEAD,STEEL,PLASTIC,METAL,MATERIAL,

And finally, it must be noted that the proposed context recognition approach is not deterministic, that is, it is always possible that the system fails to find an existing adaptable, if any, exemplar to use in the interpretation of an input context. As discussed in section 3.5, two cases must be considered: 1) either an adaptable exemplar (or view-set) exists but the system has failed to locate it, or 2) the input is unprecedented. In either situation, the intervention of the trainer is needed.

If the input context does correspond to an existing exemplar, all the trainer is expected to do is to pick this exemplar, or in fact its associated view-set, as the pattern for processing the input. Conversely, if the input context constitutes an unprecedented context, then the model automatically generates a new view-set and proceeds in the way described earlier (see Section 5.12 for an example).

that demonstrate the ability of the system to account for sense selection and various types of sense-modulations. These tests are presented in sections 5.5 and 5.6 (sense modulation: highlighting and backgrounding), 5.7 (sense selection), 5.8 (selection of multiple word senses), 5.9 and 5.10 (sense modulation: promotion and demotion). Section 5.11 reports a situation where the model fails to interpret an input context, for which the user intervention is required. The last group of tests concerns the treatment of unknown words (5.12), idiosyncratic, and diachronic interpretation (5.13 and 5.14), and polysemic mappings (5.15).

5.2 Method

The purpose of the conducted tests was two-folded. First, they were concerned with the validity and applicability of the proposed model. The goal was to investigate whether contextual effects of the sort discussed in this dissertation can be accounted for by the proposed two-tier word senses, and whether sense-views can evolve out of a number of exemplars. In other words, prototyping and experimenting with the system constituted a feasibility study for the proposed computational model of the lexicon.

Second, the tests were also aimed at verifying and adjusting the implemented case-based context recognition approach. More specifically, throughout my discussion of the ontology and of the lexicon networks, a number of parameters and threshold values were mentioned, most of them being directly related to the adopted context recognition approach. As in most localist networks, these parameters and thresholds form the “magic numbers” (Hendler, 1989) of such systems. But the values of these magic numbers are not totally arbitrary. Instead, they are ‘tweaked’ (i.e., fine-tuned) through numerous runs of the tests. That is, as with most connectionist models, the exact settings largely proceed

from trial-and-error over a large number of executions. Figure 5-1 lists these parameters and their final values, as used in the tests reported in this chapter.

<u>NonLocalityFactor</u>	0.9
<u>LocalityThreshold</u>	0.6
<u>ActivationLossFactor</u>	0.6
<u>ActivationThreshold</u>	0.3
<u>ViewWeightThreshold</u>	0.5
<u>MinimumAlikeness</u>	0.65
λ (used in similarity measurement formula)	0.3
<u>CompatibilityThreshold</u>	0.48
<u>StabilityThreshold</u>	10

Figure 5-1: System parameters and their given values.

For the tests discussed in sections 5.3 through 5.12, I used a lexicon with approximately 50 words (including 7 verbs), 50 concepts, and 12 exemplars. Exemplars were grouped into 9 view-sets; hence the majority of view-sets were exemplified with only one exemplar. The set of exemplars and their related view-sets follow:

Ex1. The rock broke the glass.	(Breaking)
Ex2. The bullet smashed the window.	(Breaking)
Ex3. John broke the law.	(Violating-Law)
Ex4. The man broke the computer with a rock.	(Breaking)
Ex5. John moved the stove.	(Moving-Object)
Ex6. The musician played the guitar.	(Playing-Music)
Ex7. The jury announced the verdict.	(Announcing)
Ex8. Mary read the article.	(Read-Material)
Ex9. The student studies chemistry.	(Studying)
Ex10. John read the newspaper.	(Read-Material)

Ex11. Mary went to the office.

(Going-to-Place)

Ex12. A shark attacked the man.

(Shark-Attack)

Here is the content of the view-set BREAKING just before the system begins processing the first input. This view-set consists of four sense-views, namely Agent, Object, Action and Instrument. Recall that sense-views are characterized by their thematic roles, number of connected sense-concepts, syntactic markers and associated properties. Three exemplars, Ex1, Ex2 and Ex4 correspond to the BREAKING view-set at this point. Notice that the role Agent occurs only in one of them (Ex4), and thus the number of exemplars for this sense-view is 1. As discussed earlier, the set of properties associated with a sense-view is represented as a frame in which slots represent individual properties. In this output, as well as in all the other results reported in this chapter, properties are specified by their names, weights and possible values²⁰.

ViewSet:Breaking

[Sense View]

Thematic Role: Agent - No. of Exemplars: 1 - Marker(s): p-subj,
GENDER(1)---->:MALE,
IsKindOf-HUMAN(0.9)---->:
IsKindOf-PRIMATE(0.81)---->:
IsKindOf-MAMMAL(0.729)---->:
IsKindOf-VERTEBRATE(0.6561)---->:

[Sense View]

Thematic Role: Object - No. of Exemplars: 3 - Marker(s): p-obj,
STATE-OF-MATTER(0.513)---->:SOLID,
FRAGILITY(0.666667)---->:FRAGILE,
MADE-OF(0.903333)---->:GLASS,

²⁰ Only properties with a weight greater than ViewWeightThreshold are displayed.

```

[Sense View]
Thematic Role: Action - No. of Exemplars: 3 - Marker(s): verb,
SYNTHESIS-DECOMPOSITION-ATTRIBUTE(1)---->:DECOMPOSITION,
PRECONDITION(1)---->:HIT,
IsKindOf-CHANGE-IN-QUALITY(0.9)---->:
TYPE-OF-CHANGE(0.9)---->:PHYSICAL,
IsKindOf-CHANGE-STATE(0.81)---->:
IsKindOf-PHYSICAL-EVENT(0.729)---->:
IsKindOf-EVENT(0.6561)---->:

[Sense View]
Thematic Role: Instrument - No. of Exemplars: 3 - Marker(s): with-pp,p-subj,
OPACITY(0.666667)---->:OPAQUE,
IsKindOf-EARTH-MATERIAL(0.6)---->:
IsKindOf-SOLID(0.54)---->:
STATE-OF-MATTER(0.7587)---->:SOLID,
MADE-OF(0.770733)---->:MATERIAL,

```

5.3 Sense-View Development (Object-of-Moving)

The goal of this test is to demonstrate the effects of including new sense-concepts into a sense-view (object-of-moving in this case), and therefore, show how sense-views evolve as the result of this process. To start with, the sense-view has only one sense-concept (*stove* in Ex5) and looks like this:

```

[Sense View]
Thematic Role: Object - No. of Exemplars: 1 - Marker(s): p-obj,
WEIGHT(1)---->:heavy,
IsKindOf-COMPLEX-MECHANISM(0.9)---->:
IsKindOf-MECHANISM(0.81)---->:
IsKindOf-DEVICE(0.729)---->:
STATE-OF-MATTER(0.729)---->:SOLID,
MADE-OF(0.729)---->:PLASTIC,METAL,
IsKindOf-ARTIFACT(0.6561)---->:
COLOR(0.6561)---->:
AGE(0.6561)---->:
OPERATED-BY(0.6561)---->:HUMAN,
IsKindOf-COOKING-APPLIANCE(0.9)---->:
IsKindOf-COOKING-EQUIPMENT(0.81)---->:
IsKindOf-EVERYDAY-ARTIFACT(0.729)---->:

```

After processing the following input contexts and consequent inclusions of *refrigerator* and *computer*, the sense-view changes to reflect properties common to these sense-concepts. The resultant sense-view is shown below, as a part of the sense generated for the input word *computer*.

S1. Mary moved the refrigerator.

S2. John moved the computer.

```
SENSE Generated for Input Word Computer
[Sense-Concept(s)] COMPUTER
[Sense View]
Thematic Role: Object - No. of Exemplars: 2 - Marker(s): p-obj,
IsKindOf-DEVICE(0.7695) ---->:
STATE-OF-MATTER(0.7695) ---->: SOLID,
MADE-OF(0.7695) ---->: PLASTIC, METAL,
IsKindOf-ARTIFACT(0.69255) ---->:
COLOR(0.69255) ---->:
AGE(0.69255) ---->:
OPERATED-BY(0.69255) ---->: HUMAN,
```

Notice the number of exemplars is 2 and not, as possibly expected, 3. This is due to the fact that entities in Mikrokosmos are essentially specified taxonomically, and since *stove* and *refrigerator* in this respect are similar (both are instances of COOKING-APPLIANCE), there is no need for the inclusion of *refrigerator* in this sense-view, for it bears no new information (see section 4.4). Clearly, this would not have been the case if *stove* and *refrigerator* had their own distinctive properties.

5.4 Sense-View Development (Destination-of-Going-to)

This test provides another example of sense-view development. The destination sense-view, exemplified only by *office* in Ex11, initially includes following properties:

```
[Sense View]
Thematic Role: Destination - No. of Exemplars: 1 - Marker(s): to-pp,
```

```

IsKindOf-BUILDING(0.9)---->:
IsKindOf-PLACE(0.81)---->:
IsKindOf-PHYSICAL-OBJECT(0.729)---->:
MADE-OF(0.729)---->:MATERIAL,
WEIGHT(0.729)---->:
SIZE(0.729)---->:
IsKindOf-OBJECT(0.6561)---->:
IsKindOf-BUILDING-ARTIFACT(0.81)---->:
IsKindOf-ARTIFACT(0.729)---->:
STATE-OF-MATTER(0.729)---->:SOLID,
COLOR(0.729)---->:
AGE(0.729)---->:
OPERATED-BY(0.729)---->:HUMAN,
IsKindOf-INANIMATE(0.6561)---->:

```

Notice that both IsKindOf-Building and IsKindOf-Place are relatively central to *office* and therefore to this sense-view. The above set of properties shrinks after processing S3:

S3. The student went to the stadium.

```

SENSE Generated for Input Word Stadium
[Sense-Concept(s)] STADIUM
[Sense View]
Thematic Role: Destination - No. of Exemplars: 2 - Marker(s): to-pp,
IsKindOf-BUILDING(0.8145)---->:
IsKindOf-PLACE(0.73305)---->:
IsKindOf-BUILDING-ARTIFACT(0.73305)---->:

```

This trend continues with S4 that produces the following sense for *park*.

