Semantic Relations in Information Science

Christopher S. G. Khoo and Jin-Cheon Na
Nanyang Technological University

Introduction

This chapter examines the nature of semantic relations and their main applications in information science. The nature and types of semantic relations are discussed from the perspectives of linguistics and psychology. An overview of the semantic relations used in knowledge structures such as thesauri and ontologies is provided, as well as the main techniques used in the automatic extraction of semantic relations from text. The chapter then reviews the use of semantic relations in information extraction, information retrieval, question-answering, and automatic text summarization applications.

Concepts and relations are the foundation of knowledge and thought. When we look at the world, we perceive not a mass of colors but objects to which we automatically assign category labels. Our perceptual system automatically segments the world into concepts and categories.[1] Concepts are the building blocks of knowledge; relations act as the cement that links concepts into knowledge structures. We spend much of our lives identifying regular associations and relations between objects, events, and processes so that the world has an understandable structure and predictability. Our lives and work depend on the accuracy and richness of this knowledge structure and its web of relations. Relations are needed for reasoning and inferencing.

Chaffin and Herrmann (1988b, p. 290) noted that "relations between ideas have long been viewed as basic to thought, language, comprehension, and memory." Aristotle's *Metaphysics* (Aristotle, 1961; McKeon, 1941/2001) expounded on several types of relations. The majority of the 30 entries in a section of the *Metaphysics* known today as the *Philosophical Lexicon* referred to relations and attributes, including cause, part-whole, same and opposite, quality (i.e., attribute) and kind-of, and defined different types of each relation. Hume (1955) pointed out that there is a connection between successive ideas in our minds, even in our dreams, and that the introduction of an idea in our mind automatically recalls an associated idea. He argued that all the objects of human reasoning are divided into relations of ideas and matters of fact and that factual reasoning is founded on the cause-effect relation. His *Treatise of Human Nature* identified seven kinds of relations: resemblance, identity, relations of time and place, proportion in quantity or number, degrees in quality, contrariety, and causation. Mill (1974, pp. 989–1004) discoursed on several types of relations, claiming that all things are either feelings, substances, or attributes, and that attributes can be a quality (which belongs to one object) or a relation to other objects.

Linguists in the structuralist tradition (e.g., Lyons, 1977; Saussure, 1959) have asserted that concepts cannot be defined on their own but only in relation to other concepts. Semantic relations appear to reflect a logical structure in the fundamental nature of thought (Caplan & Herrmann, 1993). Green, Bean, and Myaeng (2002) noted that semantic relations play a critical role in how we represent knowledge psychologically, linguistically, and computationally, and that many systems of knowledge representation start with a basic distinction between entities and relations. Green (2001, p. 3) said that "relationships are involved as we combine simple entities to form more complex entities, as we compare entities, as we group entities, as one entity performs a process on another entity, and so forth. Indeed, many things that we might initially regard as basic and elemental are revealed upon further examination to involve internal structure, or in other words, internal relationships."

Concepts and relations are often expressed in language and text. Language is used not just for communicating concepts and relations, but also for representing, storing, and reasoning with concepts and relations. We shall examine the nature of semantic relations from a linguistic and psychological perspective, with an emphasis on relations expressed in text. The usefulness of semantic relations in information science, especially in ontology construction, information extraction, information retrieval, question-answering, and text summarization is discussed.

Research and development in information science have focused on concepts and terms, but the focus will increasingly shift to the identification, processing, and management of relations to achieve greater effectiveness and refinement in information science techniques. Previous chapters in *ARIST* on natural language processing (Chowdhury, 2003), text mining (Trybula, 1999), information retrieval and the philosophy of language (Blair, 2003), and query expansion (Efthimiadis, 1996) provide a background for this discussion, as semantic relations are an important part of these applications.

What Are Semantic Relations?

Semantic Relations in Language and Logic

Semantic relations are meaningful associations between two or more concepts, entities, or sets of entities. They can be viewed as directional links between the concepts/entities that participate in the relation. The concepts/entities are an integral part of the relation as a relation cannot exist by itself. Associations between concepts/entities can be categorized into different types, abstracted, conceptualized and distinguished from other associations, and can thus be assigned meaning. The meaning or type of an association can sometimes but not always be derived from the meanings of the concepts involved. Psychologists and philosophers have attempted to identify the main types of relations and their features.

Two concepts connected by a relation are often represented as a concept-relation-concept triplet: [concept1] → (relation) → [concept2].[2] The link is labeled to indicate the type or meaning of the relation. A relation can thus be viewed as containing two places or slots that need to be filled. A relation exerts selectional restrictions on the slots that constrain the kind of concepts or entities that can occupy them. A valid participant of a relation may need to have certain semantic features or belong to a semantic category. For example, in the relation [John] → (is-father-of) → [Mary], the entity represented by "John" has to belong to the category of human beings and have the gender feature *male*. A relation can also constrain the slot filler to a concept, an entity (i.e., instance of a concept), a set of entities, or a mass concept (denoting a set of entities).

Although most relations are binary relations having two slots, a relation may have three or more slots. For example, the *buy* relation may relate four participants: the buyer, the seller, the thing that is bought, and the price. The number of slots of a relation is called its arity or valence. *Buy* is a 4-ary relation, and the four participants in the relation are assigned the roles *agent* (buyer), *source* (seller), *patient* (thing bought), and *price* to distinguish between them. It is, however, well known that relations with arity higher than two can be decomposed into a set of more primitive binary relations. For example, the *buy* relation can be converted to a *buy* concept, which can be linked to the four participants with the binary relations agent, source, patient, and price. Sowa (1984) proposed the generic *link* relation as the most primitive relation. All other relations can be defined in terms of concepts combined with the *link* relation. For example, the *eat* relation in [John] → (eat) → [apple], can be decomposed into the concept *eat* and the case relations *agent* and *patient*: [John] → (agent) → [eat] → (patient) → [apple]. The *agent* relation can be further reduced to the concept *agent* and the *link* relation: [John] → (link) → [agent] → (link) → [eat].

Sowa (1984) further suggested that tenses and modalities, such as *possibility*, *necessity*, *permission*, and *negation*, be treated as 1-ary or "monadic" relations. For example, the PAST relation can indicate that a proposition was true in the past: (PAST) → [PROPOSITION].

Semantic relations can refer to relations *between concepts in the mind* (called conceptual relations), or relations *between words* (lexical relations) or text segments. However, concepts and relations are inextricably bound with language and text and it is difficult to analyze the meaning of concepts and relations apart from the language that expresses them. Wittgenstein (1953, p. 107) said, "When I think in language, there aren't 'meanings' going through my mind in addition to the verbal expressions: the language is itself the vehicle of thought." Often the distinction between conceptual relations and lexical relations are unimportant and authors use the term *lexical-semantic relations* (Evens, 1988, p. 2) to refer to relations between lexical concepts—concepts denoted by words. They are also sometimes called *sense relations*, as some linguists maintain that they relate particular senses of words (Lyons, 1977).

In addition to words, semantic relations can occur at higher levels of text—between phrases, clauses, sentences, and larger text segments, as well as between documents and sets of documents. The analysis of semantic relations can be carried out at the textual level, close to the words that express the meaning, or at a logical level, focusing on the meaning expressed by the text or concepts in the mind.

Let us now consider some properties of relations. Murphy (2003) listed the following general properties of lexical-semantic relations that have been identified by linguists:

1. Productivity—new relations can be created easily.
2. Binarity—some relations, for example *antonymy*, are binary in the sense that a word can have only one true antonym, whereas other relations, for example *synonymy*, can relate a set of words (i.e., a word can have many synonyms).
3. Variability—relations between words vary with the sense of the word used and the context.
4. Prototypicality and canonicity—some word pairs are better exemplars of a relation than others, and some word pairs have special status as canonical examples of a relation (particularly for *antonyms*).
5. Semisemanticity—nonsemantic properties, such as grammatical category, co-occurrence in text, and similarity in morphological form, can affect whether a particular relation is considered to hold between two words.
6. Uncountability—semantic relations are an open class and they cannot all be listed or counted.

7. Predictability—semantic relations follow certain general patterns and rules.
8. Universality—the same types of semantic relations are used in any language and the same concepts are related by the same semantic relations in different languages.

A semantic relation can have one or more of the following logical properties (Cruse, 2004; Sowa, 1984, p. 381):

- Reflexivity: A relation R is reflexive if it can relate an entity to itself; $[x] \rightarrow (R) \rightarrow [x]$ is true for every x (e.g., the part-whole relation).
- Symmetry: A relation R is symmetric if the two participants of the relation can occupy either slot; $[x] \rightarrow (R) \rightarrow [y]$ implies $[y] \rightarrow (R) \rightarrow [x]$ (e.g., synonymy).
- Transitivity: A relation R is transitive if $[x] \rightarrow (R) \rightarrow [y]$ and $[y] \rightarrow (R) \rightarrow [z]$ implies $[x] \rightarrow (R) \rightarrow [z]$ (e.g., IS-A relation, and ancestor-descendent relation).
- One-to-one relation: A relation R is one-to-one if, when one participant of the relation is known, the other participant is fixed; $[x] \rightarrow (R) \rightarrow [y]$ and $[z] \rightarrow (R) \rightarrow [y]$ implies $x = z$.

A relation can be related to another relation by *similarity* (i.e., the two relations are the same) or by an *inverse* relation. A relation R is the inverse of a relation S if both can accept the same pair of participants or slot fillers but the direction of the two relations is reversed; $[x] \rightarrow (R) \rightarrow [y]$ implies $[y] \rightarrow (S) \rightarrow [x]$ (e.g., *broader* versus *narrower* relation, *parent* versus *child* relation). One relation can be a *subrelation*, or more specific type of relation, of another, and relations can thus be organized into a relation hierarchy.

The variety of semantic relations and their properties plays an important role in human comprehension and reasoning. Spellman, Holyoak, and Morrison (2001) said that conceptual relations and the role bindings that they impose on the participant objects are central to such cognitive tasks as discourse comprehension, inference, problem solving, and analogical reasoning. Chaffin and Herrmann (1984) noted that the variety of relations is important both to general models of comprehension and to semantic models. For general models of comprehension, the relations differ in their logical properties and thus permit different kinds of inferences. The different relations also call into play different sets of decision criteria in decision making (Herrmann, Chaffin, Conti, Peters, & Robbins, 1979). Relations have also been found to be important in analogical reasoning and in the use of metaphors, which involve cross-domain mapping within a conceptual system (Lakoff, 1993, p. 203). In analogical reasoning, people map connected systems of relations, in particular cause-effect relations, rather than individual features (Gentner, 1983, 1989; Holyoak & Thagard, 1995; Lakoff, 1993; Turner, 1993).

Comprehensive treatments of semantic relations in language and text can be found in Cruse (1986, 2004), Lyons (1977, 1995), and Murphy (2003).

The Psychological Reality of Semantic Relations

Are semantic relations real, or are they just an abstract theoretical construct of linguists and psychologists? Do people really perceive, recognize, and process semantic relations? There is substantial evidence from experimental psychology that, to human beings, semantic relations are endowed with psychological reality.

Chaffin and Herrmann (1984, 1987, 1988b) and Glass, Holyoak, and Kiger (1979) carried out a series of studies to demonstrate that people can distinguish between different types of relations, identify instances of similar relations, express relations in words, recognize instances of relation ambiguity, and create new relations. The evidence comes from sorting experiments in which subjects were asked to sort relations (represented by pairs of terms) into groups of similar relations, analogy tests in which subjects were asked to assess the similarity of pairs of terms representing different relations, and tasks of relating term pairs to relation names indicating the type of relation exemplified by each term pair.

Psychologists have determined that some types of semantic relations, for example antonymy, are easier for adults and children to comprehend and process than others (Chaffin & Herrmann, 1987; Herrmann & Chaffin, 1986). Landis, Herrmann, and Chaffin (1987) studied children's developmental rates in understanding five types of semantic relations (antonymy, class inclusion, part-whole, syntactic relations, and synonymy), reaching the conclusion that the ability to match relations developed faster for antonymy and part-whole relations than for others and that comprehension of class inclusion developed least rapidly.

Researchers in anthropology and psychology have also found substantial cross-cultural agreement on the meanings, and in the use, of semantic relations (Chaffin & Herrmann, 1984; Hermann & Raybeck, 1981; Hudon, 2001; Romney, Moore, & Rusch, 1997). Raybeck and Herrmann (1990) found that some types of relations (particularly antonymy, part-whole, and cause-effect relations) are recognized equally easily and used with equal frequency and accuracy by diverse groups of people from different cultural backgrounds.

Psychologists consider semantic relations to be important in explaining the coherence and structure of concepts and categories. A *category* is not just a random set of entities—the entities in a category must belong together in some way. A category or concept is *coherent*—it must make meaningful sense. Psychologists have investigated several theoretical models for explaining conceptual coherence and structure. Initial studies focused on similarity of features, but this was found to be inadequate in explaining why certain features are more important than others in determining category membership. Researchers now believe that relations between the features of the category members, the functions of the features, and the configuration of features are important. For example, Markowitz (1988) has suggested that the *modification*, *part-whole*, *function*, *agent*, and *object* relations are important in determining category membership ranking. *Modification,* particularly *size,* is used in the definitions of most categories and many categories have a specific range of acceptable sizes. The *part-whole* relation is important in natural categories, whereas *function* is important in those pertaining to manufactured objects.

Some psychologists have espoused an explanation-based or theory-based model of categorization that accounts for conceptual coherence in terms of theories people have about the relations between attributes in a concept and about the relations between concepts (Ahn & Kim, 2001; Keil, 1989, 2003; Murphy & Medin, 1985). Wattenmaker, Nakamura, and Medin (1988) argued that categories derive their coherence not from overlapping attributes but from the complex web of causal and theoretical relationships in which these attributes participate. Ahn (1999) and Rehder (2003) found that causal relations appear to determine the importance of specific attributes in human evaluation of category membership. Rehder and Hastie (2001) showed that attributes occupying a central position in a network of causal relationships (either as a common cause or a common effect) dominate category membership judgment. Ahn and Kim (2001) found that the deeper an attribute is in a causal chain, the more dominant it is in category membership judgments.

Are semantic relations *concepts*? Chaffin and Herrmann (1988b) and Chaffin (1992) found that relations have the main characteristics of concepts and concluded that they are abstract concepts. They identified four characteristics that relational concepts share with concrete concepts: (a) relations can be analyzed into more basic elements or features; (b) a new relation may be an elaboration or combination of other relations; (c) relations have graded structure (i.e., some instances of relations, represented by word pairs, are more typical of a particular relation than others); and (d) relations vary in the ease with which they can be expressed.

Linguists and psychologists have shown that the antonym, synonym, IS-A, part-whole, and case relations, often taken as primitive relations, can be decomposed into simpler relational elements (Chaffin, 1992; Chaffin & Herrmann, 1987, 1988a, 1988b; Cruse, 1986; Klix, 1986; Lyons, 1977). Murphy (2003) stated that most lexical-semantic relations have some kind of similarity and contrast element. For example, synonyms are similar in meaning but different in lexical form, and antonyms have contrasting positions on the same dimension. Chaffin and Herrmann (1984) found that subjects distinguished relations in terms of three features: contrasting/noncontrasting, logical/pragmatic, and inclusion/noninclusion. Shared features can also account for perceptions of similarity between relations (Caplan & Herrmann, 1993; Chaffin, 1992; Chaffin & Herrmann, 1984, 1987, 1988a, 1988b).

Categories of relation instances (expressed as word pairs) also differ in the extent to which their memberships are graded (Caplan & Barr, 1991). Some relations can be defined "classically" in terms of necessary and sufficient features, whereas others have "fuzzy" boundaries with many partial members. Semantic relations, like concepts, can be organized into taxonomies with broader and narrower relations (Chaffin & Herrmann, 1987; Green, 2002; Stasio, Herrman, & Chaffin, 1985).

Semantic Relations in Semantic Memory

In addition to semantic relations expressed in text, semantic relations are also encoded in knowledge structures in our brains. Psychologists working in the area of *semantic memory* have attempted to characterize the nature of these knowledge structures and the semantic relations that support them. *Semantic memory* has been characterized as our mental storehouse of knowledge about language as well as general knowledge about the world (McNamara & Holbrook, 2003; Smith, 1978).

Semantic memory is usually modeled as a network, with nodes representing concepts and labeled directional links representing relations. This semantic network model was first proposed by Quillian (1967) and Collins and Quillian (1969). In Quillian's theory (1967, 1968), words are stored in memory as configurations of pointers to other words and each configuration of pointers represents the meaning of a word. The use of semantic memory for memory recall and comprehension is modeled as spreading activation—activation that spreads from one node to neighboring nodes along the links (Collins & Loftus, 1975). A major debate in semantic memory research is the structure vs. process question: Are semantic relations prestored in semantic memory or computed dynamically from the representation of concepts (Kounios & Holcomb, 1992)? Experimental evidence suggests that at least some relations, for example the *ownership* relation, are computed as needed (Kounios, Montgomery, & Smith, 1994).

Klix (1980, 1986) distinguished between *intraconcept relations* and *interconcept relations*. Interconcept relations, also called *event relations*, are based on associations between words, concepts, and events that have been observed and experienced (e.g., *knife* is for *cutting*), and are hypothesized as being stored directly in memory. Intraconcept relations, or feature-based relations between concepts, are based on common features or feature relationships within the concepts. These relations are not stored explicitly in memory but are hypothesized to be computed from concept features using cognitive procedures stored in the brain (Kukla, 1980).

These two types of relations have been found to have different effects on memory recall and analogy recognition (Hoffmann & Trettin, 1980). Murphy (2003) argued that paradigmatic relations (discussed later), which are mainly feature-based relations, are generated using cognitive rules because new instances of the relations can be easily produced at any time. She hypothesized that paradigmatic relations are represented as "metalinguistic knowledge" about words rather than hard-coded in the lexicon and that this explains why semantic relations are determined partly by context.

Herrmann (1987) suggested another possibility: A relation between two words may be represented in semantic memory as simpler relations or relation elements between aspects of the meanings of the two words. He further proposed an *alternative-form model* of relation comprehension in which different ways of representing relations in semantic memory are employed in relation comprehension, each form providing an alternative way of processing relations under different conditions.

General knowledge in human memory has also been modeled as being organized into structures of relations called a *schema* (Alba & Hasher, 1983). One implementation of the schema introduced by Minsky (1975) is a *frame*—basically a set of labeled slots, each indicating the role of a participant in the frame. Frames with a temporal element indicating a sequence of subevents in an event type are called *scripts* (Rumelhart & Ortony, 1977; Schank, 1982; Schank & Abelson, 1977). Frames, scripts, and story schemas play a major role in models of human comprehension (Brewer & Nakamura, 1984; Butcher & Kintsch, 2003; Whitney, Budd, Bramucci, & Crane, 1995).

Types of Semantic Relations

Overview

This section surveys the types of semantic relations that have been identified by researchers: lexical-semantic relations, case relations, and relations at a higher level of text.

Can a comprehensive list of semantic relations be constructed? What are the main types of relations? There are two broad approaches to constructing a list of semantic relations: the minimalist approach and the elaborate approach. Evens (1988) referred to the two groups of researchers as "lumpers" and "splitters." The lumpers, or minimalists, define a small number of general relations based on philosophical or logical principles (e.g., Sowa, 1984, 2000; Werner, 1988). Werner (1988) used only three relations: *modification*, *taxonomy*, and *queuing*. Other researchers have a much more elaborate list of specific relations, often based on lexical-semantic relations and words found in a text (e.g., Calzolari, 1988). Lexical-oriented models often group relations into families of relations with the same core meaning or function.

Most researchers recognize two broad categories of relations: paradigmatic and syntagmatic relations. This distinction can be traced to Ferdinand de Saussure (1959). [3] Paradigmatic relations are relations between pairs of words or phrases that can occur in the same position in the same sentence (Asher, 1994, v.10, p. 5153). The words often are instances of the same part of speech, belong to the same semantic class, and are to some extent grammatically substitutable. Examples include IS-A (broader-narrower), part-whole, and synonym relations. These relations tend to be part of our semantic memory, and are typically used in thesauri. Lancaster (1986) characterized paradigmatic relations as *a priori* or permanent relations.

Syntagmatic relations refer to relations between words that co-occur (often in close syntactic positions) in the same sentence or text (Asher, 1994, v. 10, p. 5178). It is a linear or sequence relation that is synthesized and expressed between two words or phrases when we construct a sentence. The relations are governed partly by the syntactic and grammatical rules of a language. Lancaster (1986) characterized syntagmatic relations as *a posteriori* or transient relations. Green (2001) suggested that paradigmatic relations are a closed, enumerable class of relations, whereas syntagmatic relations are an open class that cannot be fully enumerated, as a new relation is invented whenever a new verb is coined.

The distinction between paradigmatic and syntagmatic relations is fuzzy. Evens, Litowitz, Markowitz, Smith, and Werner (1980) pointed out that paradigmatic relations can be expressed syntagmatically. However, they also noted that

> we seem to receive paradigmatic information typically in generic (always true) sentences, while syntagmatic relationships come to us in occasional sentences. A generic or standing sentence contains a piece of permanent information about the world, such as 'Food is edible.' An occasional sentence contains information about a particular context. (Evens et al., 1980, pp. 10–11)

Syntagmatic relations between two words can become part of our semantic memory if the words co-occur frequently enough in text or discourse to be associated (Harris, 1987). Gardin (1965) argued that paradigmatic data should be derived from accumulated syntagmatic data. As we shall see later, researchers performing corpus-based linguistic analysis have found that paradigmatically related words, especially antonyms, do, indeed, often co-occur in text.

