Chapter 19
Lexical Semantics

"When I use a word", Humpty Dumpty said in rather a scornful tone, "it means just what I choose it to mean – neither more nor less."

Lewis Carroll, Alice in Wonderland

How many legs does a dog have if you call its tail a leg?
Four.
Calling a tail a leg doesn't make it one.

Attributed to Abraham Lincoln

The previous two chapters focused on the representation of meaning representations for entire sentences. In those discussions, we made a simplifying assumption by representing *word meanings* as unanalyzed symbols like EAT or JOHN or RED. But representing the meaning of a word by capitalizing it is a pretty unsatisfactory model. In this chapter we introduce a richer model of the semantics of words, drawing on the linguistic study of word meaning, a field called **lexical semantics**.

Lexical semantics

Before we try to define *word meaning* in the next section, we first need to be clear on what we mean by *word*, since we have used the word *word* in many different ways in this book.

Lexeme
Lexicon

We can use the word **lexeme** to mean a pairing of a particular form (orthographic or phonological) with its meaning, and a **lexicon** is a finite list of lexemes. For the purposes of lexical semantics, particularly for dictionaries and thesauruses, we represent a lexeme by a **lemma**. A **lemma** or **citation form** is the grammatical form that is used to represent a lexeme. This is often the base form; thus *carpet* is the lemma for *carpets*. The lemma or citation form for *sing*, *sang*, *sung* is *sing*. In many languages the infinitive form is used as the lemma for the verb; thus in Spanish *dormir* 'to sleep' is the lemma for verb forms like *duermes* 'you sleep'. The specific forms *sung* or *carpets* or *sing* or *duermes* are called **wordforms**.

Lemma
Citation form
Wordform
Lemmatization

The process of mapping from a wordform to a lemma is called **lemmatization**. Lemmatization is not always deterministic, since it may depend on the context. For example, the wordform *found* can map to the lemma *find* (meaning 'to locate') or the lemma *found* ('to create an institution'), as illustrated in the following WSJ examples:

(19.1) He has looked at 14 baseball and football stadiums and **found** that only one – private Dodger Stadium – brought more money into a city than it took out.

(19.2) Culturally speaking, this city has increasingly displayed its determination to **found** the sort of institutions that attract the esteem of Eastern urbanites.

In addition, lemmas are part-of-speech specific; thus the wordform *tables* has two possible lemmas, the noun *table* and the verb *table*.

One way to do lemmatization is via the morphological parsing algorithms of Ch. 3. Recall that morphological parsing takes a surface form like *cats* and produces *cat +PL*. But a lemma is not necessarily the same as the stem from the morphological parse. For example, the morphological parse of the word *celebrations* might produce the stem **celebrate** with the affixes *-ion* and *-s*, while the lemma for *celebrations* is the longer form **celebration**. In general lemmas may be larger than morphological stems (e.g., *New York* or *throw up*). The intuition is that we want to have a different lemma whenever we need to have a completely different dictionary entry with its own meaning representation; we expect to have *celebrations* and *celebration* share an entry, since the difference in their meanings is mainly just grammatical, but not necessarily to share one with *celebrate*.

In the remainder of this chapter, when we refer to the meaning (or meanings) of a 'word', we will generally be referring to a lemma rather than a wordform.

Now that we have defined the locus of word meaning, we will proceed to different ways to represent this meaning. In the next section we introduce the idea of **word sense** as the part of a lexeme that represents word meaning. In following sections we then describe ways of defining and representing these senses, as well as introducing the lexical semantic aspects of the events defined in Ch. 17.

19.1 Word Senses

The meaning of a lemma can vary enormously given the context. Consider these two uses of the lemma *bank*, meaning something like 'financial institution' and 'sloping mound', respectively:

(19.3) Instead, a *bank* can hold the investments in a custodial account in the client's name.

(19.4) But as agriculture burgeons on the east *bank*, the river will shrink even more.

We represent some of this contextual variation by saying that the lemma *bank* has two **senses**. A **sense** (or **word sense**) is a discrete representation of one aspect of the meaning of a word. Loosely following lexicographic tradition, we will represent each sense by placing a superscript on the orthographic form of the lemma as in **bank**$^{\mathbf{1}}$ and **bank**$^{\mathbf{2}}$. [1]

Word sense

The senses of a word might not have any particular relation between them; it may be almost coincidental that they share an orthographic form. For example, the *financial institution* and *sloping mound* senses of bank seem relatively unrelated. In such cases we say that the two senses are **homonyms**, and the relation between the senses is one of **homonymy**. Thus **bank**$^{\mathbf{1}}$ ('financial institution') and **bank**$^{\mathbf{2}}$ ('sloping mound') are homonyms.

Homonym

Homonymy

Sometimes, however, there is some semantic connection between the senses of a word. Consider the following WSJ 'bank' example:

(19.5) While some *banks* furnish sperm only to married women, others are much less restrictive.

[1] Confusingly, the word "lemma" is itself very ambiguous; it is also sometimes used to mean these separate senses, rather than the citation form of the word. You should be prepared to see both uses in the literature.

Although this is clearly not a use of the 'sloping mound' meaning of *bank*, it just as clearly is not a reference to a promotional giveaway at a financial institution. Rather, *bank* has a whole range of uses related to repositories for various biological entities, as in *blood bank*, *egg bank*, and *sperm bank*. So we could call this 'biological repository' sense **bank**^{3}. Now this new sense **bank**^{3} has some sort of relation to **bank**^{1}; both **bank**^{1} and **bank**^{3} are repositories for entities that can be deposited and taken out; in **bank**^{1} the entity is money, where in **bank**^{3} the entity is biological.

Polysemy

When two senses are related semantically, we call the relationship between them **polysemy** rather than homonymy. In many cases of polysemy the semantic relation between the senses is systematic and structured. For example consider yet another sense of *bank*, exemplified in the following sentence:

(19.6) The bank is on the corner of Nassau and Witherspoon.

This sense, which we can call **bank**^{4}, means something like 'the building belonging to a financial institution'. It turns out that these two kinds of senses (an organization, and the building associated with an organization) occur together for many other words as well (*school*, *university*, *hospital*, etc). Thus there is a systematic relationship between senses that we might represent as

BUILDING ↔ ORGANIZATION

Metonymy

This particular subtype of polysemy relation is often called **metonymy**. Metonymy is the use of one aspect of a concept or entity to refer to other aspects of the entity, or to the entity itself. Thus we are performing metonymy when we use the phrase *the White House* to refer to the administration whose office is in the White House. Other common examples of metonymy include the relation between the following pairings of senses:

Author (*Jane Austen wrote Emma*) ↔ Works of Author (*I really love Jane Austen*)
Animal (*The chicken was domesticated in Asia*) ↔ Meat (*The chicken was overcooked*)
Tree (*Plums have beautiful blossoms*) ↔ Fruit (*I ate a preserved plum yesterday*)

While it can be useful to distinguish polysemy from homonymy, there is no hard threshold for 'how related' two senses have to be to be considered polysemous. Thus the difference is really one of degree. This fact can make it very difficult to decide how many senses a word has, i.e., whether to make separate senses for closely related usages. There are various criteria for deciding that the differing uses of a word should be represented as distinct discrete senses. We might consider two senses discrete if they have independent truth conditions, different syntactic behavior, independent sense relations, or exhibit antagonistic meanings.

Consider the following uses of the verb *serve* from the WSJ corpus:

(19.7) They rarely *serve* red meat, preferring to prepare seafood, poultry or game birds.

(19.8) He *served* as U.S. ambassador to Norway in 1976 and 1977.

(19.9) He might have *served* his time, come out and led an upstanding life.

The *serve* of *serving red meat* and that of *serving time* clearly have different truth conditions and presuppositions; the *serve* of *serve as ambassador* has the distinct subcategorization structure *serve as NP*. These heuristic suggests that these are probably three distinct senses of *serve*. One practical technique for determining if two senses are distinct is to conjoin two uses of a word in a single sentence; this kind of conjunction of antagonistic readings is called **zeugma**. Consider the following ATIS examples:

Zeugma

(19.10) Which of those flights serve breakfast?

(19.11) Does Midwest Express serve Philadelphia?

(19.12) ?Does Midwest Express serve breakfast and Philadelphia?

We use (?) to mark example those that are semantically ill-formed. The oddness of the invented third example (a case of zeugma) indicates there is no sensible way to make a single sense of *serve* work for both breakfast and Philadelphia. We can use this as evidence that *serve* has two different senses in this case.