S4. John went to the park.

```

SENSE Generated for Input Word Park
[Sense-Concept(s)] PARK
[Sense View]
Thematic Role: Destination - No. of Exemplars: 3 - Marker(s): to-pp,
IsKindOf-BUILDING(0.543)---->:
IsKindOf-PLACE(0.7074)---->:

```

Here, unlike the previous case, IsKindOf-Place is more prominent than IsKindOf-Building. This is because *park* is not a building but, nonetheless, its effect is not enough to completely eliminate IsKindOf-Building from the sense-view. The next input,

however, strengthens *IsKindOf-Building* and weakens *IsKindOf-Place*, mainly because this property is not immediately present for *auditorium* (Place is 4 levels away from *auditorium* in the ontology).

S5. The musician went to the auditorium.

```
SENSE Generated for Input Word Auditorium
[Sense-Concept(s)] AUDITORIUM
[Sense View]
Thematic Role: Destination - No. of Exemplars: 4 - Marker(s): to-pp,
IsKindOf-BUILDING(0.571275)---->:
IsKindOf-PLACE(0.53055)---->:
```

5.5 Sense Modulation (piano)

This test demonstrates the system's capability to modulate semantic features of sense-concepts. More specifically these examples show how, in each context, certain properties of *piano* become highlighted. This is achieved by assigning different sense-views to the same concept, PIANO. Here is the first input context:

S6. The musician moved the piano.

The output sense generated for *piano* is displayed below. Because of the assigned sense-view, only those properties of PIANO that portray it as a heavy physical object, will be highlighted. Also note how the sense-view changes after inclusion of *piano* (compare it with the output generated for S2 in section 5.2).

```
SENSE Generated for Input Word Piano
[Sense-Concept(s)] PIANO
[Sense View]
Thematic Role: Object - No. of Exemplars: 3 - Marker(s): p-obj,
WEIGHT(0.666667)---->:heavy,
IsKindOf-DEVICE(0.513)---->:
STATE-OF-MATTER(0.756)---->:SOLID,
MADE-OF(0.513)---->:PLASTIC,METAL,
IsKindOf-ARTIFACT(0.7047)---->:
```

```
COLOR(0.7047)---->:
AGE(0.7047)---->:
OPERATED-BY(0.7047)---->:HUMAN,
```

It is also interesting to look at the selected sense-view for *musician* in which all properties specific to *musician* are suppressed, since they are irrelevant in the present context.

```
SENSE Generated for Input Word Musician
[Sense-Concept(s)] MUSICIAN
[Sense View]
Thematic Role: Agent - No. of Exemplars: 3 - Marker(s): p-subj,
GENDER(0.885367)---->:MALE,
IsKindOf-HUMAN(0.8187)---->:
IsKindOf-PRIMATE(0.54)---->:
```

Here is the second input context:

S7. The musician played the piano.

In contrast, here the properties portraying *piano* as a musical instrument and *musician* as a social role are relevant.

```
SENSE Generated for Input Word Musician
[Sense-Concept(s)] MUSICIAN
[Sense View]
Thematic Role: Agent - No. of Exemplars: 2 - Marker(s): p-subj,
WORK-EQUIPMENT(1)---->:MUSICAL-INSTRUMENT,
AREA-OF-INTEREST(1)---->:FIELD-OF-MUSIC,
IsKindOf-ENTERTAINMENT-ROLE(0.9)---->:
IsKindOf-ARTISTIC-ROLE(0.81)---->:
IsKindOf-SOCIAL-ROLE(0.729)---->:
IsKindOf-HUMAN(0.6561)---->:
GENDER(0.6561)---->:
IsKindOf-SOCIAL-OBJECT(0.6561)---->:
```

```
SENSE Generated for Input Word Piano
[Sense-Concept(s)] PIANO
[Sense View]
Thematic Role: Object - No. of Exemplars: 2 - Marker(s): p-obj,
IsKindOf-MUSICAL-INSTRUMENT(0.81)---->:
WORK-EQUIPMENT-OF(0.81)---->:MUSICIAN,
IsKindOf-ARTIFACT(0.729)---->:
```

```

STATE-OF-MATTER(0.729)---->:SOLID,
COLOR(0.729)---->:
AGE(0.729)---->:
OPERATED-BY(0.729)---->:HUMAN,
IsKindOf-INANIMATE(0.6561)---->:

```

Here is the last input context:

S8. The man broke the piano.

In this context *piano* is portrayed as an object of breaking. Properties in this sense-view are those inherited from ARTIFACT. This input also illustrates the system's disambiguation ability (*break* as Separate-into-Pieces versus Law-Violation). Word senses generated for *break* and *piano* are displayed below.

```

SENSE Generated for Input Word Break
[Sense-Concept(s)] BREAK-SEPARATE INTO PIECES
[Sense View]
Thematic Role: Action - No. of Exemplars: 4 - Marker(s): verb,
SYNTHESIS-DECOMPOSITION-ATTRIBUTE(1)---->:DECOMPOSITION,
PRECONDITION(1)---->:HIT,
IsKindOf-CHANGE-IN-QUALITY(0.9)---->:
TYPE-OF-CHANGE(0.9)---->:PHYSICAL,
IsKindOf-CHANGE-STATE(0.81)---->:
IsKindOf-PHYSICAL-EVENT(0.729)---->:
IsKindOf-EVENT(0.6561)---->:

SENSE Generated for Input Word Piano
[Sense-Concept(s)] PIANO
[Sense View]
Thematic Role: Object - No. of Exemplars: 4 - Marker(s): p-obj,
IsKindOf-ARTIFACT(0.54675)---->:
STATE-OF-MATTER(0.567)---->:SOLID,
COLOR(0.54675)---->:
AGE(0.54675)---->:
OPERATED-BY(0.54675)---->:HUMAN,
MADE-OF(0.6775)---->:PLASTIC,GLASS,

```


5.6 Sense Modulation (book)

This test provides further evidence on how a single concept in this model can be viewed from different perspectives. Notice how the generated output for *book* changes in each of the following cases.

S9. The book broke the window.

```
SENSE Generated for Input Word Book
[Sense-Concept(s)] BOOK
[Sense View]
Thematic Role: Instrument - No. of Exemplars: 4 - Marker(s): with-pp,p-subj,
STATE-OF-MATTER(0.73305)---->:SOLID,
MADE-OF(0.82805)---->: MATERIAL,
```

S10. The student read the book.

This is also another case of lexical disambiguation, or sense selection: *read* could mean read-in-mind, announce or study for an academic degree (see next section).

```
SENSE Generated for Input Word Read
[Sense-Concept(s)] READ-IN-MIND
[Sense View]
Thematic Role: Action - No. of Exemplars: 3 - Marker(s): verb,
MODE(1)---->:IN-MIND,
IsKindOf-ACTIVE-COGNITIVE-EVENT(0.9)---->:
IsKindOf-COGNITIVE-EVENT(0.81)---->:
IsKindOf-MENTAL-EVENT(0.729)---->:
IsKindOf-EVENT(0.6561)---->:
```

```
SENSE Generated for Input Word Book
[Sense-Concept(s)] BOOK
[Sense View]
Thematic Role: Theme - No. of Exemplars: 3 - Marker(s): p-obj,
CONTAINS(0.666667)---->:INFORMATION,
MADE-OF(0.666667)---->:PAPER, INK,
IsKindOf-PRINTED-MEDIA(0.6)---->:
LOCATION(0.6)---->:ACADEMIC-BUILDING,
IsKindOf-VISUAL-MEDIA-ARTIFACT(0.54)---->:
IsKindOf-DOCUMENT(0.57)---->:
PRODUCED-BY(0.57)---->:HUMAN,
IsKindOf-LANGUAGE-RELATED-OBJECT(0.513)---->:
REPRESENTS(0.756)---->:OBJECT, EVENT, LANGUAGE,
```

And finally *book* as an object of moving.

S11. The musician moved the book.

```
SENSE Generated for Input Word Book
[Sense-Concept(s)] BOOK
[Sense View]
Thematic Role: Object - No. of Exemplars: 4 - Marker(s): p-obj,
STATE-OF-MATTER(0.731025)---->:SOLID,
MADE-OF(0.63475)---->:PLASTIC,METAL,
IsKindOf-ARTIFACT(0.69255)---->:
COLOR(0.69255)---->:
AGE(0.69255)---->:
OPERATED-BY(0.69255)---->:HUMAN,
```

5.7 Word Sense Selection

The goal in this test is to demonstrate the model's ability to deal with lexical ambiguity, where a word can have a number of contrastive readings. Consider the following input:

S12. The judge read the verdict.

The word *read* is ambiguous in my lexicon: it can be Read-in-Mind, Study toward an academic degree, and Announce. This input shows how, based on the context, one of these reading is favored by the system. Here are parts of the activation report generated by the lexicon after inputting S12. Notice view-set **LawViolation** has been activated by *verdict* because in Mikrokosmos both *law* and *verdict* are instances of Legal-Object. Therefore, the concept *verdict* had triggered *law* as a close concept, which in turn activated Ex7.

```
***** Entering ViewSet ReadMaterial
Compatibility of Sense-View (ReadMaterial, Theme) with VERDICT is: 0.10397
Compatibility of Sense-View (ReadMaterial, Action) with READ-IN-MIND is: 1
***** Entering ViewSet Study
Compatibility of Sense-View (Study, Theme) with VERDICT is: 0.447514
```

```

Compatibility of Sense-View (ReadMaterial, Agent) with JUDGE is: 0.522504
Compatibility of Sense-View (Study, Action) with STUDY is: 1
Compatibility of Sense-View (Study, Agent) with JUDGE is: 0.517795
***** Entering ViewSet LawViolation
Compatibility of Sense-View (LawViolation, Theme) with VERDICT is: 1
Compatibility of Sense-View (LawViolation, Action) with READ-IN-MIND is: 0.164226
***** Entering ViewSet Announce
Compatibility of Sense-View (Announce, Theme) with VERDICT is: 1
####Alike ness:0.381914 (with ReadMaterial)
####Alike ness:0.534389 (with Study)
Compatibility of Sense-View (LawViolation, Agent) with JUDGE is: 0.320432
####Alike ness:0.263424 (with LawViolation)
Compatibility of Sense-View (Announce, Action) with ANNOUNCE is: 1
Compatibility of Sense-View (Announce, Agent) with JUDGE is: 1
####Alike ness:1 (with Announce)

```

As shown, the view-set **Announce** has obtained the greatest likeness with the input context, and therefore, the following output sense for *read* (5.10 gives a complete trace of the system activities for this input):

```

SENSE Generated for Input Word Read
[Sense-Concept(s)] ANNOUNCE
[Sense View]
Thematic Role: Action - No. of Exemplars: 1 - Marker(s): verb,
IsKindOf-ASSERTIVE-ACT(0.9)---->:
IsKindOf-ILLOCUTIONARY-ACT(0.81)---->:
IsKindOf-SPEECH-ACT(0.729)---->:
IsKindOf-COMMUNICATIVE-EVENT(0.6561)---->:

```

5.8 Possibility of Multiple Word Senses

Unlike the previous case, there are cases in which context does not favor any of the alternative readings of a word, and therefore the ambiguity must be maintained in the output. This test demonstrates the system's ability to handle such cases. In S13, both readings of *bank* are compatible, to some degree, with the destination sense-view in GOING-TO view-set.