Many authors have attempted to enumerate semantic relations—generally, either of a particular type or for a particular purpose. Warren (1921) identified 13 other classification systems proposed before 1911. Evens et al. (1980) surveyed the sets of lexical-semantic relations that had been studied by researchers in anthropology, linguistics, psychology, and computer science before 1980. Lists of semantic relations can be found in Chaffin and Herrmann (1987, 1988b), Myaeng and McHale (1992), Neelameghan (1998, 2001), Neelameghan and Maitra (1978), Smith (1981), and Sowa (1984, 2000). Vickery (1996) has provided a summary history of associative relationships in information retrieval over the past few decades.

Lexical-Semantic Relations

Lexical-semantic relations are an important group of relations because they provide structure to lexicons, thesauri, taxonomies, and ontologies. The *structure* vs. *process* debate in semantic memory research is also present in lexical semantics. Are semantic relations stored in semantic memory as part of the meaning of a word, or are words defined in terms of their features and relations between words inferred dynamically from word meanings?

Lyons (1995) and other structural linguists hold that words cannot be defined independently of other words. A word's relationship with other words is a part of its meaning. The vocabulary of a language is thus viewed as a web of nodes, each representing a sense of a word and labeled links representing relations between the word senses. As Lyons (1977, pp. 231–232) put it,

> We cannot first identify the units [i.e., words] and then, at a subsequent stage of the analysis, enquire what combinatorial or other relations hold between them: we simultaneously identify both the units and their interrelations. Linguistic units are but points in a system, or network, of relations; they are the terminals of these relations, and they have no prior and independent existence.

Ferdinand de Saussure, generally regarded as the founder of modern structural linguistics, argued that "language is a system of interdependent terms in which the value of each term results solely from the simultaneous presence of the others" (Saussure, 1959, pp. 114–116).

Other linguists maintain that the lexical representation of a word is mainly a set of semantic features based on semantic primitives and that semantic relations are derivable from the semantic features of the words through the use of some basic relational rules (Clark, 1973; Katz, 1972; Murphy, 2003).

The main lexical-semantic relations are the paradigmatic relations of hyponymy (IS-A or broader-narrower term), part-whole relation, synonymy, and antonymy, which are discussed later. However, frequently occurring syntagmatic relations between a pair of words can be part of our linguistic knowledge and considered lexical-semantic relations. As Firth (1957, p. 195; 1968, p. 179) put it, "you shall know a word by the company it keeps." Pairs of words that co-occur in a sentence more often than pure chance would allow are referred to, broadly, as *collocations* (Smadja, 1993), although some writers prefer narrower definitions of the term.

There are different degrees of syntagmatic word association. At the extreme are idioms (e.g., "kick the bucket") whose meanings cannot be derived from the meanings of the component words. Other word sequences are less strongly associated—their meaning is related to the meaning of the component words but not completely derivable from them. Hausmann (1985) divided word associations into fixed (i.e., idiom) and nonfixed combinations, subdividing the latter into counter-affine, affine, and free combinations.

Some word pairs are so strongly associated that the presence of one word almost determines that the other word will also appear in a particular context. Mel'cuk (1988) introduced the idea of *lexical functions* (LFs) in the framework of his Meaning-Text Theory. Wanner (1996) has referred to lexical functions as "institutionalized" lexical relations. A lexical function is a mapping or relation between two terms—term1 and term2—denoted "LF(term1) = term2" for a particular meaning context. So, if a term, *term2*, is to be selected to express a particular meaning or relation, the choice of term2 is predetermined if term1 is given. An example is LF("aircraft") = "crew." The value of a lexical function can also be a set of words, e.g., LF("flock") = {"birds", "sheep"}. Institutionalized lexical relations are directed and asymmetrical, as well as language-specific. For example, LF("aircraft") = "crew" does not imply LF("crew") = "aircraft."

There are many LF relations. Mel'cuk (1996) listed 27 paradigmatic and 37 syntagmatic lexical functions. Examples of paradigmatic lexical functions are: *Syn* (synonym), *Anti* (antonym), *Conv* (converse), *Contr* (contrastive), and *Gener* (genus). Syntagmatic lexical functions include:

- Center/culmination: Centr("crisis") = "the peak" [of the crisis]
- Very/intensely: Magn("naked") = "stark"
- More: Plus("prices") = {"soar", "skyrocket"}
- Less: Minus("pressure") = "decreases"

A good introduction to lexical functions and Meaning-Text Theory is given by Wanner (1996).

The most extensive lexical-semantic network that has been constructed for the English language is WordNet (http://www.cogsci.princeton.edu/~wn) (Fellbaum, 1998; Miller, 1995; Miller & Fellbaum, 1991). WordNet is a lexical database comprising about 150,000 English nouns, verbs, adjectives, and adverbs, organized into sets of synonymous words called *synsets*, each of which represents a lexical concept. Its design is based on psycholinguistic theories of human lexical memory. Its construction has given insight into how the lexicon is structured by lexical-semantic relations. For example, nouns are structured mainly by IS-A and part-whole relations; nouns are linked to adjectives with the *attribute* link and to verbs with the *function* link; adjectives are linked primarily by *antonymy*; the most frequent relation among verbs is *troponymy*, which expresses a *manner* elaboration. Other relations among verbs encoded in WordNet are lexical entailment (e.g., *snoring* entails *sleeping*), causal relation (e.g., show/see, feed/eat, have/own), and antonymy.

Following the success of WordNet, EuroWordNet (http://www.illc.uva.nl/EuroWordNet)—a multilingual lexical database covering several European languages—was constructed (Alonge, Calzolari, Vossen, Bloksma, Castellon, Marti, et al., 1998; Vossen, 1998). EuroWordNet is patterned after WordNet but uses a richer set of lexical-semantic relations. For example, causal relations are divided into *non-factive causal relations* (i.e., one event is likely to cause another event but not necessarily so, e.g., *search* → *find*) and *factive causal relations* (the causal relation necessarily holds, e.g., *kill* → *die*). A causal relation can also be labeled with the property of *intention* to cause the result (e.g., *search* → *find*) in order to distinguish it from inadvertent causal relations. Near-synonymy, near-antonymy, and five types of part-whole relations are also used in EuroWordNet. Furthermore, sets of relations can be labeled with the properties of *conjunction* and *disjunction* to indicate relationships among sets of concepts: for example, an airplane is typically composed of several parts—wings, nose, tail AND door (conjunction), but may have only one of several possible means of propulsive power—propellers OR jet engines (disjunction). WordNets for other languages are being constructed, and these projects are listed on the Global WordNet Association Web site (http://www.globalwordnet.org).

Case Relations

Case relations, also called case roles, thematic relations, and theta roles, are the primary syntagmatic relations between the main verb in a clause and the other syntactic constituents of the clause (Fillmore, 1968; Somers, 1987). According to case grammar theory, verbs assign semantic roles to the various clause constituents—subject, direct object, indirect object, prepositional phrase, and so forth—which are sometimes termed the arguments of verbs. For example, in the sentence "Mary bought a watch for John," the case relations between the verb *buy* and the other clause constituents are:

buy → (agent) → [Mary]

→ (patient) → [watch]

→ (recipient) → [John]

Each verb sense is associated with a case frame with slots, each slot having a case role. A case frame specifies the number of entities the verb expects in the clause, the case roles assigned to these entities, whether each role is obligatory (i.e., must be filled) or optional, selectional restrictions specifying the semantic category of the entity filling a role, and the syntactic realization of each role in the clause (whether expressed as subject, direct object, etc.).

Somers (1987, p. 111) said that "a recurring problem for Case grammarians has always been the definition of a 'comfortable' set of cases." Rosner and Somers (1980) stressed that a case system should be tailored to the particular application. The rationale for using case roles is to classify and generalize the semantic roles between a verb and its arguments, and so the set of case roles should be at a level of abstraction that is appropriate for the application.

Fillmore (1971) produced a "case hierarchy" with eight roles: *agent, experiencer, instrument, object, source, goal, location,* and *time*. Cook's (1989) case frame matrix had five case roles: *agent, experiencer, benefactive, object,* and *locative*. He also listed additional "modal cases": *time, manner, instrument, cause, result*, and *purpose*. Somers's (1987) case grid defined 24 case roles using a combination of two dimensions: a spatial/temporal-orientation dimension comprising the values *source, path, goal,* and *neutral* and a second, mostly verb-type dimension with the values *active, objective, dative psychological / possessive, locative, temporal,* and *ambient*. A case role is thus considered a bundle of more primitive features. The *experiencer* role, for example, is represented as a combination of *dative psychological + goal* features. Sets of case roles have been constructed by many authors. Longacre (1996) presented 10 case roles. Myaeng, Khoo, and Li (1994) identified 46 case roles in the process of constructing case frames for all the verb senses in the *Longman Dictionary of Contemporary English* (1987). Various case grammar systems have been reviewed by Cook (1989) and Somers (1987).

Dowty (1991), however, has argued that case roles are not discrete roles but cluster concepts that have fuzzy boundaries. An individual verb-specific semantic role can belong to a case role to a greater or lesser extent. A case role is thus seen as a category or type of semantic role, which includes a cluster of more specific roles with overlapping sets of features. Each semantic role can be decomposed into features that Dowty called verbal entailments. He proposed two large clusters of case roles called *proto-agent* and *proto-patient* roles. Examples of entailments for the proto-agent role include *volitional involvement, perception, causing an event or change of state*, and *movement relative to the position of another participant*.

Case grammar theory can be extended to other parts of speech, such as nouns and adjectives. Verb case frames are applicable to nominalized verbs and gerunds, formed by adding one of several possible suffixes, such as -ing and -ion, to verbs. Case frames for these nouns can be derived from the case frames of their associated verbs, although the process is not straightforward. It has been suggested that some adjectives and nouns also have valency, inasmuch as they expect certain prepositional phrases and certain kinds of complements (e.g., Somers, 1987).

Constructing case frames for a comprehensive set of verbs is a difficult task. Automatic construction methods using text mining and corpus statistics are described later in the chapter. A major manual effort to construct a comprehensive set of case frames for English verb senses as well as predicative nouns and adjectives is being undertaken in the Berkeley FrameNet project (http://www.icsi.berkeley.edu/~framenet) (Baker, Fillmore, & Cronin, 2003; Baker, Fillmore, & Lowe, 1998). The project does not define a small number of case roles to use in all the case frames. Instead, a set of case roles called "frame elements" is defined for each "frame." A frame in the FrameNet project is a schematic representation of a particular type of situation involving various participants. Example frames are *action*, *awareness*, and *transaction* frames. To construct case frames for individual word senses, the words are clustered into groups corresponding to situations or frames and the case roles for each word sense are selected from the frame elements defined for the situation.

Many natural language processing applications make use of case frames because they correspond quite closely to the surface structure of clauses and it is thus relatively easy to label clause constituents with case roles by means of a computer program. This serves as a useful intermediate processing step when converting the text to a semantic representation. Indeed, instantiated case frames with slots filled by terms/concepts extracted from the text are often used as the intermediate representation or interlingua in natural language understanding systems (e.g., Chan & Franklin, 2003; Minker, Bennacef, & Gauvain, 1996), question answering and dialogue systems (e.g., Takemura & Ashida, 2002; Xu, Araki, & Niimi, 2003), and machine translation systems (e.g., Dorr, Levow, & Lin, 2002).

Relations Between Larger Text Segments

We turn now to semantic relations between larger units of text. Relations between sentences can be analyzed from a logical or textual perspective. Logical relations between sentences are dealt with in the fields of formal semantics (e.g., Cann, 1993), logic and philosophy (e.g., Quine, 1982), and knowledge representation (e.g., Ringland & Duce, 1988; Sowa, 1984, 2000). Often, sentences and clauses are represented as propositions or predicates, and inferencing is performed using propositional, predicate, and other kinds of logics. The main semantic relations used are entailment (or implication or consequence), presupposition, equivalence, and contradiction (Cann, 1993; Lyons, 1995; Van Dijk, 1972). The most important relation is entailment. When we say that *a sentence* S *entails a sentence* S', we mean that if *S* is true then *S'* is true. Van Dijk (1972) presented other semantic relations: time, place, cause, purpose, result, condition, concession, and topic (theme)-comment (rheme). Crombie's (1985) semantic relations between propositions were grouped under the headings *temporal, matching, cause-effect, truth and validity, alternation, bonding, paraphrase, amplification,* and *setting/conduct.* Other lists of propositional relations can be found in Beekman, Callow, and Kopesec (1981), Hobbs (1985), and Longacre (1996).

At the textual level, sentences and clauses are linked by relations of cohesion and coherence. Halliday and Hasan (1976) analyzed relations between adjacent sentences and clauses, which they termed *cohesive relations*. They emphasized that cohesion is a semantic relation and that "cohesion occurs where the interpretation of some element in the discourse is dependent on that of another" (Halliday & Hassan, 1976, p. 4). Their work focused on the linguistic devices that writers use to effect "cohesive ties" between two proximate items, usually words and phrases, in the text. They divided cohesive devices into grammatical devices (anaphoric reference, substitution, ellipsis, and conjunction) and lexical devices (use of vocabulary and repetition of words).

Cohesion is often contrasted with coherence relations. Dooley and Levinsohn (2001, p. 27) characterized text coherence as "in essence, a question of whether the hearer can make it 'hang together' conceptually, that is, interpret it within a single mental representation." Eggins (1994, p. 87) said that coherence refers to the way a group of clauses or sentences relates to the context. Cohesion emphasizes local relations between two adjacent text units, whereas coherence focuses on networks of related units and larger structures as well as on the argumentative and pragmatic purposes of the text unit.

At an even higher level of text are discourse relations and macrostructure. Van Dijk (1988) argued that syntax and semantics can be applied to sequences of clauses, sentences, or whole texts. An influential discourse structure model in information science derives from the *Rhetorical Structure Theory* of Mann and Thompson (1988, 1989; see also Mann, Matthiessen, & Thompson, 1992), which uses a set of rhetorical relations to model text structure. Rhetorical relations include *evidence, elaboration, motivation, volitional cause, evaluation,* and *background*. Each relation links two text segments, one of which is considered the nuclear, or more central, segment and the other, the satellite or peripheral segment. A few relations, for example *sequence* and *contrast*, are "multinuclear" in that the linked text segments are both considered nuclear. The rhetorical structure is recursive—a text is decomposed into a sequence of segments linked by rhetorical relations and each segment can be further decomposed into smaller segments linked by the same or other rhetorical relations.

Van Dijk (1980) maintained that a text has an overall macrolevel syntactic structure, called a *superstructure*, that is governed by a rule-based schema. He suggested the following hierarchical schema for news articles:

Situation

- Episode (subdivided into Main Events and Consequences)
- Background
    - Context (circumstances, previous events)
    - History

Comments

- Verbal reactions
- Conclusions (subdivided into Expectations and Evaluations)

Though Van Dijk regarded these as syntactic units, the unit labels suggest semantic roles. The segments can perhaps be considered to have a semantic relation to the overall content of the text. In fact, Van Dijk (1988) postulated the existence of summarizing macrorules, which relate lower-level propositions to higher-level macropropositions—topics or themes derived from the meanings of a text. A more recent discussion of the discourse structure of news articles can be found in Bell (1998).

The macrolevel structures of stories, called *story schemas* and *story grammars*, have been studied by several authors (e.g., Mandler, 1987; Mandler & Johnson, 1977; Rumelhart, 1975; Schneider & Winship, 2002), and are used in the teaching of reading comprehension, literary analysis, and story-writing in schools (see Dimino, Taylor, & Gersten, 1995; Olson & Gee, 1988). A recent review of the theory can be found in Lang (2003).

At the document level, relations between documents may be structural (e.g., an article in a journal, a chapter in a book) or associative (e.g., articles by the same author, cited articles, hyperlinked Web pages). The documents can be linked by various kinds of semantic relations—two articles may be on the same topic, one article may be a condensed version of another, an article could report a follow-up study or refute the results of another study, and so forth. Topical semantic relations can be indicated using controlled subject terms taken from a thesaurus or subject headings list, or class numbers taken from a classification scheme. Another type of document-level semantic relation can be derived from an author's citation of other works and his or her reasons for citing them. Liu (1993) reviewed many citation studies and compiled a list of possible reasons for citing another work. Green (2001) noted that little is known about the range of semantic relations between citing and cited documents. Relatively little work has been done on identifying semantic relations between documents. The main semantic relations at the document level appear to be those provided by thesauri and classification schemes.

Finally, an important type of semantic relation in information science is the relevance relation—the relevance of a document to a query or to the information need of a user. Researchers have identified many factors, in addition to topical relevance, that affect a user's judgment of the relevance of a document (Barry, 1994; Park, 1997; Schamber, 1991, 1994; Tang & Solomon, 2001). Green (2001) suggested that there may be several types of semantic relations underlying these factors, which have not been studied in depth. Green and Bean (1995) and Bean and Green (2001) have explored some of the relations underlying topical relevance.

Selected Semantic Relations

This section takes a close look at five well-known paradigmatic relations used in thesauri and ontologies, as well as the cause-effect relation, which is an important syntagmatic relation in human knowledge structures. These relations are often treated as unitary primitive relations. We show that they are complex relations that can be divided into subtypes with different properties.

Hyponym-Hyperonym Relation

The hyponymy relation has been referred to in the literature under various names, including IS-A (is-a), a-kind-of, taxonymic, superordinate-subordinate, genus-species, and class-subclass relations. *Hyponym* refers to the narrower term/concept (e.g., Alsatian) and *hyperonym* to the broader term/concept (dog). The relation implies class inclusion, e.g., all instances of Alsatians are dogs, the set of Alsatian instances is a subset of dogs, and the meaning of "Alsatian" is included in the meaning of "dog" (Cruse, 2000). Different logical definitions of the hyponymy relation are presented and discussed by Cruse (2000). Related to hyponym is the *co-hyponym* or *coordinate*—another hyponym of the same hyperonym, such as siblings with the same parent.

Lyons (1968, p. 453) called the hyponymy relation the most fundamental paradigmatic relation of sense in terms of which the vocabulary is structured. Together with the *part-whole* relation, it is a hierarchical relation often found in thesauri, taxonomies, and ontologies. Cruse (2002) asserted that of all the sense relations, it occurs across the widest range of grammatical categories and domains.

There is some question whether the hyponymy relation relates word senses, lexemes (root words), or concepts. Most linguists take the hyponymy relation to relate word senses. Cruse (2002) has argued that, in some cases, even senses can be subdivided into "facets" (e.g., the physical *book* versus the abstract text of a *book*), and that sense relations relate facets. However, the form of a word has been found to affect human judgment of relations. For example, Cruse found that people considered *cat* to be a better hyponym of *animal* than *pussy*, suggesting that people are influenced by word forms.

The hyponymy relation exhibits different linguistic behavior when expressed by means of different terms. Cruse (2002) pointed out that the expression "An X is *a kind / type of* Y" is more discriminating than "an X *is a* Y." Cruse (1986) called the first relation *taxonymy* and the second relation *simple hyponymy*. He claimed that taxonymy is not just a logical class inclusion relation—the terms used to represent the classes are important. He gave the following examples of logical hyponymy relations that do not sound correct when expressed as "a kind of":[4]

> ?A stallion/mare/foal is a kind/type of horse.
>
> A stallion is a horse.
>
> ?A blonde/queen/actress is a kind of woman.
>
> An actress is a woman.

The expression "a kind/type of" exerts selectional restrictions on the pair of terms. Cruse has suggested the existence of a "principle of taxonomic subdivision" that selects only good categories that are internally cohesive, externally distinctive, and maximally informative. Good taxonyms tend to be natural kinds that cannot be defined in terms of a few necessary and sufficient features. Cruse (1986) suggested that single-feature category division may be the reason that *stallion*, *kitten*, and *blonde* are not satisfactory taxonyms of *horse*, *cat*, and *woman*. Another possible reason is that a term may "highlight" a particular semantic feature. The word *prostitute* highlights the *sexual activity* feature so that "A prostitute is a kind of sex-worker" is better than "A prostitute is a kind of woman." The hyponymy relation is generally taken to be a transitive relation. However, Cruse (2004, p. 152) cited the following example where transitivity breaks down:

> A car seat is a type of seat.
>
> A seat is a type of furniture.
>
> * A car seat is a type of furniture.

Fellbaum (2002) suggested that the hyponymy relation works best between closely related terms, and less well between terms far apart in the hierarchy.

Troponymy Relation

Troponymy refers to broader-narrower relations between verbs. Fellbaum (2002) pointed out that the expressions "a kind of" and "IS-A" sound odd when applied to verbs, for example "(To) yodel is a kind of (to) sing" and "To murmur is to talk." She said that the main relation between verb senses is the *manner* relation, which Fellbaum and Miller

(1991) termed "troponymy." For example, the *Longman Dictionary of Contemporary English* (1995) defines *run* and *fly* as to move in some manner (*to move quickly on foot* in the case of *run*, and *to move through the air* for *fly*). The manner relation involves several dimensions. Motion verbs differ along the dimension of speed (e.g., *walk* vs. *run*) or the means of transportation. Verbs of impact (e.g., *hit*) vary along the dimension of degree of force (e.g., *chop* and *slam*). In addition to the manner relation, troponyms include the *function* and *result* relations.

Fellbaum and Chaffin (1990) determined in a psychological study that people were able to recognize and process troponymy relations: Subjects had no trouble labeling verb pairs with the type of troponymy relation, sorting verbs into related pairs, responding with related verbs in an association task, and accomplishing an analogy task. Finally, Fellbaum (2002) has observed that verb hierarchies are flatter and more "bushy" than noun hierarchies, since they generally do not exceed three or four levels.