Dictionaries tend to use many fine-grained senses so as to capture subtle meaning differences, a reasonable approach given that traditional role of dictionaries in aiding word learners. For computational purposes, we often don't need these fine distinctions and so we may want to group or cluster the senses; we have already done this for some of the examples in this chapter.

We generally reserve the word **homonym** for two senses which share both a pronunciation and an orthography. A special case of multiple senses that causes problems for speech recognition and spelling correction is **homophones**. **Homophones** are senses that are linked to lemmas with the same pronunciation but different spellings, such as *wood*/*would* or *to*/*two*/*too*. A related problem for speech synthesis are **homographs** (Ch. 8). **Homographs** are distinct senses linked to lemmas with the same orthographic form but different pronunciations, such as these homographs of *bass*:

Homophone

Homograph

(19.13) The expert angler from Dora, Mo., was fly-casting for **bass** rather than the traditional trout.

(19.14) The curtain rises to the sound of angry dogs baying and ominous **bass** chords sounding.

How can we define the meaning of a word sense? Can we just look in a dictionary? Consider the following fragments from the definitions of *right*, *left*, *red*, and *blood* from the *American Heritage Dictionary* (Morris, 1985).

right	*adj.* located nearer the right hand esp. being on the right when facing the same direction as the observer.
left	*adj.* located nearer to this side of the body than the right.
red	*n.* the color of blood or a ruby.
blood	*n.* the red liquid that circulates in the heart, arteries and veins of animals.

Note the amount of circularity in these definitions. The definition of *right* makes two direct references to itself, while the entry for *left* contains an implicit self-reference in the phrase *this side of the body*, which presumably means the *left* side. The entries for *red* and *blood* avoid this kind of direct self-reference by instead referencing each other in their definitions. Such circularity is, of course, inherent in all dictionary definitions; these examples are just extreme cases. For humans, such entries are still useful since the user of the dictionary has sufficient grasp of these other terms.

For computational purposes, one approach to defining a sense is to make use of a similar approach to these dictionary definitions; defining a sense via its relationship with other senses. For example, the above definitions make it clear that *right* and *left* are similar kinds of lemmas that stand in some kind of alternation, or opposition, to one another. Similarly, we can glean that *red* is a color, it can be applied to both *blood* and *rubies*, and that *blood* is a *liquid*. **Sense relations** of this sort are embodied in on-line databases like **WordNet**. Given a sufficiently large database of such relations, many applications are quite capable of performing sophisticated semantic tasks (even if they do not *really* know their right from their left).

A second computational approach to meaning representation is to create a small finite set of semantic primitives, atomic units of meaning, and then create each sense definition out of these primitives. This approach is especially common when defining aspects of the meaning of *events* such as *semantic roles*.

We will explore both of these approaches to meaning in this chapter. In the next section we introduce various relations between senses, followed by a discussion of WordNet, a sense relation resource. We then introduce a number of meaning representation approaches based on semantic primitives such as semantic roles.

19.2 Relations between Senses

This section explores some of the relations that hold among word senses, focusing on a few that have received significant computational investigation: **synonymy**, **antonymy**, and **hypernymy**, as well as a brief mention of other relations like **meronymy**.

19.2.1 Synonymy and Antonymy

Synonym

When the meaning of two senses of two different words (lemmas) are identical or nearly identical we say the two senses are **synonyms**. Synonyms include such pairs as:

couch/sofa *vomit/throw up* *filbert/hazelnut* *car/automobile*

Propositional meaning

A more formal definition of synonymy (between words rather than senses) is that two words are synonymous if they are substitutable one for the other in any sentence without changing the truth conditions of the sentence. We often say in this case that the two words have the same **propositional meaning**.

While substitutions between some pairs of words like *car*/*automobile* or *water*/H_2O are truth-preserving, the words are still not identical in meaning. Indeed, probably no two words are absolutely identical in meaning, and if we define synonymy as identical meanings and connotations in all contexts, there are probably no absolute synonyms. Many other facets of meaning that distinguish these words are important besides propositional meaning. For example H_2O is used in scientific contexts, and would be inappropriate in a hiking guide; this difference in genre is part of the meaning of the word. In practice the word *synonym* is therefore commonly used to describe a relationship of approximate or rough synonymy.

Instead of talking about two *words* being synonyms, in this chapter we will define synonymy (and other relations like hyponymy and meronymy) as a relation between senses rather than between words. We can see the usefulness of this by considering the words *big* and *large*. These may seem to be synonyms in the following ATIS sentences, since we could swap *big* and *large* in either sentence and retain the same meaning:

(19.15) How big is that plane?

(19.16) Would I be flying on a large or small plane?

But note the following WSJ sentence where we cannot substitute *large* for *big*:

(19.17) Miss Nelson, for instance, became a kind of big sister to Benjamin.

(19.18) ?Miss Nelson, for instance, became a kind of large sister to Benjamin.

That is because the word *big* has a sense that means being older, or grown up, while *large* lacks this sense. Thus it will be convenient to say that some senses of *big* and *large* are (nearly) synonymous while other ones are not.

Antonym

Synonyms are words with identical or similar meanings. **Antonyms**, by contrast, are words with opposite meaning such as the following:

long/short	*big/little*	*fast/slow*	*cold/hot*	*dark/light*
rise/fall	*up/down*	*in/out*		

It is difficult to give a formal definition of antonymy. Two senses can be antonyms if they define a binary opposition, or are at opposite ends of some scale. This is the case for *long*/*short*, *fast/slow*, or *big*/*little*, which are at opposite ends of the *length* or *size* scale. Another groups of antonyms is **reversives**, which describe some sort of change or movement in opposite directions, such as *rise/fall* or *up/down*.

From one perspective, antonyms have very different meanings, since they are opposite. From another perspective, they have very similar meanings, since they share almost all aspects of their meaning except their position on a scale, or their direction. Thus automatically distinguishing synonyms from antonyms can be difficult.

19.2.2 Hyponymy

Hyponym

Hypernym

Superordinate

One sense is a **hyponym** of another sense if the first sense is more specific, denoting a subclass of the other. For example, *car* is a hyponym of *vehicle*; *dog* is a hyponym of *animal*, and *mango* is a hyponym of *fruit*. Conversely, we say that *vehicle* is a **hypernym** of *car*, and *animal* is a hypernym of *dog*. It is unfortunate that the two words (hypernym and hyponym) are very similar and hence easily confused; for this reason the word **superordinate** is often used instead of **hypernym**.

Superordinate	vehicle	fruit	furniture	mammal
Hyponym	car	mango	chair	dog

We can define hypernymy more formally by saying that the class denoted by the superordinate extensionally includes the class denoted by the hyponym. Thus the class of animals includes as members all dogs, and the class of moving actions includes all walking actions. Hypernymy can also be defined in terms of entailment. Under this definition, a sense A is a hyponym of a sense B if everything that is A is also B

and hence being an A entails being a B, or $\forall x \; A(x) \Rightarrow B(x)$. Hyponymy is usually a transitive relation; if A is a hyponym of B and B is a hyponym of C, then A is a hyponym of C.

The concept of hyponymy is closely related to a number of other notions that play central roles in computer science, biology, and anthropology and computer science. The term **ontology** usually refers to a set of distinct objects resulting from an analysis of a domain, or **microworld**. A **taxonomy** is a particular arrangement of the elements of an ontology into a tree-like class inclusion structure. Normally, there are a set of well-formedness constraints on taxonomies that go beyond their component class inclusion relations. For example, the lexemes *hound*, *mutt*, and *puppy* are all hyponyms of *dog*, as are *golden retriever* and *poodle*, but it would be odd to construct a taxonomy from all those pairs since the concepts motivating the relations is different in each case. Instead, we normally use the word **taxonomy** to talk about the hypernymy relation between *poodle* and *dog*; by this definition **taxonomy** is a subtype of hypernymy.

Ontology
Taxonomy

19.2.3 Semantic Fields

So far we've seen the relations of synonymy, antonymy, hypernomy, and hyponymy. Another very common relation is **meronymy**, the **part-whole** relation. A *leg* is part of a *chair*; a *wheel* is part of a *car*. We say that *wheel* is a **meronym** of *car*, and *car* is a **holoynm** of *wheel*.