S13. John went to the bank.

Here is the output word sense for *bank*.

```
SENSE Generated for Input Word Bank
[Sense-Concept(s)] RIVER-BANK, BANK-BRANCH
[Sense View]
Thematic Role: Destination - No. of Exemplars: 6 - Marker(s): to-pp,
IsKindOf-BUILDING(0.51585)---->:
IsKindOf-PLACE(0.5967)---->:
```

Also notice, how the property *IsKindOf-Place*, after the inclusion of these two new members in the destination sense-view, again becomes more prominent than *IsKindOf-Building* (see 5.3).

5.9 Property Promotion/Demotion (Shark)

This test is inspired by Anderson et al. (1976) who found that *shark* was a better cue than *fish* for subjects in remembering a sentence like S14. They concluded that *fish* was instantiated to, and encoded accordingly as, *shark* in the subjects' memory.

In this test, the StabilityThreshold for SHARK-ATTACK view-set was set to 1, meaning just one context (Ex12 in this case) was enough to stabilize this view-set.

S14. The fish attacked the man.

Notice how in the output sense *fish* is associated with properties specific to *shark* (Aggressive and Color).

```
SENSE Generated for Input Word Fish
[Sense-Concept(s)] FISH
[Sense View]
Thematic Role: Agent - No. of Exemplars* 1 - Marker(s : p-subj,
COLOR(1)---->:BLACK,
AGGRESSIVE(1)---->:
IsKindOf-FISH(0.9)---->:
IsKindOf-VERTEBRATE(0.81)---->:
IsKindOf-ANIMAL(0.729)---->:
GENDER(0.729)---->:
IsKindOf-ANIMATE(0.6561)---->:
```

5.10 Property Promotion/Demotion (Student)

This test provides an example of how one concept appearing in a context can be associated with properties not necessary present in its original representation. Here is the input context:

S15. Mary reads physics.

Because of its appearance in this context, *Mary* (in fact, its corresponding concept, Female-Human) will be depicted as *student*. In other words, through the assigned sense-view, properties specific to student (e.g., being a social/academic role) will be associated with *Mary* in this context. The test also provides another example of the system's lexical disambiguation ability (*read* is ambiguous). The StabilityThreshold for this view set was also set to 1. Here is the output word sense for *Mary*.

```
SENSE Generated for Input Word Mary
[Sense-Concept(s)] HUMAN-FEMALE
[Sense View]
Thematic Role: Agent - No. of Exemplars: 1 - Marker(s): p-subj,
IsKindOf-ACADEMIC-ROLE(0.9)---->:
IsKindOf-SOCIAL-ROLE(0.81)---->:
IsKindOf-HUMAN(0.729)---->:
GENDER(0.729)---->:
IsKindOf-PRIMATE(0.6561)---->:
IsKindOf-SOCIAL-OBJECT(0.729)---->:
IsKindOf-OBJECT(0.6561)---->:
```

5.11 Failure to Interpret Input

This is an example of a situation where the system fails to find any interpretation for the input. As discussed earlier, such a situation arises when no adaptable exemplar for the input is found. The input context is as follows:

S16. The boy read the story.

Here are *alikeness* measurement values between the input context and the activated view-sets. As displayed, the view-set *ReadMaterial* has the highest degree of adaptability to the input. However, it is not close enough.

```
***** Entering ViewSet MovingObject
Compatibility of Sense-View (MovingObject, Object) with STORY is: 0
Compatibility of Sense-View (MovingObject, Action) with READ-IN-MIND is: 0.076961
***** Entering ViewSet LawViolation
Compatibility of Sense-View (LawViolation, Theme) with STORY is: 0.164226
***** Entering ViewSet Study
Compatibility of Sense-View (Study, Theme) with STORY is: 0.21198
***** Entering ViewSet ReadMaterial
Compatibility of Sense-View (ReadMaterial, Theme) with STORY is: 0.231471
Compatibility of Sense-View (MovingObject, Agent) with BOY is: 1
####Alikeness:0.161569 (with MovingObject)
***** Entering ViewSet Announce
Compatibility of Sense-View (Announce, Theme) with STORY is: 0.101972
Compatibility of Sense-View (LawViolation, Action) with READ-IN-MIND is: 0.164226
***** Entering ViewSet Breaking
***** Entering ViewSet GoingtoPlace
Compatibility of Sense-View (Study, Action) with STUDY is: 1
***** Entering ViewSet SharkAttack
Compatibility of Sense-View (SharkAttack, Object) with STORY is: 0
Compatibility of Sense-View (ReadMaterial, Action) with READ-IN-MIND is: 1
Compatibility of Sense-View (ReadMaterial, Agent) with BOY is: 0.971184
Compatibility of Sense-View (Announce, Action) with ANNOUNCE is: 1
Compatibility of Sense-View (Breaking, Object) with STORY is: 0
Compatibility of Sense-View (GoingtoPlace, Action) with READ-IN-MIND is: 0.076961
Compatibility of Sense-View (GoingtoPlace, Agent) with BOY is: 0.996963
Compatibility of Sense-View (Study, Agent) with BOY is: 0.40584
####Alikeness:0.378019 (with Study)
Compatibility of Sense-View (LawViolation, Agent) with BOY is: 0.755806
####Alikeness:0.223384 (with LawViolation)
Compatibility of Sense-View (SharkAttack, Action) with READ-IN-MIND is: 0.21198
Compatibility of Sense-View (SharkAttack, Agent) with BOY is: 0.12517
####Alikeness:0.641195 (with ReadMaterial)
Compatibility of Sense-View (Announce, Agent) with BOY is: 0.247244
Compatibility of Sense-View (Breaking, Action) with READ-IN-MIND is: 0.109439
####Alikeness:0.161265 (with GoingtoPlace)
####Alikeness:0.182101 (with SharkAttack)
####Alikeness:0.257148 (with Announce)
Compatibility of Sense-View (Breaking, Instrument) with BOY is: 0
Compatibility of Sense-View (Breaking, Agent) with BOY is: 0.755806
####Alikeness:0.163132 (with Breaking)
```

A closer look at the individual compatibility values for this view-set reveals that the compatibility between the Theme of reading and *story* is too low. Here is the content of this sense-view at this point:

```
[Sense View]
Thematic Role: Theme - No. of Exemplars: 3 - Marker(s): p-obj,
CONTAINS(0.666667) ---->: INFORMATION,
MADE-OF(0.666667) ---->: PARCHMENT, NEWSPRINT, PAPER, INK,
IsKindOf-PRINTED-MEDIA(0.6) ---->:
LOCATION(0.6) ---->: ACEDOMIC-BUILDING,
IsKindOf-VISUAL-MEDIA-ARTIFACT(0.54) ---->:
IsKindOf-DOCUMENT(0.57) ---->:
PRODUCED-BY(0.57) ---->: HUMAN,
IsKindOf-LANGUAGE-RELATED-OBJECT(0.513) ---->:
REPRESENTS(0.756) ---->: OBJECT, EVENT, LANGUAGE,
```

Being exemplified with only exemplars Ex8 and Ex10 and the output generated for S10, this sense-view requires the Theme of reading to be a kind of printed media and so on. This is why the compatibility between this sense-view and *story* is so low. Here is the output generated for this input:

```
Unable to Interpret the Input.
```

In such cases user intervention is required (for more details, see sections 3.5 and 4.6). For this test, all the user is required to do is to direct the system to classify the input under the view-set READ-MATERIAL. Consequently, the sense-view Theme will properly adjust as the result of including *story*. Here is its content afterwards:

```
[Sense View]
Thematic Role: Theme - No. of Exemplars: 4 - Marker(s): p-obj,
IsKindOf-LANGUAGE-RELATED-OBJECT(0.58725) ---->:
REPRESENTS(0.7695) ---->: OBJECT, EVENT, LANGUAGE,
IsKindOf-REPRESENTATIONAL-OBJECT(0.528525) ---->:
CONVEYS(0.528525) ---->: INFORMATION,
```

5.12 Unknown Words

This test demonstrates the use of sense-views in dealing with unknown words: sense-views can provide clues about an unknown word according to its embedding context. Here is an example:

S17. Mary went to the palladium.

The lexicon does not have any entry for the word *palladium*. However, the system succeeds to interpret the input as an act of GoingtoPlace, therefore, suggesting classifying *palladium* as a place-to-go. Here is the actual output:

```
SENSE Generated for Input Word Palladium
[Sense-Concept(s)] ???
[Sense View]
Thematic Role: Destination - No. of Exemplars: 6 - Marker(s): to-pp,
IsKindOf-BUILDING(0.51585)---->:
IsKindOf-PLACE(0.5967)---->:
```

5.13 Idiosyncrasy

In chapter 3, I argued that the exemplar-based nature of the model allows it to naturally account for idiosyncrasy. In this simple test I demonstrate how presenting the system with different set of exemplars can result in differences in sense-views.

For this test, two sets of exemplars were presented to the system. The first set is as follows:

Ex13. John moved the stove.

Ex14. Mary moved the piano.

Ex15. John moved the bed.

Since these objects are all heavy, this property would have a strong presence in the sense-view moving-object. Here is the content of this sense-view, with weight property highlighted.


```

ViewSet:MovingObject
[Sense View]
Thematic Role: Object - No. of Exemplars: 3 - Marker(s): p-obj,
WEIGHT(1)---->:HEAVY,
STATE-OF-MATTER(0.7047)---->:SOLID,
IsKindOf-ARTIFACT(0.6804)---->:
COLOR(0.6804)---->:
AGE(0.6804)---->:
OPERATED-BY(0.6804)---->:HUMAN,

```

However, if I present the system with a different set of exemplars in which not all objects of moving are heavy, the heaviness property would not show up in the sense-view. In addition to the above three exemplars, this time the system was presented with the following exemplars:

Ex16. John moved the book.

Ex17. Mary moved the glass.