Meronym-Holonym Relation

The meronymy relation, also known as the part-whole relation and partonymy, refers to the relation between a concept/entity and its constituent parts. The distinction between meronymy and hyponymy relations is clear for concrete concepts but fuzzy for abstract concepts. Hyponymy relations can be said to exist within concepts, but meronymy relations are between concepts. Pribbenow (2002) has pointed out that both are logically asymmetric and transitive relations. Hyponyms inherit features from hyperonyms, but parts do not inherit features from wholes, although there is upward inheritance for some attributes, such as color, material, and function (Tversky, 1990).

Lyons (1977, v. 1, p. 313) demonstrated that the part-whole relation is intransitive at the level of linguistic expression:

> The door has a handle.
>
> The house has a door.
>
> ? The house has a handle.

Cruse (1979) attempted to resolve the problem by characterizing the functional context of the relation. He claimed that when we say X is a (functional) component of Y, we usually mean that X is a major component of Y.

Iris, Litowitz, and Evens (1988) found that the part-whole relation is really a family of relations, divided into four main types:

- Functional component of a whole (e.g., wheel of a bicycle)

- The segmented whole (the whole divided into pieces like a pie)
- Members of a collection of elements
- Subsets of sets (set inclusion, e.g., fruits and apples)

Winston, Chaffin, and Herrmann (1987) identified six types of part-whole relations, including the following three additional types: stuff-object (steel-car), feature-activity (paying-shopping), and place-area.

Gerstl and Pribbenow (1995) divided part-whole relations broadly into those relating to the natural structure of the whole (e.g., functional components of an object) and partitions of the whole by construction (i.e., artificial partitions based on attributes, e.g., dividing objects by color). These were further divided into subtypes.

Within the Meaning-Text Theory, Wanner (1996) listed the following meronymic relations:

- LF Mult (member-collection), e.g., Mult("dog")="pack"
- LF Equip (social whole-staff), e.g., Equip("aircraft")="crew"
- LF Cap (organization and its head), e.g., Cap("ship")="captain"
- LF Sing (a whole and its uniform unit), e.g., Sing("sand")="grain"
- LF Centr (a whole and its center or culmination), e.g., Centr("mountain")="peak" [of the mountain].

Other classifications of the part-whole relation have been developed by Barriere (1997, 2002); Markowitz, Nutter, and Evens (1992); and Sattler (1995) specifically for an engineering application; Uschold (1996) for ecological information systems; and Bernauer (1996) for the medical domain.

Synonymy

Lyons (1995) has noted that absolute synonymy is very rare. Two expressions are absolutely synonymous if all their meanings are identical in all linguistic contexts. Synonymy can be analyzed from a logical point of view or from the linguistic expression point of view. Logically synonymous terms have been called *logical synonyms* (Murphy, 2003) and *propositional synonyms* (Cruse, 2004).

Common types of synonyms are *sense-synonyms* (which share one or more senses), *near-synonyms* (which have no identical senses but are close in meaning), and *partial synonyms* (which share some senses but differ in some aspect, e.g., in the way they are used or in some dimension of meaning) (Cruse, 1986; Lyons, 1995). Sense-synonyms that share at least one sense and match in every other property for that sense are *complete synonyms* (Lyons, 1981). Church, Gale, Hanks, Hindle, and Moon (1994) discussed *gradient synonyms*—sets of synonyms in which one core term is considered prototypical and the other synonyms differ from the prototype in various ways, often giving additional information.

Synonyms are usually treated as reflexive, symmetrical, and transitive, although Murphy (2003) has argued that they are not always so.

Antonymy

Antonymy, or opposition, is one of the best-studied relations and the one that people find easiest to learn and process (Jones, 2002). Cruse (1986, p. 197) called it the most readily apprehended of sense relations, hedged with magical properties in the eyes of many people:

> Indeed, there is a widespread idea that the power of uniting or reconciling opposites is a magical one, an attribute of the Deity, or a property of states of mind brought about by profound meditation, and so on. ... Philosophers and others from Heraclitus to Jung have noted the tendency of things to slip into their opposite states; and many have remarked on the thin dividing line between love and hate, genius and madness, etc.

Evens et al. (1980) observed that antonymy is irreflexive, symmetric, and intransitive. Of the many different types of antonymy, *canonical antonymy* is the best studied. Canonical antonyms constitute a special class of opposites that are stable and enjoy wide cultural currency. For example, *hot/cold* is a better example of antonymy than *steamy/frigid*, even though both pairs indicate opposite ends of the temperature scale (Murphy, 2003). Such antonym pairs, for example *big/small, good/bad, good/evil*, are automatically recalled by subjects in free word association tasks and are taught to children (Murphy, 2003, p. 10).

Justeson and Katz (1991, 1992) and Jones (2002) found that antonymous adjectives tend to co-occur within the same sentences in texts, often linked by the conjunctions *and* and *or*, as in, for example, the phrases "rich or poor" and "large and small." They also often substitute for each other in parallel, essentially identical, phrases such as "am I right or am I wrong?" and "new lamps for old ones." Justeson and Katz (1992, p. 181) concluded that "the patterns [of phrasal substitution] are so pervasive, that there is simply no chance for a genuine antonym pair to fail to show up in them, at a reasonable rate. So those that do not, cannot be antonymic." They suggested that the frequent co-occurrence of antonyms in text and discourse reinforces people's knowledge of antonymous pairs, which partly explains how antonymous pairs are learned and why antonym relations are graded.

Frequently co-occurring antonymous words are more likely to be judged as good antonyms than less frequently co-occurring antonyms.

Many types of antonymy have been identified (Cruse, 1986; Lehrer & Lehrer, 1982; Lyons, 1977; Murphy, 2003; Ogden, 1967). Jones (2002) examined how antonyms are used in a newspaper corpus and identified several antonym classes based on their linguistic behavior.

Cause-Effect Relation

The concept of causation is complex and surprisingly difficult to define. Since Aristotle, philosophers have grappled with the concept (Ehring, 1997; Mellor, 1995; Owens, 1992; Sosa & Tooley, 1993). Reviews of the concept from a philosophical and psychological perspective can be found in Khoo, Chan, and Niu (2002) and Khoo (1995).

One can distinguish between *necessary* and *sufficient* causes. An event A is a *sufficient*, although not a *necessary*, condition for event B if, when A occurs, B always follows, but when A does not occur, B sometimes occurs and sometimes not. A is a *necessary*, although not a *sufficient*, condition for B if, when A does not occur, B never occurs, but when A occurs, B sometimes occurs and sometimes not. An often cited definition of causation is Mackie's (1980) INUS condition, which defined a cause as an *Insufficient* but *Necessary* part of an *Unnecessary* but *Sufficient* condition for an event. Psychologists Jaspars, Hewstone, and Fincham (1983) and Jaspars (1983) found evidence that whether a cause is a necessary and/or sufficient condition varies with the type of entity being considered for causal status. Cause is likely to be attributed to a person if the person is a sufficient condition, whereas cause is likely to be attributed to the circumstances or situation if the situation is a necessary condition. Cause is ascribed to a stimulus when it is both a necessary and a sufficient condition. So, "a personal cause is seen more as a sufficient condition, whereas situational causes are conceived primarily as necessary conditions" (Jaspars et al., 1983, pp. 16–17).

However, Mackie (1980) pointed out that our concept of causation also includes some presumption of continuity from the cause to the effect, a causal mechanism by which the cause generates the effect. The concept of probabilistic causation has also gained popularity (Eells, 1991; Salmon, 1984). This view recognizes the possibility of indeterministic causation—instances where the causal mechanism is inherently probabilistic, as in the field of quantum mechanics.

Aristotle (1996) identified four kinds of cause: material cause (i.e., the material of an object causes its existence), formal cause (i.e., the form or structure of an object causes its existence), efficient cause (i.e., an entity acting mechanically upon an object causes it to change, move, or come to rest), and final or teleological cause (the intended future effect is the ultimate cause of a present action undertaken to bring about that future event).

Recently, Barriere (1997, 2002) has presented a classification of general cause-effect relations:

- Existence dependency
  - Creation
  - Prevention
  - Destruction
  - Maintenance
- Influence dependency
  - Preservation
  - Modification
    - Increase
    - Decrease

Taking a somewhat different approach, Terenziani and Torasso (1995) have categorized cause and effect on the basis of temporal considerations. Their analysis yielded five categories:

- *One-shot causation*: The presence of the cause is required only momentarily to allow the action to begin.
- *Continuous causation*: The continued presence of the cause is required to sustain the effect.
- *Mutually sustaining causation*: Each bit of cause causes a slightly later bit of the effect.
- *Culminated event causation*: The effect comes about only by achieving the culmination of the causal event (e.g., "run a mile in less than 4 minutes" causes "receive a prize").
- *Causal connection with a threshold*: There is a delay between the beginning of the cause and the beginning of the effect, and the effect is triggered only when some kind of threshold is reached.

Warren, Nicholas, and Trabasso (1979) identified four types of cause-effect relations in narrative texts: motivation, psychological causation, physical causation, and enablement. Dick (1997), in attempting to model the causal situation in a legal case, distinguished between the following types of cause and effect: distant vs. direct cause, animate vs. inanimate agent, animate agent vs. instrument, volitive vs. nonvolitive cause, active vs. passive cause, central vs. peripheral (or abstract) cause, explicit vs. implicit cause, and aims vs. actual effect.

Khoo (1995) analyzed the verb entries in the *Longman Dictionary of Contemporary English* (1987) and came up with a total of 2,082 causative verbs (verbs with a causal component in their meaning), which he grouped into 47 types of effects. Levin (1993) provided a systematic and extensive classification of verbs based on their syntactic behavior. Many of the verb classes were found to have a causal component in their meaning.

Semantic Relations in Knowledge Structures

Semantic Relations in Thesauri

A thesaurus is a set of terms structured using a small set of semantic relations to indicate the controlled (or preferred) term for each concept and relationships between the terms/concepts. It is designed to support consistent subject indexing of documents and effective information retrieval. The relations between terms help both indexers and searchers to navigate thesauri in order to identify various kinds of related terms.

The American National Standards Institute/National Information Standards Organization (ANSI/NISO) Z39.15-1993 standard "Guidelines for the Construction, Format, and Management of Monolingual Thesauri" (National Information Standards Organization, 1994) and the International Organization for Standardization (ISO) 2788 standard "Guidelines for the Establishment and Development of Monolingual Thesauri" (International Organization for Standardization, 1986) recognize three types of semantic relations: equivalence (*use* and *use for*), hierarchical (*broader term* and *narrower term*), and associative (*related term*).

The ANSI/NISO standard lists seven types of synonym relations: terms of different linguistic origins, popular term-scientific name, generic noun-trade name, variant names, current-outdated term, common nouns-slang/jargon, and dialectical variants. It also describes other kinds of equivalence relation: lexical variants and quasi-synonyms. Hierarchical relations include generic (IS-A), part-whole, and instantiation relations. Part-whole relations include organs of the body; geographic locations; subject disciplines; and hierarchical organizational, corporate, social, or political structures. Nine types of associative relations are also identified.

Associative relations in thesauri have been analyzed by several authors. Aitchison, Gilchrist, and Bawden (1997) listed 14 categories. Lancaster (1986) and Raitt (1980) each listed 10 categories. In an analysis of hierarchical relations in Medical Subject Headings (MeSH), Bean (1998) identified 67 types of relations other than the generic and instantiation relations. Most of the relations could be considered associative. Aitchison et al. (1997) cited a 1965 study by Perreault, who found 120 types of relations. Discussion of thesaural relations in general can be found in Aitchison et al. (1997), Clarke (2001), and Milstead (2001).

Semantic Relations in Indexing Languages

In the subject indexing of a document, controlled terms from a thesaurus, subject headings list, or classification scheme are assigned to the document to reflect the main subjects and concepts in the content of the document. The index terms are generally not precoordinated, in other words, the index terms are assigned as separate terms and there is no indication whether two or more concepts are related in a particular way in the document. For example, if a document is assigned the terms *information retrieval*, *user interface,* and *evaluation*, there is no indication whether *evaluation* is related to *information retrieval* or to *user interface*. During retrieval, the user may specify the Boolean query "information retrieval AND evaluation," which requires the system to search for the two terms separately and to combine the two sets of documents retrieved to identify documents containing both terms. There is no assurance that the documents retrieved will discuss *evaluation of information retrieval*—only that these two concepts occur in the same document. Such an indexing approach is called *postcoordinate* indexing.

In some indexing languages, the index terms are *precoordinated*, in other words, the human indexer indicates an association between two or more concepts in the document by using the syntax of the language and placing the terms in a particular order. However, the type of association is not specified explicitly but is implied by the context. Such is the case with the Library of Congress Subject Headings (LCSH) and faceted classification schemes like Ranganathan's Colon Classification (Kishore, 1986; Ranganathan, 1965). Precoordinate indexing allows the user to search for some kind of association between two or more index terms.

Farradane (1967) advocated the use of explicitly specified relations in indexing systems. He pointed out that implied relations in precoordinate indexing are unambiguous only in a narrow domain. More recently, Green (1995a, 1995b) has called for the inclusion of syntagmatic relations in indexing languages, examined the issues involved, and suggested a frame-based representation of syntagmatic relations.

Two indexing systems that make explicit use of relations are Farradane's (1950, 1952, 1967) relational classification system and the SYNTOL model (Gardin, 1965; Levy, 1967). Farradane's system had nine types of relations: concurrence, self-activity, association, equivalence, dimensional (time, space, state), appurtenance, distinctness, reaction, and functional dependence (causation). The SYNTOL project used four main types of relations: coordinative, consecutive, predicative, and associative. The associative relation was subdivided into finer relations. There is no experimental evidence yet that the use of explicitly specified relations in indexing yields better retrieval results compared to postcoordinate indexing or precoordinate indexing with implied relations.

Semantic Relations in Ontologies

A thesaurus lists the main concepts/terms in a particular domain and specifies relations between the concepts/terms using only a small number of relation types. This small set of relations may be adequate for information retrieval applications because the focus of a thesaurus is on indexing and searching, but it is not sufficient for more complex or intelligent applications that require knowledge-based inferencing and a detailed representation of domain knowledge.

A more detailed representation of domain knowledge is called an ontology. Many definitions of ontology from different perspectives have been put forward. The following definition by Berners-Lee, Hendler, and Lassila (2001, p. 40) alludes to some of the different aspects of ontology:

> In philosophy, an ontology is a theory about the nature of existence, of what types of things exist; ontology as a discipline studies such theories. Artificial-intelligence and Web researchers have co-opted the term for their own jargon, and for them an ontology is a document or file that formally defines the relations among terms. The most typical kind of ontology for the Web has a taxonomy and a set of inference rules.

Ontology, with an uppercase "O," refers to a branch of philosophy dealing with the nature of being or existence—what categories of things exist and what their features are (Guarino & Giaretta, 1995; Sowa, 2000). This is often contrasted with *Epistemology,* which deals with the nature and sources of knowledge. Ontology with a lowercase "o" can refer to the conceptual framework or knowledge of a particular domain shared by a group of people—for instance, something that exists in people's minds. Or it can refer to the symbolic representation of this conceptual frame, perhaps in the form of a "logical theory" that can be used by a computer program (Guarino & Giaretta, 1995).

An often quoted definition is that of Gruber (1993, p. 199): "An ontology is an explicit specification of a conceptualization." In practice, an ontology is expressed as a taxonomy of concepts linked by IS-A, part-whole, and attribute-value relations, sometimes enriched by other kinds of relations as well as additional rules or constraints called *axioms*. One major difference between an ontology and a thesaurus is the richer set of relations used in an ontology. Guarino and Giaretta (1995), Guarino (1997), and Gómez-Pérez, Fernandez-Lopez, and Corcho (2004) analyzed various definitions of ontology. A collection of definitions can be found at http://www.aaai.org/AITopics/html/ontol.html.

Ontologies come in many types and flavors, depending on the domain, application, representation scheme used, philosophical principles adopted by the authors, and the construction method and tools used. Those functioning as online search aids are more lexically oriented and may not contain nontaxonomic relations or axioms, whereas others supporting inferencing may be formally represented in a logic representation and have many axioms. Gómez-Pérez et al. (2004) outlined the different typologies of ontologies that have been proposed and suggested that even thesauri can be considered light-weight ontologies.

There is growing interest in ontologies because of their potential for encoding knowledge in a way that allows computer programs and agent software to perform intelligent tasks on the Web.

> Ontologies provide support in integrating heterogeneous and distributed information sources. This gives them an important role in areas such as knowledge management and electronic commerce. ... Ontologies enable machine-understandable semantics of data, and building this data infrastructure will enable completely new kinds of automated services. (Fensel, 2001, p. 8)

The OWL Web Ontology Language Use Cases and Requirements (World Wide Web Consortium, 2004a) lists the following areas where ontologies are expected to be useful: Web portals, multimedia collections, corporate Web site management, design documentation, agents and services, and ubiquitous computing.

Ontologies are seen as the backbone of the Semantic Web. The Semantic Web was characterized by Berners-Lee et al. (2001, p. 37) as "an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation." The World Wide Web Consortium (2004c) views the Semantic Web as providing "a common framework that allows data to be shared and reused across application, enterprise, and community boundaries." A fundamental technology for realizing the Semantic Web is the Web service. Web services are self-contained computer programs that can be accessed on the Internet by other computer programs through public interfaces and bindings that are defined by means of Extensible Markup Language (XML) (World Wide Web Consortium, 2004d). Because the interface definition of a Web service can be discovered by other computer programs, this allows computer programs to dynamically locate and interact with one another in an automated and unattended way. To help people and software agents locate appropriate information, objects, and Web services on the Internet, ontologies are needed.

The World Wide Web Consortium (2004b) has developed the Resource Description Framework (RDF) and the Web Ontology Language (OWL) for encoding ontologies with XML. OWL can be used to specify types of relations between concept instances, called *properties*. The following relations between user-defined *properties* can be specified: *equivalentProperty* (synonymous property), *inverseOf* (e.g., "hasChild" is the inverse of "hasParent"), and *subPropertyOf* (a kind of hyponymy relation). The user-defined *properties* can also be labeled with the following attributes: *TransitiveProperty*, *SymmetricProperty, Functional Property* (i.e., each instance has no more than one value for this property), and *InverseFunctionalProperty*.

Well-known ontologies include:

- CYC (http://www.cyc.com), which contains about 40,000 concepts and 300,000 axioms (inter-concept relations and constraints) and is built for commonsense reasoning.

- Suggested Upper Merged Ontology (SUMO) (http://suo.ieee.org, http://ontology.teknowledge.com, http://www.ontol ogyportal.org), a standard upper ontology developed by the Institute of Electrical and Electronic Engineers (IEEE) Standard Upper Ontology Working Group. SUMO and its several domain ontologies contain altogether about 20,000 terms and 60,000 axioms.

- Unified Medical Language System (UMLS) (http://www. nlm.nih.gov/research/umls), which contains 135 semantic types, 54 semantic relations, and about 250,000 concepts.

- MIKROKOSMOS (http://crl.nmsu.edu/users/sb/papers/thesis/node26.html), which contains about 4,800 concepts and is built to support machine translation.

- Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) (http://www.loa-cnr.it/DOLCE.html), which aims at capturing ontological categories underlying natural language and human common sense and has been developed to serve as a starting point for comparing and analyzing other ontologies.

- WordNet (http://www.cogsci.princeton.edu/~wn), a lexical database, often considered a lexical or terminological ontology, containing approximately 150,000 English nouns, verbs, adjectives, and adverbs that are grouped into 115,000 synonym sets (synsets), each of which represents an underlying lexical concept.

- The Enterprise Ontology (http://www.aiai.ed.ac.uk/project/enterprise), a collection of terms and definitions relevant to business enterprises to assist in the acquisition, representation, and manipulation of enterprise knowledge.

- Toronto Virtual Enterprise (TOVE) (http://www.eil.utoronto.ca/enterprise-modelling/tove/index.html), used to model the structure, activities, processes, information, resources, people, behavior, goals, and constraints of an enterprise.

Ontologies vary widely in the number and types of relations used. They can range from a simple taxonomy structured by IS-A and part-whole relations, to a small number of relations as in WordNet, to thousands of relations in CYC (Lenat, Miller, & Yokoi, 1995). Relations in ontologies are often structured in a relation hierarchy or grouped into major categories. Many relation hierarchies have been proposed in the literature. For example, Sowa (2000, 2001) divided his role concepts into two groups: roles pertaining to the *PrehendingEntity* (the subject of the relation, e.g., *whole*) and the *PrehendedEntity* (the object of the relation, e.g., *part*). *PrehendedEntity* is subdivided into *Correlative* and *Component*, the latter further subdivided as follows:

- Component
  - Part
    - Piece
    - Participant
    - Stage
  - Property
    - Attribute
    - Manner

A role concept is converted to a relation by combining the concept with the *has* relation. For example, the *part* role concept can be converted to the *has-part* relation.

A more extensive relation hierarchy is presented in the Generalized Upper Model (GUM) ontology (Bateman, Fabris, & Magnini, 1995). At the top level, relations are arranged into four categories: *participant*, *circumstance*, *process*, and *logical-relation*. The CGKAT system (Martin, 1995, 1996) has a default hierarchy of about 200 relations. Relations are organized into nine classes at the top-level: *attributive_relation*, *component_relation*, *constraint_or_measure_relation*, *relation_from_a_situation*, *relation_to_a_situation, relation_from_a_proposition, relation_referring_to_a_process, relation_with_a_special_property*, and *relation_used_by_an_agent*. The Unified Medical Language System (UMLS) relation hierarchy (U.S. National Library of Medicine, 2004) contains 54 relations grouped broadly into *IS-A* and *associated_with* relation types, the latter being subdivided into *physically_related_to, spatially_related_to, functionally_related_to, temporally_related_to*, and *conceptually_related_to*. Markowitz et al. (1992) presented a hierarchy of lexical relations containing nearly 100 relations as leaf nodes.

Some researchers have developed methods to formally represent relations in a knowledge representation scheme for use in data modeling and knowledge-based inferencing. This usually involves explicitly representing the attributes of semantic relations, modeling the hierarchical relationships between semantic relations, and defining axioms or rules for reasoning with the relations. Priss (1999) developed a mathematical formalism for representing a network of semantic relations in a lattice structure by analyzing the relations using *formal concept analysis* (Ganter & Wille, 1997) and identifying relational components (see the chapter by Priss in the present volume). Wille (2003) also showed how commonsense logical relations between concepts can be represented using a concept lattice. Methods for representing and reasoning with semantic relations have been developed in the *conceptual graph* formalism (Sowa, 1984, 2000) as well as in description logics (Baader, Calvanese, McGuinness, Nardi, & Patel-Schneider, 2003)—a family of knowledge representation languages that focuses on expressing knowledge about concepts, concept hierarchies, roles and instances, and reasoning about them. A collection of papers describing various formalisms for modeling concepts and relations can be found in Lehmann (1992). Several authors have examined the issues involved in formalizing the part-whole relation in data modeling and inferencing systems (e.g., Artale, Franconi, Guarino, & Pazzi, 1996; Lambrix, 2000; Lee, Chan, & Yeung, 1995). Some issues involved in organizing semantic relations in a knowledge base were examined by Stephens and Chen (1996).

To our knowledge, no systematic analysis of the types of semantic relations used in ontologies has been reported in the literature. Such an analysis should be carried out in the context of the domain and application for which the ontology was constructed. Little is known about what constitutes an appropriate set of semantic relations for a domain or application, or the most effective way to structure the relations into a relation hierarchy. Although much has been written about the potential uses of ontologies and methods for their construction, and although small case studies of applications have been reported, there has not been any systematic evaluation of the effectiveness of ontologies or the various types of semantic relations occurring in real applications.

One possible exception is the part-whole relation. Researchers have analyzed the different types of part-whole relations used to model data and objects for various purposes (e.g., Artale et al., 1996; Gerstl &

Pribbenow, 1995). Many ontologies specify a few types of part-whole relations. Nevertheless, in a review of 10 well-known ontologies, Noy and Hafner (1997) found that different ontologies represented the part-whole relation in dissimilar ways and that, in many cases, the ontologies did not deal adequately with the distinctions between different types of part-whole relations.

Automatic Identification of Semantic Relations

Overview

Automatic identification and extraction of semantic relations in text is a difficult task. The accuracy rate varies widely and depends on many factors: the type of semantic relation to be identified, the domain or subject area, the type of text/documents being processed, the amount of training text available, whether knowledge-based inferencing is used, and the accuracy of the syntactic preprocessing of the text. Furthermore, because there are different types of semantic relations at different text levels, no system can identify semantic relations accurately at all levels. This is a major barrier to widespread use of semantic relations in information science.

In this section, we consider the automatic identification and extraction of semantic relations between words/phrases and the concepts that they represent. Identification of higher-level relations such as cohesion relations (including anaphor and co-reference resolution), rhetorical relations, and text macrostructure is important, but the literature is too broad to cover in this survey. A general introduction to information extraction technology is given by Appelt and Israel (1999). The three major applications of automatic identification of relations in text are in information extraction, ontology construction/knowledge acquisition, and information retrieval. This section examines the main techniques used to extract relations in information extraction and ontology construction. Information retrieval applications are discussed later in the chapter.

In information extraction applications, concepts and relations are extracted from text to fill predefined templates representing various kinds of information about an event (e.g., terrorist attack or corporate merger), entity (e.g., company), or process. The slots in a template are labeled and can be considered roles related to the event/entity/process. In the 1980s and early 1990s, artificial intelligence researchers used sophisticated natural language processing and knowledge-based inferencing to extract concepts and relations from text and represent them in a semantic representation or knowledge representation scheme (e.g., Berrut, 1990; Mauldin, 1991; Rau, 1987; Rau, Jacobs, & Zernik, 1989). Unfortunately, such complex systems could be built only for narrow domains. In the 1990s, it was found that simple methods of relation extraction using shallow text processing and pattern matching utilizing many simple patterns were equally effective. However, constructing a good set of extraction patterns for an application still involves considerable manual effort. Current research is focused on automatic pattern construction, which requires a large training set of documents and manually filled templates representing the associated answer key. For information extraction technology to become widely used, automatic pattern construction techniques that are effective with small training sets need to be developed, together with good interfaces that help the end-user to construct the training examples and to guide the process of pattern construction.

Whereas information extraction applications seek to extract every instance of concepts and relations relevant to the domain or application, automatic ontology construction focuses on well-established knowledge, namely, concepts and relations that occur with some frequency in the text collection. Hence, corpus statistics techniques incorporating co-occurrence statistics, machine learning, and data mining can be employed together with pattern matching techniques to extract frequently occurring concept-relation-concept triples from a corpus. These triples can then be used to build a knowledge base of facts or connected together to form a semantic network or an ontology.

Automatic Identification of Semantic Relations Using Pattern Matching

Automatic identification of semantic relations in text involves looking for certain linguistic patterns in the text that indicate the presence of a particular relation. For example, a simple linear pattern for identifying some *cause-effect* information is:

[cause] *is a cause of* [effect]

The tokens in square brackets represent slots to be filled by words/phrases in the text. The slots indicate which part of the sentence represents the *cause* and which part represents the *effect* in the *cause-effect* relation. The following sentence contains a match for this pattern:

<u>Smoking</u> *is a cause of* <u>lung cancer</u>

An extraction pattern is thus a sequence of tokens, each token representing a literal word to be matched in the text, a wildcard that can match any word, or a slot to be filled. The following selectional restrictions can be specified for each token: the syntactic category (e.g., part of speech), type of phrase, syntactic role (e.g., subject, direct object, etc.), and the voice of the verb. Semantic restrictions can also be specified using concept categories from an ontology or type of entity (e.g., organization name, person name, date, amount of money). Pattern-matching is performed to identify the segments of the text that match each pattern.

A major component of any information extraction system is its set of extraction patterns. Construction of patterns can be done manually or automatically by analyzing sample relevant texts and the associated answer keys indicating the information to be extracted. The answer keys are typically constructed by trained human analysts. Pattern construction thus entails constructing patterns that will extract the same information from the text as the human analysts did. The patterns should not be too general lest information be extracted from non-relevant text fragments or incorrect information from relevant text fragments.

Two approaches can be used in the pattern construction: a top-down approach, in which general patterns are first constructed and then gradually specialized to reduce errors, or a bottom-up approach, in which specific patterns are first constructed and then gradually combined to reduce the number of patterns or generalized to cover more situations. Before pattern construction and pattern matching, the text is usually subjected to some amount of preprocessing, which can include tokenizing, stemming or conversion of words to their base forms, syntactic tagging (to identify the part of speech), chunking (to identify particular types of phrases), and semantic tagging (to identify the semantic class, e.g., *inanimate object* and *organization name*, to which the word/phrase belongs). Some information extraction systems make use of a thesaurus or ontology to infer the semantic classes of text tokens and to generalize two or more concepts to a single broader concept.

Automatic Construction of Extraction Patterns

Because manual construction of good extraction patterns is a difficult and time-consuming task, there is a need for automatic or machine-aided pattern construction. Researchers have developed effective techniques for automatic pattern construction. To perform automatic pattern construction, the system needs well-defined heuristics for constructing the initial patterns, for generalizing and specializing the patterns based on positive and negative examples, for selecting which generalization/specialization methods to use in which situation, and for deciding on the order in which the methods are tried. Typically, a variation of the inductive learning algorithm described by Mitchell (1997) is used for pattern learning.

Our survey focuses on information extraction from free text, rather than from structured or semi-structured documents, as our interest is in semantic relations expressed in free text. The learning of patterns for extracting information from structured documents, such as Web pages, is called wrapper induction and it relies on structure identification using HTML tags (Muslea, 1999). An example is Information Extraction based on Pattern Discovery (IEPAD) (Chang & Lui, 2001), a wrapper induction system that generates extraction patterns for Web documents without needing user-labeled examples.

Some well-known systems that learn extraction patterns from free text are AutoSlog (Riloff, 1993), PALKA (Kim, 1996; Kim & Moldovan, 1995), CRYSTAL (Soderland, 1997), WHISK (Soderland, 1999), and

Table 5.1 Pattern templates and examples of instantiated patterns in AutoSlog

Pattern template	Example of pattern constructed
<subj:slot> <passive-verb>	[victim] was murdered
<subj:slot > <active-verb>	[perpetrator] bombed
<subj:slot > <verb> <infinitive-phrase>	[perpetrator] attempted to kill
<subj:slot > <auxiliary-verb> <noun>	[victim] was victim
<passive-verb> <direct-obj:slot>	killed [victim]
<active-verb> <direct-obj:slot>	bombed [target]

RAPIER (Califf & Mooney, 2003). The patterns constructed by these systems generally perform sentence-level extraction, leaving co-reference resolution and merging of extracted information across sentences to later modules, such as discourse parsing modules (Soderland, 1999). A survey of the various types of extraction patterns generated by machine learning algorithms was carried out by Muslea (1999).

AutoSlog (Riloff, 1993) is the earliest system developed to learn text extraction patterns from training examples. It uses partial case frames as linear patterns. Each pattern has only one slot and usually includes a verb and a noun phrase (a subject or direct object). A set of pattern templates define the linear patterns that the system will construct. Each pattern is thus an instantiation of a pattern template. Table 5.1 lists the pattern templates used in AutoSlog and an illustration for each template. The pattern template "<passive-verb> <direct-obj:slot>" was included because a sentence analyzer called CIRCUS (Lehnert, 1991) occasionally confused active and passive constructions.

Before pattern construction, the training corpus is preprocessed by CIRCUS to identify clause boundaries and the major syntactic constituents: subject, verb, direct object, noun phrases, and prepositional phrases. Relevant text segments that contain the semantic relations of interest are identified and answer keys are constructed to indicate which noun phrase should be extracted and what its semantic role is. If the domain of interest is terrorist activities, the semantic roles would include *perpetrators*, *targets*, *victims*, and so forth.

During pattern construction, pattern matching is used to match the pattern templates with the training text segments. If a pattern template matches a relevant text segment, then a pattern is constructed by replacing the tokens in the template with the words in the text. If a token in the template indicates a slot, this token is allowed to match a noun phrase in the text only if the noun phrase appears in the answer key (i.e., a human analyst has indicated that this is the information to be extracted). A slot token is placed in the pattern being constructed, and the semantic role for the slot is taken from the answer key. The constructed pattern is thus an instantiation of the pattern template. Finally, a human analyst inspects each pattern and decides which ones should be accepted or rejected. AutoSlog-TS (Riloff, 1996), an extension of AutoSlog, creates dictionaries of extraction patterns using only untagged text. A user needs to provide training texts (relevant and irrelevant texts) and to filter and label the resulting extraction patterns. Generally, extraction patterns occurring in irrelevant texts are filtered out. The accuracy rates come close to those of AutoSlog, in which tagged text is used. However, AutoSlog cannot learn rules that extract values for multiple slots (such as *[victim] was killed by [attacker]*) and does not adjust the patterns by generalizing or specializing them once they are constructed.

In the PALKA system (Kim, 1996; Kim & Moldovan, 1995), the patterns involve the whole clause. Sentences in the training text are first converted to simple clauses. The clauses containing a semantic relation of interest are processed one at a time. If the set of patterns already constructed does not match a clause, then a new pattern is constructed for the clause. This initial pattern covers the main verb, the subject, the object, and the words to be extracted (i.e., the slot). Each of these constituents in the clause is represented by a token in the pattern. Each token is assigned a semantic category from a conceptual hierarchy. Generalizations and specializations are applied only to the semantic constraints. When two similar patterns sharing the same target slots and literal strings are generated, their semantic constraints are generalized by locating a broader concept or ancestor in the conceptual hierarchy that is common to both semantic categories.

The CRYSTAL system (Soderland, Fisher, Aseltine, & Lehnert, 1996) uses a similar approach but is more complex. CRYSTAL learns rules that can extract values for multiple slots. Initially, CRYSTAL constructs a very specific pattern for every sentence in the training text. The sentences are not simplified into simple clauses. The constraints in the initial patterns are gradually relaxed to increase their coverage and to merge similar patterns. CRYSTAL identifies possible generalizations by locating pairs of highly similar patterns. This similarity is measured by counting the number of relaxations required to unify the two patterns. A new pattern is created when constraints are relaxed just enough to merge the two patterns by dropping constraints that the two do not share and finding a common ancestor for their semantic constraints. The new pattern is tested against the training corpus to make sure it does not extract information not specified in the answer keys. If the new pattern is valid, all the patterns subsumed by the new pattern are deleted. This generalization continues until a pattern that exceeds a specified error threshold is generated.

WHISK (Soderland, 1999) induces rules in a top-down manner, first finding the most general rule that covers the seed (i.e., hand-tagged training examples), then constraining the rule by adding terms one at a time. The learned rules are in the form of regular expressions that can extract either single or multiple slots.

RAPIER (Califf & Mooney, 2003) is a bottom-up learning algorithm that incorporates techniques from several inductive logic programming systems (Lavrac & Dzeroski, 1994). Its algorithm starts with initial specific rules created from the input corpus and then incrementally replaces the rules with more general rules using automatic rule evaluation. The rule learning is done separately for each slot; thus, RAPIER cannot learn rules that extract values for multiple slots.

SRV (Freitag, 2000) employs a top-down rule learner that uses a covering algorithm. As each rule is learnt, all positive examples covered by the new rule are removed from consideration for the creation of future rules. Rule learning ends when all positive examples have been covered. SRV utilizes the length of a fragment, the location of a particular token, the relative locations of two tokens, and various user-defined token features, such as capitalization, digits, and word length. SNoW-IE (Roth & Yih, 2001) learns extraction patterns by means of propositional learning mechanisms. Ciravegna (2001) developed a pattern learner that makes use of rule induction and generalization.

Text Mining for Semantic Relations

Text mining for semantic relations is concerned with the extraction of new and implicit relationships between different concept entities from large collections of textual data. Although some semantic relations are clearly expressed through the use of well-defined syntactic structures, other semantic relations are not, and only a multistep sequence of reasoning based on semantic analysis of the text collection can extract them. Most semantic extraction systems take advantage of an existing domain knowledge source (i.e., semantic information) and make use of cue words and syntactic tags provided by a syntactic parser.

Various approaches to automatic semantic extraction from corpus documents have been developed. Girju, Badulescu, and Moldovan (2003) worked on the discovery of semantic relations, especially part-whole relations, from text. They used rich syntactic and semantic features to discover useful and implicit relations from text. The C4.5 decision tree learning algorithm (Quinlan, 1993) was used to learn semantic constraints so as to detect part-whole relations, while WordNet served as the domain knowledge base to identify and disambiguate target concepts (i.e., part and whole components). Girju (2002) also investigated the extraction of causal relations in her dissertation work. The Artequakt system (Alani, Kim, Millard, Weal, Hall, Lewis, et al., 2003) automatically extracts knowledge about artists from the Web, populates a knowledge base, and uses it to generate personalized biographies. Artequakt links a knowledge extraction tool with an ontology to identify entity relationships using ontology relation declarations, such as "[Person] – *place of birth* – [Place]," where [Person] and [Place] are concepts and *"place of birth"* is a semantic relation between them. Dyvik (2004) investigated a method for deriving semantic relations in WordNet from data extracted from the English-Norwegian Parallel Corpus (Johansson, 1997), which comprises around 2.6 million words. The method was based on the hypotheses that semantically closely related words have strongly overlapping sets of translations, and words with a wide range of meanings have a higher number of translations than words with few meanings. The implementation took words with their sets of translations from the corpus as input and returned thesaurus-like entries containing senses, synonyms, hyperonyms, and hyponyms. Calzolari and Picchi (1989; see also Calzolari, 1992) looked into the acquisition of semantic information from machine-readable dictionaries, in which semantic information is implicitly contained. They aimed at reorganizing free-text definitions in natural language form into informationally equivalent structured forms in a lexical knowledge base.

In the medical area, semantic tagging that uses domain knowledge is important for effective text mining. Many studies make use of the Unified Medical Language System (UMLS) (Humphreys, Lindberg, Schoolman, & Barnett, 1998; U.S. National Library of Medicine, 2004) as the domain knowledge base. Blake and Pratt (2001) mined for semantic relationships between medical concepts from medical texts. They mapped the terms in the texts to concepts in UMLS in order to reduce the number of features for data mining. Blake and Pratt focused on *Breast Cancer Treatment* using association rule mining (Borgelt & Kruse, 2002) to find associated concept pairs like *magnesium–migraines*. They were mainly interested in mining the *existence* of relationships between medical concepts (i.e., finding associated concept pairs in breast cancer treatment), not in identifying the specific semantic relations for the associated concept pairs. Lee, Na, and Khoo (2003) carried out a small experiment using a sample of medical abstracts from MEDLINE, a biomedical bibliographic database maintained by the U.S. National Library of Medicine, to identify concept pairs related to *Colon Cancer Treatment*. The semantic relations between the concepts in each pair were then inferred using the UMLS semantic network. They were able to infer semantic relations between concepts automatically from the UMLS semantic network 68 percent of the time, although the method could not distinguish between a few possible relation types.

The Semantic Knowledge Representation (SKR) project at the National Library of Medicine has developed programs that extract usable semantic information from biomedical texts (Rindflesch & Aronson, 2002). Two programs, MetaMap (Aronson, 2001) and SemRep (Rindflesch, Jayant, & Lawrence, 2000), play a major role in semantic information extraction. MetaMap maps noun phrases in free text to concepts in the UMLS Metathesaurus, while SemRep uses the Semantic Network in UMLS to infer possible relationships between those concepts. Consider the input phrase "ablation of pituitary gland." SemRep looks up a semantic rule (i.e., extraction pattern), which declares that the preposition *of* matches the Semantic Network relation *location_of*, and also notes that one of the relationships in the Semantic Network with this predicate is "[Body Part, Organ, or Organ Component] – LOCATION_OF – [Therapeutic or Preventive Procedure]." The

Metathesaurus concept for *ablation* is *Excision, NOS,* found by MetaMap. The semantic type for this concept is *Therapeutic or Preventive Procedure*; the type for *Pituitary Gland* is *Body Part, Organ, or Organ Component.* Because these semantic types match those found in the relationship indicated by the preposition *of* (*location_of*), "Pituitary Gland – *location_of* – Excision, NOS" is extracted as a new semantic relation.

Srinivasan and Rindflesch (2002) have used SemRep in combination with MeSH index terms to find potentially interesting semantic relationships in large sets of MEDLINE abstracts. Rindflesch et al. (2000) built ARBITER (Assess and Retrieve BInding TErminology), which uses UMLS as domain knowledge and relies on syntactic cues (such as the single verb *bind*) provided by a syntactic parser, to identify and extract molecular binding semantic relations from MEDLINE records. Rindflesch, Libbus, Hristovski, Aronson, and Kilicoglu (2003) also built a natural language processing program, called SemGen, to identify and extract causal relations between genetic phenomena and diseases from MEDLINE records. They were able to achieve 76 percent precision with sample sentences.

Automatic Construction of Case Frames

Text mining using co-occurrence statistics is employed in the automatic construction of case frames. The process has three main stages:

- Constructing "subcategorization frames" (Chomsky, 1965) (i.e., identifying the combination of syntactic constituents or arguments that the verb expects).
- Identifying the selectional restriction for each syntactic constituent or slot (e.g., which semantic class of nouns can be the direct object of the verb).
- Assigning a case role to each syntactic constituent or slot in the case frame.

Typically, statistical collocations are mined from the text collection as a first step to finding the words/phrases and types of words/phrases that tend to co-occur with each verb. Some syntactic preprocessing—part-of-speech tagging, chunking to identify types of phrases, or syntactic parsing—is first performed. Associations between verbs and types of co-occurring syntactic constituents can be used to build subcategorization frames (Basili, Pazienza, & Vindigni, 1997; Brent, 1993; Manning, 1993; Nedellec, 2000). The head nouns of the constituent phrases can be generalized to a semantic class so as to identify the selectional restriction for a slot. This semantic generalization is performed with the aid of a thesaurus or ontology (Framis, 1994; Li & Abe, 1998). If a thesaurus is not available, nouns in the text collection can be clustered according to the context in which they tend to appear. For example, clusters of nouns that tend to co-occur as direct objects of the same verbs can be identified. The noun clusters can be accepted as semantic classes, or a similarity measure between the nouns can be used to generalize the selectional restrictions (Grishman & Sterling, 1994).

Automatic assignment of case role labels to case frame slots is more difficult. To some extent it can be determined by examining the semantic classes of nouns filling the roles. Verbs in the text collection can also be clustered to identify sets of verbs that tend to co-occur with the same nouns. This can help to identify clusters of verbs with similar semantics, an aid to identifying the semantic roles assigned by the verbs (Pereira, Tishby, & Lee, 1993). A more promising approach is to use a machine-learning technique to learn the characteristics of verb-noun combinations for each case role. New verb-noun combinations can be assigned a case role label based on their similarity to prototypical verb-noun combinations for each case role. Wanner (2004) used this approach to extract verb-noun collocations from text and categorize them into one or more of 20 lexical functions. A centroid was computed for each lexical function through the use of training verb-noun examples for each lexical function and the utilization of concept classes in EuroWordNet as features.

Finally, dictionary definitions have also been mined to construct case frames (e.g., Calzolari, 1992).

Semantic Relations in Information Retrieval

Overview

To date, research and development in information retrieval have focused on term and concept matching. Some researchers have, however, explored the possibility of using semantic relations to enhance recall and precision. Recall enhancement—increasing the number of relevant documents retrieved—is usually accomplished through query expansion, in other words, adding alternative terms to the query. Typically, paradigmatic relations, especially synonyms and partial synonyms, are used for query expansion, although syntagmatic relations can be used as well. Terms that are semantically related to each query term are added to the search query using the Boolean disjunction operator *OR*.

Precision enhancement—reducing the proportion of nonrelevant documents retrieved—is accomplished through relation matching. This involves specifying additional relational criteria for retrieval, in other words, the documents retrieved must contain not only the terms/concepts specified in the query but must also express the same relations between the concepts as expressed in the query. The relations are in a sense added to the search by means of the Boolean conjunction operator *AND*. Typically, syntagmatic relations are used in relation matching.

A more precise form of information retrieval is question-answering—answering a user's question with facts or text passages extracted from documents. This requires identifying specific semantic relations between document concepts and concepts in the user's question. The appropriate semantic relation to be used for identifying potential answers in documents is determined by the question type (e.g., definition question, list question, and so forth).

Automatic text summarization extracts the most important information from a document or set of documents, then generates an abridged version for a particular user or task (Mani & Maybury, 1999). This helps users skim through a set of retrieved documents to determine their relevance and potential usefulness. Semantic relations are useful for identifying within a document related concepts and statements that can be compressed, as well as for analyzing the document's discourse structure, which can then be used to identify its central concepts. Multidocument summarization can provide an overview of a set of documents, pointing out information that is common to the document set, information unique to each document, and contradictory statements found in the set. Semantic relations between concepts and statements across the documents (cross-document discourse structure) are useful for multidocument summarization.

In this section, we survey research applying semantic relations to query expansion, precision enhancement, question-answering, and automatic text summarization.

Semantic Relations in Query Expansion

Query expansion with related terms is important for improving recall, although it can improve information retrieval precision as well (Wang, Vandendorpe, & Evens, 1985). Related terms can be taken from a knowledge structure, such as a thesaurus, a taxonomy, a semantic network or an ontology, or from a more informal term association list. As explained earlier, knowledge structures such as thesauri and ontologies distinguish between a few types of semantic relations: minimally, the synonymy relation, the hierarchical relations (IS-A and part-whole), and the associative relation (related term). Such knowledge structures are usually manually constructed, although some are constructed semi-automatically. On the other hand, informal term association lists are often constructed using corpus analysis and co-occurrence statistics. (Two terms are associated if they co-occur in the same document or in close proximity in text more often than can be attributed to pure chance alone.) A commonly used term association measure is the mutual information measure (Church & Hanks, 1989).

Query expansion can be performed either automatically without user intervention or manually by a user selecting appropriate related terms from a thesaurus. The usefulness of query expansion depends on many factors: the size and type of the document collection, whether the searching is performed "free text" or on an indexing field using controlled vocabulary, whether the thesaurus is domain-specific or generic, whether the system is a Boolean or best-match search system, and so forth. Most of the large-scale studies have been conducted on TREC (Text REtrieval Conference) corpora (http://trec.nist.gov), using free-text best-match systems and automatic query expansion (see the chapter on TREC in the present volume). However, manual query expansion on a Boolean search system, featuring controlled vocabulary searches conducted with terms drawn from a domain-specific thesaurus, has been performed by generations of librarians and there is perhaps less doubt as to its usefulness.

Query Expansion Using Term Association

Automatic query expansion using term associations derived from a corpus using co-occurrence statistics has not produced promising results. Sparck Jones (1971) even obtained a decrease in retrieval performance. Peat and Willett (1991) demonstrated that the effectiveness of term association is limited because the similar terms identified by co-occurrence data tend to occur very frequently in the database, and frequently occurring terms are poor at discriminating between relevant and nonrelevant documents.

Some researchers managed to obtain positive results with variations of the standard term association method. Qiu and Frei (1993) obtained positive results with their concept-based query expansion method, in which the query is expanded with terms that are strongly related to *all* the query terms. They suggested that the usual term association methods fail because these tend to add terms that are strongly related only to *individual* query terms.

Chen and Lynch (1992) developed a different association measure and "cluster algorithm" for constructing term association lists. Their work was not strictly on automatic query expansion because their term association file was used to display related terms for the user to select. However, they showed that a word co-occurrence algorithm can produce terms that are semantically related. Ruge (1992) introduced a term association method that made use of head/modifier relations (a kind of syntactic relation). She combined linguistic knowledge and co-occurrence in her experiments to produce linguistically based thesaurus relations.

Grefenstette (1992) and Strzalkowski (1995) made use of second-order term association, specifically, they regarded two terms as related if they each tended to co-occur with a third term and bore the same syntactic relation to it. Grefenstette obtained a small improvement in retrieval effectiveness on a collection of medical abstracts.

Information retrieval researchers participating in the TREC series of conferences have carried out large-scale experiments investigating the usefulness of query expansion for full-text searching in large heterogeneous document collections using state-of-the-art, best-match information retrieval systems. From the TREC experiments, researchers have learned that the most effective method of query expansion using associated terms is pseudo-relevance feedback (also called blind, or local, feedback). This involves using the original query to retrieve an initial ranked list of documents. The terms in the top-ranked documents are weighted in some way and added to the original query, and the retrieval process is repeated with this expanded query (Belkin, Head, Jeng, Kelly, Lin, Park, et al., 1999; Buckley, Singhal, Mitra, & Salton, 1996; Hawking, Thistlewaite, & Craswell, 1998; Kwok & Chan, 1998; Xu & Croft, 1996). In this way, the terms added to the query are related to the query as a whole and not just to individual terms within the query.

More recent work has focused on selecting which documents and words to use. Usually, only the most frequently occurring words are used (Buckley et al., 1996). From the top-ranked documents, Buckley, Mitra, Walz, and Cardie (1998) identified clusters of documents corresponding to different query concepts, selected high frequency words from each cluster, and weighted them appropriately. Xu and Croft (1996) retrieved a ranked list of passages instead of whole documents to make pseudo-relevance feedback more precise. In a later study, Xu and Croft (2000) used an additional criterion: The terms selected from the top-ranked passages should co-occur with query terms in those passages. Terms that co-occur with more query terms are preferred. We hypothesize that even better results can be obtained by considering the semantic relations between the associated terms in these top-ranked documents/passages and the query terms found in the documents.

Query Expansion Using Lexical-Semantic Relations

Lexical-semantic relations can be used to distinguish between different kinds of term associations to use for query expansion. Some researchers have investigated what types of semantic relations are useful for query expansion.

Fox (1980) used 73 classes of lexical relations for query expansion. The lexically related words for each query term were identified manually. Some of the relations (e.g., those between *dog* and *bark,* and *lion* and *Africa*) were syntagmatic and associative relations. Using the SMART best-match retrieval system, Fox found that the most effective retrieval was accomplished by using all categories of relations except the antonym relation. In a follow-up study, Wang et al. (1985) used 44 relations, a different weighting scheme, and a different document collection, as well as constructing a relational thesaurus—something not explicitly done by Fox. The results were comparable to Fox's (1980), indicating that the synonym relation and the broader-narrower term relation are not the only relations that can be employed for query expansion. However, these studies involved only very small document collections using single-domain thesauri.

Using the MEDLINE database and MeSH, Rada and Bicknell (1989) found that automatic query expansion using broader-narrower term relations as well as non-hierarchical relations could improve retrieval effectiveness, if the semantic relations were selected carefully. In another study using the Excerpta Medica database and an enriched

EMTREE thesaurus, Rada, Barlow, Potharst, Zanstra, and Bijstra (1991) found that only when the query explicitly mentioned a particular non-hierarchical relation could the retrieval system make use of the specific relation in the thesaurus to improve document ranking.

Wan, Evens, Wan, and Pao (1997) used a relational thesaurus for automatic indexing in a Chinese information retrieval system. They reported that their relational thesaurus, which employed 11 types of semantic relations, did improve retrieval effectiveness in terms of average precision with both manual and automatic indexing. However, the experiment was based on a small database of only 555 Chinese abstracts in computer and information science, with retrieval based on the index field. However, the thesaurus could be used interactively—users could select terms for query expansion. Abu-Salem (1992) also used an interactive relational thesaurus to improve recall in an Arabic retrieval system.

Greenberg (2001) investigated the effect of different thesaural relationships for query expansion using the ProQuest Controlled Vocabulary on the ABI/Inform database, which was searched via a Boolean retrieval system (the Dialog system). She found that synonyms and narrower terms increased relative recall with a nonsignificant decrease in precision, whereas related terms and broader terms increased relative recall with a statistically significant decrease in precision.

Using the TREC-2 test collection and a best-match retrieval system, Voorhees (1994) performed query expansion with various types of semantic relations encoded in WordNet. Even in a best-case scenario with the expanded terms selected by hand, query expansion did not improve retrieval results for long queries that were relatively complete. On the other hand, short queries, consisting of a single sentence describing the topic of interest, obtained significantly better results with the expansion.

Mandala, Tokunaga, and Tanoka (1999), carried out query expansion with a combination of three different types of thesauri—WordNet, a co-occurrence-based thesaurus, and one based on head-modifier relations. Head-modifier relations include four syntactic relations—subject-verb, verb-object, adjective-noun, and noun-noun relations. The expanded terms were also weighted based on their similarity to all the terms in the original query and to those in all three thesauri. Using the TREC-7 test collection, Mandala et al. found that query expansion with a combination of the three thesauri gave better average precision than was the case when no expansion was used or when it involved only one thesaurus.

Working with a Finnish full-text newspaper database and a Boolean information retrieval system, Kristensen and Järvelin (1990) found that expanding a query with synonyms and near-synonyms improved recall substantially with a small loss of precision. Kristensen (1993) experimented with broader-term, narrower-term, related-term, and synonym relations, and concluded that automatic query expansion using all these relations together improved recall twofold with a small reduction in precision. Using a best-match, full-text retrieval system (INQUERY) and the Finnish newspaper database, Kekalainen and Järvelin (1998) showed that the effect of query expansion depended on how the query was structured. Query expansion worked well with strongly structured queries but was detrimental to weakly structured queries in which, for example, the query terms and the expanded terms were treated as one list of weighted terms. The best results were obtained by expanding with all the relations.

It is clear that query expansion with related terms is crucial for improving information retrieval effectiveness and that in addition to the IS-A or broader-narrower term relations, associative relations are useful for query expansion. However, available experimental results do not suggest that it is beneficial to distinguish between specific types of associative relations. It is possible that different types of semantic relations will prove useful for expanding different queries. Rada et al. (1991) suggested that if a particular associative relation is mentioned in the query, then that relation may be useful for expanding the query. More research is needed to investigate whether specific types of semantic relations are useful for expanding specific types of queries.

A literature survey of the use of thesaural relations in information retrieval was carried out by Evens (2002).

Relation Matching for Precision Enhancement

Relation matching in information retrieval can be performed by the use of either syntactic or semantic relations. A *syntactic relation* is the relation between two words derived from the syntactic structure of the sentence; a *semantic relation* is only partly dependent on the syntactic structure of the sentence. As a semantic relation can be expressed in many syntactic forms, semantic relation matching involves matching across different syntactic relations and can yield more matches than syntactic relation matching.

Most studies on relation matching are on syntactic relations. Croft (1986), Croft, Turtle, and Lewis (1991), Dillon and Gray (1983), Hyoudo, Niimi, and Ikeda (1998), Smeaton and van Rijsbergen (1988) all recorded a small improvement in retrieval effectiveness when syntactic relations in documents and queries were taken into account in the retrieval process. Strzalkowski, Carballo, and Marinescu (1995) obtained an improvement of 20 percent, but their system included other enhancements as well. Smeaton, O'Donnell, and Kelledy (1995) obtained worse results from relation matching (using a tree-matching procedure) than from keyword matching. The retrieval results from syntactic relation matching appear to be no better than the results obtainable using index phrases generated through statistical methods, such as those described by Fagan (1989).

Metzler and Haas (1989), Metzler, Haas, Cosic, and Weise (1990), Schwarz (1990), and Ruge, Schwarz, and Warner (1991) performed syntactic processing to produce dependency trees that indicate which terms modify which other terms. Smeaton and van Rijsbergen (1988) found that the premodifier-head noun relation (e.g., adjective-noun) has a greater impact on retrieval than other relations.

In the 1980s and early 1990s, some researchers developed *conceptual information retrieval systems* that made use of complex linguistic processing and knowledge-based inferencing to extract information from text to store in a semantic representation or knowledge representation system. Examples of such systems are RIME (Berrut, 1990), the patent-claim retrieval system described by Nishida and Takamatsu (1982), SCISOR (Rau, 1987; Rau et al., 1989), and FERRET (Mauldin, 1991). Information retrieval was performed by comparing the information in the store with the semantic representation of the user's query. These systems required extensive domain knowledge, much of which was stored in case frames that specified the participant roles in an event, what types of entities could fill those roles, and what syntactic function each participant would have in the sentence (Fillmore, 1968; Somers, 1987). Because the domain knowledge had to be constructed manually, such systems were necessarily restricted to narrow domains.

The DR-LINK project (Liddy & Myaeng, 1993; Myaeng et al., 1994) investigated general methods for extracting semantic relations for information retrieval using machine-readable versions of the *Longman Dictionary of Contemporary English* (2nd ed.) (1987) and *Roget's International Thesaurus* (3rd ed.) (1962). Case frames were constructed semimanually for all verb entries and senses in the *Longman Dictionary*. However, researchers found few relation matches between queries and documents.

Lu (1990) also did not obtain good retrieval results with case relation matching. Case relations exist between words that occur close together within the same clause. Semantic relations between terms occurring in such close proximity can probably be inferred from their co-occurrence; hence, explicit semantic relation identification probably confers no advantage to retrieval effectiveness.

Gay and Croft (1990) focused on the identification of semantic relations between the members of compound nouns. The knowledge base they used included case frames and associations between entities and events. Although their system correctly interpreted compound nouns 76 percent of the time, it was not deemed likely to yield a substantial improvement in retrieval effectiveness.

Liu (1997) investigated partial relation matching. Instead of trying to match the whole concept-relation-concept triple, he sought to match each individual concept together with the semantic role that the concept has in the sentence. Instead of trying to find matches for "word1 → (relation) → word2", his system sought to find matches for "word1 → (relation)" and "(relation) → word2" separately. Liu used case roles and was able to obtain positive results only for long queries (i.e., abstracts used as queries).

Khoo, Myaeng, and Oddy (2001) developed an automatic method to identify causal relations in text and attempted to match causal relations in documents with those in queries. Causal relation matching did not perform better than word proximity matching within the same sentence. Causal relation matching worked best when one member of the causal relation (either the cause or the effect) was represented as a wildcard that could match any word.

In reviewing six years of TREC experiments (1992–1997), Sparck Jones (2000) and Perez-Carballo and Strzalkowski (2000) concluded that sophisticated natural language processing was not helpful for full-text retrieval. They noted that extracting normalized syntactic phrases (e.g., head-modifier pairs) did not give better results than statistical phrases defined by adjacency and proximity. Sparck Jones (2000) commented that there was a lack of clear evidence that a thesaurus helped in manual query construction because many other factors were involved. "It is therefore impossible to determine whether, for example, a good result is attributable to the use of vocabulary aids or just to spending a lot of time on query formation" (Sparck Jones, 2000, p. 65). She further noted that the use of elaborately structured thesauri had not been proven to be better than using a term association database.

Overall, the use of specific semantic relations either for query expansion or relation matching does not appear to be useful for document retrieval. Perhaps document retrieval is too coarse-grained to require the subtlety of semantic relations, which may be more useful for more refined kinds of information retrieval, such as question-answering.

Question-Answering with Full-Text Documents

The technology for question-answering based on full-text documents is still immature. Current approaches in TREC are focused on term matching and passage extraction. Voorhees (2003) has outlined the general approach to question-answering as comprising three steps: a) determining the expected answer type of the question, b) employing information retrieval methods to retrieve documents or passages likely to contain the answers, and c) performing more refined matching to extract the answer or trim nonrelevant text.

Some researchers have applied information extraction techniques such as pattern matching to extract the final answer from the shortlisted document passages. Paranjpe, Ramakrishnan, and Srinivasan (2004) used WordNet to score document passages using Bayesian inferencing and then used different regular expression patterns to select text segments for different kinds of questions. Harabagiu, Moldovan, Clark, Bowden, Williams, and Bensley (2004) also employed WordNet and information extraction using pattern matching. Gaizauskas, Greenwood, Hepple, Roberts, Saggion, et al. (2004) passed the top-ranked passages retrieved by an information retrieval system to an information extraction system, which converted sentences to a predicate-argument logical form.

Different patterns were used to extract answers for different kinds of questions. Litkowski (2001, 2002) extracted concept-relation-concept triples from both documents and questions, and used relational matches as one of the criteria for ranking sentences.

Semantic Relations in Automatic Text Summarization

Mani and Maybury (1999) have provided a good overview of the use of various kinds of relations in text summarization. Summarization includes three kinds of condensation operations: selection of salient or non-redundant information, aggregation of information, and generalization or abstraction. Each of these operations makes use of relations between terms/concepts and between text passages. They further identified three main approaches to text summarization:

- *The surface-level features approach*, which relies on term frequency statistics, location of a sentence within a text, presence of terms from title or user query, cue words indicating summarizing sentences or important concepts.
- *The entity-level approach*, which models the terms/concepts in the text and their relationships as a semantic network, with relations between concepts based on similarity, proximity in the text, co-occurrence, thesaural relations, co-reference, syntactic relations, and logical relations.
- *The discourse-level approach*, which models the structure of the text.

Some researchers have adapted information extraction systems for text summarization. Others have used sophisticated natural language processing to convert the text to a semantic representation and then performed summarization using knowledge-based inferencing—similar to the approach used in conceptual information retrieval systems. Text summarization can be performed on individual documents, namely, *single document summarization*, or on a set of documents, namely, *multi-document summarization*.

As Radev, Hovy, and McKeown (2002) noted, most summarization systems perform sentence extraction or passage extraction—identifying sentences/passages in the document containing important information based on surface-level features. Paice (1990) provided an overview of this approach, and argued that both the processing of anaphoric and rhetorical relations in the document and analysis of the text structure are necessary for generating high-quality abstracts. Kupiec, Pedersen, and Chen (1999) and Myaeng and Jang (1999) developed statistical models for assigning a probabilistic score to each document sentence based on the presence of surface features. The models were developed on the basis of a collection of training documents, in which sentences had been manually tagged to indicate good summary sentences. Passage extraction methods have also been applied to multidocument summarization (e.g., Goldstein, Mittal, Carbonell, & Callan, 2000).

Entity-level approaches were adopted by Hovy and Lin (1999), who used WordNet as a thesaurus to generalize the terms, and Boguraev and Kennedy (1999), who made use of cohesion relations (including anaphoric references) between terms. Barzilay and Elhadad (1999) linked up the terms in the text into lexical chains, based on cohesion relations of synonymy, repetition, hypernymy, antonymy, and holonymy. Some of the term relations were derived from WordNet. Sentences were then extracted on the basis of "strong" chains with the aid of a number of heuristics.

Entity-level approaches have also been applied to multidocument summarization. Salton, Singhal, Mitra, and Buckley (1999) constructed a network of related paragraphs based on information retrieval similarity measures. Text units that were strongly connected to other units were considered salient and good candidates for extraction. Mani and Bloedorn (1999) constructed a network of terms and text units based on cohesion relations. Spreading activation was used to identify salient nodes on the basis of connectivity and the strengths of the links. Commonalities and differences between documents were then computed on the basis of the salient nodes for each document.

Marcu (1999, 2000) developed a parser to identify rhetorical relations in text to form a rhetorical structure tree, which was then used to identify important clauses. Each rhetorical relation links two text segments—one text segment is considered the nucleus node representing the central information and the other, the satellite node representing secondary information. Nucleus nodes are considered more salient than satellite nodes, with nucleus nodes linked to higher-level nucleus nodes at the top of the tree considered the most salient. Salience scores were computed for the nodes of the rhetorical tree and used to extract corresponding sentences or clauses to form summaries.

Teufel and Moens (1999) made use of macrolevel text structure, focusing on sections of the document that they called the *argumentative structure* of the text. The document sections were also identified with "global rhetorical relations"—relations of the text segment with respect to the content of the whole document. They used the following roles: background, topic, related work, purpose/problem, solution/method, result, and conclusion/claim. The abstract they created also used this argumentative template and sentences were extracted from the corresponding document section to fill the abstract template.

Strzalkowski, Stein, Wang, and Wise (1999) used a discourse structure of news summaries to combine query-relevant information with related but "out-of-context information." They made use of background-main news relations to identify such out-of-context information.

Radev (2000) introduced a theory of cross-document structure, which can be used to describe the rhetorical structure of a set of related documents. Cross-document structure theory makes use of a multidocument graph to represent text simultaneously at different levels of granularity (words, phrases, sentences, paragraphs, and documents). It contains links representing cross-document semantic relationships among text units, such as equivalence, cross-reference, contradiction, and historical background. Different summaries can be generated from the graph according to user needs by preserving some links in the graph while removing others.

Information extraction techniques have also been applied to text summarization. The SUMMONS system (McKeown & Radev, 1999) used information extraction for multidocument summarization. Information was first extracted from each document to fill a template. When the templates for different documents were merged, operations were performed to identify the following logical relations between templates—change of perspective, contradiction, addition, refinement, agreement, superset, trend, and no information.

The RIPTIDES system (White, Korelsky, Cardie, Ng, Pierce, & Wagstaff, 2001) also used an information extraction system to fill templates for summarization in the natural disasters domain. However, additional potentially relevant information not found in the templates were also extracted from selected sentences and added to the summary to round it off.

Knowledge-based approaches to summarization using a semantic representation of the text were adopted in the SUSY (Fum, Guida, & Tasso, 1985), SCISOR (Rau et al., 1989), and TOPIC systems (Hahn & Reimer, 1999; Reimer & Hahn, 1988). The TOPIC system converted the text into a terminological logic representation scheme. From this representation, "salience operators" extracted concepts, relations, and properties, which were then synthesized into a hierarchical text graph incorporating discourse and concept relations.

Lehnert (1999) proposed an inference-based technique for summarizing narratives based on structural relations around plot units. Primitive plot units, including *problem, success, failure, hidden blessing*, and *mixed blessing,* are building blocks for more complex plot units. The method focuses on affect or emotional states, and the relations between events and affect states. Lehnert listed three affect states: *positive event, negative event,* and *mental state* (neutral affect). The relations between events and affect states include *motivation, actualization, termination,* and *equivalence.* These can be used to build primitive plot units, from which more complex plot units can be derived.

Conclusion

Information retrieval in the 20th century focused on terms, especially nouns, and concepts. We seem to be approaching the limit of what term-based and concept-based approaches can accomplish. For example, in the TREC series of conferences, the ad hoc information retrieval track, once considered the main retrieval task, has been discontinued because of the lack of improvement in participating systems.

We believe that natural language processing and semantic relations, in particular, point the way forward for information retrieval in the 21st century. But as we have seen, semantic relations are subtle things. They are difficult for computer programs to identify and process. Yet human minds process semantic relations effortlessly. Our facility with symbolic processing and semantic relations certainly distinguishes us from machines.

Two factors have retarded progress in the effective use of semantic relations in information processing applications. One is the difficulty of automatically identifying semantic relations in text with accuracy. The other is the difficulty of identifying suitable application areas that require the subtlety of semantic relations. Ad-hoc, full-text document retrieval does not appear to require the use of semantic relations. Coarse-grained methods of term matching, appropriate term weighting and document length normalization, and query expansion with term associations based on term co-occurrence statistics, seem to yield as good a retrieval result as we are likely to get. More promising applications for the use of semantic relations are question-answering, document summarization, and information extraction. Effective text processing and text mining tools for identifying semantic relations in text will help to promote more research in their use.

Further studies of relevance relationships between documents and user information needs can also yield deeper insights into how information retrieval effectiveness can be improved. Although several studies have identified different types of relevance relations and factors that affect relevance judgments, we know little about the thought processes, inferencing mechanisms, and domain knowledge used by humans to judge relevance. We need more in-depth studies of the types of relationships between the user's information need, task, situation, and the document content that determine the relevance and usefulness of the document.

It is also not known whether making fine distinctions between the different types of semantic relations and their properties is useful in information processing applications. Because such fine distinctions are found in both language and human information processing, we hypothesize that they are important in information processing, but it is not clear in what way and for what applications they might be used.

Two exciting new areas for research are the manual and automatic construction of ontologies for various applications and the development of methods for exploiting ontologies effectively in different real-life applications. With the availability of vast quantities of textual documents on the World Wide Web, mining the Web for concepts and relations to build relational knowledge bases and ontologies will become increasingly important.

Other promising research areas not covered in this survey are user profiling and personalization (e.g., Jung, Rim, & Lee, 2004), and special types of text categorization and automated content analysis. For example, in the area of automatic sentiment categorization (categorizing documents into those expressing positive or favorable sentiment versus negative or unfavorable sentiment), Nasukawa and Yi (2003) and Na, Sui, Khoo, Chan, and Zhou (2004) found that it was not sufficient to consider just the sentiment-bearing terms in the text; it was important to determine the subject and object to which the sentiments were linked.

Lack of understanding of semantic relations among information science researchers and practitioners has also retarded its use in information science. One purpose of this *ARIST* chapter has been to pull together information about semantic relations from several disciplines to provide a deeper understanding of their nature and types, as well as to suggest some possible applications.

Acknowledgments

We are very grateful to three anonymous reviewers for their close reading of an earlier draft of this chapter. Any errors that remain are ours alone.

We also gratefully acknowledge the excellent editorial work of Ms. Soon-Kah Liew.

Endnotes

1. *Categories* refer to sets of objects, whereas *concepts* refer to the mental representations of the categories. The terms are often used interchangeably when it is not necessary to distinguish between them.
2. A word within square brackets is a label for a concept. A word within round brackets is a label for a relation. Arrows indicate the direction of the relation.
3. Saussure used the term *associative relations* for what is now known as *paradigmatic relations*.
4. We follow the convention that a question mark indicates a sentence is grammatically or semantically odd; an asterisk indicates a sentence is grammatically or semantically abnormal.

References

Abu-Salem, H. (1992). *A microcomputer-based Arabic bibliographic information retrieval system with relational thesauri*. Unpublished doctoral dissertation, Illinois Institute of Technology.

Ahn, W.-K. (1999). Effect of causal structure on category construction. *Memory & Cognition, 27*(6), 1008–1023.

Ahn, W.-K., & Kim, N. S. (2001). The causal status effect in categorization: An overview. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory, Vol. 40* (pp. 23–65). San Diego, CA: Academic Press.

Aitchison, J., Gilchrist, A., & Bawden, D. (1997). *Thesaurus construction and use: A practical manual* (3rd ed.). London: Aslib.

Alani, H., Kim, S., Millard, D. E., Weal, M. J., Hall, W., Lewis, P. H., et al. (2003). Automatic ontology-based knowledge extraction from Web documents. *IEEE Intelligent Systems, 18*(1), 14–21.

Alba, J. W., & Hasher L. (1983). Is memory schematic? *Psychological Bulletin, 93*, 203–231.

Alonge, A., Calzolari, N., Vossen, P., Bloksma, L., Castellon, I., Marti, M. A., et al. (1998). The linguistic design of the EuroWordNet. *Computers and the Humanities, 32*(2–3), 91–115.

Appelt, D. E., & Israel, D. J. (1999). *Introduction to information extraction technology: IJCAI-99 tutorial*. Retrieved December 31, 2004, from http://www.ai.sri.com/~appelt/ie-tutorial

Aristotle. (1961). *Aristotle's metaphysics* (J. Warrington, Ed. & Trans.). London: J. M. Dent.

Aristotle. (1996). *Physics* (R. Waterfield, Trans.). Oxford, UK: Oxford University Press.

Aronson, A. R. (2001). Effective mapping of biomedical text to the UMLS Metathesaurus: The MetaMap program. In S. Bakken (Ed.), *Proceedings of the 2001 AMIA Annual Symposium* (pp. 17–21). Bethesda, MD: American Medical Informatics Association.

Artale, A., Franconi, E., Guarino, N., & Pazzi, L. (1996). Part–whole relations in object-centered systems: An overview. *Data & Knowledge Engineering, 20*, 347–383.

Asher, R. E. (Ed.). (1994). *The encyclopedia of language and linguistics*. Oxford, UK: Pergamon Press.

Baader, F., Calvanese, D., McGuinness, D., Nardi, D., & Patel-Schneider, P. (2003). *The description logic handbook: Theory, implementation and applications*. Cambridge, UK: Cambridge University Press.

Baker, C. F., Fillmore, C. J., & Cronin, B. (2003). The structure of the Framenet database. *International Journal of Lexicography, 16*(3), 281–296.

Baker, C. F., Fillmore, C. J., & Lowe, J. B. (1998). The Berkeley FrameNet project. *Proceedings of the 17th International Conference on Computational Linguistics* (vol. 1), 86–90.

Barriere, C. (1997). *From a children's first dictionary to a lexical knowledge base of conceptual graphs*. Unpublished doctoral dissertation, Simon Fraser University, British Columbia, Canada.

Barriere, C. (2002). Hierarchical refinement and representation of the causal relation. *Terminology, 8*(1), 91–111.

Barry, C. L. (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science, 45*, 149–159.

Barzilay, R., & Elhadad, M. (1999). Using lexical chains for text summarization. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 111–122). Cambridge, MA: MIT Press.

Basili, R., Pazienza, M. T., & Vindigni, M. (1997). Corpus-driven unsupervised learning of verb subcategorization frames. In *AI*IA 97: Advances in Artificial Intelligence: 5th Congress of the Italian Association for Artificial Intelligence* (Lecture Notes in Artificial Intelligence, Vol. 1321, pp. 159–170). Berlin: Springer-Verlag.

Bateman, J. A., Fabris, G., & Magnini, B. (1995). The Generalized Upper Model knowledge base: Organization and use. In N. Mars (Ed.), *Second International Conference on Building and Sharing of Very Large-Scale Knowledge Bases (KBKS '95)* (pp. 60–72). Amsterdam: IOS Press.

Bean, C. A. (1998). The semantics of hierarchy: Explicit parent-child relationships in MeSH tree structures. *Proceedings of the Fifth International ISKO Conference*, 133–138.

Bean, C. A., & Green, R. (2001). Relevance relationships. In C. A. Bean & R. Green (Eds.), *Relationships in the organization of knowledge* (pp. 115–132). Dordrecht, NL: Kluwer.

Beekman, J., Callow, J., & Kopesec, M. (1981). *The semantic structure of written communication* (5th ed.). Dallas, TX: Summer Institute of Linguistics.

Belkin, N. J., Head, J., Jeng, J., Kelly, D., Lin, S., Park, S. Y., et al. (1999). Relevance feedback versus local context analysis as term suggestion devices: Rutgers' TREC-8 Interactive Track experience. *Eighth Text Retrieval Conference (TREC-8)* (NIST Special Publication 500–246, pp. 565–574). Retrieved September 1, 2004, from http://trec.nist.gov/pubs/trec8/t8_proceedings.html

Bell, A. (1998). The discourse structure of news stories. In A. Bell & P. Garrett (Eds.), *Approaches to media discourse* (pp. 64–104). Oxford, UK: Blackwell.

Bernauer, J. (1996). Analysis of part–whole relation and subsumption in the medical domain. *Data and Knowledge Engineering, 20*, 405–415.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American, 284*(5/May), 34–43. Retrieved March 11, 2004, from http://www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21

Berrut, C. (1990). Indexing medical reports: The RIME approach. *Information Processing & Management, 26*(1), 93–109.

Blair, D. C. (2003). Information retrieval and the philosophy of language. *Annual Review of Information Science and Technology, 37*, 3–50.

Blake, C., & Pratt, W. (2001). Better rules, fewer features: A semantic approach to selecting features from text. *Proceedings of the 2001 IEEE International Conference on Data Mining*, 59–66.

Boguraev, B., & Kennedy, C. (1999). Salience-based content characterization of text documents. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 99–110). Cambridge, MA: MIT Press.

Borgelt, C., & Kruse, R. (2002). Induction of association rules: Apriori implementation. *15th Conference on Computational Statistics (Compstat 2002)*. Retrieved December 31, 2004, from http://fuzzy.cs.uni-magdeburg.de/~borgelt/papers/cstat_02.pdf

Brent, M. R. (1993). From grammar to lexicon: Unsupervised learning of lexical syntax. *Computational Linguistics, 19*(2), 243–262.

Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 119–160). Hillsdale, NJ: Erlbaum.

Buckley, C., Mitra, M., Walz, J., & Cardie, C. (1998). Using clustering and superconcepts within SMART: TREC 6. *The Sixth Text REtrieval Conference (TREC-6)* (NIST Special Publication 500–240, pp. 107–124). Retrieved September 1, 2004, from http://trec.nist.gov/pubs/trec6/t6_proceedings.html

Buckley, C., Singhal, A., Mitra, M., & Salton, G. (1996). New retrieval approaches using SMART: TREC 4. *The Fourth Text Retrieval Conference (TREC-4)* (NIST Special

Publication 500–236, pp. 25–48). Retrieved September 1, 2004, from http://trec.nist.gov/pubs/trec4/t4_proceedings.html

Butcher, K. R., & Kintsch, W. (2003). Text comprehension and discourse processing. In I. B. Weiner (Ed.), *Handbook of psychology* (Vol. 4, pp. 575–595). Hoboken, NJ: Wiley.

Califf, M., & Mooney, R. (2003). Bottom-up relational learning of pattern matching rules for information extraction. *Journal of Machine Learning Research, 4*, 177–210.

Calzolari, N. (1988). The dictionary and the thesaurus can be combined. In M. W. Evens (Ed.), *Relational models of the lexicon: Representing knowledge in semantic networks* (pp. 75–96). Cambridge, UK: Cambridge University Press.

Calzolari, N. (1992). Acquiring and representing semantic information in a lexical knowledge base. In *Lexical Semantics and Knowledge Representation: First SIGLEX Workshop: Proceedings* (pp. 235–243). Berlin: Springer-Verlag.

Calzolari, N., & Picchi, E. (1989). Acquisition of semantic information from an on-line dictionary. *Proceedings of the 12th International Conference on Computational Linguistics* (Vol. 1), 87–92.

Cann, R. (1993). *Formal semantics: An introduction*. Cambridge, UK: Cambridge University Press.

Caplan, L. J., & Barr, R. A. (1991). The effects of feature necessity and extrinsicity on gradedness of category membership and class inclusion. *British Journal of Psychology, 82*, 427–440.

Caplan, L. J., & Herrmann, D. J. (1993). Semantic relations as graded concepts. *Zeitschrift für Psychologie, 201*, 85–97.

Chaffin. R. (1992). The concept of a semantic relation. In A. Lehrer & E. F. Kittay (Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization* (pp. 253–288). Hillsdale, NJ: Erlbaum.

Chaffin, R., & Herrmann, D. J. (1984). The similarity and diversity of semantic relations. *Memory & Cognition, 12*(2), 134–141.

Chaffin, R., & Herrmann, D. J. (1987). Relation element theory: A new account of the representation and processing of semantic relations. In D. Gorfein & R. Hoffman (Eds.), *Learning and memory: The Ebbinghaus Centennial Conference* (pp. 221–251). Hillsdale, NJ: Erlbaum.

Chaffin, R., & Herrmann, D. J. (1988a). Effects of relation similarity on part–whole decisions. *Journal of General Psychology, 115*,131–139.

Chaffin, R., & Herrmann, D. J. (1988b). The nature of semantic relations: A comparison of two approaches. In M. W. Evens (Ed.), *Relational models of the lexicon: Representing knowledge in semantic networks* (pp. 289–334). Cambridge, UK: Cambridge University Press.

Chan, S. W. K., & Franklin, J. (2003). Dynamic context generation for natural language understanding: A multifaceted knowledge approach. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 33*(1), 23–41.

Chang, C.-H., & Lui, S.-C. (2001). IEPAD: Information extraction based on pattern discovery. *Proceedings of the 10th International Conference on World Wide Web*, 681–688.

Chen, H., & Lynch, K. J. (1992). Automatic construction of networks of concepts characterizing document databases. *IEEE Transactions on Systems, Man and Cybernetics, 22*(5), 885–902.

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Chowdhury, G. G. (2003). Natural language processing. *Annual Review of Information Science and Technology, 37,* 51–90.

Church, K. W., Gale, W., Hanks, P., Hindle, D., & Moon, R. (1994). Lexical substitutability. In B. T. S. Atkins & A. Zampolli (Eds.), *Computational approaches to the lexicon* (pp. 153–177). Oxford, UK: Oxford University Press.

Church, K. W., & Hanks, P. (1989). Word association norms, mutual information and lexicography. *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, 76–83.

Ciravegna, F. (2001). Adaptive information extraction from text by rule induction and generalization. *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, 1251–1256.

Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 65–110). New York: Academic Press.

Clarke, S. G. D. (2001). Thesaural relationships. In C. A. Bean & R. Green (Eds.), *Relationships in the organization of knowledge* (pp. 37–52). Dordrecht, NL: Kluwer.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240–247.

Cook, W. A. (1989). *Case grammar theory*. Washington, DC: Georgetown University Press.

Croft, W. B. (1986). Boolean queries and term dependencies in probabilistic retrieval models. *Journal of the American Society for Information Science, 37*(2), 71–77.

Croft, W. B., Turtle, H. R., & Lewis, D. D. (1991). The use of phrases and structured queries in information retrieval. *Proceedings of the Fourteenth Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval*, 32–45.

Crombie, W. (1985). *Process and relation in discourse and language learning*. London: Oxford University Press.

Cruse, D. A. (1979). On the transitivity of the part–whole relation. *Journal of Linguistics, 15*, 29–38.

Cruse, D. A. (1986). *Lexical semantics*. Cambridge, UK: Cambridge University Press.

Cruse, D. A. (2000). *Meaning in language: An introduction to semantics and pragmatics*. Oxford, UK: Oxford University Press.

Cruse, D. A. (2002). Hyponymy and its varieties. In R. Green, C. Bean, & S. H. Myaeng (Eds.), *The semantics of relationships: An interdisciplinary perspective* (pp. 3–22). Dordrecht, NL: Kluwer.

Cruse, D. A. (2004). *Meaning in language: An introduction to semantics and pragmatics* (2nd ed.). Oxford, UK: Oxford University Press.

Dick, J. P. (1997). Modeling cause and effect in legal text. In D. Lukose, H. Delugach, M. Keeler, L. Searle, & J. Sowa (Eds.), *Conceptual Structures: Fulfilling Peirce's Dream: Fifth International Conference on Conceptual Structures, ICCS'97* (pp. 244–259). Berlin: Springer-Verlag.

Dillon, M., & Gray, A. S. (1983). FASIT: A fully automatic syntactically based indexing system. *Journal of the American Society for Information Science, 34*(2), 99–108.

Dimino, J. A., Taylor, R. M., & Gersten, R. M. (1995). Synthesis of the research on story grammar as a means to increase comprehension. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 11*(1), 53–72.

Dooley, R. A., & Levinsohn, S. H. (2001). *Analyzing discourse: A manual of basic concepts*. Dallas, TX: SIL International.

Dorr, B. J., Levow, G.-A., & Lin, D. (2002). Construction of a Chinese-English verb lexicon for machine translation and embedded multilingual applications. *Machine Translation, 17*, 99–137.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language, 67*(3), 547–619.

Dyvik, H. (2004). Translations as semantic mirrors: From parallel corpus to WordNet. *Language and Computers, 49*(1), 311–326.

Eells, E. (1991). *Probabilistic causality*. Cambridge, UK: Cambridge University Press.

Efthimiadis, E. N. (1996). Query expansion. *Annual Review of Information Science and Technology, 31*, 121–188.

Eggins, S. (1994). *An introduction to systemic functional linguistics*. London: Pinter Publishers.

Ehring, D. (1997). *Causation and persistence: A theory of causation*. New York: Oxford University Press.

Evens, M. W. (Ed.). (1988). *Relational models of the lexicon: Representing knowledge in semantic networks*. Cambridge, UK: Cambridge University Press.

Evens, M. W. (2002). Thesaural relations in information retrieval. In R. Green, C. Bean, & S. H. Myaeng (Eds.), *The semantics of relationships: An interdisciplinary perspective* (pp. 143–160). Dordrecht, NL: Kluwer.

Evens, M. W., Litowitz, B. E., Markowitz, J. A., Smith, R. N., & Werner, O. (1980). *Lexical-semantic relations: A comparative survey*. Edmonton, AB: Linguistic Research.

Fagan, J. L. (1989). The effectiveness of a nonsyntactic approach to automatic phrase indexing for document retrieval. *Journal of the American Society for Information Science, 40*(2), 115–132.

Farradane, J. E. L. (1950). A scientific theory of classification and indexing and its practical applications. *Journal of Documentation, 6*(2), 83–99.

Farradane, J. E. L. (1952). A scientific theory of classification and indexing: Further considerations. *Journal of Documentation, 8*(2), 73–92.

Farradane, J. E. L. (1967). Concept organization for information retrieval. *Information Storage & Retrieval, 3*(4), 297–314.

Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.

Fellbaum, C. (2002). On the semantics of troponymy. In R. Green, C. Bean, & S. H. Myaeng (Eds.), *The semantics of relationships: An interdisciplinary perspective* (pp. 23–34). Dordrecht, NL: Kluwer.

Fellbaum, C., & Chaffin, R. (1990). Some principles of the organization of the verb lexicon. *Proceedings of the 12th Annual Conference of the Cognitive Science Society*, 420–428.

Fellbaum, C., & Miller, G. A. (1991). Semantic networks of English. *Cognition, 41*, 197–229.

Fensel, D. (2001). Ontologies and electronic commerce (editor's introduction). *IEEE Intelligent Systems, 16*(1), 8–14.

Fillmore, C. J. (1968). The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1–88). New York: Holt, Rinehart and Wilson.

Fillmore, C. J. (1971). Some problems for case grammar. In R. J. O'Brien (Ed.), *22nd Annual Round Table. Linguistics: Developments of the Sixties—Viewpoints of the Seventies* (Monograph Series on Language and Linguistics, 24) (pp. 35–56). Washington, DC: Georgetown University Press.

Firth, J. R. (1957). *Papers in linguistics 1934–1951.* London: Oxford University Press.

Firth, J. R. (1968). *Selected papers of J. R. Firth 1952–1959* (F. R. Palmer, Ed.). London: Longman.

Fox, E. A. (1980). Lexical relations: Enhancing effectiveness of information retrieval systems. *SIGIR Forum, 15*(3), 5–36.

Framis, F. R. (1994). An experiment on learning appropriate selectional restrictions from a parsed corpus. *Proceedings of the Fifteenth International Conference on Computational Linguistics* (Vol. 2), 769–774.

Freitag, D. (2000). Machine learning for information extraction in informal domains. *Machine Learning, 39*, 169–202.

Fum, D., Guida, G., & Tasso, C. (1985). Evaluating importance: A step towards text summarization. *Proceedings of the 9th International Joint Conference on Artificial Intelligence (IJCAI'85)* (Vol. 2), 840–844.

Gaizauskas, R., Greenwood, M. A., Hepple, M., Roberts, I., Saggion, H., et al. (2004). The University of Sheffield's TREC 2003 Q&A experiments. *The Twelfth Text Retrieval Conference (TREC 2003)* (NIST Special Publication 500–255, pp. 782–790). Retrieved December 31, 2004, from http://trec.nist.gov/pubs/trec12/t12_proceedings.html

Ganter, B., & Wille, R. (1997). *Formal concept analysis: Mathematical foundations* (C. Franzke, Trans.). Berlin: Springer.

Gardin, J.-C. (1965). *SYNTOL.* New Brunswick, NJ: Graduate School of Library Service, Rutgers University.

Gay, L. S., & Croft, W. B. (1990). Interpreting nominal compounds for information retrieval. *Information Processing & Management, 26*(1), 21–38.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199–241). Cambridge, UK: Cambridge University Press.

Gerstl, P., & Pribbenow, S. (1995). Midwinters, end games, and body parts: A classification of part–whole relations. *International Journal of Human–Computer Studies, 43*, 865–889.

Girju, R. (2002). *Text mining for semantic relations.* Unpublished doctoral dissertation, University of Texas at Dallas.

Girju, R., Badulescu, A., & Moldovan, D. (2003). Learning semantic constraints for the automatic discovery of part–whole relations. *Proceedings of the Human Language Technology Conference*, 80–87.

Glass, A. L., Holyoak, K. J., & Kiger, J. I. (1979). Role of antonymy relations in semantic judgments. *Journal of Experimental Psychology: Human Learning & Memory, 5*, 598–606.

Goldstein, J., Mittal, V., Carbonell, J., & Callan, J. (2000). Creating and evaluating multi-document sentence extract summaries. *Proceedings of the 9th International Conference on Information and Knowledge Management*, 165–172.

Gómez-Pérez, A., Fernandez-Lopez, M., & Corcho, O. (2004). *Ontological engineering*. London: Springer-Verlag.

Green, R. (1995a). The expression of conceptual syntagmatic relationships: A comparative survey. *Journal of Documentation, 51*, 315–338.

Green, R. (1995b). Syntagmatic relationships in index languages: A reassessment. *Library Quarterly, 65*, 365–385.

Green, R. (2001). Relationships in the organization of knowledge: An overview. In C. A. Bean & R. Green (Eds.), *Relationships in the organization of knowledge* (pp. 3–18). Dordrecht, NL: Kluwer.

Green, R. (2002). Internally-structured conceptual models in cognitive semantics. In R. Green, C. Bean, & S. H. Myaeng (Eds.), *The semantics of relationships: An interdisciplinary perspective* (pp. 73–90). Dordrecht, NL: Kluwer.

Green, R., & Bean, C. A. (1995). Topical relevance relationships: An exploratory study and preliminary typology. *Journal of the American Society for Information Science, 46*, 654–662.

Green, R., Bean, C., & Myaeng, S. H. (Eds.). (2002). *The semantics of relationships: An interdisciplinary perspective*. Dordrecht, NL: Kluwer.

Greenberg, J. (2001). Automatic query expansion via lexical-semantic relationships. *Journal of the American Society for Information Science and Technology, 52*(5), 402–415.

Grefenstette, G. (1992). Use of syntactic context to produce term association lists for text retrieval. *Proceedings of the Fifteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 89–97.

Grishman, R., & Sterling, J. (1994). Generalizing automatically generated selectional patterns. *Proceedings of the Fifteenth International Conference on Computational Linguistics*, 742–747.

Gruber, T. R. (1993). A translation approach to portable ontology specification. *Knowledge Acquisition, 5*(2), 199–220.

Guarino, N. (1997). Understanding, building and using ontologies. *International Journal of Human–Computer Studies, 46*, 293–310.

Guarino, N., & Giaretta, P. (1995). Ontologies and knowledge bases: Towards a terminological clarification. In N. J. I. Mars (Ed.), *Towards very large knowledge bases: Knowledge building and knowledge sharing* (pp. 25–32). Amsterdam: IOS Press.

Hahn, U., & Reimer, U. (1999). Knowledge-based text summarization: Salience and generalization operators for knowledge base abstraction. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 215–232). Cambridge, MA: MIT Press.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Harabagiu, S., Moldovan, D., Clark, C., Bowden, M., Williams, J., & Bensley, J. (2004). Answer mining by combining extraction techniques with abductive. *The Twelfth Text Retrieval Conference (TREC 2003)* (NIST Special Publication 500–255, pp. 375–382). Retrieved December 31, 2004, from http://trec.nist.gov/pubs/trec12/t12_proceedings.html

Harris, R. (1987). *Reading Saussure: A critical commentary on the Cours de linguistique generale*. La Salle, IL: Open Court.

Hausmann, F. J. (1985). Kollokationen im deutschen Wörterbuch: Ein Beitrag zur Theorie des lexikographischen Beispiels. In H. Bergenholtz & J. Mugdan (Eds.), *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch* (pp. 118–129). Tübingen, Germany: Niemeyer.

Hawking, D., Thistlewaite, P., & Craswell, N. (1998). ANU/ACSys TREC-6 experiments. *The Sixth Text Retrieval Conference (TREC-6)* (NIST Special Publication 500–240, pp. 275–290). Retrieved September 1, 2004, from http://trec.nist.gov/pubs/trec6/t6_proceedings.html

Herrmann, D. J. (1987). Representational forms of semantic relations and the modeling of relation comprehension. In E. van der Meer & J. Hoffmann (Eds.), *Knowledge-aided information processing* (pp. 13–29). Amsterdam: North-Holland.

Herrmann, D. J., & Chaffin, R. (1986). Comprehension of semantic relations as a function of the definitions of relations. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performances* (pp. 311–319). Amsterdam: North-Holland.

Herrmann, D. J., Chaffin, R. J. S., Conti, G., Peters, D., & Robbins, P. H. (1979). Comprehension of antonymy and the generality of categorization models. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 585–597.

Herrmann, D. J., & Raybeck, D. (1981). Similarities and differences in meaning in six cultures. *Journal of Cross-Cultural Psychology, 12*, 194–206.

Hobbs, J. (1985). *On the coherence and structure of discourse* (Technical Report CSLI–85–37) Stanford, CA: Center for the Study of Language and Information, Stanford University.

Hoffmann, J., & Trettin, M. (1980). Organizational effects of semantic relations. In F. Klix & J. Hoffmann (Eds.), *Cognition and memory* (pp. 95–102). Amsterdam: North-Holland.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Hovy, E., & Lin, C. (1999). Automated text summarization in SUMMARIST. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 81–94). Cambridge, MA: MIT Press.

Hudon, M. (2001). Relationships in multilingual thesauri. In C. A. Bean & R. Green (Eds.), *Relationships in the organization of knowledge* (pp. 67–80). Dordrecht, NL: Kluwer.

Hume, D. (1955). *An enquiry concerning human understanding, and selections from a treatise of human nature.* La Salle, IL: The Open Court Publishing Company

Humphreys, B. L., Lindberg, D. A. B., Schoolman, H. M., & Barnett, G. O. (1998). The Unified Medical Language System: An informatics research collaboration. *Journal of the American Medical Informatics Association, 5*(1), 1–13.

Hyoudo, Y., Niimi, K., & Ikeda, T. (1998). Comparison between proximity operation and dependency operation in Japanese full-text retrieval. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 341–342.

International Organization for Standardization. (1986). *Guidelines for the establishment and development of monolingual thesauri* (2nd ed.) (ISO 2788–1986(E)). Geneva: International Standards Organization.

Iris, M. A., Litowitz, B. E., & Evens, M. (1988). Problems of the part–whole relation. In M. W. Evens (Ed.), *Relational models of the lexicon: Representing knowledge in semantic networks* (pp. 261–288). Cambridge, UK: Cambridge University Press.

Jaspars, J. (1983). The process of attribution in common-sense. In M. Hewstone (Ed.), *Attribution theory: Social and functional extensions* (pp. 28–44). Oxford, UK: Blackwell.

Jaspars, J., Hewstone, M., & Fincham, F. D. (1983). Attribution theory and research: The state of the art. In J. Jaspars, F. D. Fincham, & M. Hewstone (Eds.), *Attribution theory and research: Conceptual, developmental and social dimensions* (pp. 3–36). London: Academic Press.

Johansson, S. (1997). Using the English-Norwegian Parallel Corpus—a corpus for contrastive analysis and translation studies. In B. Lewandowska-Tomaszczyk & P. J. Melia (Eds.), *Practical applications in language corpora* (pp. 282–296). Lodz, Poland: Lodz University.

Jones, S. (2002). *Antonymy: A corpus-based perspective*. London: Routledge.

Jung, K.-Y., Rim, K.-W., & Lee, J.-H. (2004). Automatic preference mining through learning user profile with extracted information. In *Structural, Syntactic and Statistical Pattern Recognition: Joint IAPR International Workshops SSPR 2004 and SPR 2004* (Lecture Notes in Computer Science 3138, pp. 815–823). Berlin: Springer-Verlag.

Justeson, J. S., & Katz, S. M. (1991). Co-occurrence of antonymous adjectives and their contexts. *Computational Linguistics, 17*, 1–19.

Justeson, J. S., & Katz, S. M. (1992). Redefining antonymy: The textual structure of a semantic relation. *Literary and Linguistic Computing, 7*(3), 176–184.

Katz, J. J. (1972). *Semantic theory.* New York: Harper and Row.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA: MIT Press.

Keil, F. C. (2003). Categorisation, causation, and the limits of understanding. *Language and Cognitive Processes, 18*(5/6), 663–692.

Kekalainen, J., & Järvelin, K. (1998). The impact of query structure and query expansion on retrieval performance. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 130–135.

Khoo, C. S. G. (1995). *Automatic identification of causal relations in text and their use for improving precision in information retrieval.* Unpublished doctoral dissertation, Syracuse University.

Khoo, C., Chan, S., & Niu, Y. (2002). The many facets of the cause–effect relation. In R. Green, C. Bean, & S. H. Myaeng (Eds.), *The semantics of relationships: An interdisciplinary perspective* (pp. 51–70). Dordrecht, NL: Kluwer.

Khoo, C., Myaeng, S. H., & Oddy, R. (2001). Using cause–effect relations in text to improve information retrieval precision. *Information Processing & Management, 37*(1), 119–145.

Kim, J.-T. (1996). Automatic phrasal pattern acquisition for information extraction from natural language texts. *Journal of KISS (B), Software and Applications, 23*(1), 95–105.

Kim, J.-T., & Moldovan, D. I. (1995). Acquisition of linguistic patterns for knowledge-based information extraction. *IEEE Transactions on Knowledge and Data Engineering, 7*(5), 713–724.

Kishore, J. (1986). *Colon classification: Enumerated & expanded schedules along with theoretical formulations*. New Delhi, India: Ess Publications.

Klix, F. (1980). On structure and function of semantic memory. In F. Klix & J. Hoffmann (Eds.), *Cognition and memory* (pp. 11–25). Amsterdam: North-Holland.

Klix, F. (1986). On recognition processes in human memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performances* (pp. 321–338). Amsterdam: North-Holland.

Kounios, J., & Holcomb, P. J. (1992). Structure and process in semantic memory: Evidence from event-related brain potentials and reaction times. *Journal of Experimental Psychology: General, 121*, 459–479.

Kounios, J., Montgomery, E. C., & Smith, R. W. (1994). Semantic memory and the granularity of semantic relations: Evidence from speed-accuracy decomposition. *Memory & Cognition, 22*(6), 729–741.

Kristensen, J. (1993). Expanding end-user's query statements for free text searching with a search-aid thesaurus. *Information Processing & Management, 29*(6), 733–744.

Kristensen, J., & Järvelin, K, (1990). The effectiveness of a searching thesaurus in free-text searching in a full-text database. *International Classification, 17*(2), 77–84.

Kukla, F. (1980). Componential analysis of the recognition of semantic relations between concepts. In F. Klix & J. Hoffmann (Eds.), *Cognition and memory* (pp. 169–176). Amsterdam: North-Holland.

Kupiec, J., Pedersen, J., & Chen, F. (1999). Trainable document summarizer. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 55–60). Cambridge, MA: MIT Press.

Kwok, K. L., & Chan, M. (1998). Improving two-stage ad-hoc retrieval for short queries. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 250–256.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed.) (pp. 202–251). Cambridge, UK: Cambridge University Press.

Lambrix, P. (2000). *Part–whole reasoning in an object-centered framework* (Lecture Notes in Artificial Intelligence 1771). Berlin: Springer-Verlag.

Lancaster, F. W. (1986). *Vocabulary control for information retrieval*. Arlington, VA: Information Resources Press.

Landis, T. Y., Herrmann, D. J., & Chaffin, R. (1987). Developmental differences in the comprehension of semantic relations. *Zeitschrift für Psychologie, 195*, 129–139.

Lang, R. R. (2003). Story grammars: Return of a theory. In M. Mateas & P. Sengers (Eds.), *Narrative intelligence* (pp. 199–212). Amsterdam: John Benjamins.

Lavrac, N., & Dzeroski, S. (1994). *Inductive logic programming: Techniques and applications*. New York: Ellis Horwood.

Lee, C. H., Na, J. C., & Khoo, C. (2003). Ontology learning for medical digital libraries. *6th International Conference on Asian Digital Libraries, ICADL 2003* (Lecture Notes in Computer Science 2911), 302–305.

Lee, J. W. T., Chan, S. C. F., & Yeung, D. S. (1995). Modelling constraints in part structures. *Computers & Industrial Engineering, 28*(3), 645–657.

Lehmann, F. (Ed.). (1992). *Semantic networks in artificial intelligence*. Oxford, UK: Pergamon Press.

Lehnert, W. (1991). Symbolic/subsymbolic sentence analysis: Exploiting the best of two worlds. In J. Barnden & J. Pollack (Eds.), *Advances in connectionist and neural computation theory* (Vol. 1, pp. 135–164). Norwood, NJ: Ablex.

Lehnert, W. G. (1999). Plot units: A narrative summarization strategy. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 177–214). Cambridge, MA: MIT Press.

Lehrer, A., & Lehrer, K. (1982). Antonymy. *Linguistics and Philosophy, 5*, 483–501.

Lenat, D., Miller, G., & Yokoi, T. (1995). CYC, WordNet, and EDR: Critiques and responses. *Communications of the ACM, 38*(11), 45–48.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Levy, F. (1967). On the relative nature of relational factors in classifications. *Information Storage & Retrieval, 3*(4), 315–329.

Li, H., & Abe, N. (1998). Generalizing case frames using a thesaurus and the MDL principle. *Computational Linguistics, 24*(2), 217–244.

Liddy, E. D., & Myaeng, S. H. (1993). DR-LINK's linguistic-conceptual approach to document detection. *The First Text REtrieval Conference (TREC-1)* (NIST Special Publication 500-207, pp. 1–20).

Litkowski, K. C. (2001). Syntactic clues and lexical resources in question-answering. *Ninth Text REtrieval Conference (TREC-9)* (NIST special publication 500-249, pp. 157–166).

Litkowski, K. C. (2002). CL research experiments in TREC-10 question answering. *Tenth Text Retrieval Conference (TREC 2001)* (NIST special publication 500-250, pp. 122–131). Retrieved December 31, 2004, from http://trec.nist.gov/pubs/trec10/t10_proceedings.html

Liu, G. Z. (1997). Semantic vector space model: Implementation and evaluation. *Journal of the American Society for Information Science, 48*(5), 395–417.

Liu, M. (1993). Progress in documentation: The complexities of citation practice: A review of citation studies. *Journal of Documentation, 49*, 370–408.

Longacre, R. E. (1996). *The grammar of discourse* (2nd ed.). New York: Plenum Press.

Longman dictionary of contemporary English. (1987). (2nd ed.) Edinburgh Gate, Harlow: Longman.

Longman dictionary of contemporary English. (1995). (3rd ed.) Edinburgh Gate, Harlow: Pearson Education.

Lu, X. (1990). *An application of case relations to document retrieval*. Unpublished doctoral dissertation, University of Western Ontario.

Lyons, J. (1968). *Introduction to theoretical linguistics*. Cambridge, UK: Cambridge University Press.

Lyons, J. (1977). *Semantics*. Cambridge, UK: Cambridge University Press.

Lyons, J. (1981). *Language and linguistics*. Cambridge, UK: Cambridge University Press.

Lyons, J. (1995). *Linguistic semantics: An introduction*. Cambridge, UK: Cambridge University Press.

Mackie, J. L. (1980). *The cement of the universe: A study of causation*. Oxford, UK: Oxford University Press.

Mandala, R., Tokunaga, T., & Tanoka, H. (1999). Combining multiple evidence from different types of thesaurus for query expansion. *Proceedings of the 22nd Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval*, 191–197.

Mandler, J. M. (1987). On the psychological reality of story structure. *Discourse Processes, 10*, 1–29.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9*(1), 111–151.

Mani, I., & Bloedorn, E. (1999). Summarizing similarities and differences among related documents. In I. Mani, & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 357–380). Cambridge, MA: MIT Press.

Mani, I., & Maybury, M. T. (Eds.). (1999). *Advances in automatic text summarization*. Cambridge, MA: MIT Press.

Mann, W. C., Matthiessen, C. M. I. M., & Thompson, S. A. (1992). Rhetorical structure theory and text analysis. In W. C. Mann & S. A. Thompson (Eds.), *Discourse description: Diverse linguistic analyses of a fund-raising text* (pp. 39–78). Amsterdam: John Benjamins.

Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text, 8*(3), 243–281.

Mann, W. C., & Thompson, S. A. (1989). Rhetorical structure theory: A theory of text organization. (Information Sciences Institute Research Report 87-190). Marina del Rey, CA: The Institute.

Manning, C. D. (1993). Automatic acquisition of a large subcategorization dictionary from corpora. *31st Annual Meeting of the Association for Computational Linguistics*, 235–242.

Marcu, D. (1999). Discourse trees are good indicators of importance in text. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 123–136). Cambridge, MA: MIT Press.

Marcu, D. (2000). *The theory and practice of discourse parsing and summarization*. Cambridge, MA: MIT Press.

Markowitz, J. (1988). An exploration into graded set membership. In M. W. Evens (Ed.), *Relational models of the lexicon: Representing knowledge in semantic networks* (pp. 239–260). Cambridge, UK: Cambridge University Press.

Markowitz, J., Nutter, T., & Evens, M. (1992). Beyond IS-A and part–whole: More semantic network links. In F. Lehmann & E. Y. Rodin (Eds.), *Semantic networks in artificial intelligence* (pp. 377–390). Oxford, UK: Pergamon.

Martin, P. (1995). Using the WordNet concept catalog and a relation hierarchy for knowledge acquisition. In *Proceedings of Peirce '95: 4th International Workshop on Peirce, University of California, Santa Cruz*. Retrieved December 31, 2004, from http://www.inria.fr/acacia/Publications/1995/peirce95phm.ps.Z

Martin, P. (1996). *The CGKAT top-level relation ontology*. Retrieved December 31, 2004, from http://www-sop.inria.fr/acacia/personnel/phmartin/ontologies/reCGKAT.html

Mauldin, M. L. (1991). Retrieval performance in FERRET: A conceptual information retrieval system. *Proceedings of the Fourteenth Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval*, 347–355.

McKeon, R. (Ed.). (1941/2001). *The basic works of Aristotle*. New York: Random House.

McKeown, K., & Radev, D. R. (1999). Generating summaries of multiple news articles. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 381–390). Cambridge, MA: MIT Press.

McNamara, T. P., & Holbrook, J. B. (2003). Semantic memory and priming. In I. B. Weiner (Ed.), *Handbook of psychology* (Vol. 4, pp. 447–474). Hoboken, NJ: Wiley.

Mel'cuk, I. (1996). Lexical functions: A tool for the description of lexical relations in a lexicon. In L. Wanner (Ed.), *Lexical functions in lexicography and natural language processing* (pp. 37–102). Amsterdam: John Benjamins.

Mel'cuk, I. A. (1988). *Dependency syntax: Theory and practice*. Albany: State University of New York Press.

Mellor, D. H. (1995). *The facts of causation*. London: Routledge.

Metzler, D. P., & Haas, S. W. (1989). The Constituent Object Parser: Syntactic structure matching for information retrieval. *ACM Transactions on Information Systems, 7*(3), 292–316.

Metzler, D. P., Haas, S. W., Cosic, C. L., & Weise, C. A. (1990). Conjunction, ellipsis, and other discontinuous constituents in the constituent object parser. *Information Processing & Management, 26*(1), 53–71.

Mill, J. S. (1974). *A system of logic: Ratiocinative and inductive* (books 1V–V1 and appendices). Toronto: University of Toronto Press.

Miller, G. A. (1995). WordNet: A lexical database. *Communications of the ACM, 38*(11), 39–41.

Miller, G. A., & Fellbaum, C. (1991). Semantic networks of English. *Cognition, 41*, 197–229.

Milstead, J. L. (2001). Standards for relationships between subject indexing terms. In C. A. Bean & R. Green (Eds.), *Relationships in the organization of knowledge* (pp. 53–66). Dordrecht, NL: Kluwer.

Minker, W., Bennacef, S., & Gauvain, J.-L. (1996). A stochastic case frame approach for natural language understanding. *Proceedings ICSLP 96: Fourth International Conference on Spoken Language Processing* (Vol. 2), 1013–1016.

Minsky, M. (1975). A framework for representing knowledge. In P. Winston (Ed.), *The psychology of computer vision* (pp. 211–280). New York: McGraw-Hill.

Mitchell, T. M. (1997). *Machine learning*. New York: McGraw-Hill.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*(3), 289–316.

Murphy, M. L. (2003). *Semantic relations and the lexicon: Antonymy, synonymy, and other paradigms*. Cambridge, UK: Cambridge University Press.

Muslea, I. (1999). Extraction patterns for information extraction tasks: A survey. In *AAAI–99 Workshop on Machine Learning for Information Extraction* (AAAI Workshop Technical Report WS-99-11, pp. 1–6). Menlo Park, CA: AAAI Press. Retrieved December 31, 2004, from http://www.isi.edu/muslea/ RISE/ML4IE

Myaeng, S. H., & Jang, D. (1999). Development and evaluation of a statistically based document summarization system. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 61–70). Cambridge, MA: MIT Press.

Myaeng, S. H., Khoo, C., & Li, M. (1994). Linguistic processing of text for a large-scale conceptual information retrieval system. *Conceptual Structures: Current Practices: Second International Conference on Conceptual Structures, ICCS '94*, 69–83.

Myaeng, S. H., & McHale, M. L. (1992). Toward a relation hierarchy for information retrieval. *Advances in Classification Research: Proceedings of the 2nd ASIS SIG/CR Classification Research Workshop*, 101–113.

Na, J. C., Sui, H., Khoo, C., Chan, S., & Zhou, Y. (2004). Effectiveness of simple linguistic processing in automatic sentiment classification of product reviews. *Proceedings of the Eighth International ISKO Conference*, 49–54.

Nasukawa, T., & Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the International Conference on Knowledge Capture* (pp. 70–77). New York: ACM.

National Information Standards Organization. (1994). *Guidelines for the construction, format, and management of monolingual thesauri*. Bethesda, MD: NISO Press.

(ANSI/NISO Z39.19–1993). Retrieved September 30, 2004, from http://www.niso.org/standards/resources/Z39-19.html

Nedellec, C. (2000). Corpus-based learning of semantic relations by the ILP system, Asium. In *Learning Language in Logic* (Lecture Notes in Artificial Intelligence, Vol. 1925, pp. 259–278). Berlin: Springer-Verlag.

Neelameghan, A. (1998). Lateral relations and links in multi-cultural, multimedia databases in the spiritual and religious domains: Some observations. *Information Studies, 4*, 221–246.

Neelameghan, A. (2001). Lateral relationships in multicultural, multilingual databases in the spiritual and religious domains: The OM Information Service. In C. A. Bean & R. Green (Eds.), *Relationships in the organization of knowledge* (pp. 185–198). Dordrecht, NL: Kluwer.

Neelameghan, A., & Maitra, R. (1978). *Non-hierarchical associative relationships among concepts: Identification and typology* (part A of FID/CR report no. 18). Bangalore, India: FID/CR Secretariat.

Nishida, F., & Takamatsu, S. (1982). Structured-information extraction from patent-claim sentences. *Information Processing & Management, 18*(1), 1–13.

Noy, N. F., & Hafner, C. D. (1997). The state of the art in ontology design: A survey and comparative review. *AI Magazine, 18*(3), 53–74. Retrieved March 11, 2004 from http://www.aaai.org/Library/Magazine/Vol18/18-03/vol18-03.html

Ogden, C. K. (1967). *Opposition: A linguistic and psychological analysis.* Bloomington: Indiana University Press.

Olson, M. W., & Gee, T. C. (1988). Understanding narratives: A review of story grammar research. *Childhood Education, 64*(5), 302–306.

Owens, D. (1992). *Causes and coincidences.* Cambridge, UK: Cambridge University Press.

Paice, C. D. (1990). Constructing literature abstracts by computer: Techniques and prospects. *Information Processing & Management, 26*(1), 171–186.

Paranjpe, D., Ramakrishnan, G., & Srinivasan, S. (2004). Passage scoring for question answering via Bayesian inference on lexical relations. *Twelfth Text Retrieval Conference (TREC 2003)* (NIST Special Publication 500-255, pp. 305–310). Retrieved December 31, 2004, from http://trec.nist.gov/pubs/trec12/t12_proceedings.html

Park, H. (1997). Relevance of science information: Origins and dimensions of relevance and their implications to information retrieval. *Information Processing & Management, 33*(3), 339–352.

Peat, H. J., & Willett, P. (1991). The limitations of term co-occurrence data for query expansion in document retrieval systems. *Journal of the American Society for Information Science, 42*(5), 378–383.

Pereira, F., Tishby, N., & Lee, L. (1993). Distributional clustering of English words. *31st Annual Meeting of the Association for Computational Linguistics*, 183–190.

Perez-Carballo, J., & Strzalkowski, T. (2000). Natural language information retrieval: Progress report. *Information Processing & Management 36*, 155–178.

Pribbenow, S. (2002). Meronymic relationships: From classical mereology to complex part–whole relations. In R. Green, C. Bean, & S. H. Myaeng (Eds.), *The semantics of relationships: An interdisciplinary perspective* (pp. 35–50). Dordrecht, NL: Kluwer.

Priss, U. (1999). Efficient implementation of semantic relations in lexical databases. *Computational Intelligence, 15*(1), 79–87.

Qiu, Y., & Frei, H. P. (1993). Concept Based Query Expansion. *Proceedings of the Sixteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 160–169.

Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science, 12*, 410–430.

Quillian, M. R. (1968). Semantic memory. In M. L. Minsky (Ed.), *Semantic information processing* (pp. 227–270). Cambridge, MA: MIT Press.

Quine, W. V. (1982). *Methods of logic* (4th ed.). Cambridge, MA: Harvard University Press.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann.

Rada, R., Barlow, J., Potharst, J., Zanstra, P., & Bijstra, D. (1991). Document ranking using an enriched thesaurus. *Journal of Documentation, 47*(3), 240–253.

Rada, R., & Bicknell, E. (1989). Ranking documents with a thesaurus. *Journal of the American Society for Information Science, 40*(5), 304–310.

Radev, D. (2000). A common theory of information fusion from multiple text sources step one: Cross-document structure. *Proceedings of the 1st SIGdial Workshop on Discourse and Dialogue*. Retrieved December 31, 2004, from http://www.sigdial.org/workshops/workshop1/proceedings/radev.pdf

Radev, D. R., Hovy, E., & McKeown, K. (2002). Introduction to the special issue on summarization. *Computational Linguistics, 28*(4), 399–408.

Raitt, D. I. (1980). Recall and precision devices in interactive bibliographic search and retrieval systems. *Aslib Proceedings, 37*, 281–301.

Ranganathan, S. R. (1965). *The Colon Classification*. New Brunswick, NJ: Graduate School of Library Service, Rutgers University.

Rau, L. (1987). Knowledge organization and access in a conceptual information system. *Information Processing & Management, 23*(4), 269–283.

Rau, L. F., Jacobs, P. S., & Zernik, U. (1989). Information extraction and text summarization using linguistic knowledge acquisition. *Information Processing & Management, 25*(4), 419–428.

Raybeck, D., & Herrmann, D. (1990). A cross-cultural examination of semantic relations. *Journal of Cross-Cultural Psychology, 21*(4), 452–473.

Rehder, B. (2003). Categorization as causal reasoning. *Cognitive Science, 27*, 709–748.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General, 130*(3), 323–360.

Reimer, U., & Hahn, U. (1988). Text condensation as knowledge base abstraction. *Proceedings of the 4th Conference on Artificial Intelligence Applications (CAIA'88)*, 338–344.

Riloff, E. (1993). Automatically constructing a dictionary for information extraction tasks. *Proceedings of the Eleventh National Conference on Artificial Intelligence*, 811–816.

Riloff, E. (1996). Automatically generating extraction patterns from untagged text. *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI-96)* (Vol. 2), 1044–1049.

Rindflesch, T. C., & Aronson, A. R. (2002). Semantic processing for enhanced access to biomedical knowledge. In V. Kashyap & L. Shklar (Eds.), *Real world Semantic Web applications* (pp. 157–172). Amsterdam: IOS Press. Retrieved December 31, 2004, from http://nls5.nlm.nih.gov/pubs/semwebapp.5a.pdf

Rindflesch, T. C., Jayant, R., & Lawrence, H. (2000). Extracting molecular binding relationships from biomedical text. *Proceedings of the 6th Applied Natural Language Processing Conference*, 188–195.

Rindflesch, T. C., Libbus, B., Hristovski, D., Aronson, A. R., & Kilicoglu, H. (2003). Semantic relations asserting the etiology of genetic diseases. *Proceedings of 2003 AMIA Annual Symposium*, 554–558. Retrieved December 31, 2004, from http://skr.nlm.nih.gov/papers/references/sgn_amia03.final.pdf

Ringland, G. A., & Duce, D. A. (Eds.). (1988). *Approaches to knowledge representation: An introduction*. Letchworth, Herts.: Research Studies Press.

Roget's international thesaurus (3rd ed.). (1962). New York: Thomas Y. Crowell.

Romney, A. K., Moore, C. C., & Rusch, C. D. (1997). Cultural universals: Measuring the semantic structure of emotion terms in English and Japanese. *Proceedings of the National Academy of Sciences of the United States, 94*, 5489–5494.

Rosner, M., & Somers, H. L. (1980). Case in linguistics and cognitive science. *UEA Papers in Linguistics 13*, 1–29.

Roth, D., & Yih, W. (2001). Relational learning via propositional algorithms: An information extraction case study. *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, 1257–1263.

Ruge, G. (1992). Experiments on linguistically-based term associations. *Information Processing & Management, 28*(3), 317–332.

Ruge, G., Schwarz, C., & Warner, A. J. (1991). Effectiveness and efficiency in natural language processing for large amounts of text. *Journal of the American Society for Information Science, 42*(6), 450–456.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211–236). New York: Academic Press.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99–135). Hillsdale, NJ: Erlbaum.

Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton, NJ: Princeton University Press.

Salton, G., Singhal, A., Mitra, M., & Buckley, C. (1999). Automatic text structuring and summarization. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 341–356). Cambridge, MA: MIT Press.

Sattler, U. (1995). A concept language for an engineering application with part–whole relations. In A. Borgida, M. Lenzerini, D. Nardi, & B. Nebel (Eds.), *Proceedings of the 1995 International Workshop on Description Logics* (pp. 119–123). Rome: Universita degli Studi di Roma "La Sapienza", Dipartimento di Informatica e Sistemistica.

Saussure, F. de. (1959). *Course in general linguistics* (C. Bally & A. Sechehaye, Eds.; W. Baskin, Trans). New York: McGraw-Hill.

Schamber, L. (1991). *Users' criteria for evaluation in multimedia information seeking and use situations*. Unpublished doctoral dissertation, Syracuse University.

Schamber, L. (1994). Relevance and information behavior. *Annual Review of Information Science and Technology, 29*, 3–48.

Schank, R. C. (1982). *Dynamic memory*. New York: Cambridge University Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.

Schneider, P., & Winship, S. (2002). Judgments of fictional story quality. *Journal of Speech, Language, & Hearing Research, 45*(2), 372–383.

Schwarz, C. (1990). Automatic syntactic analysis of free text. *Journal of the American Society for Information Science, 41*(6), 408–417.

Smadja, F. (1993). Retrieving collocations from text: Xtract. *Computational Linguistics, 19*(1), 143–177.

Smeaton, A. F., O'Donnell, R., & Kelledy, F. (1995). Indexing structures derived from syntax in TREC-3: System description. *Overview of the Third Text REtrieval Conference (TREC-3)* (NIST Special Publication 500-225, pp. 55–67).

Smeaton, A. F., & van Rijsbergen, C. J. (1988). Experiments on incorporating syntactic processing of user queries into a document retrieval strategy. *11th Annual International ACM / SIGIR Conference on Research and Development in Information Retrieval*, 31–51.

Smith, E. E. (1978). Theories of semantic memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes. Vol. 4, Linguistic functions in cognitive theory* (pp. 1–56). Potomac, MD: Erlbaum.

Smith, R. N. (1981). On defining adjectives, part III. *Dictionaries: Journal of the Dictionary Society of North America, 3*, 28–38.

Soderland, S. (1997). *Learning text analysis rules for domain-specific natural language processing*. Unpublished doctoral dissertation (Technical report UM–CS–1996–087), University of Massachusetts, Amherst.

Soderland, S. (1999). Learning information extraction rules for semi-structured and free text. *Machine Learning, 34* (1–3), 233–272.

Soderland, S., Fisher, D., Aseltine, J., & Lehnert, W. (1996). Issues in inductive learning of domain-specific text extraction rules. In S. Wermter, E. Riloff, & G. Scheler (Eds.), *Connectionist, Statistical and Symbolic Approaches to Learning for Natural Language Processing* (pp. 290–301). Berlin: Springer-Verlag.

Somers, H. L. (1987). *Valency and case in computational linguistics*. Edinburgh, UK: Edinburgh University Press.

Sosa, E., & Tooley, M. (Eds.). (1993). *Causation*. Oxford, UK: Oxford University Press.

Sowa, J. F. (1984). *Conceptual structures: Information processing in mind and machine*. Reading, MA: Addison-Wesley.

Sowa, J. F. (2000). *Knowledge representation: Logical, philosophical, and computational foundations*. Pacific Grove, CA: Brooks/Cole.

Sowa, J. F. (2001). *Roles and relations*. Retrieved December 31, 2004, from http://www.jfsowa.com/ontology/roles.htm

Sparck Jones, K. (1971). *Automatic keyword classification for information retrieval*. London: Butterworth.

Sparck Jones, K. (2000). Further reflections on TREC. *Information Processing & Management, 36*, 37–85.

Spellman, B. A., Holyoak, K., & Morrison, R. G. (2001). Analogical priming via semantic relations. *Memory & Cognition, 29*(3), 383–393.

Srinivasan, P., & Rindflesch, T. C. (2002). Exploring text mining from MEDLINE. In *Proceedings of the 2002 AMIA Annual Symposium* (pp. 722–726). Bethesda, MD:

American Medical Informatics Association. Retrieved December 31, 2004 from http://nls5.nlm.nih.gov/pubs/amia02-TR2.pdf

Stasio, T., Herrmann, D. J., & Chaffin, R. (1985). Predictions of relation similarity according to relation definition theory. *Bulletin of the Psychonomic Society, 23*, 5–8.

Stephens, L. M., & Chen, Y. F. (1996). Principles for organizing semantic relations in large knowledge bases. *IEEE Transactions on Knowledge and Data Engineering, 8*(3), 492–496.

Strzalkowski, T. (1995). Natural language information retrieval. *Information Processing & Management, 31*(3), 397–417.

Strzalkowski, T., Carballo, J. P., & Marinescu, M. (1995). Natural language information retrieval: TREC-3 report. *Overview of the Third Text REtrieval Conference (TREC-3)* (NIST Special Publication 500–225, pp. 39–53).

Strzalkowski, T., Stein, G., Wang, J., & Wise, B. (1999). A robust practical text summarizer. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 137–154). Cambridge, MA: MIT Press.

Takemura, T., & Ashida, N. (2002). A study of the medical record interface to natural language processing. *Journal of Medical Systems, 26*(2), 79–87.

Tang, R., & Solomon, P. (2001). Use of relevance criteria across stages of document evaluation: On the complementarity of experimental and naturalistic studies. *Journal of the American Society for Information Science and Technology, 52*(8), 676–685.

Terenziani, P., & Torasso, P. (1995). Time, action-types, and causation: An integrated analysis. *Computational Intelligence, 11*(3), 529–552.

Teufel, S., & Moens, M. (1999). Argumentative classification of extracted sentences as a first step towards flexible abstracting. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 155–175). Cambridge, MA: MIT Press.

Trybula, W. J. (1999). Text mining. *Annual Review of Information Science and Technology, 34*, 385–419.

Turner, M. (1993). An image-schematic constraint on metaphor. In R. A. Geiger & B. Rudzka-Ostyn (Eds.), *Conceptualizations and mental processing in language* (pp. 291–306). Berlin: Mouton de Gruyter.

Tversky, B. (1990). Where partonomies and taxonomies meet. In S. L. Tsohatzidis (Ed.), *Meanings and prototypes: Studies in linguistic categorization* (pp. 334–344). New York: Routledge.

U.S. National Library of Medicine. (2004). *Unified Medical Language System: Documentation*. Retrieved December 31, 2004 from http://www.nlm.nih.gov/research/umls/documentation.html

Uschold, M. (1996). The use of the typed lambda calculus for guiding naïve users in the representation and acquisition of part–whole knowledge. *Data and Knowledge Engineering, 20*, 385–404.

Van Dijk, T. A. (1972). *Some aspects of text grammars: A study in theoretical linguistics and poetics*. The Hague: Mouton.

Van Dijk, T. A. (1980). *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition*. Hillsdale, NJ: Erlbaum.

Van Dijk, T. A. (1988). *News as discourse*. Hillsdale, NJ: Erlbaum.

Vickery, B. (1996). Conceptual relations in information systems [letter to the editor]. *Journal of Documentation, 52*, 198–200.

Voorhees, E. M. (1994). Query expansion using lexical-semantic relations. *Proceedings of the Seventeenth Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval*, 61–69.

Voorhees, E. M. (2003). Overview of the TREC 2003 Question Answering Track. *Twelfth Text Retrieval Conference (TREC 2003)* (NIST Special Publication, SP 500-255, pp. 54–68). Retrieved December 31, 2004, from http://trec.nist.gov/pubs/trec12/t12_proceedings.html

Vossen, P. (1998). Introduction to EuroWordNet. *Computers and the Humanities, 32*(2–3), 73–89.

Wan, T. L., Evens, M., Wan, Y. W., & Pao, Y. Y. (1997). Experiments with automatic indexing and a relational thesaurus in a Chinese information retrieval system. *Journal of the American Society for Information Science, 48*(12), 1086–1096.

Wang, Y.-C., Vandendorpe, J. & Evens, M. (1985). Relational thesauri in information retrieval. *Journal of the American Society for Information Science, 36*(1), 15–27.

Wanner, L. (1996). Introduction. In L. Wanner (Ed.), *Lexical functions in lexicography and natural language processing* (pp. 1–36). Amsterdam: John Benjamins.

Wanner, L. (2004). Towards automatic fine-grained semantic classification of verb-noun collocations. *Natural Language Engineering, 10*(2), 95–143.

Warren, H. C. (1921). *A history of the association psychology*. New York: Scribner's.

Warren, W. H., Nicholas, D. W., & Trabasso, T. (1979). Event chains and inferences in understanding narratives. In R.O. Freedle (Ed.), *Advances in discourse processes, Vol. 2: New directions in discourse processing* (pp. 23–52). Norwood, NJ: Ablex.

Wattenmaker, W. D., Nakamura, G. V., & Medin, D. L. (1988). Relationships between similarity-based and explanation-based categorization. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 204–240). Washington Square, NY: New York University Press.

Werner, O. (1988). How to teach a network: Minimal design features for a cultural acquisition device or C-KAD. In M. W. Evens (Ed.), *Relational models of the lexicon: Representing knowledge in semantic networks* (pp. 141–166). Cambridge, UK: Cambridge University Press.

White, M., Korelsky, T., Cardie, C, Ng, V., Pierce, D., & Wagstaff, K. (2001). Multi-document summarization via information extraction. *First International Conference on Human Language Technology Research (HLT-01)*. Retrieved March 25, 2005, from http://www.hlt.utdallas.edu/~vince/papers/hlt01.html

Whitney, P., Budd, D., Bramucci, R. S., & Crane, R. S. (1995). On babies, bathwater, and schemata: A reconsideration of top-down processes in comprehension. *Discourse Processes, 20*, 135–166.

Wille, R. (2003). Truncated distributive lattices: Conceptual structures of simple-implicational theories. *Order, 20*, 229–238.

Winston, M., Chaffin, R., & Herrmann, D. (1987). A taxonomy of part–whole relations. *Cognitive Science, 11*(4), 417–444.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). New York: Macmillan.

World Wide Web Consortium (2004a). *OWL Web Ontology Language: Use cases and requirements* (W3C Recommendation 10 February 2004). Retrieved March 11, 2004, from http://www.w3.org/TR/2004/REC-webont-req-20040210

World Wide Web Consortium. (2004b). *OWL Web Ontology Language overview* (W3C Recommendation 10 Feb 2004). Retrieved March 18, 2004, from http://www.w3.org/TR/2004/REC-owl-features-20040210

World Wide Web Consortium (2004c). *Semantic Web*. Retrieved September 1, 2004, from http://www.w3.org/2001/sw

World Wide Web Consortium. (2004d). *Web services architecture requirements* (W3C Working Group Note 11 February 2004). Retrieved December 1, 2004, from http://www.w3.org/TR/wsa-reqs

Xu, J., & Croft, W. B. (1996). Query expansion using local and global document analysis. *Proceedings of the19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 4–11.

Xu, J., & Croft, W. B. (2000). Improving the effectiveness of information retrieval with local context analysis. *ACM Transactions on Information Systems, 18*(1), 79–112.

Xu, Y., Araki, M., & Niimi, Y. (2003). A multilingual-supporting dialog system across multiple domains. *Acoustical Science and Technology, 24*(6), 349–357.