Meronymy
Part-whole
Meronym
Holonym

But there is a more general way to think about sense relations and word meaning. Where the relations we've defined so far have been binary relations between two senses, a **semantic field** is a model of a more integrated, or holistic, relationship among entire sets of words from a single domain. Consider the following set of words:

Semantic field

reservation, flight, travel, buy, price, cost, fare, rates, meal, plane

We could assert individual lexical relations of hyponymy, synonymy, and so on between many of the words in this list. The resulting set of relations does not, however, add up to a complete account of how these words are related. They are clearly all defined with respect to a coherent chunk of common sense background information concerning air travel. Background knowledge of this kind has been studied under a variety of frameworks and is known variously as a **frame** (Fillmore, 1985), **model** (Johnson-Laird, 1983), or **script** (Schank and Albelson, 1977), and plays a central role in a number of computational frameworks.

We will discuss in Sec. 19.4.5 the **FrameNet** project (Baker et al., 1998), which is an attempt to provide a robust computational resource for this kind of frame knowledge. In the FrameNet representation, each of the words in the frame is defined with respect to the frame, and shares aspects of meaning with other frame words.

19.3 WordNet: A Database of Lexical Relations

WordNet

The most commonly used resource for English sense relations is the **WordNet** lexical database (Fellbaum, 1998). WordNet consists of three separate databases, one each

The noun "bass" has 8 senses in WordNet.
1. bass^{1} - (the lowest part of the musical range)
2. bass^{2}, bass part^{1} - (the lowest part in polyphonic music)
3. bass^{3}, basso^{1} - (an adult male singer with the lowest voice)
4. sea bass^{1}, bass^{4} - (the lean flesh of a saltwater fish of the family Serranidae)
5. freshwater bass^{1}, bass^{5} - (any of various North American freshwater fish with lean flesh (especially of the genus Micropterus))
6. bass^{6}, bass voice^{1}, basso^{2} - (the lowest adult male singing voice)
7. bass^{7} - (the member with the lowest range of a family of musical instruments)
8. bass^{8} - (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

The adjective "bass" has 1 sense in WordNet.
1. bass^{1}, deep^{6} - (having or denoting a low vocal or instrumental range) *"a deep voice"; "a bass voice is lower than a baritone voice"; "a bass clarinet"*

Figure 19.1 A portion of the WordNet 3.0 entry for the noun *bass*.

for nouns and verbs, and a third for adjectives and adverbs; closed class words are not included. Each database consists of a set of lemmas, each one annotated with a set of senses. The WordNet 3.0 release has 117,097 nouns, 11,488 verbs, 22,141 adjectives, and 4,601 adverbs. The average noun has 1.23 senses, and the average verb has 2.16 senses. WordNet can be accessed via the web or downloaded and accessed locally.

Gloss

Parts of a typical lemma entry for the noun and adjective *bass* are shown in Fig. 19.1. Note that there are eight senses for the noun and one for the adjective, each of which has a **gloss** (a dictionary-style definition), a list of synonyms for the sense (called a **synset**), and sometimes also usage examples (shown for the adjective sense). Unlike dictionaries, WordNet doesn't represent pronunciation, so doesn't distinguish the pronunciation [b ae s] in **bass**$^{\mathbf{4}}$, **bass**$^{\mathbf{5}}$, and **bass**$^{\mathbf{8}}$ from the other senses pronounced [b ey s].

Synset

The set of near-synonyms for a WordNet sense is called a **synset** (for **synonym set**); synsets are an important primitive in WordNet. The entry for *bass* includes synsets like *bass*^{1}, *deep*^{6}, or *bass*^{6}, *bass voice*^{1}, *basso*^{2}. We can think of a synset as representing a **concept** of the type we discussed in Ch. 17. Thus instead of representing concepts using logical terms, WordNet represents them as a lists of the word-senses that can be used to express the concept. Here's another synset example:

{chump^{1}, fool^{2}, gull^{1}, mark^{9}, patsy^{1}, fall guy^{1}, sucker^{1}, soft touch^{1}, mug^{2}}

The gloss of this synset describes it as *a person who is gullible and easy to take advantage of*. Each of the lexical entries included in the synset can, therefore, be used to express this concept. Synsets like this one actually constitute the senses associated with WordNet entries, and hence it is synsets, not wordforms, lemmas or individual senses, that participate in most of the lexical sense relations in WordNet.

Let's turn now to these lexical sense relations, some of which are illustrated in Fig. 19.2 and Fig. 19.3. WordNet hyponymy relations correspond to the notion of

Relation	Also called	Definition	Example
Hypernym	Superordinate	From concepts to superordinates	$breakfast^1 \longrightarrow meal^1$
Hyponym	Subordinate	From concepts to subtypes	$meal^1 \longrightarrow lunch^1$
Member Meronym	Has-Member	From groups to their members	$faculty^2 \longrightarrow professor^1$
Has-Instance		From concepts to instances of the concept	$composer^1 \longrightarrow Bach^1$
Instance		From instances to their concepts	$Austen^1 \longrightarrow author^1$
Member Holonym	Member-Of	From members to their groups	$copilot^1 \longrightarrow crew^1$
Part Meronym	Has-Part	From wholes to parts	$table^2 \longrightarrow leg^3$
Part Holonym	Part-Of	From parts to wholes	$course^7 \longrightarrow meal^1$
Antonym		Opposites	$leader^1 \longrightarrow follower^1$

Figure 19.2 Noun relations in WordNet.

Relation	Definition	Example
Hypernym	From events to superordinate events	$fly^9 \longrightarrow travel^5$
Troponym	From a verb (event) to a specific manner elaboration of that verb	$walk^1 \longrightarrow stroll^1$
Entails	From verbs (events) to the verbs (events) they entail	$snore^1 \longrightarrow sleep^1$
Antonym	Opposites	$increase^1 \Longleftrightarrow decrease^1$

Figure 19.3 Verb relations in WordNet.

immediate hyponymy discussed on page 628. Each synset is related to its immediately more general and more specific synsets via direct hypernym and hyponym relations. These relations can be followed to produce longer chains of more general or more specific synsets. Fig. 19.4 shows hypernym chains for **bass**$^{\mathbf{3}}$ and **bass**$^{\mathbf{7}}$.

In this depiction of hyponymy, successively more general synsets are shown on successive indented lines. The first chain starts from the concept of a human bass singer. Its immediate superordinate is a synset corresponding to the generic concept of a singer. Following this chain leads eventually to concepts such as *entertainer* and *person*. The second chain, which starts from musical instrument, has a completely different path leading eventually to such concepts as musical instrument, device and physical object. Both paths do eventually join at the very abstract synset *whole*, *unit*, and then proceed together to *entity* which is the top (root) of the noun hierarchy (in WordNet this root is generally called the **unique beginner**).

Unique beginner

19.4 Event Participants: Semantic Roles and Selectional Restrictions

An important aspect of lexical meaning has to do with the semantics of events. When we discussed events in Ch. 17, we introduced the importance of predicate-argument structure for representing an event, and the use of Davidsonian reification of events to represent each participant distinctly from the event itself. We turn in this section to representing the meaning of these event *participants* or *arguments*. We introduce two kinds of semantic constraints on the arguments of event predicates: **semantic roles** and **selectional restrictions**. We begin with a particular model of semantic roles called

```
Sense 3
bass, basso --
(an adult male singer with the lowest voice)
=> singer, vocalist, vocalizer, vocaliser
   => musician, instrumentalist, player
      => performer, performing artist
         => entertainer
            => person, individual, someone...
               => organism, being
                  => living thing, animate thing,
                     => whole, unit
                        => object, physical object
                           => physical entity
                              => entity
               => causal agent, cause, causal agency
                  => physical entity
                     => entity

Sense 7
bass --
(the member with the lowest range of a family of
musical instruments)
=> musical instrument, instrument
   => device
      => instrumentality, instrumentation
         => artifact, artefact
            => whole, unit
               => object, physical object
                  => physical entity
                     => entity
```

Figure 19.4 Hyponymy chains for two separate senses of the lemma *bass*. Note that the chains are completely distinct, only converging at the very abstract level *whole, unit*.

thematic roles.

19.4.1 Thematic Roles

Consider how we represented the meaning of arguments in Ch. 17 for sentences like these:

(19.19) Sasha broke the window.

(19.20) Pat opened the door.