Ex18. John moved the computer.

Here is the new content of the moving-object sense-view, which does not include the property weight:

```

ViewSet:MovingObject
[Sense View]
Thematic Role: Object - No. of Exemplars: 6 - Marker(s): p-obj,
STATE-OF-MATTER(0.5967)---->:SOLID,
MADE-OF(0.573167)---->:,
IsKindOf-ARTIFACT(0.57105)---->:
COLOR(0.57105)---->:
AGE(0.57105)---->:
OPERATED-BY(0.57105)---->:HUMAN,

```

5.14 Diachrony

This test aims to demonstrate the model's ability to account for changes that may occur in the interpretation of a single context over time. The system is initially trained with the following exemplars.

Ex19. John played the trumpet.

Ex20. The musician played the piano.

At this point in time, the system generates the following interpretation for *music-player* in an input such as:

S18. Mary played the piano.

```
SENSE Generated for Input Word Mary
[Sense-Concept(s)] HUMAN-FEMALE
[Sense View]
Thematic Role: Agent - No. of Exemplars: 2 - Marker(s): p-subj,
GENDER(0.82804996)---->:MALE(0.5),
IsKindOf-HUMAN(0.77804995)---->:
```

As displayed, this sense-view portrays *Mary* (the player) simply as a human being. However, when presented with more exemplars, this would change. Here are the new exemplars:

Ex21. The guitarist played the guitar.

Ex22. The pianist played the piano.

Now, presenting the same input will result in a different interpretation. Here is the generated word sense, in which *Mary* is presented as a musician.

```
SENSE Generated for Input Word Mary
[Sense-Concept(s)] HUMAN-FEMALE
[Sense View]
Thematic Role: Agent - No. of Exemplars: 5 - Marker(s): p-subj,
WORK-EQUIPMENT(0.58000004)---->:MUSICAL-INSTRUMENT,
AREA-OF-INTEREST(0.56)---->:FIELD-OF-MUSIC,
IsKindOf-ENTERTAINMENT-ROLE(0.504)---->:
GENDER(0.53121996)---->:
```

5.15 Polysemic/Metonymic Mappings

The simple tests reported in this section describe how the proposed model could account for polysemic, as well as, metonymic mappings. For instance, as argued by

Pustejovsky (1995), the two occurrences of *fig* in the following sentences are related through a Fruit/Plant mapping.

John ate the fig.

John watered the fig.

Similarly, *book* in the following sentence is metonymically used for its author.

The book broke the law.

Such mappings can be specified in the system. Here is an example:

```
TestLex1.AddMapping("TREE-FRUIT", "TREE", "PRODUCED-BY") ;
```

The first two arguments describe the mapping. In this case, the concept TREE-FRUIT could be mapped to TREE. The last argument pertains to the implementation of the mapping, as explained below.

In general, for a concept to be mapped to another concept, first, there must be a mapping rule that can be applied to it. For example consider the APPLE-FRUIT concept. Since it is a kind-of TREE-FRUIT, the above rule is applicable. The third argument of AddMapping specifies that concept APPLE-FRUIT can be mapped to the value of property PRODUCED-BY. The value of this property for APPLE-FRUIT is APPLE-TREE, indicating *apple* (fruit) is produced by *apple tree*. Notice that such information is available in the ontology network. Subsequently, the ontology node APPLE-TREE is located and its associated concept in the lexicon network is returned.

In order to allow the agents to consider such mappings in measuring the likeness of their associated exemplars with input contexts, Algorithm 4-4 (section 4.5.2) was enhanced to not only consider concepts associated with an input word, but also to try to

map each concept according to the mappings available in the system. The algorithm was then tested with three inputs.

Here are the mapping patterns used for this test:

```
TestLex1.AddMapping("TREE-FRUIT", "TREE", "PRODUCED-BY")
TestLex1.AddMapping("BOOK-DOCUMENT", "AUTHOR", "PRODUCED-BY")
```

The system was initially trained with the following exemplars:

Ex23. The man ate the apple.

Ex24. The man watered the apple.

Ex25. The man broke the law.

Ex26. Mary broke the glass.

Here is the first input, followed by the word-sense generated for *cherry*.

S19. Mary ate the cherry.

```
SENSE Generated for Input Word Cherry
{Sense-Concept(s)} CHERRY-FRUIT
{Sense View}
Thematic Role: Object - No. of Exemplars: 1 - Marker(s): p-obj,
PRODUCED-BY(1.0)---->:APPLE-TREE(1.0),
IsKindOf-TREE-FRUIT(0.9)---->:
IsKindOf-FRUIT-FOOD-STUFF(0.80999994)---->:
IsKindOf-PLANT-FOOD-STUFF(0.7289999)---->:
IsKindOf-FOOD-STUFF(0.6560999)---->:
```

No mapping has been applied, since the word *cherry* in the lexicon is connected to only one concept, CHERRY-FRUIT, which is compatible with the sense-view object-of-eating. However, this will not work in a sentence like S20:

S20. Mary watered the cherry.

Here is the generated word-sense:

```
SENSE Generated for Input Word Cherry
{Sense-Concept(s)} CHERRY-FRUIT
```

```

Mapped To CHERRY-TREE
[Sense View]
Thematic Role: Object - No. of Exemplars: 1 - Marker(s): p-obj,
IsKindOf-TREE(0.9)---->:
IsKindOf-PLANT(0.80999994)---->:
IsKindOf-ANIMATE(0.7289999)---->:
IsKindOf-SEPARABLE-ENTITY(0.6560999)---->:

```

As displayed, the concept CHERRY-FRUIT is mapped to CHERRY-TREE, and then used in this sense. Here is now an example of a metonymic mapping:

S21. The book broke the law.

Due to the ambiguity of *break*, two view-sets, namely LAW-VIOLATION and BREAKING (corresponding to exemplars Ex25 and Ex26), will be activated. The concept associated with the word *law* is not compatible with the sense-view breaking-object (corresponding to *glass* in Ex26), and there is no mapping available. Thus, the view-set BREAKING is rejected. Similarly, *book* is not compatible with law-violator (corresponding to *man* in Ex25). However, according to the Book-Author mapping, *book* can be mapped to *author*. Here is the sense generated for *book*:

```

SENSE Generated for Input Word Book
[Sense-Concept(s)] BOOK
Mapped To AUTHOR
[Sense View]
Thematic Role: Agent - No. of Exemplars: 2 - Marker(s): p-subj,
GENDER(0.905)---->:MALE(0.5),
IsKindOf-HUMAN(0.85499996)---->:
IsKindOf-PRIMATE(0.7694999)---->:
IsKindOf-MAMMAL(0.69254994)---->:

```

The simple processing of such mappings suggests that two-tier senses should be useful in dealing with the understanding of metaphors. This topic is investigated next.

Chapter Six

Understanding Metaphors

6.1 Introduction

This chapter discusses one application of the proposed model of lexicon in order to emphasize the importance and usefulness of decoupling patterns of contextual effects from the representation of word senses. This decoupling, resulting from the two tiers I propose for sense representation, contrasts with connectionist lexicons (which cannot manifest such effects independently despite their potential ability to account for contextual effects (see section 2.5)). The application I consider here is the understanding of metaphors of the form *Topic is Vehicle*. Others could be the processing of *etc.* and of *ad hoc* categories.

Metaphors are often explained on a comparison ground. That is, given a metaphor of the form *Topic is Vehicle*, the similarities and dissimilarities, in terms of shared

properties, of the topic and vehicle are contrasted and then it is decided: a) how 'good' a metaphor is, and b) how it should be interpreted according to those properties (see below). Section 6.2 discusses this contrast-based strategy and some critical views of it. Computational metaphor understanding models can be divided into two categories: 1) models based on semantic networks, and 2) feature-based models. These approaches are discussed in section 6.3. I will argue that such models basically face the same criticism that contrast-based theories do. My work instead subscribes to a different strategy, one in which metaphors are explained as class inclusion assertions (see Glucksberg & Keysar, 1990). Section 6.4 argues that the proposed sense-views can be used to implement Glucksberg and Keysar's theory. Details pertaining to the implementation as well as tests conducted to demonstrate the system's ability are described in 6.5. Section 6.5 also revisits the incremental development of concepts (i.e., concept evolution) and illustrates how new properties can be learned for a concept, as it becomes associated with different sense-views.

6.2 Theories of Metaphor Understanding

Various theories have been proposed to explain metaphors and the way people understand them (see Tourangeau & Sternberg, 1982 for a review). All these theories, to some extent, draw on a feature comparison element, in which metaphors are explained on the basis of similarity between topic and vehicle. For instance, a metaphor like *John is an encyclopedia* is explained based on similarities between *John* (topic) and *encyclopedia* (vehicle). According to one theory, topic and vehicle are both abstracted to reduce the dissimilarities. This abstraction process continues until a category shared by both topic

and vehicle is found. Such a category will constitute the understanding ground for that metaphor (*ibid.*).

Tversky (1977) and later Ortony (1979) elaborated on the comparison task. They both assumed that metaphors are essentially implicit comparisons. In Ortony's proposal, the similarity between topic and vehicle not only depends on the overlapping (as well as non-overlapping) features but also on the *salience* of such features in both topic and vehicle. Ortony succeeded in distinguishing between literal similarity and metaphorical similarity on the basis of *salience imbalance*. Consider these examples:

1. An encyclopedia is a book.
2. The boy is an encyclopedia.

In (1) the shared properties are of equal salience for both topic and vehicle, while in (2) properties salient for vehicle (*encyclopedia*) are not salient for topic (*the boy*) and this is what makes (1) a literal similarity and (2) a metaphoric comparison.

In general, such contrast-based models have been criticized specially for their presupposition of shared features. The first problem is that it is always possible to find some common features for any two things (Tourangeau & Sternberg, 1982). For instance consider the following:

- 3.* The encyclopedia is a boy.

Though (3) can hardly be considered a good metaphor or even a literal similarity, it is always possible to find some semantic overlap between *boy* and *encyclopedia*, such as both can be moved or both are physical objects.

Moreover, as empirically shown by Camac and Glucksberg (1984), word pairs that form topic and vehicle of good metaphors need not be associatively related prior to

their appearance in a metaphor. And finally, it has been found (including by Ortony, 1979) that "many metaphoric comparisons seem to involve properties that are not present in the listener's mental representation of the topic concept at all *until the metaphor is uttered and understood*" (Glucksberg & Keysar, 1990:7). Clearly, such evidence seriously questions the existence of shared features as the comparison ground.

Glucksberg and Keysar (1990) proposed an alternative view in which metaphors are understood as implicit class inclusion assertions. More specifically, a metaphor of the form *topic is vehicle* does not constitute a comparison between topic and vehicle. Instead it is an assertion to include topic in a category represented by vehicle. For example, (2) is interpreted as the inclusion of *the boy* in a category represented by *encyclopedia* (*the boy IS-A encyclopedia*).

The key to their argument is that both topic and vehicle can belong to various categories. For instance, *encyclopedia* can be cross-classified under different categories such as a physical object, a printed matter, something flammable, and a information repository. Also, *boy* can be a young male human, a student, a son, or a person with vast information. In this proposal, a metaphor is understood as a category to which both the topic and vehicle belong. Indeed, in (2), bearing vast knowledge (or knowledge repository) constitutes the ground for understanding this metaphor.

6.3 Review of Computational Models

Most existing computational models rely on some kind of semantic network to interpret metaphors. Metaphors in such models are recognized as constraint violations: resolving them requires some mapping between the two inconsistent domains. Martin's

metaphor understanding model (1990, 1991) is the most representative of such models. In this model, a sentence such as “How can I *enter* Lisp?” is considered to be a metaphor because *enter* and Lisp are incompatible (hence the constraint violation). Through its knowledge base, his model finds a path from Lisp to an existing mapping between computer program and enclosure. This pattern is then used to map *enter* onto *invoke*. Therefore, the metaphor is understood as “How can I *invoke* Lisp?”²¹

Veale and Kean (1992) also adopted the same approach. In their model, however, first a scaffolding for the core interpretation of a metaphor is constructed (again, based on some already available mappings for conventional metaphors). This core interpretation is later expanded based on world knowledge and context.

Obviously, such models require extensive and prior hand-coding of conceptual hierarchies and basic mappings, which makes them quite problematic. Besides, in such models, the similarity between two concepts is determined mainly based on structural closeness, which is again static and inflexible (Sun, 1995). And finally, such models have difficulty in accounting for the introduction of new properties in the topic by the vehicle, as discussed earlier.

Sun (1995) takes a radically different approach by proposing a microfeature-based model based on Ortony’s similarity imbalance theory. The model is implemented as a hybrid system, consisting of a localist network for representing domain concepts (e.g., *ugly*, *billboard*, *wart*, etc.) and a distributed network of microfeatures. Once presented with a metaphor such as *Billboards are warts*, nodes corresponding to *billboard* and *wart*

²¹ It can be argued that there is a subtle difference between “entering a program” and “invoking a

are activated in the first network. This leads to the activation of their related microfeatures in the second network, where, through a non-linear interaction, a proper interpretation emerges (e.g., *ugly*) from the microfeatures that remain activated.

Implicit in this model is the fact that all domain concepts are coded along a universal set of microfeatures: a problematic assumption (Corriveau, 1995). Also, it is not clear how well the microfeature network can distinguish between metaphors and literal (as well as abnormal) similarities.

In summary, the two major difficulties facing existing models of metaphor understanding are: 1) selection of the understanding ground for a metaphor from possibly a multitude of properties shared by both topic and vehicle, and 2) accounting for properties that the vehicle introduces to the topic. The model presented in the next section attempts to address these issues.

6.4 Metaphors as Class-Inclusion Assertions

The thrust of the proposed implementation is to use the sense-views associated with topic and vehicle as definitions of the various categories they belong to. For example, suppose *encyclopedia* appears in the following exemplars:

4. John looked for the name in the encyclopedia
5. The encyclopedia broke the window.
6. The student read the encyclopedia.
7. The encyclopedia was printed in September.

These exemplars will associate *encyclopedia* with different sense-views, namely those labeled for convenience as information repository (4), breaking instrument (5), reading

program”, something not preserved by this mapping.

material (6) and printed matter (7). Recall once more that, in fact, sense-views are identified only by their properties and the view-set(s) they belong to. The cross-classification of *encyclopedia* is shown in Figure 6-1.

Figure 6-1: *Encyclopedia* and its associated sense-views.

As previously explained, these sense-views can be shared by other concepts. For example, the following exemplars introduce new members to the above mentioned sense-views (see Figure 6-2):

8. The rock broke the glass.	for <u>Breaking-instrument</u>
9. The bullet smashed the window.	
10. The man broke the computer with a rock.	
11. Mary read the article.	for <u>Reading-material</u>
12. John read the newspaper.	
13. The student searched the Internet.	for <u>Information-repository</u>
14. Mary searched the archive.	
15. The student printed the article.	for <u>Printed-matter</u>
16. The newspaper is printed in Ontario.	

In Figure 6-2, from the viewpoint breaking-instrument, *encyclopedia* and *rock* are alike. From a different viewpoint, for instance information-repository, *encyclopedia* and the *Internet* are alike; and so on.

My contention is that every sense-view in fact represents a category of concepts deemed alike from a certain perspective. Such an *ad hoc* category (Barsalou, 1983) brings together concepts that share a similar classificatory role. In other words, from a certain perspective, concepts connected to a sense-view are alike, and this perspective is their shared context and the role they play in it.

Figure 6-2: *Encyclopedia* and its sense-views after exemplars 8 to 16.

This strategy can be used for metaphor understanding. However, two issues must be addressed: 1) how such contextually similar concepts can be classified together, and 2) how the understanding ground for a metaphor is selected? By defining the *alike* relationship, I have already shown that one can classify contextually similar concepts together. Here I intend to address the second question.

Algorithm 6-1 describes an algorithm that interprets an input metaphor, presented as *Topic is Vehicle*, by including the topic in one of the many sense-views that are

potentially associated with the vehicle. The algorithm first checks if the input is a literal similarity, that is, whether *topic* is a kind of *vehicle*. If this is not the case, the topic will be compared with all the sense-views associated with the vehicle. If the topic is not compatible with any of the sense-views, then the algorithm fails to find any interpretation for the input.

If more than one compatible sense-view is found, the most *typical* one will be selected. The typicality of the selected sense-view with the vehicle also determines the *metaphoricity*, that is, the goodness of the given metaphor.

1. If topic and vehicle concepts are conceptually close, then the input is best interpreted as a literal similarity; otherwise:
2. Find all sense-views associated with vehicle.
3. Select only those sense-views that are compatible with topic.
4. If there is no compatible sense-view, the input is anomalous (neither a literal similarity nor a metaphoric one).
5. If there is more than one compatible sense-view, then find vehicle's most typical sense-views.
6. The input metaphor is interpreted as an assertion to include topic as a member (i.e. Is-A) in the category represented by the selected sense-view of the vehicle.
7. The typicality of this sense-view for the vehicle indicates the *metaphoricity* of the input metaphor.

Algorithm 6-1: Understanding metaphors as class-inclusion assertions.

The typicality of a sense-view for a concept is assessed based on the centrality of the sense-view's properties for the concept. Recall that the centrality of a property is defined

based on the role that property will play in characterizing as well as distinguishing it from other concepts. As explained in chapter 4, in this implementation, properties derived immediately from a concept's corresponding node in the ontology have the highest centrality. Properties inherited from parent nodes have less centrality, depending on their distance from the concept's node. For example, considering the hierarchy given in Figure 6-3, one can assume that physical object properties are less central for humans than mammal properties.

Therefore, the typicality of a sense-view for a concept can be determined based on the number of central properties of this concept that are present in the sense-view. For example, breaking-instrument is not a very typical sense-view for *encyclopedia* since, of all of *encyclopedia's* properties, this sense-view contains only those portraying it as a physical object, which are not particularly central to this concept. The same sense-view would be more typical for *rock*.

6.5 Tests

Algorithm 6-1 was tested on a number of simple metaphors, which will be discussed in this section. The metaphors used in these tests are all nominal metaphors. However, I believe this approach is probably easily applicable to other forms of metaphor. Nevertheless, this is a subject of further study.

Inputs to the suggested algorithm are in the form *topic is a vehicle*, and for simplicity, it is assumed that neither topic nor vehicle concept are lexically ambiguous.

6.5.1 Test One

The lexicon was trained with exemplars 4 to 16 presented earlier in this chapter. Consequently *encyclopedia* was associated with four sense-views whose contents are shown below:

Breaking Instrument

STATE-OF-MATTER(0.569)---->:SOLID,
MADE-OF(0.803)---->: MATERIAL,

Reading Material

CONTAINS(0.666)---->:INFORMATION,
MADE-OF(0.633)---->:PAPER, INK,
IsKindOf-PRINTED-MEDIA(0.57)---->:
LOCATION(0.57)---->:ACEDEMIC-BUILDING,
IsKindOf-VISUAL-MEDIA-ARTIFACT(0.513)---->:
IsKindOf-DOCUMENT(0.543)---->:
PRODUCED-BY(0.543)---->:HUMAN,
REPRESENTS(0.7317)---->:OBJECT, EVENT, LANGUAGE,

Printed Matter

CONTAINS(0.666)---->:INFORMATION,
MADE-OF(0.633)---->: PAPER, INK,
IsKindOf-PRINTED-MEDIA(0.57)---->:
LOCATION(0.57)---->:ACEDEMIC-BUILDING,
IsKindOf-VISUAL-MEDIA-ARTIFACT(0.513)---->:
IsKindOf-DOCUMENT(0.543)---->:
PRODUCED-BY(0.543)---->:HUMAN,
REPRESENTS(0.7317)---->:OBJECT, EVENT, LANGUAGE,

Information Repository

INFORMATION-AMOUNT(1.0)---->:PLENTY,
CONTAINS(1.0)---->:INFORMATION,
MADE-OF(0.761)---->:

Six other exemplars were also used. They are:

17. The teacher teaches Physics.
18. The man teaches Chemistry.
19. John moved the stove.
20. Mary moved the computer.
21. The man moved the book.
22. John moved the man.

To understand why these contexts are needed, we should first look at the way *human* is defined in Mikrokosmos. Figure 6-3 illustrates the *human* taxonomy in Mikrokosmos.

As displayed, nowhere in this hierarchy is *human* associated with the capability of bearing information. Even properties concerning the physical aspect of *human* are hardly visible because of the distance between *physical-object* and *human*²².

Figure 6-3: Human hierarchy in Mikrokosmos.

One way to associate *human* with these properties is to manually hand-code them. However, this is unacceptable mainly because there are many other such *inferable* properties for *human*, and for any other concept for that matter. Besides, as argued by Barsalou (1982) and other researchers, there are always *ad hoc* properties for any concept that normally are not considered as parts of its characteristic until the appearance of the

²² Remember that two parameters of the ontology network, NonLocalityFactor and LocalityThreshold, determine the number of superclasses a concept inherits properties from.

concept in contexts implying such properties. Therefore, it would be impossible to hand-code all possible properties of a concept.