A neo-Davidsonian event representation of these two sentences would be:

$$\exists e,x,y\ Isa(e,Breaking) \wedge Breaker(e,Sasha) \wedge BrokenThing(e,y) \wedge Isa(y,Window)$$
$$\exists e,x,y\ Isa(e,Opening) \wedge Opener(e,Pat) \wedge OpenedThing(e,y) \wedge Isa(y,Door)$$

Thematic Role	Definition
AGENT	The volitional causer of an event
EXPERIENCER	The experiencer of an event
FORCE	The non-volitional causer of the event
THEME	The participant most directly affected by an event
RESULT	The end product of an event
CONTENT	The proposition or content of a propositional event
INSTRUMENT	An instrument used in an event
BENEFICIARY	The beneficiary of an event
SOURCE	The origin of the object of a transfer event
GOAL	The destination of an object of a transfer event

Figure 19.5 Some commonly-used thematic roles with their definitions.

Deep roles

In this representation, the roles of the subjects of the verbs *break* and *open* are *Breaker* and *Opener* respectively. These **deep roles** are specific to each possible kind of event; *Breaking* events have *Breakers*, *Opening* events have *Openers*, *Eating* events have *Eaters*, and so on.

If we are going to be able to answer questions, perform inferences, or do any further kinds of natural language understanding of these events, we'll need to know a little more about the semantics of these arguments. *Breakers* and *Openers* have something in common. They are both volitional actors, often animate, and they have direct causal responsibility for their events.

Thematic role
Agent
Theme

Thematic roles are one attempt to capture this semantic commonality between *Breakers* and *Eaters*. We say that the subjects of both these verbs are **agents**. Thus AGENT is the thematic role which represents an abstract idea such as volitional causation. Similarly, the direct objects of both these verbs, the *BrokenThing* and *OpenedThing*, are both prototypically inanimate objects which are affected in some way by the action. The thematic role for these participants is **theme**.

Thematic roles are one of the oldest linguistic models, proposed first by the Indian grammarian Panini sometime between the 7th and 4th centuries BCE. Their modern formulation is due to Fillmore (1968) and Gruber (1965). Although there is no universally agreed-upon set of thematic roles, Fig. 19.5 and Fig. 19.6 present a list of some thematic roles which have been used in various computational papers, together with rough definitions and examples.

19.4.2 Diathesis Alternations

The main reason computational systems use thematic roles, and semantic roles in general, is to act as a shallow semantic language that can let us make simple inferences that aren't possible from the pure surface string of words, or even the parse tree. For example, if a document says that *Company A acquired Company B*, we'd like to know that this answers the query *Was Company B acquired?* despite the fact that the two sentences have very different surface syntax. Similarly, this shallow semantics might act as a useful intermediate language in machine translation.

Thus thematic roles are used in helping us generalize over different surface real-

Thematic Role	**Example**
AGENT	*The waiter* spilled the soup.
EXPERIENCER	*John* has a headache.
FORCE	*The wind* blows debris from the mall into our yards.
THEME	Only after Benjamin Franklin broke *the ice*...
RESULT	The French government has built a *regulation-size baseball diamond*...
CONTENT	Mona asked *"You met Mary Ann at a supermarket"?*
INSTRUMENT	He turned to poaching catfish, stunning them *with a shocking device*...
BENEFICIARY	Whenever Ann Callahan makes hotel reservations *for her boss*...
SOURCE	I flew in *from Boston*.
GOAL	I drove *to Portland*.

Figure 19.6 Some prototypical examples of various thematic roles.

izations of predicate arguments. For example while the AGENT is often realized as the subject of the sentence, in other cases the THEME can be the subject. Consider these possible realizations of the thematic arguments of the verb *break*:

(19.21) *John* *broke the window.*
AGENT THEME

(19.22) *John* *broke the window with a rock.*
AGENT THEME INSTRUMENT

(19.23) *The rock* *broke the door.*
INSTRUMENT THEME

(19.24) *The window broke.*
THEME

(19.25) *The window was broken by John.*
THEME AGENT

The examples above suggest that *break* has (at least) the possible arguments AGENT, THEME, and INSTRUMENT. The set of thematic role arguments taken by a verb is often called the **thematic grid**, θ-grid, or **case frame**. We can see that there are (among others) the following possibilities for the realization of these arguments of *break*:

Thematic grid
Case frame

- AGENT:Subject, THEME:Object
- AGENT:Subject, THEME:Object , INSTRUMENT:PP_{with}
- INSTRUMENT:Subject, THEME:Object
- THEME:Subject

It turns out that many verbs allow their thematic roles to be realized in various syntactic positions. For example, verbs like *give* can realize the THEME and GOAL arguments in two different ways:

(19.26) a. *Doris* *gave the book to Cary.*
AGENT THEME GOAL

b. *Doris* *gave Cary the book.*
AGENT GOAL THEME

These multiple argument structure realizations (the fact that *break* can take AGENT, INSTRUMENT, or THEME as subject, and *give* can realize its THEME and GOAL in either order) are called **verb alternations** or **diathesis alternations**. The alternation we showed above *give*, the **dative alternation**, seems to occur with particular semantic classes of verbs, including "verbs of future having" (*advance*, *allocate*, *offer*, *owe*), "send verbs" (*forward*, *hand*, *mail*), "verbs of throwing" (*kick*, *pass*, *throw*), and so on. Levin (1993) is a reference book which lists for a large set of English verbs the semantic classes they belong to and the various alternations that they participate in. These lists of verb classes have been incorporated into the online resource VerbNet (Kipper et al., 2000).

Verb alternation
Diathesis alternation
Dative alternation

19.4.3 Problems with Thematic Roles

Representing meaning at the thematic role level seems like it should be useful in dealing with complications like diathesis alternations. But despite this potential benefit, it has proved very difficult to come up with a standard set of roles, and equally difficult to produce a formal definition of roles like AGENT, THEME, or INSTRUMENT.

For example, researchers attempting to define role sets often find they need to fragment a role like AGENT or THEME into many specific roles. Levin and Rappaport Hovav (2005) summarizes a number of such cases, such as the fact there seem to be at least two kinds of INSTRUMENTS, *intermediary* instruments that can appear as subjects and *enabling* instruments that cannot:

(19.27) The cook opened the jar with the new gadget.
The new gadget opened the jar.

(19.28) Shelly ate the sliced banana with a fork.
*The fork ate the sliced banana.

In addition to the fragmentation problem, there are cases where we'd like to reason about and generalize across semantic roles, but the finite discrete lists of roles don't let us do this.

Finally, it has proved very difficult to formally define the semantic roles. Consider the AGENT role; most cases of AGENTS are animate, volitional, sentient, causal, but any individual noun phrase might not exhibit all of these properties.

These problems have led most research to alternative models of semantic roles. One such model is based on defining **generalized semantic roles** that abstract over the specific thematic roles. For example PROTO-AGENT and PROTO-PATIENT are generalized roles that express roughly agent-like and roughly patient-like meanings. These roles are defined, not by necessary and sufficient conditions, but rather by a set of heuristic features that accompany more agent-like or more patient-like meanings. Thus the more an argument displays agent-like properties (intentionality, volitionality, causality, etc) the greater likelihood the argument can be labeled a PROTO-AGENT. The more patient-like properties (undergoing change of state, causally affected by another participant, stationary relative to other participants, etc), the greater likelihood the argument can be labeled a PROTO-PATIENT.

Generalized semantic role
Proto-agent
Proto-patient

In addition to using proto-roles, many computational models avoid the problems with thematic roles by defining semantic roles that are specific to a particular verb, or specific to a particular set of verbs or nouns.

In the next two sections we will describe two commonly used lexical resources which make use of some of these alternative versions of semantic roles. **PropBank** uses both proto-roles and verb-specific semantic roles. **FrameNet** uses frame-specific semantic roles.

19.4.4 The Proposition Bank

PropBank

The **Proposition Bank**, generally referred to as **PropBank**, is a resource of sentences annotated with semantic roles. The English PropBank labels all the sentences in the Penn TreeBank; there is also a Chinese PropBank which labels sentences in the Penn Chinese TreeBank. Because of the difficulty of defining a universal set of thematic roles, the semantic roles in PropBank are defined with respect to an individual verb sense. Each sense of each verb thus has a specific set of roles, which are given only numbers rather than names: **Arg0**, **Arg1** **Arg2**, and so on. In general, **Arg0** is used to represent the PROTO-AGENT, and **Arg1** the PROTO-PATIENT; the semantics of the other roles are specific to each verb sense. Thus the **Arg2** of one verb is likely to have nothing in common with the **Arg2** of another verb.