Instead, the approach adopted here stems from the fact that sense-views are more than a simple generalization for a number of concepts with overlapping properties. A sense-view may introduce new properties to some of its associated concepts (see sections 5.9 and 5.10). Therefore, in this model, a comprehensive set of properties for a concept can be obtained by looking at both 1) the concept's base definition obtained from the ontology, and 2) all sense-views associated over time with the concept.

Exemplar (17) sets a precedent in which *teacher* presumably has a property like <Contains: information>. Appearance of *man* in a similar context, exemplar (18) in this case, will lead to the classification of *man* under the same category as *teacher*. Consequently <Contains: Information> becomes a possible, or *inferable*, property for *human* (the concept underlying the word *man*). Similarly, exemplar (22) makes properties concerning movable objects, used in exemplars (20) and (21), *inferable* for the concept *human*. This is because exemplar (22) relates *human* (i.e., the concept associated with the input word *man*) with the sense-view object of moving.

After training the system with exemplars 4 to 22, it was presented with the following metaphoric and non-metaphoric inputs.

Input 1: An encyclopedia is a book.

This is clearly a literal similarity. Here is the generated output in which the topic is correctly recognized as a kind-of the vehicle (literal similarity).

```
Topic ENCYCLOPEDIA
Vehicle BOOK
Topic(Encyclopedia) is Is-Kind-of Book
```

Input 2: John is an encyclopedia.

Here is a complete trace for this input. Note that *John* (concept *human male*) is less compatible with the sense-view information repository (the recipient of action *search*, as in exemplars (4), (13) and (14)) than with breaking instrument. But later, when it comes to deciding which of the compatible sense-views is more typical, the former one is selected. The two other sense-views, printed matter and reading material are already excluded due to their low compatibility with the topic. Finally, the system recognizes the input as a very good metaphor (metaphoricity is 1) and interprets it as an assertion to include *John* in one of the categories *encyclopedia* can represent (i.e., its associated sense-views), namely information repository. Notice the property <Information-Amount: Plenty> is a new property for *John*, introduced by this sense-view.

```
Topic HUMAN-MALE
Vehicle ENCYCLOPEDIA
Compatibility of HUMAN-MALE with sense-view Search(Recipient) is:0.6379
Compatibility of HUMAN-MALE with sense-view Breaking(Instrument) is:1.0
Compatibility of HUMAN-MALE with sense-view ReadMaterial(Theme) is:0.2724
Compatibility of HUMAN-MALE with sense-view Printing(Object) is:0.2724

Typicality of Sense-View Breaking (Instrument) for ENCYCLOPEDIA is:0.5852
Typicality of Sense-View Search (Recipient) for ENCYCLOPEDIA is:1.0

Topic(John) IsA Encyclopedia
      [Metaphoricity:1]
  ViewSet Search
sense-View--> Recipient - No of Exemplars:3 - Marker(s)-->p-obj,in-pp,
INFORMATION-AMOUNT(1.0)---->:PLENTY,
CONTAINS(1.0)---->:INFORMATION,
MADE-OF(0.76169)---->:
```

Input 3: *The encyclopedia is John.

This input is presented to further test the algorithm. One typical behavior of metaphors is their irreversibility. In the few cases in which topic and vehicle of a metaphor are

interchangeable, the understanding ground changes completely (see the second test below).

Here is the output produced by the system. As shown, *John* is associated with seven sense-views, which all, except for the last one, are incompatible with the new topic, *encyclopedia*. Therefore, the input is understood on the basis of *encyclopedia* being like *John* with respect to being an object of moving. Remember because of (22), *human* can be moving objects.

```
Topic ENCYCLOPEDIA
Vehicle HUMAN-MALE

Compatibility of ENCYCLOPEDIA with sense-view Breaking(Agent) is:0.0
Compatibility of ENCYCLOPEDIA with sense-view ReadMaterial(Agent) is:0.0
Compatibility of ENCYCLOPEDIA with sense-view Search(Agent) is:0.0
Compatibility of ENCYCLOPEDIA with sense-view Teaching(Agent) is:0.0
Compatibility of ENCYCLOPEDIA with sense-view MovingObject(Agent) is:0.0
Compatibility of ENCYCLOPEDIA with sense-view MovingObject(Agent) is:0.0
Compatibility of ENCYCLOPEDIA with sense-view MovingObject(Object) is:1.0

Topic(Encyclopedia) IsA John
      [Metaphoricity:0]
ViewSet MovingObject

[Sense View]
Thematic Role: Object - No. of Exemplars: 4 - Marker(s): p-obj,
STATE-OF-MATTER(0.548775)---->:SOLID,
MADE-OF(0.63475)---->:PLASTIC, METAL,
IsKindOf-ARTIFACT(0.5103)---->:
COLOR(0.5103)---->:
AGE(0.5103)---->:
OPERATED-BY(0.5103)---->:HUMAN,
```

However, the system recognizes the input as an extremely poor metaphor (metaphoricity is zero). As argued by Ortony (1979), for a metaphor to be considered “good” the understanding ground must comprise properties that are not salient for topic, yet are salient for the vehicle. This will guarantee the informativeness of metaphors. In this case,

being a moveable object is not a salient aspect of human beings. Therefore, the input is not a good metaphor.

6.5.2 Test Two

The test reported in this section demonstrates the ability of the system to account for reversible metaphors. As mentioned earlier, in the few cases that the topic and vehicle of a metaphor can be interchanged, the understanding ground also changes completely.

Examples of such cases are:

- 23. My surgeon is a butcher.
- 24. My butcher is a surgeon.
- 25. Warts are billboards.
- 26. Billboards are warts.

For instance, 23 describes the incompetence of the *surgeon*, whereas 24 is indicative of the skill and competence of the *butcher*.

In order to demonstrate the ability of Algorithm 6-1 to deal with such cases, the system was trained with the following exemplars:

- 27. Mary saw the billboard.
- 28. John spotted the lighthouse.
- 29. The man broke the billboard.
- 30. The wart marred the face.
- 31. The smog spoiled the scene.
- 32. The surgeon operated on the wart.

As the result, billboard and wart were each associated with two sense-views, namely, Obvious-Objects and Object-of-Breaking for *billboard*, and Blemish and Object-of-Surgery for *wart*. The system was then presented with the following inputs.

Input 1: Warts are billboards.

Here is the output generated for this metaphor. As displayed, *warts* are understood on the ground of *obviousness*, that is, one of the two sense-views associated with *billboards*. It is also judged as a good metaphor.

```

Topic WART
Vehicle BILLBOARD
Compatibility of WART with view Seeing(Theme) is:0.5
Compatibility of WART with view Breaking(Object) is:0.211674
Topic(Wart) IsA Billboard
                        [Metaphoricity:1]
ViewSet Seeing
  [Sense View]
Thematic Role: Theme - No. of Exemplars: 2 - Marker(s): p-obj,
STANDS-OUT(1)---->:
OBVIOUS(1)---->:

```

Input 2: Billboards are Warts.

This is the reverse metaphor and we expect a different ground to be selected. Here is the output:

```

Topic BILLBOARD
Vehicle WART
Compatibility of BILLBOARD with view Blemishing(Agent) is:0.5
Compatibility of BILLBOARD with view Surgery(Object) is:0.281885
Topic(Billboard) IsA Wart
                        [Metaphoricity:1]
ViewSet Blemishing
  [Sense View]
Thematic Role: Agent - No. of Exemplars: 2 - Marker(s): p-subj,
VISIBLE(1)---->:
UGLY(1)---->:

```

As shown, the input is interpreted as the inclusion of *billboard* in a category of ugly objects, or agents of blemishing (one of the two sense-views associated with *wart*). The input is also recognized as a good metaphor, since properties in the selected sense-view are strongly typical for *warts*.

Thus, in summary, my initial tests seem to suggest that the separation of sense representations into two tiers indeed offers a promising avenue to deal with metaphor understanding.

Chapter Seven

Discussion and Future Work

7.1 Introduction

Accounting for effects of context in highlighting/dimming certain properties of concepts has been the primary motivation behind this work. Throughout this dissertation, I have argued for an alternative model of lexicon in which sense selection and sense modulation go hand in hand. In my opinion, this sets the proposed model apart from most other models of lexicon. The role of the lexicon in the proposed model exceeds the conventional usage of a lexicon as a repository of word senses. Through the exemplars it maintains, the proposed lexicon participates in the process of interpreting an input context. Moreover, by dealing with variations in the encoding and storage of concepts through the use of sense-views, it goes beyond mere representations of words, and takes a

step towards becoming a model of memory and the mapping of linguistic representation of meaning to some longer storage.

This last chapter intends to reflect on these claims through a review of the initial motivations and results of the proposed model with respect to these claims. Also some future extensions will be discussed.

7.2 Accomplishments

The contributions of this thesis can be summarized as follows:

- It exploits a two-tiered word sense structure that separates representation of concepts from the way they are viewed in various contexts in order to deal with semantic flexibility.
- It proposes an exemplar-based approach to the gradual development of sense-views from alike concepts, i.e., concepts that play similar roles in similar contexts.
- It suggests exemplars as examples of typical use. Concepts in the proposed lexicon are partially defined by these exemplars.
- It puts forth a distributed process that integrates sense-selection and sense-modulation. This process is controlled and carried out by individual exemplars.

Each of these issues will be briefly reviewed in the following subsections.

7.2.1 Sense-Views vs. Sense-Concepts

This distinction was primarily advanced to account for semantic flexibility. The idea was to allow a single concept to be viewed differently in various contexts. The

dichotomy corresponds to the representational and classificatory functions of concepts, as suggested by Franks (1995). I argued that by allowing word senses to consist of two components, namely sense-concepts and sense-views, one can account for changes that context causes in highlighting or backgrounding certain properties of concepts. The idea is that while the core characteristics for occurrences of a concept in various contexts (i.e., sense-concepts) remain the same, contextual effects are accounted for by the sense-views that will accompany sense-concepts in each context.

I showed that sense-views can be used to account for various forms of sense-modulation, namely, property highlighting/backgrounding (see sections 5.5 and 5.6) and promotion/demotion (sections 5.9 and 5.10). I also showed how sense-views can establish clues for unknown words (see section 5.12). Simple examples of the model's ability to address polysemic and metonymic mappings were also presented (section 5.15). Since the implemented prototype is currently restricted to sentential contexts, no specific example for juxtaposition was presented. However, every test reported in chapters 5 and 6 clearly emphasizes the model's ability to separate between 1) what a concept typically represents and 2) an instantiated use of this concept. And thus, I believe that my model will scale up to the processing of juxtaposition sentences.

7.2.2 Evolution of Sense-Views

I have argued that different concepts become alike if they appear in similar contexts. I defined *alike* concepts as concepts with the same thematic roles in similar contexts. In other words, alike concepts share the same sense-view. I contended that the characteristics of a sense-view can be dynamically extracted from the characteristics of the sense-concepts that are associated with it. This argument rests on the observation that

alike concepts represent a category of objects that can be characterized by the sense-view they all share.

To accommodate new members (sense-concepts) as they become associated, this sense-view should constantly change to reflect a “generalization” of its associated sense-concepts. This explains the evolution of sense-views: At the beginning, a sense-view has the same properties as the single sense-concept that defines it. Over time, the sense-view evolves as new sense-concepts become associated with it.

7.2.3 Exemplars

Traditionally concepts are typically defined by means of central properties, that is properties used to distinguish one concept from others. In the proposed lexicon, however, this definition is augmented with examples of typical use, or *exemplars*. Therefore, to know a concept is to know how it is used. This approach complies with Wittgenstein’s admonition: “do not tell me the meaning, tell me the use” (in Berwick, 1989).

It also agrees with the exemplar view of conceptual structure (for a review of these different theories, see Medin & Goldstone, 1990). From this standpoint, a concept is represented by the individual instances (exemplars) that constitute it (e.g., see Hintzman, 1986; 1988). An exemplar-based approach clearly has crucial implications regarding meaning representation. Balota remarks:

[The] meaning of a given word is not simply represented as a list of isolable semantic features or a prototype but rather should be conceived as a cumulation of individual episodic experiences with that word. Each of the experiences (i.e., the word and its embedded context) contributes to the evolving meaning that we attribute to a given word. (1990:26)

Another advantage of such an approach is that it eliminates the need for extensive and prior encoding of concepts. Many properties of a concept can be *inferred* as it appears in various contexts. Exemplars also offer support for *ad hoc* properties (see the example in 6.5.1).

Moreover, this approach makes the definition of a concept partially dependent on the exemplars maintained by the lexicon for that concept. This dependency has some important implications. In particular, the definition of a single concept can vary as its set of exemplar changes. This implies that concept definitions can become personalized. Such idiosyncratic definitions comply with the reader-based position to language comprehension promoted by Corriveau (1995) in computational linguistics, and many in reception theory (e.g., Holub, 1984). The ability of the model to address issues such as idiosyncrasy and diachrony was demonstrated in section 5.14 and 5.15 respectively.

7.2.4 Augmenting Sense selection with Sense modulation

Unlike most conventional lexicons, which support only sense selection, the lexicon discussed in this dissertation augments sense selection with sense modulation. In contrast with other lexicons, which are used to look up single words, the proposed lexicon works with contexts. A context is presented to the lexicon as a number of words along with their syntactic markers. As output, the model is expected to not only select proper senses for each word, but also to modulate the selected senses according to the surrounding context.

The key criterion in selecting senses for the input words is that the selected senses interact with each other. Again using exemplars, I presented a computational approach to

determine the interactability between word senses. Put simply, a sense can replace another sense in an exemplar if it can be associated with its corresponding sense-view (see 3.5).

As for sense modulation, it occurs by associating a selected sense with a sense-view, which may result in changes in this sense-view.

Finally, sense selection and sense modulation in the proposed lexicon are carried out by agents attached to individual exemplars. More specifically, once a word is presented to the model, any exemplar that this word appears in is activated. Each exemplar represents one usage of the word. The agent associated with each exemplar concurrently and independently verifies whether senses can be derived for other input words so that they can interact with the sense it exemplifies for that word. If so, such a set of interactable senses will represent one possible interpretation for the input context.

7.3 Future Work

In this section I consider some extensions and future research.

7.3.1 Beyond Simple Nouns/Verbs

Processes and algorithms presented throughout the dissertation were designed for simple sentential contexts. More specifically, sense-views were assigned to simple nouns and verbs. The first step in scaling up this work is to extend these processes and algorithms to also account for compound nouns, as well as phrases and clauses. Examples of such sentences are:

1. The old book with the leather cover broke the window.
2. The book, which my friend Berthold gave me for my birthday, burnt in the fire.

3. The stone lion scared the boy.

This requires the ability to parse such compound structures, as well as to devise a mechanism to take into account the effects of elements of a compound structure on the overall making of this structure's semantic properties. Such an account could be as simple as property alteration, as in (1) in which the adjective *old* and the *with* prepositional phrase alter properties already attached to *book*. But this is not the case in (2) in which the given clause does not directly modify any of *book's* properties. And in (3), the combination of *stone* and *lion* is an example of how concept combination can be difficult to handle. As discussed in Franks (1995), combinations such as *stone lion* and *fake gun* (also known as “privative combinations”) are usually considered as extreme test-cases for theories of concepts and concept combination because how the two concepts are ultimately combined depends on the context. For example, in (3) the “scariness” of *lion* will be maintained in the combination, but not in a context such as *The kids were keen to sit on the stone lion*.

7.3.2 Further Evolution of Sense-Views

In the proposed model, a sense-view is initialized by a sense-concept and evolves as more alike sense-concepts are connected to it. Because the alike relationship is not fine-grained, the range of objects associated with a sense-view can be very broad. Consequently, some information can be lost as the result of generalization. For instance, consider the moving object sense-view (section 5.3). Sense-concepts such as *book*, *piano*, *table*, *vase* etc. can be associated with this sense-view. Some of these objects are heavy, while others are not. This information will be lost as the result of the generalization process, though intuitively, weight is an important property to consider in moving.

Another example is the difference in properties describing attitude in each of the instances of *talk* in the following examples, a modulation that will be lost through the generalization.

John talked to his wife.

John talked to his boss.

John talked to his dog.

One solution is to allow a sense-view to split into finer categories. For example, moving objects could be divided into heavy and non-heavy objects. This could be done automatically, using simple classifier algorithms (see Quinlan, 1993) when the cardinality of sense-view exceeds a certain limit. This method would allow us to refine the alike relationship while maintaining its exemplar-based nature. Consequently, the model would be able to make a fine-grained distinction between alike sense-concepts (e.g., to recognize that moving a *book* is more like moving a *bag* and less like moving a *piano*).

7.3.3 Enhancing the Implemented Prototype

The implemented prototype could be enhanced in a number of ways. The first enhancement that comes to mind is the enhancement of the user interface. Functionality was the major goal for the implemented prototype, therefore, very little work has been done to provide user friendly interfaces. Examples of such interfaces include an interface for the ontology network to allow users to define and inspect concepts and their properties, and an interface to inspect the contents of various levels of the lexicon network.

Also as mentioned before (section 4.4), the Mikrokosmos ontology is more concerned with concept taxonomy and less with defining concepts based on properties.

To have a more accurate sense-views, in terms of property specification, it is imperative to either improve this situation or find another ontology.

And, currently no relationship is assumed between properties. In other words, properties are in no way related to each other. Therefore, property matching is very primitive. For example, there is no way the system can determine *metal* is a kind of *material*! This simplistic strategy complies with microfeature approaches, as used in connectionist models. But in order to allow the system to match properties in a more sophisticated way, the ontology network should be expanded to include properties too. That is to say, much like concepts, properties need to be organized in a taxonomic structure. Clearly, the system's algorithm would also need to be adjusted.

Finally, the entire lexicon network is built and maintained in dynamic memory, that is, at run time. Therefore, whatever the system learns (in terms of new exemplars that get added and sense-views that evolve) will not be maintained across sessions! Obviously, some approach to persistence is required to make such information available from execution to execution.

But, in conclusion, I believe that the current model, prototype, and tests do demonstrate the importance, computational feasibility, and usefulness of two-tiered sense representations for the processing of contextual effects.

References

- Alterman, R.; and Bookman, L.A. 1992. Reasoning about the Semantic Encoding of the Connectivity of Events. *Cognitive Science* 16:205-232.
- Anderson, R.; Pichert J.W.; Goetz, E.T.; Schallert, D.L.; Stevens, K.V.; and Trollip, S.R. 1976. Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior* 15:667-679.
- Balota, D.A. 1990. The role of meaning in word recognition. In Balota, D.; Flores d'Arcais G.; and Rayner, K. (Eds.) *Comprehension Process in Reading* Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Barclay, J.R.; Bransford, J.D.; Franks, J.J.; McCarrell, N.S.; and Nitsch, K. 1974. Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior* 13:471-481.
- Bareiss, R. 1989. *Exemplar-based knowledge acquisition: a unified approach to concept representation, classification, and learning*. Academic Press.
- Barsalou, L.W. 1982. Context-independent and context-dependent information in concepts. *Memory and Cognition* 10(11):82-93.
- Barsalou, L.W. 1983. Ad hoc categories. *Memory and Cognition* 11(3):211-227.
- Barsalou, L.W. 1987. The instability of graded structure: Implications for the nature of the concepts. In Neisser U. (Ed.) *Concepts and conceptual development: Ecological and intellectual factors in categorization*. Cambridge, UK: Cambridge University Press.
- Barsalou, L.W. 1993. Flexibility, structure, and linguistic vagary in concepts: manifestations of a compositional system of perceptual symbols. In Collins, A.F. (Ed.) *Theories of Memory* Sussex, UK: Lawrence Erlbaum Associates.