Here are some slightly simplified PropBank entries for one sense each of the verbs *agree* and *fall*; the definitions for each role (“Other entity agreeing”, “amount fallen”) are informal glosses intended to be read by humans, rather than formal definitions.

(19.29) Frameset **agree.01**
Arg0: Agreer
Arg1: Proposition
Arg2: Other entity agreeing
Ex1: [$_{\text{Arg0}}$ The group] *agreed* [$_{\text{Arg1}}$ it wouldn’t make an offer unless it had Georgia Gulf’s consent].
Ex2: [$_{\text{ArgM-TMP}}$ Usually] [$_{\text{Arg0}}$ John] *agrees* [$_{\text{Arg2}}$ with Mary] [$_{\text{Arg1}}$ on everything].

(19.30) **fall.01** “move downward”
Arg1: Logical subject, patient, thing falling
Arg2: Extent, amount fallen
Arg3: start point
Arg4: end point, end state of arg1
ArgM-LOC: medium
Ex1: [$_{\text{Arg1}}$ Sales] *fell* [$_{\text{Arg4}}$ to \$251.2 million] [$_{\text{Arg3}}$ from \$278.7 million].
Ex1: [$_{\text{Arg1}}$ The average junk bond] *fell* [$_{\text{Arg2}}$ by 4.2%].

Note that there is no Arg0 role for *fall*, because the normal subject of *fall* is a PROTO-PATIENT.

The PropBank semantic roles can be useful in recovering shallow semantic information about verbal arguments. Consider the verb *increase*:

(19.31) **increase.01** “go up incrementally”

Arg0: causer of increase
Arg1: thing increasing
Arg2: amount increased by, EXT, or MNR
Arg3: start point
Arg4: end point

A PropBank semantic role labeling would allow us to infer the commonality in the event structures of the following three examples, i.e., that in each case *Big Fruit Co.* is the AGENT, and *the price of bananas* is the THEME, despite the differing surface forms.

(19.32) [$_{\text{Arg0}}$ Big Fruit Co.] increased [$_{\text{Arg1}}$ the price of bananas].

(19.33) [$_{\text{Arg1}}$ The price of bananas] was increased again [$_{\text{Arg0}}$ by Big Fruit Co.]

(19.34) [$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].

19.4.5 FrameNet

While making inferences about the semantic commonalities across different sentences with *increase* is useful, it would be even more useful if we could make such inferences in many more situations, across different verbs, and also between verbs and nouns.

For example, we'd like to extract the similarity between these three sentences:

(19.35) [$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].

(19.36) [$_{\text{Arg1}}$ The price of bananas] rose [$_{\text{Arg2}}$ 5%].

(19.37) There has been a [$_{\text{Arg2}}$ 5%] rise [$_{\text{Arg1}}$ in the price of bananas].

Note that the second example uses the different verb *rise*, and the third example uses the noun rather than the verb *rise*. We'd like a system to recognize that *the price of bananas* is what went up, and that *5%* is the amount it went up, no matter whether the *5%* appears as the object of the verb *increased* or as a nominal modifier of the noun *rise*.

FrameNet

The **FrameNet** project is another semantic role labeling project that attempts to address just these kinds of problems (Baker et al., 1998; Lowe et al., 1997; Ruppenhofer et al., 2006). Where roles in the PropBank project are specific to an individual verb, roles in the FrameNet project are specific to a **frame**. A **frame** is a script-like structure, which instantiates a set of frame-specific semantic roles called **frame elements**. Each word evokes a frame and profiles some aspect of the frame and its elements. For example, the **change_position_on_a_scale** frame is defined as follows:

Frame

Frame element

> This frame consists of words that indicate the change of an Item's position on a scale (the Attribute) from a starting point (Initial_value) to an end point (Final_value).

Some of the semantic roles (frame elements) in the frame, separated into **core** roles and **non-core roles**, are defined as follows (definitions are taken from the FrameNet labelers guide (Ruppenhofer et al., 2006)).

Core Roles	
ATTRIBUTE	The ATTRIBUTE is a scalar property that the ITEM possesses.
DIFFERENCE	The distance by which an ITEM changes its position on the scale.
FINAL_STATE	A description that presents the ITEM's state after the change in the ATTRIBUTE's value as an independent predication.
FINAL_VALUE	The position on the scale where the Item ends up.
INITIAL_STATE	A description that presents the ITEM's state before the change in the ATTRIBUTE's value as an independent predication.
INITIAL_VALUE	The initial position on the scale from which the ITEM moves away.
ITEM	The entity that has a position on the scale.
VALUE_RANGE	A portion of the scale, typically identified by its end points, along which the values of the ATTRIBUTE fluctuate.
Some Non-Core Roles	
DURATION	The length of time over which the change takes place.
SPEED	The rate of change of the VALUE.
GROUP	The GROUP in which an ITEM changes the value of an ATTRIBUTE in a specified way.

Here are some example sentences:

(19.38) [$_{\text{ITEM}}$ Oil] *rose* [$_{\text{ATTRIBUTE}}$ in price] in price [$_{\text{DIFFERENCE}}$ by 2%].

(19.39) [$_{\text{ITEM}}$ It] has *increased* [$_{\text{FINAL\_STATE}}$ to having them 1 day a month].

(19.40) [$_{\text{ITEM}}$ Microsoft shares] *fell* [$_{\text{FINAL\_VALUE}}$ to 7 5/8].

(19.41) [$_{\text{ITEM}}$ Colon cancer incidence] *fell* [$_{\text{DIFFERENCE}}$ by 50%] [$_{\text{GROUP}}$ among men].

(19.42) a steady *increase* [$_{\text{INITIAL\_VALUE}}$ from 9.5] [$_{\text{FINAL\_VALUE}}$ to 14.3] [$_{\text{ITEM}}$ in dividends]

(19.43) a [$_{\text{DIFFERENCE}}$ 5%] [$_{\text{ITEM}}$ dividend] *increase*...

Note from these example sentences that the frame includes target words like *rise*, *fall*, and *increase*. In fact, the complete frame consists of the following words:

VERBS:	dwindle	move	soar	escalation	shift
advance	edge	mushroom	swell	explosion	tumble
climb	explode	plummet	swing	fall	
decline	fall	reach	triple	fluctuation	**ADVERBS:**
decrease	fluctuate	rise	tumble	gain	increasingly
diminish	gain	rocket		growth	
dip	grow	shift	**NOUNS:**	hike	
double	increase	skyrocket	decline	increase	
drop	jump	slide	decrease	rise	

FrameNet also codes relationships between frames and frame elements. Frames can inherit from each other, and generalizations among frame elements in different frames can be captured by inheritance as well. Other relations between frames like causation are also represented. Thus there is a **Cause_change_of_position_on_a_scale** frame which is linked to the **Change_of_position_on_a_scale** frame by the **cause** relation, but adds an AGENT role and is used for causative examples such as the following:

(19.44) [$_{\text{AGENT}}$ They] *raised* [$_{\text{ITEM}}$ the price of their soda] [$_{\text{DIFFERENCE}}$ by 2%].

Together, these two frames would allow an understanding system to extract the common event semantics of all the verbal and nominal causative and non-causative usages.

Ch. 20 will discuss automatic methods for extracting various kinds of semantic roles; indeed one main goal of PropBank and FrameNet is to provide training data for such semantic role labeling algorithms.

19.4.6 Selectional Restrictions

Semantic roles gave us a way to express some of the semantics of an argument in its relation to the predicate. In this section we turn to another way to express semantic constraints on arguments. A **selectional restriction** is a kind of semantic type constraint that a verb imposes on the kind of concepts that are allowed to fill its argument roles. Consider the two meanings associated with the following example:

(19.45) I want to eat someplace that's close to ICSI.

There are two possible parses and semantic interpretations for this sentence. In the sensible interpretation *eat* is intransitive and the phrase *someplace that's close to* ICSI is an adjunct that gives the location of the eating event. In the nonsensical *speaker-as-Godzilla* interpretation, *eat* is transitive and the phrase *someplace that's close to* ICSI is the direct object and the THEME of the eating, like the NP *Malaysian food* in the following sentences:

(19.46) I want to eat Malaysian food.

How do we know that *someplace that's close to* ICSI isn't the direct object in this sentence? One useful cue is the semantic fact that the THEME of EATING events tends to be something that is *edible*. This restriction placed by the verb *eat* on the filler of its THEME argument, is called a **selectional restriction**. A selectional restriction is a constraint on the semantic type of some argument.

Selectional restriction

Selectional restrictions are associated with senses, not entire lexemes. We can see this in the following examples of the lexeme *serve*:

(19.47) Well, there was the time they served green-lipped mussels from New Zealand.

(19.48) Which airlines serve Denver?

Example (19.47) illustrates the cooking sense of *serve*, which ordinarily restricts its THEME to be some kind foodstuff. Example (19.48) illustrates the *provides a commercial service to* sense of *serve*, which constrains its THEME to be some type of appropriate location. We will see in Ch. 20 that the fact that selectional restrictions are associated with senses can be used as a cue to help in word sense disambiguation.

Selectional restrictions vary widely in their specificity. The verb *imagine*, for example, imposes strict requirements on its AGENT role (restricting it to humans and other animate entities) but places very few semantic requirements on its THEME role. A verb like *diagonalize*, on the other hand, places a very specific constraint on the filler of its THEME role: it has to be a matrix, while the arguments of the adjectives *odorless* are restricted to concepts that could possess an odor:

(19.49) In rehearsal, I often ask the musicians to *imagine* a tennis game.

(19.50) Radon is an *odorless* gas that can't be detected by human senses.

(19.51) To diagonalize a matrix is to find its eigenvalues.

These examples illustrate that the set of concepts we need to represent selectional restrictions (being a matrix, being able to possess an oder, etc) is quite open-ended. This distinguishes selectional restrictions from other features for representing lexical knowledge, like parts-of-speech, which are quite limited in number.

Representing Selectional Restrictions

One way to capture the semantics of selectional restrictions is to use and extend the event representation of Ch. 17. Recall that the neo-Davidsonian representation of an event consists of a single variable that stands for the event, a predicate denoting the kind of event, and variables and relations for the event roles. Ignoring the issue of the λ-structures, and using thematic roles rather than deep event roles, the semantic contribution of a verb like *eat* might look like the following:

$$\exists e,x,y\ Eating(e) \wedge Agent(e,x) \wedge Theme(e,y)$$

With this representation, all we know about y, the filler of the THEME role, is that it is associated with an $Eating$ event via the *Theme* relation. To stipulate the selectional restriction that y must be something edible, we simply add a new term to that effect:

$$\exists e,x,y\ Eating(e) \wedge Agent(e,x) \wedge Theme(e,y) \wedge Isa(y, EdibleThing)$$

When a phrase like *ate a hamburger* is encountered, a semantic analyzer can form the following kind of representation:

$$\exists e,x,y\ Eating(e) \wedge Eater(e,x) \wedge Theme(e,y) \wedge Isa(y, EdibleThing) \\ \wedge Isa(y, Hamburger)$$

This representation is perfectly reasonable since the membership of y in the category *Hamburger* is consistent with its membership in the category *EdibleThing*, assuming a reasonable set of facts in the knowledge base. Correspondingly, the representation for a phrase such as *ate a takeoff* would be ill-formed because membership in an event-like category such as *Takeoff* would be inconsistent with membership in the category *EdibleThing*.

While this approach adequately captures the semantics of selectional restrictions, there are two practical problems with its direct use. First, using FOPC to perform the simple task of enforcing selectional restrictions is overkill. There are far simpler formalisms that can do the job with far less computational cost. The second problem is that this approach presupposes a large logical knowledge-base of facts about the concepts that make up selectional restrictions. Unfortunately, although such common sense knowledge-bases are being developed, none currently have the kind of scope necessary to the task.

A more practical approach is to state selectional restrictions in terms of WordNet synsets, rather than logical concepts. Each predicate simply specifies a WordNet synset

```
Sense 1
hamburger, beefburger --
(a fried cake of minced beef served on a bun)
=> sandwich
   => snack food
      => dish
         => nutriment, nourishment, nutrition...
            => food, nutrient
               => substance
                  => matter
                     => physical entity
                        => entity
```

Figure 19.7 Evidence from WordNet that hamburgers are edible.

as the selectional restriction on each of its arguments. A meaning representation is well-formed if the role filler word is a hyponym (subordinate) of this synset.

For our *ate a hamburger* example, for example, we could set the selectional restriction on the THEME role of the verb *eat* to the synset {**food, nutrient**}, glossed as *any substance that can be metabolized by an animal to give energy and build tissue*: Luckily, the chain of hypernyms for *hamburger* shown in Fig. 19.7 reveals that hamburgers are indeed food. Again, the filler of a role need not match the restriction synset exactly, it just needs to have the synset as one of its superordinates.

We can apply this approach to the THEME roles of the verbs *imagine*, *lift* and *diagonalize*, discussed earlier. Let us restrict *imagine*'s THEME to the synset {entity}, *lift*'s THEME to {physical entity} and *diagonalize* to {matrix}. This arrangement correctly permits *imagine a hamburger* and *lift a hamburger*, while also correctly ruling out *diagonalize a hamburger*.

Of course WordNet is unlikely to have the exactly relevant synsets to specify selectional restrictions for all possible words of English; other taxonomies may also be used. In addition, it is possible to learn selectional restrictions automatically from corpora.

We will return to selectional restrictions in Ch. 20 where we introduce the extension to **selectional preferences**, where a predicate can place probabilistic preferences rather than strict deterministic constraints on its arguments.

19.5 Primitive Decomposition

Back at the beginning of the chapter, we said that one way of defining a word is to decompose its meaning into a set of primitive semantics elements or features. We saw one aspect of this method in our discussion of finite lists of thematic roles (agent, patient, instrument, etc). We turn now to a brief discussion of how this kind of model, called **primitive decomposition**, or **componential analysis**, could be applied to the meanings of all words. Wierzbicka (1992, 1996) shows that this approach dates back at least to continental philosophers like Descartes and Leibniz.

Consider trying to define words like *hen*, *rooster*, or *chick*. These words have some-

Semantic feature

thing in common (they all describe chickens) and something different (their age and sex). This can be represented by using **semantic features**, symbols which represent some sort of primitive meaning:

hen	`+female, +chicken, +adult`
rooster	`-female, +chicken, +adult`
chick	`+chicken, -adult`

A number of studies of decompositional semantics, especially in the computational literature, have focused on the meaning of verbs. Consider these examples for the verb *kill*:

(19.52) Jim killed his philodendron.

(19.53) Jim did something to cause his philodendron to become not alive.

There is a truth-conditional ('propositional semantics') perspective from which these two sentences have the same meaning. Assuming this equivalence, we could represent the meaning of *kill* as:

(19.54) KILL(x,y) $\Leftrightarrow$ CAUSE(x, BECOME(NOT(ALIVE(y))))

thus using semantic primitives like *do*, *cause*, *become not*, and *alive*.

Indeed, one such set of potential semantic primitives has been used to account for some of the verbal alternations discussed in Sec. 19.4.2 (Lakoff, 1965; Dowty, 1979). Consider the following examples.

(19.55) John opened the door. $\Rightarrow$ (CAUSE(John(BECOME(OPEN(door)))))

(19.56) The door opened. $\Rightarrow$ (BECOME(OPEN(door)))

(19.57) The door is open. $\Rightarrow$ (OPEN(door))

The decompositional approach asserts that a single state-like predicate associated with *open* underlies all of these examples. The differences among the meanings of these examples arises from the combination of this single predicate with the primitives CAUSE and BECOME.

While this approach to primitive decomposition can explain the similarity between states and actions, or causative and non-causative predicates, it still relies on having a very large number of predicates like *open*. More radical approaches choose to break down these predicates as well. One such approach to verbal predicate decomposition is **Conceptual Dependency** (CD), a set of ten primitive predicates, shown in Fig. 19.8.

Conceptual Dependency

Below is an example sentence along with its CD representation. The verb *brought* is translated into the two primitives ATRANS and PTRANS to indicate the fact that the waiter both physically conveyed the check to Mary and passed control of it to her. Note that CD also associates a fixed set of thematic roles with each primitive to represent the various participants in the action.

(19.58) The waiter brought Mary the check.

$$\exists x,y\ Atrans(x) \wedge Actor(x,Waiter) \wedge Object(x,Check) \wedge To(x,Mary)$$
$$\wedge Ptrans(y) \wedge Actor(y,Waiter) \wedge Object(y,Check) \wedge To(y,Mary)$$

Primitive	Definition
ATRANS	The abstract transfer of possession or control from one entity to another.
PTRANS	The physical transfer of an object from one location to another
MTRANS	The transfer of mental concepts between entities or within an entity.
MBUILD	The creation of new information within an entity.
PROPEL	The application of physical force to move an object.
MOVE	The integral movement of a body part by an animal.
INGEST	The taking in of a substance by an animal.
EXPEL	The expulsion of something from an animal.
SPEAK	The action of producing a sound.
ATTEND	The action of focusing a sense organ.

Figure 19.8 A set of conceptual dependency primitives.

There are also sets of semantic primitives that cover more than just simple nouns and verbs. The following list comes from Wierzbicka (1996):

substantives:	I, YOU, SOMEONE, SOMETHING, PEOPLE
mental predicates:	THINK, KNOW, WANT, FEEL, SEE, HEAR
speech:	SAY
determiners and quantifiers:	THIS, THE SAME, OTHER, ONE, TWO, MANY (MUCH), ALL, SOME, MORE
actions and events:	DO, HAPPEN
evaluators:	GOOD, BAD
descriptors:	BIG, SMALL
time:	WHEN, BEFORE, AFTER
space:	WHERE, UNDER, ABOVE,
partonomy and taxonomy:	PART (OF), KIND (OF)
movement, existence, life:	MOVE, THERE IS, LIVE
metapredicates:	NOT, CAN, VERY
interclausal linkers:	IF, BECAUSE, LIKE
space:	FAR, NEAR, SIDE, INSIDE, HERE
time:	A LONG TIME, A SHORT TIME, NOW
imagination and possibility:	IF... WOULD, CAN, MAYBE

Because of the difficulty of coming up with a set of primitives that can represent all possible kinds of meanings, most current computational linguistic work does not use semantic primitives. Instead, most computational work tends to use the lexical relations of Sec. 19.2 to define words.

19.6 Advanced concepts: Metaphor

Metaphor

We use a **metaphor** when we refer to and reason about a concept or domain using words and phrases whose meanings come from a completely different domain.

Metaphor is similar to **metonymy**, which we introduced as the use of one aspect of a concept or entity to refer to other aspects of the entity. In Sec. 19.1 we introduced metonymies like the following,

(19.59) Author (*Jane Austen wrote Emma*) ↔ Works of Author (*I really love Jane Austen*).

in which two senses of a polysemous word are systematically related. In metaphor, by contrast, there is a systematic relation between two completely different domains of meaning.

Metaphor is pervasive. Consider the following WSJ sentence:

(19.60) That doesn't scare Digital, which has grown to be the world's second-largest computer maker by poaching customers of IBM's mid-range machines.

The verb *scare* means 'to cause fear in', or 'to cause to lose courage'. For this sentence to make sense, it has to be the case that corporations can experience emotions like fear or courage as people do. Of course they don't, but we certainly speak of them and reason about them as if they do. We can therefore say that this use of *scare* is based on a metaphor that allows us to view a corporation as a person, which we will refer to the CORPORATION AS PERSON metaphor.

This metaphor is neither novel nor specific to this use of *scare*. Instead, it is a fairly conventional way to think about companies and motivates the use of *resuscitate*, *hemorrhage* and *mind* in the following WSJ examples:

(19.61) Fuqua Industries Inc. said Triton Group Ltd., a company it helped **resuscitate**, has begun acquiring Fuqua shares.

(19.62) And Ford was **hemorrhaging**; its losses would hit $1.54 billion in 1980.

(19.63) But if it changed its **mind**, however, it would do so for investment reasons, the filing said.

Each of these examples reflects an elaborated use of the basic CORPORATION AS PERSON metaphor. The first two examples extend it to use the notion of health to express a corporation's financial status, while the third example attributes a mind to a corporation to capture the notion of corporate strategy.

Conventional metaphor

Metaphorical constructs such as CORPORATION AS PERSON are known as **conventional metaphors**. Lakoff and Johnson (1980) argue that many if not most of the metaphorical expressions that we encounter every day are motivated by a relatively small number of these simple conventional schemas.

19.7 Summary

This chapter has covered a wide range of issues concerning the meanings associated with lexical items. The following are among the highlights:

- **Lexical semantics** is the study of the meaning of words, and the systematic meaning-related connections between words.
- A **word sense** is the locus of word meaning; definitions and meaning relations are defined at the level of the word sense rather than wordforms as a whole.

- **Homonymy** is the relation between unrelated senses that share a form, while **polysemy** is the relation between related senses that share a form.
- **Synonymy** holds between different words with the same meaning.
- **Hyponymy** relations hold between words that are in a class-inclusion relationship.
- **Semantic fields** are used to capture semantic connections among groups of lexemes drawn from a single domain.
- **WordNet** is a large database of lexical relations for English words.
- **Semantic roles** abstract away from the specifics of deep semantic roles by generalizing over similar roles across classes of verbs.
- **Thematic roles** are a model of semantic roles based on a single finite list of roles. Other semantic role models include per-verb semantic roles lists and **proto-agent**/**proto-patient** both of which are implemented in **PropBank**, and per-frame role lists, implemented in **FrameNet**.
- Semantic **selectional restrictions** allow words (particularly predicates) to post constraints on the semantic properties of their argument words.
- **Primitive decomposition** is another way to represent the meaning of word, in terms of finite sets of sub-lexical primitives.

Bibliographical and Historical Notes

Cruse (2004) is a useful introductory linguistic text on lexical semantics. Levin and Rappaport Hovav (2005) is a research survey covering argument realization and semantic roles. Lyons (1977) is another classic reference. Collections describing computational work on lexical semantics can be found in Pustejovsky and Bergler (1992), Saint-Dizier and Viegas (1995) and Klavans (1995).

The most comprehensive collection of work concerning WordNet can be found in Fellbaum (1998). There have been many efforts to use existing dictionaries as lexical resources. One of the earliest was Amsler's (1980, 1981) use of the Merriam Webster dictionary. The machine readable version of Longman's Dictionary of Contemporary English has also been used (Boguraev and Briscoe, 1989). See Pustejovsky (1995), Pustejovsky and Boguraev (1996), Martin (1986) and Copestake and Briscoe (1995), inter alia, for computational approaches to the representation of polysemy. Pustejovsky's theory of the **Generative Lexicon**, and in particular his theory of the **qualia structure** of words, is another way of accounting for the dynamic systematic polysemy of words in context.

Generative Lexicon
Qualia structure

As we mentioned earlier, thematic roles are one of the oldest linguistic models, proposed first by the Indian grammarian Panini sometimes between the 7th and 4th centuries BCE. Their modern formulation is due to Fillmore (1968) and Gruber (1965). Fillmore's work had a large and immediate impact on work in natural language processing, as much early work in language understanding used some version of Fillmore's case roles (e.g., Simmons (1973, 1978, 1983)). Fillmore's extension of this work to the

FrameNet project is described in Baker et al. (1998), Narayanan et al. (1999), and Baker et al. (2003).

Work on selectional restrictions as a way of characterizing semantic well-formedness began with Katz and Fodor (1963). McCawley (1968) was the first to point out that selectional restrictions could not be restricted to a finite list of semantic features, but had to be drawn from a larger base of unrestricted world knowledge.

Lehrer (1974) is a classic text on semantic fields. More recent papers addressing this topic can be found in Lehrer and Kittay (1992).

The use of semantic primitives to define word meaning dates back to Leibniz; in linguistics, the focus on componential analysis in semantics was due to Hjelmslev (1969). See Nida (1975) for a comprehensive overview of work on componential analysis. Wierzbicka (1996) has long been a major advocate of the use of primitives in linguistic semantics; Wilks (1975a) has made similar arguments for the computational use of primitives in machine translation and natural language understanding. Another prominent effort has been Jackendoff's Conceptual Semantics work (1983, 1990), which has also been applied in machine translation (Dorr, 1993, 1992).

Computational approaches to the interpretation of metaphor include convention-based and reasoning-based approaches. Convention-based approaches encode specific knowledge about a relatively small core set of conventional metaphors. These representations are then used during understanding to replace one meaning with an appropriate metaphorical one (Norvig, 1987; Martin, 1990; Hayes and Bayer, 1991; Veale and Keane, 1992; Jones and McCoy, 1992). Reasoning-based approaches eschew representing metaphoric conventions, instead modeling figurative language processing via general reasoning ability, such as analogical reasoning, rather than as a specifically language-related phenomenon. (Russell, 1976; Carbonell, 1982; Gentner, 1983; Fass, 1988, 1991, 1997).

An influential collection of papers on metaphor can be found in Ortony (1993). Lakoff and Johnson (1980) is the classic work on conceptual metaphor and metonymy. Russell (1976) presents one of the earliest computational approaches to metaphor. Additional early work can be found in DeJong and Waltz (1983), Wilks (1978) and Hobbs (1979b). More recent computational efforts to analyze metaphor can be found in Fass (1988, 1991, 1997), Martin (1990), Veale and Keane (1992), Iverson and Helmreich (1992), Chandler (1991), and Martin (2006). Martin (1996) presents a survey of computational approaches to metaphor and other types of figurative language. Gries and Stefanowitsch (2006) is a recent collection of papers on corpus-based approaches to metaphor.

Exercises

19.1 Collect three definitions of ordinary non-technical English words from a dictionary of your choice that you feel are flawed in some way. Explain the nature of the flaw and how it might be remedied.

19.2 Give a detailed account of similarities and differences among the following set of lexemes: *imitation*, *synthetic*, *artificial*, *fake*, and *simulated*.

19.3 Examine the entries for these lexemes in WordNet (or some dictionary of your choice). How well does it reflect your analysis?

19.4 The WordNet entry for the noun *bat* lists 6 distinct senses. Cluster these senses using the definitions of homonymy and polysemy given in this chapter. For any senses that are polysemous, give an argument as to how the senses are related.

19.5 Assign the various verb arguments in the following WSJ examples to their appropriate thematic roles using the set of roles shown in Fig. 19.6.

- **a**. The intense heat buckled the highway about three feet.
- **b**. He melted her reserve with a husky-voiced paean to her eyes.
- **c**. But Mingo, a major Union Pacific shipping center in the 1890s, has melted away to little more than the grain elevator now.

19.6 Using WordNet, describe appropriate selectional restrictions on the verbs *drink*, *kiss*, and *write*.

19.7 Collect a small corpus of examples of the verbs *drink*, *kiss*, and *write*, and analyze how well your selectional restrictions worked.

19.8 Consider the following examples from (McCawley, 1968):

My neighbor is a father of three.

?My buxom neighbor is a father of three.

What does the ill-formedness of the second example imply about how constituents satisfy, or violate, selectional restrictions?

19.9 Find some articles about business, sports, or politics from your daily newspaper. Identify as many uses of conventional metaphors as you can in these articles. How many of the words used to express these metaphors have entries in either WordNet or your favorite dictionary that directly reflect the metaphor.

19.10 Consider the following example:

The stock exchange wouldn't talk publicly, but a spokesman said a news conference is set for today to introduce a new technology product.

Assuming that stock exchanges are not the kinds of things that can literally talk, give a sensible account for this phrase in terms of a metaphor or metonymy.

19.11 Choose an English verb that occurs in both FrameNet and PropBank. Compare and contrast the FrameNet and PropBank representations of the arguments of the verb.

Chapter 20
Computational Lexical Semantics

To get a single right meaning is better than a ship-load of pearls,
To resolve a single doubt is like the bottom falling off the bucket.
Yuen Mei (1785) (translation by Arthur Waley)

The asphalt that Los Angeles is famous for occurs mainly on its freeways. But in the middle of the city is another patch of asphalt, the La Brea tar pits, and this asphalt preserves millions of fossil bones from the last of the Ice Ages of the Pleistocene Epoch. One of these fossils is the *Smilodon*, or sabre-toothed tiger, instantly recognizable by its long canines. Five million years ago or so, a completely different sabre-tooth tiger called *Thylacosmilus* lived in Argentina and other parts of South America. Thylacosmilus was a marsupial where Smilodon was a placental mammal, but had the same long upper canines and, like Smilodon, had a protective bone flange on the lower jaw. The similarity of these two mammals is one of many example of parallel or convergent evolution, in which particular contexts or environments lead to the evolution of very similar structures in different species (Gould, 1980).

The role of context is also important in the similarity of a less biological kind of organism: the word. Suppose we wanted to decide if two words have similar meanings. Not surprisingly, words with similar meanings often occur in similar contexts, whether in terms of corpora (having similar neighboring words or syntactic structures in sentences) or in terms of dictionaries and thesauruses (having similar definitions, or being nearby in the thesaurus hierarchy). Thus similarity of context turns out to be an important way to detect semantic similarity. Semantic similarity turns out to play an important roles in a diverse set of applications including information retrieval, question answering, summarization and generation, text classification, automatic essay grading and the detection of plagiarism.

In this chapter we introduce a series of topics related to computing with word meanings, or **computational lexical semantics**. Roughly in parallel with the sequence of topics in Ch. 19, we'll introduce computational tasks associated with word senses, relations among words, and the thematic structure of predicate-bearing words. We'll see the role of important role of context and similarity of sense in each of these.

We begin with **word sense disambiguation**, the task of examining word tokens in context and determining which sense of each word is being used. WSD is a task with a long history in computational linguistics, and as we will see, is a non-trivial undertaking given the somewhat elusive nature of many word senses. Nevertheless, there are robust algorithms that can achieve high levels of accuracy given certain reasonable assumptions. Many of these algorithms rely on contextual similarity to help choose the proper sense.

This will lead us natural to a consideration of the computation of **word similarity** and other relations between words, including the **hypernym**, **hyponym**, and **meronym** WordNet relations introduced in Ch. 19. We'll introduce methods based purely on corpus similarity, and others based on structured resources such as WordNet.

Finally, we describe algorithms for **semantic role labeling**, also known as **case role** or **thematic role assignment**. These algorithms generally use features extracted from syntactic parses to assign semantic roles such as AGENT, THEME and INSTRUMENT to the phrases in a sentence with respect to particular predicates.

20.1 Word Sense Disambiguation: Overview

Our discussion of compositional semantic analyzers in Ch. 18 pretty much ignored the issue of lexical ambiguity. It should be clear by now that this is an unreasonable approach. Without some means of selecting correct senses for the words in an input, the enormous amount of homonymy and polysemy in the lexicon would quickly overwhelm any approach in an avalanche of competing interpretations.

word sense disambiguation
WSD

The task of selecting the correct sense for a word is called **word sense disambiguation**, or **WSD**. Disambiguating word senses has the potential to improve many natural language processing tasks. As we'll see in Ch. 25, **machine translation** is one area where word sense ambiguities can cause severe problems; others include **question-answering**, **information retrieval**, and **text classification**. The way that WSD is exploited in these and other applications varies widely based on the particular needs of the application. The discussion presented here ignores these application-specific differences and focuses on the implementation and evaluation of WSD systems as a stand-alone task.

In their most basic form, WSD algorithms take as input a word in context along with a fixed inventory of potential word senses, and return the correct word sense for that use. Both the nature of the input and the inventory of senses depends on the task. For machine translation from English to Spanish, the sense tag inventory for an English word might be the set of different Spanish translations. If speech synthesis is our task, the inventory might be restricted to homographs with differing pronunciations such as *bass* and *bow*. If our task is automatic indexing of medical articles, the sense tag inventory might be the set of MeSH (Medical Subject Headings) thesaurus entries. When we are evaluating WSD in isolation, we can use the set of senses from a dictionary/thesaurus resource like WordNet or LDOCE. Fig. 20.1 shows an example for the word *bass*, which can refer to a musical instrument or a kind of fish.[1]

WordNet Sense	Spanish Translation	Roget Category	Target Word in Context
$bass^4$	lubina	FISH/INSECT	...fish as Pacific salmon and striped **bass** and...
$bass^4$	lubina	FISH/INSECT	...produce filets of smoked **bass** or sturgeon...
$bass^7$	bajo	MUSIC	...exciting jazz **bass** player since Ray Brown...
$bass^7$	bajo	MUSIC	...play **bass** because he doesn't have to solo...

Figure 20.1 Possible definitions for the inventory of sense tags for *bass*.

[1] The WordNet database includes 8 senses; we have arbitrarily selected two for this example; we have also arbitrarily selected one of the many possible Spanish names for fishes which could be used to translate English *sea-bass*.