- Berwick, R.C. 1989. Learning word meanings from examples. In Waltz D. (Ed.) *Semantic structures: Advances in Natural Language Processing*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Bookman, L.1994. *Trajectories through knowledge space*. Kluwer Academic Publishers.
- Camac, M.K.; and Glucksberg. 1984. Metaphors do not use associations between concepts, they are used to create them. *Journal of Psycholinguistic Research*. 13:443-455.
- Clark, H.H.; and Gerrig, R.J. 1983. Understanding old words with new meanings. *Journal of Verbal Learning and Verbal Behavior* 22:591-608.
- Corriveau, J.-P. 1995. *Time-constrained memory*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Cruse, D. 1986. *Lexical semantics*, Cambridge University Press.
- Cruse, D. 1995. Polysemy and related phenomena from a cognitive linguistic viewpoint. In Saint-Dizier, P.; and Viegas, E. (Eds.) *Computational lexical semantics*. NY, Cambridge University Press.
- Cullingford, R. 1978. Script applications: computer understanding of newspaper stories (PhD thesis), Dept. of Computer Science, New Haven, CT: Yale University.
- Delisle S.; Copeck, T.; Szpakowicz, S.; and Barker, K. 1993. Pattern matching for case analysis: a computational definition of closeness. In *Proceedings of the International Conference on Computer and Information-93* 310-315.
- Delisle S.; Barker, K.; Copeck, T.; and Szpakowicz, S. 1996. Interactive semantic analysis of technical texts. *Computational Intelligence* 12(2), 273-306.
- DiMarco C. and Hirst, G. 1993. Usage notes as the basis for a representation of near-synonymy for lexical choice. In *Proceedings of the 9th Annual Conference of the*

University of Waterloo Centre for the new Oxford English Dictionary and Text Research, Oxford, 33-43.

Dyer, M.G. 1983. *In-depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension*. Cambridge, MA: MIT Press.

Fodor, J. and Lepore, E. 1996. The emptiness of the lexicon (Tech. Rep. RuCCS TR-27) New Brunswick, NJ: Rutgers University, Center for Cognitive Science.

Franks, B. 1995. Sense generation: A “quasi-classical” approach to concepts and concept combination. *Cognitive Science* 19:441-505.

Frawley, W. 1992. *Linguistic semantics*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Frazier, L.; and Rayner, K. 1990. Taking on semantic commitments: processing multiple meanings vs. multiple senses. *Journal of Memory and Language* 29:181-200.

Gerrig, R.; J. 1989. The time course of the sense creation. *Memory and Cognition* 17 (2):194-207.

Glucksberg, S.; and Keysar, B. 1990. Understanding metaphorical comparisons: beyond similarity. *Psychological Review* 97:3-18.

Greenspan, S.L. 1986. Semantic flexibility and referential specificity of concrete nouns. *Journal of Memory and Language* 25:539-557.

Hendler, J.A. 1987. *Integrating marker passing and problem solving: a spreading-activation approach to improved choice in planning*. Norwood, NJ: Ablex.

Hendler, J.A. 1989. marker passing over microfeatures: towards a hybrid symbolic/connectionist model. *Cognitive Science*, 13:79-106.

Hintzman, D. 1986. “Schema abstraction” in a multiple-trace memory model. *Psychological Review* 93(4):411-428.

- Hintzman, D. 1988. Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review* 95(4):528-551.
- Hirst, G. 1987. *Semantic interpretation and the resolution of ambiguity* Cambridge, UK: Cambridge University Press.
- Hirst, G. and St-Onge, D. 1995. Lexical chains as representations of context for detection and correction of malapropisms. In Fellbaum, C. (Ed.) *WordNet: An electronic lexical database and some of its applications* Cambridge, MA: The MIT Press.
- Holub, R. 1984. *Reception theory: A critical introduction*. NY: Methuen.
- Jones, D. 1996. *Analogical natural language processing*. London, UK: University College London Press.
- Kawamoto, A.H. 1993. Nonlinear dynamics in the resolution of lexical ambiguity: a parallel distributed processing account. *Journal of Memory and Language* 32:474-516.
- Kilgarriff, A. 1997. What is word sense disambiguation good for?. *Proceedings of Natural Language Processing in the Pacific Rim(NLPRS-97)* Phuket, Thailand.
- Kilgarriff, A. 1998. "I do not believe in word senses". *Computers and Humanities* 31(2): 91-113.
- Kintsch, W. 1988. The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review* 95(2):163-182.
- Kintsch, W.; Welsch, D.; Schmalhofer, F.; and Zimny, S. 1990. Sentence memory: a theoretical analysis. *Journal of Memory and Language* 29:133-159.
- Lakoff, G. 1987. *Women, fire and dangerous things: What categories reveal about the mind*. The University of Chicago Press.
- Landman, F. 1986. *Towards a theory of information*. Dordrecht, Netherlands: Foris.

- Lange, T.; and Dyer, M. 1989. Frame selection in a connectionist model of high-level inferencing. *Proceedings of the 11th Conference of the Cognitive Science Society*.
- Lea, D. 1996. *Concurrent Programming in Java* Reading, MA: Addison-Wesley.
- Lenat, D.; Miller, G.; and Yokoi, T. 1995. CYC, WordNet, and EDR: Critiques and responses. *Communication of ACM* 8(11):45-48.
- Mahesh K.; and Nirenburg, S. 1995. A situated ontology for practical NLP. *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing. International Joint Conference of AI. Montreal, Canada*.
- Martin, J.H. 1990. *A computational model of metaphor interpretation*. San Diego, CA: Academic Press.
- Martin, J.H. 1992. Computer understanding of conventional metaphoric language. *Cognitive Science* 16:233-270.
- McClelland, J.L. and Kawamoto, A.H. 1986. Mechanisms of sentence processing: assigning roles to constituents of sentences. In McClelland J.L.; and Rumelhart, D.E (Eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* Vol. 2, MIT press.
- McCloskey, M.E.; and Glucksberg, S. 1978. Natural categories: Well defined or fuzzy sets? *Memory and Cognition* 6(4):462-472.
- Medin, D.L.; and Goldstone, R.L. 1990. Concepts. In Eysenck M.W. (Ed.) *The Blackwell dictionary of cognitive psychology* Cambridge, Massachusetts: Basil Blackwell Ltd.
- Michalski, R.S. 1983. A theory and methodology of inductive learning. In Anderson, J.; Michalski, R.; Carbonell, J.; Mitchell, T. (Eds.) *Machine learning* Palo Alto, CA.

- Michalski, R.S. 1989. Two-tiered concept meaning, inferential matching and conceptual cohesiveness. In Vosniadou, S.; and Ortony, A. (Eds.) *Similarity and analogical reasoning*. N.Y. : Cambridge University Press.
- Miikkulainen, R. 1993. *Subsymbolic natural language processing: An integrated model of scripts, lexicon, and memory*. Cambridge, MA: MIT Press.
- Miller, G.A. 1995. *A Lexical Database for English*. *Communication of ACM* 8(11):39-41.
- Miller, G.A.; and Charles, W.G. 1991. Contextual correlates of semantic similarity. *Language and Cognitive Processes* 6 (1):1-28.
- Murphy, G.L.; and Medin, D.L. 1985. The role of theories in conceptual coherence. *Psychological Review* 92:289-316.
- Niemeyer, P. & Peck, J. (1996) *Exploring Java* Sebastopol, CA: O'Reilly & Associates.
- Oden, G.C. 1987. Concepts, knowledge, and thought. *Annual Review of Psychology* 38: 203-227.
- Ortony, A. 1979. Beyond literal similarity. *Psychological Review* 86(3):161-180.
- Pustejovsky, J.; and Boguraev, B. 1993. Lexical knowledge representation and natural language processing. *Artificial Intelligence* 63 (1-2):193-223.
- Pustejovsky, J. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.
- Quillian, M.R. 1967. Word Concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science*, 12, 403-417.
- Quinlan, J. 1993. *C4.5: Programs for machine learning*, San Mateo, CA: Morgan Kaufmann.
- Rais-Ghasem, M. & Corriveau, J.-P. (1996) Beyond concept recognition. *Proceedings of the Ninth Florida Artificial Intelligence Research Symposium (FLAIRS)*, Key West Florida.

Rais-Ghasem, M. & Corriveau, J.-P. (1998a) Exemplar-based sense modulation. To be presented to *the First International Conference on Mental Lexicon*, University of Alberta, Canada.

Rais-Ghasem, M. & Corriveau, J.-P. (1998b) An exemplar-based sense modulation, to appear in *the Proceedings of the COLING-ACL workshop on Computational Treatment of Nominals*. Montreal, Canada.

Ruhl, C. (1989) *On monosemy: A study in linguistic semantics* Albany, N.Y: State University of New York Press.

Saint-Dizier, P. & Viegas, E. (1995) *Computational Lexical Semantics* Cambridge, UK: Cambridge University Press.

Schank, R.C. 1975. *Conceptual information processing*. North Holland, Amsterdam.

Schank, R.C. 1980. Language and memory. *Cognitive Science* 4:234-284.

Schank R. C.; and Abelson, R. 1977. *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Seidenberg, M.S.; Tanenhaus, M.K.; Leiman, J.M.; and Bienkowski, M. 1982. Automatic access of the meanings of ambiguous words in context: some limitations of knowledge-based processing. *Cognitive Psychology* 14:489-537.

Simpson, G.B.; and Burgess, C. 1985. Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance* 11 (1):28-39.

Small, S.L. 1987. WordExpert parsing. In Shapiro, S.; and Eckroth, D. (Eds.) *Encyclopedia of Artificial Intelligence* New York, NY: John Wiley.

- Speelman, C.P.; and Kirsner, K. 1990. The representation of text-based and situation-based information in discourse comprehension. *Journal of Memory and Language* 29:119-132.
- Stanfill, C.; and Waltz, D. 1986. Toward memory-based reasoning. *Communication of ACM* 29(12):1213-1228.
- Sun, R. 1995. A Microfeature based approach towards metaphor interpretation. In Proceeding of the International Joint Conference of AI.
- Tulving, E. 1990. Encoding principle specificity. In Eysenck M.W. (Ed.) *The Blackwell Dictionary of Cognitive Psychology* Cambridge, Massachusetts: Basil Blackwell Ltd.
- Tversky, A. 1977. Features of Similarity. *Psychological review* 85:327-352.
- Veale, T.; and Kean, M. 1992. Conceptual scaffolding. *Computational Intelligence* 8(3): 494-519.
- Veronis, J. & Ide, N. (1995). Large neural networks for the resolution of lexical ambiguity. In Saint-Dizier, P. & Viegas, E. (Eds.) *Computational Lexical Semantics* NY, Cambridge University Press.
- Waltz, D.L.; and Pollack, J.B. 1985. Massively parallel parsing: a strongly interactive model of natural language interpretation. *Cognitive Science* 9:51-74.
- Wilensky, R. 1978. *Understanding goal-based stories* (Research Report), Dept. of Computer Science, New Haven, CT: Yale University.
- Witney, P.; McKay, T.; and Kellas, G. 1985. semantic activation of noun concepts in context. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11:126-135.
- Yokoi, T. 1995. The EDR electronic dictionary. *Communication of ACM* 8(11):42-44.

IMAGE EVALUATION TEST TARGET (QA-3)

APPLIED IMAGE, Inc.
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved