Jerry R. Hobbs

Information Sciences Institute

University of Southern California

4676 Admiralty Way

Marina del Rey, California 90292

E-mail: hobbs@isi.edu

Tel: +1-310-448-9140

Fax: +1-310-822-0751

33. Word Meaning and World Knowledge

Lexical semantics should be in part about linking the meanings of words with underlying theories of the world. But for this to be even remotely possible, the theories need to be informed by the insights of cognitive and other linguists about the conceptual structure on which language is based. They have to be axiomatizations of a kind of abstract topology that, for example, includes the domains of composite entities (things made of other things), scalar notions, change of state, and causality. Theories of each of these domains are sketched briefly, and it is shown how three very common polysemous words can be defined or characterized in terms of these theories. Finally, there is a discussion of what sort of boundary one can hope to draw between lexical knowledge and other world knowledge.

1 Introduction

We use words to talk about the world. Therefore, to understand what words mean, we should have a prior explication of how we view the world.

Suppose we have a formal logical theory of some domain, or some aspect of the world, that is, a set of predicates intended to capture the concepts in that domain and a set of axioms or rules that constrain the possible meanings of those predicates. Then a formal theory of lexical semantics in that domain would be a matter of writing axioms to relate predicates corresponding to the words in the domain to the predicates in the underlying theory of the domain. For example, the word "until" might be anchored in a formal theory of time that provides an axiomatization of intervals and a *before* relation. (See article 29 *Frame semantics* for a similar view, where frames correspond to the domain theory.)

For the last forty years researchers in artificial intelligence have made efforts to encode various aspects of world knowledge formally. These efforts have primarily been in commonsense physics in the areas of space, time, and qualitative physics, and, in commonsense psychology, in concepts related to belief and intention. A good review of this work that is old but has not lost its relevance is Davis (1990). Most of this work has focused on narrow areas of commonsense knowledge. But there have been several large-scale efforts to encode knowledge of many domains, most notably, Cyc (Lenat and Guha, 1990; Cycorp, 2008). One might think that this work could form the basis of an effort toward a formal theory of lexical semantics anchored in world knowledge. However, these theories for the most part were not designed with language in mind, and in particular what is missing is precisely some of the linguists' insights described in the previous several articles of this volume.

All of this seriously undercuts the utility for lexical semantics of Cyc and similar large ontologies, and indeed of most of the small-scale theories as well.

In trying to link words and world, there are a number of bad ways to go about it. For example, we could take our theory of the world to be quantum mechanics and attempt to define, say, verbs of motion in terms of the primitives provided by that theory. A less obviously wrong approach, and one that has sometimes been tried, is to adopt Euclidean 3-space as the underlying model of space and attempt to define, say, spatial prepositions in terms of that. More common is a serious misstep, with respect to language, that many large-scale ontologies take at the start. Cyc begins by enforcing a rigid distinction between tangible and intangible entities, and in other hierarchical ontologies, the top-level split is between physical and abstract entities. Yet this distinction plays very little role in language. We can be in a room, in a social group, in the midst of an activity, in trouble, and in politics. We can move a chair from the desk to the table, move money from one bank account to another, move a discussion from religion to politics, and move an audience to tears. A fundamental distinction between tangibles and intangibles rules out the possibility of understanding the sense of "in" or "move" common to all these uses.

Our effort, by contrast, has sought to exploit the insights of linguists such as Gruber (1965), the generative semanticists, Johnson (1987), Lakoff

(1987), Jackendoff (see 31 *Conceptual semantics*), and Talmy (see 27 *Cognitive semantics: An overview*). Johnson, Lakoff, Talmy and others have used the term "image schemas" to refer to a conceptual framework that includes topological relations but excludes, for example, Euclidean notions of magnitude and shape. We have been developing core theories that formalize something like the image schemas, and we have been using these to define or characterize words. Among the theories we have developed are theories of composite entities, or things made of other things, the figure-ground relation, scalar notions, change of state, and causality. The idea behind these abstract core theories is that they capture a wide range of phenomena that share certain features. The theory of composite entities, for example, is intended to accommodate natural physical objects like volcanos, artifacts like automobiles, complex events and processes like concerts and photosynthesis, and complex informational objects like mathematical proofs. The theory of scales captures commonalities shared by distance, time, numbers, money, and degrees of risk, severity, and happiness. The most common words in English (and other languages) can be defined or characterized in terms of these abstract core theories. Specific kinds of composite entities and scales, for example, are then defined as instances of these abstract concepts, and we thereby gain access to the rich vocabulary the abstract theories provide.

We can illustrate the link between word meaning and core theories with the rather complex verb "range". A core theory of scales provides axioms involving predicates such as *scale*, $<$, *subscale*, *top*, *bottom*, and *at*. Then we are able to define "range" by the following axiom:

$$
\begin{aligned}
&(\forall\, x, y, z) range(x, y, z) \equiv \\
&\quad (\exists\, s, s_1, u_1, u_2) scale(s) \wedge subscale(s_1, s) \wedge bottom(y, s_1) \\
&\qquad \wedge top(z, s_1) \wedge u_1 \in x \wedge at(u_1, y) \wedge u_2 \in x \wedge at(u_2, z) \\
&\qquad \wedge (\forall\, u \in x)(\exists\, v \in s_1) at(u, v)
\end{aligned}
$$

That is, x ranges from y to z if and only if there is a scale s with a subscale s_1 whose bottom is y and whose top is z, such that some member u_1 of x is at y, some member u_2 of x is at z, and every member u of x is at some point v in s_1. Then by choosing different scales and instantiating the *at* relation in different ways, we can get such uses as

> The buffalo ranged from northern Texas to southern Saskatchewan.
>
> The students' SAT scores range from 1100 to 1550.
>
> The hepatitis cases range from moderate to severe.
>
> His behavior ranges from sullen to vicious.

Many things can be conceptualized as scales, and when this is done, a large vocabulary, including the word "range", becomes available.

It may seem strange for one to embrace logic and the image-schema insight in the same framework, because the two are often taken by cognitive linguists to be contradictory. But the use of logic amounts to less than one might at first think. It can be viewed simply as a well-understood way of representing complex information. To use the notation of first-order logic is to adopt a style of representation that provides for predicate-argument relations (so we know the difference between "Dog bites man" and "Man bites dog"), conjunction (so we have the additive effect of two propositions), implication and modus ponens (so we can derive one proposition from others), and universal instantiation (so we can derive specific instances from general principles). Any adequate representation scheme for knowledge and information must give us at least these features.

The use of logic is also often taken to mean that words have strict definitions, and we know strict definitions are usually not possible. This is why I have used the phrase "define or characterize" rather than "define". In general, we cannot hope to find definitions for words. That is, for very few words p will we find necessary and sufficient conditions, giving us axioms of the form

$$(\forall x)p(x) \equiv \ldots$$

Rather, we will find many necessary conditions and many sufficient conditions.

$$(\forall x)p(x) \supset \ldots$$

$$(\forall x)\ldots \supset p(x)$$

However, the accumulation of enough such axioms will tightly constrain the possible interpretations of the predicate, and hence the meaning of the word.

This, by the way, gives us a different perspective on the notion of semantic primitives. Our theories should be as elegant as possible, and thus they will have as few "central" predicates as possible. These will give the semblance of a small set of semantic primitives, and in fact are similar to those usually proposed. But in our approach we do not attempt to reduce all concepts to undefinable primitive predicates. Rather, strictly speaking, every predicate is primitive, but its set of possible interpretations is more or less tightly constrained by the axioms it participates in (see 17 *Lexical decomposition*; 19 *Lexical conceptual structure*; 21 *Sense relations*).

A further feature required of our logic breaks down the rigidity of formal logic that cognitive linguists sometimes react against. There must be some mechanism for defeasibility; we have to be able to state inferences that are normally true but can be defeated in particular contexts. There are many such logics (e.g., McCarthy, 1980; Ginsberg, 1987; Shoham, 1987). In Hobbs et al. (1993) and Hobbs (2004) it is argued that interpretation of discourse is a matter of coming up with the best proof of the content of an utterance and the fact of its occurrence, using a method of defeasible inference known as abduction. This provides a means of evaluating possibly contradictory "proofs" to determine the best proof, or interpretation. Thus there may be a large number of possible inferences that one may draw in any given context, but only some of them will be a part of the best interpretation. The mystery of how words acquire their manifold shades of meaning in different contexts thereby translates into the problem of how we choose the best interpretation, or, in a sense, how we select the right set of inferences to draw from the use of a word in context. This is far from a solved problem, but recasting meaning and interpretation in this way gives us a formal, computational way of approaching the problem.

Defeasibility in the logic gives us an approach to prototypes (see article 28 *Prototype theory*; Rosch, 1975). Categories correspond to predicates and are characterized by a set of possibly defeasible inferences, expressed as axioms, among which are their traditional defining features. For example, bachelors are unmarried and birds fly.

$$(\forall x) bachelor(x) \supset unmarried(x)$$

$$(\forall x) bird(x) \wedge etc_1(x) \supset fly(x)$$

where $etc_1(x)$ indicates the defeasibility of the axiom. Each instance of a category has a subset of the defeasible inferences that hold in its particular case. The more prototypical, the more inferences. In the case of the penguin, which is not a prototypical bird, the defeasible inference about flying is defeated. In this view, the basic level category is the predicate with the richest set of associated axioms. For example, there is more gain in useful knowledge from learning an animal is a dog than from learning a dog is a boxer.

Similarly, defeasible inference lends itself to a treatment of novel metaphor. In metaphor, some properties are transferred from a source to a target, and some are not. When we say Pat is a pig, we draw inferences about manner and quantity of eating from “pig”, but not about four-leggedness or species membership. The latter inferences are defeated by the other things we know. Hobbs (1992) develops this idea.

Taking abstract core theories as basic may seem to run counter to a central tenet of cognitive linguistics, namely, that our understanding of many abstract domains is founded on spatial metaphor. It is certainly true that the field of spatial relationships, along with social relationships, is one of the domains babies have to figure out first. But I think that to say we figure out space first and then transfer that knowledge to other domains is to seriously underestimate the difficulty of figuring out space. There are many ways one could conceptualize space, e.g., via Euclidean geometry. But in fact it is the topological concepts which predominate in a baby’s spatial understanding. A one-year-old baby fascinated by “in” might put a necklace into a trash can and a Cheerio into a shoe, despite their very different sizes and shapes. In spatial metaphor it is generally the topological properties that get transferred from the source to the target. In taking the abstract core theories as basic, we are isolating precisely the topological properties of space that are most likely to be the basis for understanding metaphorical domains.

If one were inclined to make innateness arguments, one position would be that we are born with a instinctive ability to operate in spatial environments.

We begin to use this immediately when we are born, and when we encounter abstract domains, we tap into its rich models. The alternative, more in line with our development here, is that we are born with at least a predisposition towards instinctive abstract patterns – composite entities, scales, change, and so on – which we first apply in making sense of our spatial environment, and then apply to other, more abstract domains as we encounter them. This has the advantage over the first position that it is specific about exactly what properties of space might be in our innate repertoire. For example, the scalar notions of "closer" and "farther" are in it; exact measures of distance are not. A nicely paradoxical coda for summing up this position is that we understand space by means of a spatial metaphor. I take Talmy's critique of the "concreteness as basic" idea as making a similar point (see 27 *Cognitive semantics: An overview*).

Many of the preceding articles have proposed frameworks for linking words to an underlying conceptual structure. These can all be viewed as initial forays into the problem of connecting lexical meaning with world knowledge. The content of this work survives translation among the various frameworks that have been used for examining it, and survives recasting it as a problem of explicitly encoding world knowledge, specifically, a theory of image schemas explicating such concepts as composite entities, figure-ground, scales, change of state, causality, aggregation, and granularity shifts—an abstract theory that can be instantiated in many different, more specialized domains. The core theories we are developing are not so much theories about *particular* aspects of the world, but rather abstract frameworks that are useful in making sense of a number of different kinds of phenomena. Levin and Rappaport Hovav (see 19 *Lexical conceptual structure*) say, "All theories of event structure, either implicitly or explicitly, recognize a distinction between the primitive predicates which define the range of event types available and a component which represents what is idiosyncratic in a verb's meaning." The abstract theories presented here are an explication of the former of these.

This work can be seen as an attempt at a kind of deep lexical semantics. Not only are the words "decomposed" into what were once called primitives, but also the primitives are explicated in axiomatic theories, enabling one to reason deeply about the concepts conveyed by the text.

2 Core abstract theories

2.1 Composite entities

Composite entities are things made of other things. A composite entity is characterized by a set of components, a set of properties of these components, and a set of relations among the components and between the components and the whole. The concept of composite entity captures the minimal complexity something must have in order for it to have structure. It is hard to imagine something that cannot be conceptualized as a composite entity. For this reason, a vocabulary for talking about composite entities will be broadly applicable.

The elements of a composite entity can themselves be viewed as composite entities, and this gives us a very common example of shifting granularities. It allows us to distinguish between the *structure* and the *function* of an entity. The function of an entity as a component of a larger composite entity is its relations to the other elements of the larger composite entity, its environment, while the entity itself is viewed as indecomposable. The structure of the entity is revealed when we decompose it and view it as a composite entity itself. We look at it at a finer granularity.

An important question any time we can view an entity both functionally and structurally is how the functions of the entity are implemented in its structure. We need to spell out the structure-function articulations.

For example, a librarian might view a book as an indecomposable entity and be interested in its location in the library, its relationship to other books, to the bookshelves, and to the people who check the book out. This is a functional view of the book with respect to the library. We can also view it structurally by inquiring as to its parts, its content, its binding, and so on. In spelling out the structure-function articulations, we might say something about how its content, its size, and the material used in its cover determines its proper location in the library.

A composite entity can serve as the *ground* against which some external *figure* can be located or can move (see 27 *Cognitive semantics: An overview*). A primitive predicate *at* expresses this relation. In

$$at(x, y, s)$$

s is a composite entity, y is one of its elements, and x is an external entity. The relation says that the figure x is at a point y in the composite entity s, which is the ground.

The *at* relation plays primarily two roles in the knowledge base. First, it is involved in the "decompositions" of many lexical items. We saw this above in the definition of "range". There is a very rich vocabulary of terms for talking about the figure-ground relation. This means that whenever a relation in some domain can be viewed as an instance of the figure-ground relation, we acquire at a stroke a rich vocabulary for talking about that domain.

This gives rise to the second role the *at* predicate plays in the knowledge base. A great many specific domains have relations that are stipulated to be instances of the *at* relation. There are a large number of axioms of the form

$$(\forall\, x, y, s) r(x, y, s) \supset at(x, y, s)$$

It is in this way that many of the metaphorical usages that pervade natural language discourse are accommodated. Once we characterize some piece of the world as a composite entity, and some relation as an *at* relation, we have acquired the whole locational way of talking about it. Once this is enriched with a theory of time and change, we can import the whole vocabulary of motion. For example, in computer science, a data structure can be viewed as a composite entity, and we can stipulate that if a pointer points to a node in a data structure, then the pointer is *at* that node. We have then acquired a spatial metaphor, and we can subsequently talk about, for example, the pointer *moving around* the data structure. Space, of course, is itself a composite entity and can be talked about using a locational vocabulary.

Other examples of *at* relations are

A person at an object in a system of objects:

> John is at his desk.

An object at a location in a coordinate system:

> The post office is at the corner of 34th Street and Eighth Avenue.

A person's salary at a particular point on the money scale:

> John's salary reached $75,000 this year.

An event at a point on the time line:

> The meeting is at three o'clock.

2.2 Scales

The theory of scales was mentioned in the introduction. It provides the basic vocabulary for talking about partial orderings, including *scale*, $<$, *subscale*, total ordering, *top*, *bottom*, *reverse*, and intervals. The theory also explicates monotone-increasing scale-to-scale functions ("the more X, the more Y"), the construction of composite scales, and the characterization of qualitatively high and low regions of a scale.

A scale is a composite entity, so we can talk about an entity being *at* a point on the scale. An obvious example of a scale is the scale of nonnegative integers. The cardinality of a set can be defined in the standard way:

$$card(\phi) = 0$$

$$(\forall\, x, s) x \notin s \;\supset\; card(\{x\} \cup s) = card(s) + 1$$

We can then define cardinality to be an *at* relation, where N is the scale of nonnegative integers:

$$(\forall\, s, n) card(s) = n \;\supset\; at(s, n, N)$$

This gives us access to the rich vocabulary of spatial relationships when talking about cardinality, allowing us to say things like

> The population of Cairo *reached* 15 million this year.

Many scales are composite. A scale s is a composite of scales s_1 and s_2 if its elements are the ordered pairs $< x, y >$ where x is in s_1 and y is in s_2.

The ordering in s has to be consistent with the orderings in s_1 and s_2; if x_1 is less than x_2 in s_1, and y_1 is less than y_2 in s_2, then $< x_1, y_1 >$ is less than $< x_2, y_2 >$ in s. The converse is not necessarily true; the composite scale may have more structure than that inherited from its component scales. We need composite scales to deal with complex scalar predicates, such as *damage*. When something is damaged, it no longer fulfills its function in a goal-directed system. It needs to be repaired, and repairs cost. Thus, there are (at least) two ways in which damage can be serious, first in the degradation of its function, second in the cost of its repair. These are independent scales. Damage that causes a car not to run may cost next to nothing to fix, and damage that only causes the car to run a little unevenly may be very expensive.

It is very useful to be able to isolate the high and low regions of a scale. We can do this with operators called Hi and Lo. The Hi region of a scale includes its top; the Lo region includes its bottom. The points in the Hi region are all greater than any of the points in the Lo region. Otherwise, there are no general topological constraints on the Hi and Lo regions. In particular, the bottom of the Hi region and the top of the Lo region may be indeterminate with respect to the elements of the scale. The Hi and Lo operators provide us with a coarse-grained structure on scales, useful when greater precision is not necessary or not possible.

The absolute form of adjectives frequently isolate Hi and Lo regions of scales. A totally ordered Height Scale can be defined precisely, but frequently we are only interested in qualitative judgments of height. The word "tall" isolates the Hi region of the Height Scale; the word "short" isolates the Lo region. A Happiness Scale cannot be defined precisely. We cannot get much more structure for a Happiness Scale than what is given to us by the Hi and Lo operators. The Hi and Lo operators can be iterated, to give us the concepts "happy", "very happy", and so on.

In any given context, the Hi and Lo operators will identify different regions of the scale. That is, the inferences we can draw from the fact that something is in the Hi region of a scale are context-dependent; indeed, inferences are always context-dependent. But two important constraints on the Hi and Lo regions relate them to distributions and functionality. The Hi and Lo regions must be related to common distributions of objects on the scale in an as-yet nonexistent qualitative theory of distributions. If something is significantly above average for the relevant set, then it is in the Hi region. The regions must also be related to goal-directed behavior; often something is in the Hi region of a scale precisely because that property aids or defeats the achievement of some goal in a plan. For example, saying that a talk is long often means that it is longer than the audience's attention span, and thus the goal of conveying information is defeated. Often when we call someone tall, we mean tall enough or too tall for some purpose.

2.3 Change of state

A predicate of central importance is the predicate *change*. This is a relation between situations, or conditions, or predications, and indicates a change of state. A change from p being true of x to q being true of x, using an ontologically promiscuous notation that reifies states and events (see Hobbs, 1985; 35 *Event semantics*), can be represented

$$change(e_1, e_2) \wedge p'(e_1, x) \wedge q'(e_2, x)$$

This says that there is a change from the situation e_1 of p being true of x to the situation e_2 of q being true of x. A very common pattern involves a change of location:

$$change(e_1, e_2) \wedge at'(e_1, x, y, s) \wedge at'(e_2, x, z, s)$$

That is, there is a change from the situation e_1 of x being at y in s to the situation e_2 of x being at z in s.

When there is a change, generally there is some entity involved in both the start and end states; there is something that is changing—x in the above formulas.

The predicate *change* possesses a limited transitivity. There was a change from Bill Clinton being a law student to Bill Clinton being President, because they are two parts of the same ongoing process, even though he was governor in between. There was a change from Bill Clinton being President to George W. Bush being president. But we probably do not want to say there was a change from Bill Clinton being a law student to George W. Bush being President. They are not part of the same process.

A state cannot change into the same state without going through an intermediate different state.

The concept of *change* is linked with time in the obvious way. If state e_1 changes into state e_2, then e_2 cannot be before e_1. My view is that the relation between change and time is much deeper, cognitively. The theory of change of state suggests a view of the world as consisting of a large number of more or less independent, occasionally interacting processes, or histories, or sequences of events. x goes through a series of changes, and y goes through a series of changes, and occasionally there is a state that involves a relation between the two. We can then view the time line as an artificial construct, a regular sequence of imagined abstract events—think of them as ticks of a clock in the National Institute of Science and Technology—to which other events can be related by chains of copresence. Thus, I know I went home at six o'clock because I looked at my watch, and I had previously set my watch by going to the NIST Web site. In any case, there is no need to choose between such a view of time and one that takes time as basic. They are inter-definable in a straightforward fashion (Hobbs et al., 1987).

For convenience, we define one-argument predicates $changeFrom$ and $changeTo$, suppressing one or the other argument of *change*.

2.4 Cause

Our treatment of causality (Hobbs, 2005) rests on a distinction between causal complexes and the predicate *cause*. When we flip a switch and the light comes on, we say that flipping the switch caused the light to come on. But many other factors were involved. The wiring and the light bulb had to be intact, the power had to be on in the city, and so forth. We say that all these other states and events constitute the causal complex for the effect. A causal complex for an effect is the set of all the eventualities that must happen or hold in order for the effect to occur. The two principal properties of causal complexes are that when all the eventualities happen, the effect happens, and that every eventuality in the causal complex is required for the effect to happen. These are strictly true, and the notion of causal complex is not a defeasible one.

The "cause" of an effect, by contrast, is a distinguished element within the causal complex, one that cannot normally be assumed to hold. It is often the action that is under the agent's immediate control. It is only defeasibly true that when a cause occurs the effect also occurs. This inference can be defeated because some of the other states and events in the causal complex that normally hold do not hold in this particular case. The notion of *cause* is much more useful in commonsense reasoning because we can rarely if ever enumerate all the eventualities in a causal complex. Most of our commonsense causal knowledge is expressed in terms of the predicate *cause*.

The concept *cause* has the expected properties, such as defeasible transitivity and consistency with temporal ordering. But we should not expect to have a highly developed theory of causality *per se*. Rather we should expect to see causal information distributed throughout our knowledge base. For example, there is no axiom of the form

$$(\forall\, e_1, e_2) cause(e_1, e_2) \equiv \ldots$$

defining *cause*. But there will be many axioms of the forms

$$p'(e_1, x) \supset q'(e_2, x) \wedge cause(e_1, e_2)$$
$$r'(e_3, x) \supset p'(e_1, x) \wedge cause(e_1, e_3)$$

expressing causal connections among specific states and events; e.g., p-like events cause q-like events or r-like events are caused by p-like events. We don't know precisely what causality is, but we know lots and lots of examples of things that cause other things.

Some would urge that causes and effects can only be events, but it seems to me that we want to allow states as well, since in

> The slipperiness of the ice caused John to fall.

the cause is a state. Moreover, intentional agents are sometimes taken to be the unanalyzed causes of events. In

> John lifted his arm.

John is the cause of the change of position of his arm, and we probably don't want to have to coerce this argument into some imagined event taking place inside John. Physical forces may also act as causes, as in

Gravity causes the moon to circle the earth.

The world is laced with threads of causal connection. In general, two entities x and y are causally connected with respect to some behavior p of x, if whenever p happens to x, there is some corresponding behavior q that happens to y. Attachment of physical objects is one variety of causal connection. In this case, p and q are both *move*. If x and y are attached, moving x causes y to move. Containment is similar.

A particularly common variety of causal connection between two entities is one mediated by the motion of a third entity from one to the other. This might be called, somewhat facetiously, a "vector boson" connection. In particle physics, a vector boson is an elementary particle that transfers energy from one point to another. Photons, which really are vector bosons, mediate the causal connection between the sun and our eyes. Other examples of such causal connections are rain drops connecting a state of the clouds with the wetness of our skin and clothes, a virus transmitting disease from one person to another, and utterances passing information between people.

Containment, barriers, openings, and penetration are all with respect to paths of causal connection. Force is causality with a scalar structure (see 27 *Cognitive semantics: An overview*).

The event structure underlying many verbs exhibits causal chains. Instruments, for example, are usually vector bosons. In the sentence,

John pounded the nail with a hammer for Bill.

the underlying causal structure is that the agent John causes a change in location of the instrument, the hammer, which causes a change in location of the object or theme, the nail, which causes or should cause a change in the mental or emotional state of the beneficiary, Bill.

$$Agent \text{ –}cause\text{–}> \; change(at(Instr, x, s), at(Instr, Object, s))$$

$$\text{–}cause\text{–}> \; change(at(Object, y_1, s), at(Object, y_2, s))$$

$$\text{–}cause\text{–}> \; change(p_1(Beneficiary), p_2(Beneficiary))$$

Much of case grammar and work on thematic roles can be seen as a matter of identifying where the arguments of verbs fit into this kind of causal chain when we view the verbs as instantiating this abstract frame (see Jackendoff, 1972; 18 *Thematic roles*; 19 *Lexical conceptual structure*).

In addition, in this theory we define such concepts as *enable*, *prevent*, *help*, and *obstruct*. There are also treatments of attempts, success, failure, ability, and difficulty.

With this vocabulary, we are in a position to characterize more precisely the intuitive notions of state, event, action, and process. A state is a static property that does not involve a change (at the relevant granularity), such as an *at* relationship, $at(x, y, s)$. To be up, for example, is a state. An event is a change of state, a common variety of which is a change of location:

$$change(e_1, e_2) \wedge at'(e_1, x, y, s) \wedge at'(e_2, x, z, s)$$

For example, the verb "rise" denotes a change of location of something to a higher point. An action is the causing of an event by an intentional agent:

$$cause(a, e) \wedge change'(e, e_1, e_2) \wedge at'(e_1, x, y, s) \wedge at'(e_2, x, z, s)$$

The verb "raise" denotes an action by someone of effecting a change of location of something to a higher point. A process is a sequence of events or actions. For example, to fluctuate is to undergo a sequence of risings and fallings, and to pump is to engage in a sequence of raisings and lowerings. We can coarsen the granularity on processes so that the individual changes of state become invisible, and the result is a state. This is a transformation of perspective that is effected by the progressive aspect in English. Thus, fluctuating can be viewed as a state.

Detailed expositions of all the core theories can be found at

`http://www.isi.edu/ hobbs/csk.html`

3 Linking word meaning with the theories

Once we have in place the core theories that capture world knowledge at a sufficiently abstract level, we can begin to construct the axioms that link word meaning to the theories. We illustrate here how that would go, using the words "have", "remove", and "remain". Words have senses, and for each sense the linkage will be different. Here we examine the word senses in

WordNet (Miller, 1995) and FrameNet (Baker et al., 2003), since they are the most heavily used lexical resources in computational linguistics. The word sense numbers correspond to their order in the Web interfaces to the two resources:

`http://wordnet.princeton.edu/`

`http://framenet.icsi.berkeley.edu`

3.1 "Have"

In WordNet the verb "have" has 19 senses. But they can be grouped into three broad "supersenses". In its first supersense, X has Y means that X is in some relation to Y. The WordNet senses this covers are as follows:

1. a broad sense, including have a son, having a condition hold and having a college degree
2. having a feature or property, i.e., the property holding of the entity
3. a sentient being having a feeling or internal property
4. a person owning a possession

7. have a person related in some way: have an assistant

9. have left: have three more chapters to write

12. have a disease: have influenza

17. have a score in a game: have three touchdowns

The supersense can be characterized by the axiom

$$have\text{-}s1(x,y) \supset relatedTo(x,y)$$

In these axioms, supersenses are indexed with s, WordNet senses with w, and FrameNet senses with f. Unindexed predicates are from core theories.

The individual senses are then specializations of the supersense where more domain-specific predicates are explicated in more specialized domains. For example, sense 4 relates to the supersense as follows:

$$have\text{-}w4(x,y) \equiv possess(x,y)$$

$$have\text{-}w4(x,y) \supset have\text{-}s1(x,y)$$

where the predicate *possess* would be explicated in a commonsense theory of economics, relating it to the priveleged use of the object. Similarly, $have\text{-}w3(x,y)$ links with the supersense but has the restrictions that x is sentient and that the "relatedTo" property is the predicate-argument relation between the feeling and its subject.

The second supersense of "have" is "come to be in a relation to". This is our *changeTo* predicate. Thus, the definition of this supersense is

$$have\text{-}s2(x,y) \equiv changeTo(e) \wedge have\text{-}s1'(e,x,y)$$

The WordNet senses this covers are as follows:

10. be confronted with: we have a fine mess
11. experience: the stocks had a fast run-up
14. receive something offered: have this present
15. come into possession of: he had a gift from her

16. undergo, e.g., an injury: he had his arm broken in the fight

18. have a baby

In these senses the new relation is initiated but the subject does not necessarily play a causal or agentive role. The particular change involved is specialized in the WordNet senses to a confronting, a receiving, a giving birth, and so on.

The third supersense of "have" is "cause to come to be in a relation to". The axiom defining this is

$$have\text{-}s3(x, y) \equiv cause(x, e) \wedge have\text{-}s2'(e, x, y)$$

The WordNet senses this covers are

5. cause to move or be in a certain position or condition: have your car ready

6. consume: have a cup of coffee

8. organize: have a party

13. cause to do: she had him see a doctor

19. have sex with

In all these cases the subject initiates the change of state that occurs.

FrameNet has five simple transitive senses for "have". Their associated frames are

1. Have associated

2. Possession

3. Ingestion

4. Inclusion

5. Birth

The first sense corresponds to the first WordNet supersense:

$$have\text{-}f1(x,y) \equiv have\text{-}s1(x,y)$$

The second sense is WordNet sense 4.

$$have\text{-}f2(x,y) \equiv have\text{-}w4(x,y)$$

The third sense is WordNet sense 6. The fourth sense is a *partOf* relation. It is a specialization of WordNet sense 2.

$$have\text{-}f4(x,y) \equiv partOf(x,y)$$

$$have\text{-}f4(x,y) \supset have\text{-}w2(x,y)$$

The fifth sense is WordNet sense 18.

3.2 "Remove"

If x removes y from z, then x causes a change from the state in which y is at z.

$$remove(x,y,z) \supset cause(x,e_1) \wedge changeFrom'(e_1,e_2) \wedge at'(e_2,y,z,s)$$

This is the "supersense" covering all of the WordNet and FrameNet senses of "remove".

WordNet lists 8 senses of "remove". In WordNet sense 1, *at* is instantiated as physical location. In sense 2, *at* is instantiated as position in an organization, as in "The board removed the VP of operations." In sense 3, y is somehow dysfunctional, as in removing trash. In sense 4, *at* is instantiated as the membership relation in a set; y is removed from set z. In sense 5, the change is functional or strategic, as in a general removing his troops from a vulnerable position. In sense 6, x and y are identical, as in "He removed himself from the contest." In sense 7, *at* is instantiated as "alive", as in "The Mafia don removed his enemy." In sense 8, y is abstract and dysfunctional, as in removing an obstacle.

FrameNet has two senses of the word. The first is the general meaning, our supersense. In the second sense, x is a person, y is clothes, and z is a body.

Note that the supersense gives the topological structure of the meaning of the verb. The various senses are then generated from that by instantiating the *at* relation to something more specific, or by adding domain constraints to the arguments x, y and z.

3.3 "Remain"

There are four WordNet senses of the verb "remain":

1. Not change out of an existing state: He remained calm.
2. Not change out of being at a location: He remained at his

post.

3. Entities in a set remaining after others are removed: Three problems remain.

4. A condition remains in a location: Some smoke remained after the fire was put out.

The first sense is the most general and subsumes the other three. We can characterize it by the axiom

$$remain\text{-}w1(x, e) \supset arg(x, e) \wedge \neg changeFrom(e)$$

That is, if x remains in condition e, then e is a property of x (or x is an argument of e), and there is no change from state e holding. By the properties of $changeFrom$ it follows that x is in state e, as is presupposed.

In the second sense, the property e of x is being in a location.

$$remain\text{-}w2(x, e) \equiv remain\text{-}w1(x, e) \wedge at'(e, x, y)$$

The fourth sense is a specialization of the second sense in which the entity x that remains is a state or condition.

$$remain\text{-}w4(x, e) \equiv remain\text{-}w2(x, e) \wedge state(x)$$

The third sense is the most interesting to characterize. As in the fourth WordNet sense of "remove", there is a process that removes elements from a set, and what remains is the set difference between the original and the set of elements that are removed. In this axiom x remains after process e.

$$remain\text{-}w3(x,e) \equiv remove\text{-}w4'(e,y,s_2,s_1) \wedge setdiff(s_3,s_1,s_2) \wedge member(x,s_3)$$

That is, x remains after e if and only if e is a removal event by some agent y of a subset s_2 from s_1, s_3 is the set difference between s_1 and s_2, and x is a member of s_3.

There are four FrameNet senses of "remain". The first is the same as WordNet sense 1. The second is the same as WordNet sense 3. The third and fourth are two specializations of WordNet sense 3, one in which the removal process is destructive and one in which it is not.

There are two nominalizations of the verb "remain"—"remainder" and "remains". All of their senses are related to WordNet sense 3. The first WordNet noun sense is the most general.

$$remainder\text{-}w1(x,e) \equiv remain\text{-}w3(x,e)$$

That is, x is the remainder after process e if and only if x remains after e. The other three senses result from specialization of the removal process to arithmetic division, arithmetic subtraction, and the purposeful cutting of a piece of cloth. The noun "remains" refers to what remains ($w3$) after a process of consumption or degradation.

3.4 The nature of word senses

The most common words in a language are typically the most polysemous. They often have a central meaning indicating their general topological structure. Each new sense introduces inferences that cannot be reliably determined just from a core meaning plus contextual factors. They tend to build up along what Brugmann (1981), Lakoff (1987) and others have called a radial category structure (see 28 *Prototype theory*). Sense 2 may be a slight modification of sense 1, and senses 3 and 4 different slight modifications of sense 2. It is easy to describe the links that take us from one sense to an adjacent one in the framework presented here. Each sense corresponds to a predicate which is characterized by one or more axioms involving that predicate. A move to an adjacent sense happens when incremental changes are made to the axioms. As we have seen in the examples of this section, the changes are generally additions to the antecedents or consequents of the axioms. The principal kinds of additions are embedding in *change* and *cause*, as we saw in the supersenses of "have"; the instantiation of general predicates like *relatedTo* and *at* to more specific predicates in particular domains, as we saw in all three cases; and the addition of domain-specific constraints on arguments, as in restricting y to be clothes in *remove-f2*.

A good account of the lexical semantics of a word should not just catalog various word senses. It should detail the radial category structure of the word senses, and for each link, it should say what incremental addition or modification resulted in the new sense. Note that radial categories provide us with a logical structure for the lexicon, and also no doubt a historical one, but not a developmental one. Children often learn word senses independently and only later if ever realize the relation among the senses. See article 28 *Prototype theory* for further discussion of issues with respect to radial categories.

4 Distinguishing lexical and world knowledge

It is perhaps natural to ask whether a principled boundary can be drawn between linguistic knowledge and knowledge of the world. To make this issue more concrete, consider the following seven statements:

(1) If a string w_1 is a noun phrase and a string w_2 is a verb phrase, then the concatenation $w_1 w_2$ is a clause.

(2) The transitive verb "moves" corresponds to the predication $move_2(x, y)$, providing a string describing x occurs as its subject and a string describing y occurs as its direct object.

(3) If an entity x moves (in sense $move_2$) an entity y, then x causes a change of state or location of y.

(4) If an entity y changes to a new state or location, it is no longer in its old state or location.

(5) If a physical object x moves a physical object y through a fluid medium, then x must apply force to y against the resistance of the medium.

(6) The function of a barge is to move freight across water.

(7) A barge moved the wreckage of Flight 1549 to New Jersey.

Syntax consists in part of rules like (1), or generalizations of them. One could view the lexicon as consisting of axioms expressing information like (2), specifying for each word sense and argument realization pattern what predication is conveyed, perhaps together with some generalizations of such statements. (Lexical knowledge of other languages would be encoded as similar axioms, sometimes linking to the same underlying predicates, sometimes different.) Axioms expressing information like (3) link the lexical predicates with underlying domain theories, in this case, theories of the abstract domains of causality and change of state. Axioms expressing facts like (4) are internal to domain theories, in this case, the theory of the abstract domain of change of state. Axioms expressing general facts like (5) are part of a commonsense or scientific theory of physics, which can be viewed as a specialization and elaboration of the abstract theories. Axioms expressing facts like (6) encode telic information about artifacts. Statement (7) is a specific, accidental fact about the world.

Many have felt that the viability of lexical semantics as a research enterprise requires a principled distinction between lexical knowledge and world knowledge, presumably somewhere below axioms like (2) and above facts like (7). Many of those who have believed that no such distinction is possible have concluded that lexical semantics is impossible, or at least can only be very limited in its scope.

For example, in his discussion of meaning, Bloomfield (1933, pp. 139-140) rules out the possibility of giving definitions of most words.

> In order to give a scientifically accurate definition of meaning of every form of a language, we should have to have a scientifically accurate knowledge of everything in the speakers' world. While this may be possible for certain scientifically well-understood terms like "salt", we have no precise way of defining words like "love" or "hate" which concern situations that have not been accurately classified – and these latter are in the great majority.

He concludes that

> The statement of meanings is therefore a weak point in language-study, and will remain so until human knowledge advances very far beyond its present state.

Lexical semantics is impossible because we would need a theory of the world. Bloomfield goes on to talk about such phenomena as synonymy and antonymy, and leaves issues of meaning at that.

More recently, Fodor (1980) similarly argued that lexical semantics would need a complete and correct scientific theory of the world to proceed, and is consequently impossible in the foreseeable future.

A counterargument is that we don't need a scientifically correct theory of the world, because people don't have that as they use language to convey meaning. We rather need to capture people's commonsense theories of the world. In fact, there are a number of interesting engineering efforts to encode commonsense and scientific knowledge needed in specific applications or more broadly. Large ontologies of various domains, such as biomedicine and geography, are being developed for the Semantic Web and other computational uses. Cyc (Lenat and Guha, 1990) has been a large-scale effort to encode commonsense knowledge manually since the middle 1980s; it now contains millions of rules. The Open Mind Common Sense project (Singh, 2002) aims at accumulating huge amounts of knowledge rapidly by marshaling millions of "netizens" to make contributions; for example, a participant might be asked to complete the sentence "Water can ..." and reply with "Water can put out fires." Many of these projects, including Cyc, involve a parallel effort in natural language processing to relate their knowledge of the world to the way we talk about the world. Might we do lexical semantics by explicating the meanings of words in terms of such theories?

Fodor (1983) can be read as responding to this possibility. He argues that peripheral processes like speech recognition and syntactic processing are encapsulated in the sense that they require only limited types of information. Central processes like fixation of belief, by contrast, can require any knowledge from any domain. He gives the example of the power of analogical reasoning in fixation of belief. The body of knowledge that can be appealed to in analogies can not be circumscribed; analogies might involve mappings from anything to anything else. Scientific study of modular processes is feasible, but scientific study of global processes is not. No scientific account of commonsense reasoning is currently available or likely to be in the foreseeable future; by implication reasoning about commonsense world knowledge is not currently amenable to scientific inquiry, nor is a lexical semantics that depends on it. Syntax *is* amenable to scientific study, but only, according to Fodor, because it is informationally encapsulated.

Thus, the debate on this issue often centers on the modularity of syntax. Do people do syntactic analysis of utterances in isolation from world knowledge? Certainly at time scales at which awareness functions, there is no distinction in the processing of linguistic and world knowledge. We rarely if ever catch ourselves understanding the syntax of a sentence we hear without understanding much about its semantics. For example, in Chomsky's famous grammatical sentence, "Colorless green ideas sleep furiously," there is no stage in comprehension at which we are aware that "colorless" and "green" are adjectives, but haven't yet realized they are contradictory.

Moreover, psychological studies seem to indicate that syntactic processing and the use of world knowledge are intricately intertwined. Much of this work has focused on the use of world knowledge to resolve references and disambiguate ambiguous prepositional phrase attachments. Tanenhaus and Brown-Schmidt (2008) review some of this research that makes use of methods of monitoring eye movements to track comprehension. For example, they present evidence that subjects access the current physical context while they are processing syntactically ambiguous instructions and integrate it with the language immediately. In terms of our examples, they are using facts like (1) and facts like (7) together. The authors contend that their results “are incompatible with the claim that the language processing includes subsystems (modules) that are informationally encapsulated, and thus isolated from high-level expectations.”

Often the line between linguistic and world knowledge is drawn to include selectional constraints within language. Hagoort et al. (2004) used electroencephalogram and functional magnetic resonance imaging data to investigate whether there was any difference between the temporal course of processing true sentences like “Dutch trains are yellow and very crowded”, factually false but sensible sentences like “Dutch trains are white and very crowded”, and sentences that violate selectional constraints like “Dutch trains are sour and very crowded.” The false sentences and the selectionally anomalous sentences showed a virtually identical peak of activity in the left inferior prefrontal cortex. The authors observed that there is “strong empirical evidence that lexical semantic knowledge and general world knowledge are both integrated in the same time frame during sentence interpretation, starting at $\sim$300ms after word onset.” However, there is a difference in frequency profile between the two conditions, consisting of a measurable increase in activity in the 30-70 Hz range (gamma frequency) for the false sentences, and an increase in the 4-7Hz range (theta frequency) in the anomalous condition. The authors conclude that "semantic interpretation is not separate from its integration with nonlinguistic elements of meaning," but that nevertheless "the brain keeps a record of what makes a sentence hard to interpret, whether this is word meaning or world knowledge."

Thus, if the brain makes a distinction between linguistic and world knowledge, it does not appear to be reflected in the temporal course of processing language.

The most common argument in linguistics and related fields for drawing a strict boundary between lexicon and world is a kind of despair that a scientific study of world knowledge is possible. Others have felt it is possible to identify lexically relevant domains of world knowledge that are accessible to scientific study.

Linguists investigating "lexical conceptual structure" (e.g., see article 19 *Lexical conceptual structure*) are attempting to discover generalizations in how the way an entity occurs in the underlying description of a situation or event in terms of abstract topological predicates influences the way it is realized in the argument structure in syntax. For example, do verbs that undergo dative alternation all have a similar underlying abstract structure? Does the causative always involve embedding an event as the effect in a causal relation, where the cause is the agent or an action performed by the agent? The hypothesis of this work is that facts like (2), which are linguistic, depend crucially on facts like (3), which have a more world-like flavor. However, this does not mean that we have identified a principled boundary between linguistic and world knowledge. One could just as well view this as a strategic decision about how to carve out a tractable research problem.

Pustejovsky (1995) pushes the line between lanaguage and world farther into the world. He advocates representing what he calls the "qualia structure" of words, which includes facts about the constituent parts of an entity (Constitutive), its place in a larger domain (Formal), its purpose and function (Telic), and the factors involved in its origin (Agentive). One can then, for example, use the Telic information to resolve a metonymy like "She began a cigarette" into its normal reading of "She began smoking a cigarette," rather than any one of the many other things one could do with a cigarette—eating it, rolling it, tearing it apart, and so on. His framework is an attempt to relate facts like (2) about what arguments can appear with what predicates with facts like (6) about the functions and other properties of things. Several places in his book, Pustejovsky suggests that it is important to see his qualia structures as part of lexical semantics, and hence linguistics, as opposed to general commonsense knowledge that is not linguistic. But he never makes a compelling argument to this effect. All of his qualia structures and coercion mechanisms are straightforward to express in a logical framework, so there are no formal reasons for the distinction.

I think it is best to see this particular carving out of knowledge and interpretation processes, as with the study of lexical conceptual stuctures, as a strategic decision to identify a fruitful and tractable research problem.

Pustejovsky's work is an attempt to specify the knowledge that is required for interpreting at least the majority of nonstandard uses of words. Kilgarriff (2001) tests this hypothesis by examining the uses of nine particular words in a 20-million word corpus. 41 of 2276 instances were judged to be nonstandard since they did not correspond to any of the entries for the word in a standard dictionary. Of these, only two nonstandard uses were derivable from Pustejovsky's qualia structures. The others required deeper commonsense knowledge or previous acquaintance with collocations. Kilgarriff's conclusion is that "Any theory that relies on a distinction between general and lexical knowledge will founder." (p. 325)

Some researchers in natural language processing have argued that lexical knowledge should be distinguished from other knowledge because it results in more efficient computation or more efficient comprehension and production. One example concerns hyperonymy relations, such as that $car(x)$ implies $vehicle(x)$. It is true that some kinds of inferences lend themselves more to efficient computation than others, and inferences involving only monadic predicates are one example. But where this is true, it is a result not of their content but of structural properties of the inferences, and these cut across the lexical-world distinction. Any efficiency realized in inferring $vehicle(x)$

can be realized in inferring $expensive(x)$ as well.

All of statements (1)-(7) are facts about the world, because sentences and their structure and words and their roles in sentences are things in the world, as much as barges, planes, and New Jersey. There is certainly knowledge we have that is knowledge about words, including how to pronounce and spell words, predicate-argument realization patterns, alternation rules, subcategorization patterns, grammatical gender, and so on. But words are part of the world, and one might ask why this sort of knowledge should have any special cognitive status. Is it any different in principle from the kind of knowledge one has about friendship, cars, or the properties of materials? In all these cases, we have entities, properties of entities, and relations among them. Lexical knowledge is just ordinary knowledge where the entities in question are words. There are no representational reasons for treating linguistic knowledge as special, providing we are willing to treat the entities in our subject matter as first-class individuals in our logic (cf. Hobbs, 1985). There are no procedural reasons for treating linguistic knowledge as special, since parsing, argument realization, lexical decomposition, the coercion of metonymies, and so on can all be implemented straightforwardly as inference. The argument that parsing and lexical decomposition, for example, can be done efficiently on present-day computers, whereas commonsense reasoning cannot, does not seem to apply to the human brain; psycholinguistic studies show that the influence of world knowledge kicks in as early as syntactic and lexical knowledge, and yields the necessary results just as quickly.

We are led to the conclusion that any drawing of lines is for the strategic purpose of identifying a coherent, tractable and fruitful area of research. Statements (1)-(6) are examples from six such areas. Once we have identified and explicated such areas, the next question is what connections or articulations there are among them; Pustejovskky's research and work on lexical conceptual structures are good examples of people addressing this question.

However, all of this does not mean that linguistic insights can be ignored. The world can be conceptualized in many ways. Some of them lend themselves to a deep treatment of lexical semantics, and some of them impede it. Put the other way around, looking closely at language leads us to a particular conceptualization of the world that has proved broadly useful in everyday life. It provides us with topological relations rather than with the precision of Euclidean 3-space. It focuses on changes of state rather than on correspondences with an *a priori* time line. A defeasible notion of causality is central in it. It provides means for aggregation and shifting granularities. It encompasses those properties of space that are typically transferred to new target domains when what looks like a spatial metaphor is invoked.

More specific domains can then be seen as instantiations of these abstract theories. Indeed, Euclidean 3-space itself is such a specialization. Language provides us with a rich vocabulary for talking about the abstract domains. The core meanings of many of the most common words in language can be defined or characterized in these core theories. When the core theory is instantiated in a specific domain, the vocabulary associated with the abstract domain is also instantiated, giving us a rich vocabulary for talking about and thinking about the specific domain. Conversely, when we encounter general words in the contexts of specific domains, understanding how the specific domains instantiate the abstract domains allows us to determine the specific meanings of the general words in their current context.

We understand language so well because we know so much. Therefore, we will not have a good account of how language works until we have a good account of what we know about the world and how we use that knowledge. In this article I have sketched a formalization of one very abstract way of conceptualizing the world, one that arises from an investigation of lexical semantics and is closely related to the lexical decompositions and image schemas that have been argued for by other lexical semanticists. It enables us to capture formally the core meanings of many of the most common words in English and other languages, and it links smoothly with more precise theories of specific domains.

I have profited from discussions of this work with Gully Burns, Peter Clark, Tim Clausner, Christiane Fellbaum, and Rutu Mulkar-Mehta. This work was performed in part under the IARPA (DTO) AQUAINT program,

contract N61339-06-C-0160.

5 References

Baker, Colin F., Charles J. Fillmore, & Beau Cronin 2003. The structure of the Framenet database. *International Journal of Lexicography* 16:281-296.

Bloomfield, Leonard 1933. *Language.* New York: Holt, Rinehart and Winston.

Brugmann, Claudia 1981. *The story of "over"*, MA thesis, University of California, Berkeley, California.

Cycorp 2008. http://www.cyc.com/.

Davis, Ernest 1990. *Representations of commonsense knowledge.* San Mateo: Morgan Kaufmann Publishers.

Fodor, Jerry A. 1980. Methodological solipsism considered as a research strategy in cognitive science. *Behavioral and Brain Sciences* 3:63-109.

Fodor, Jerry A. 1983. *The modularity of mind: An Essay on Faculty Psychology.* Cambridge, MA: The MIT Press.

Ginsberg, Matthew L., editor 1987. *Readings in nonmonotonic reasoning*, Los Altos: Morgan Kaufmann Publishers, Inc.

Gruber, Jeffrey C. 1965. *Studies in lexical relations.* Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hagoort, Peter, Lea Hald, Marcel Bastiaansen, & Karl Magnus Petersson 2004. Integration of word meaning and world knowledge in language comprehension. *Science* 304:438-441.

Hobbs, Jerry R. 1985. Ontological promiscuity." *Proceedings*, 23rd Annual Meeting of the Association for Computational Linguistics, pp. 61-69. Chicago, Illinois, July 1985.

Hobbs, Jerry R. 1992. Metaphor and abduction. In: A. Ortony, J. Slack, & O. Stock (eds.). *Communication from an Artificial Intelligence Perspective: Theoretical and Applied Issues*. Berlin: Springer-Verlag, 35-58.

Hobbs, Jerry R. 2004. Abduction in natural language understanding. In: L. Horn & G. Ward (eds.). *Handbook of Pragmatics*, 724-741, Malden, MA: Blackwell.

Hobbs, Jerry R. 2005. Toward a useful notion of causality for lexical semantics. *Journal of Semantics* 22:181-209.

Hobbs, Jerry R., William Croft, Todd Davies, Douglas Edwards, & Kenneth Laws 1987. Commonsense metaphysics and lexical semantics. *Computational Linguistics* 13:241-250.

Hobbs, Jerry R., Mark Stickel, Douglas Appelt, & Paul Martin 1993. Interpretation as abduction. *Artificial Intelligence* 63:69-142.

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: The MIT Press.

Johnson, Mark 1987. *The body in the mind: The bodily basis of meaning,* *imagination, and reason.* Chicago: University of Chicago Press.

Kilgariff, Adam 2001. Generative lexicon meets corpus data: The case of nonstandard word uses. In: P. Bouillion & F. Busa (eds.), *The Language of Word Meaning.* Cambridge, England: Cambridge University Press, 312-328.

Lakoff, George 1987. *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press,

Lenat, Douglas, & R. V. Guha 1990. *Building large knowledge-based systems: Representation and inference in the Cyc project.* Reading, MA: Addison-Wesley.

McCarthy, John 1980. Circumscription: A form of nonmonotonic reasoning. *Artificial Intelligence* 13:27-39.

Miller, George 1995. WordNet: a lexical database for English. *Communications of the ACM* 38:39-41.

Pustejovsky, James 1995. *The generative lexicon.* Cambridge, MA: The MIT Press, Cambridge, Massachusetts.

Rosch, Eleanor 1975. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* 104:192-233.

Shoham, Yoav 1987. Nonmonotonic logics: Meaning and utility. *Proceedings*, International Joint Conference on Artificial Intelligence, 388-393. Milano, Italy, August 1987.

Singh, Push 2002. The public acquisition of commonsense knowledge. *Pro-*

ceedings, AAAI Spring Symposium on Acquiring (and Using) Linguistic (and World) Knowledge for Information Access. Palo Alto, California, March 2002.

Tanenhaus Michael K., & Sarah Brown-Schmidt, 2008. Language processing in the natural world. In: B.C.M.Moore, L.K. Tyler, & W.D. Marslen-Wilson (eds.), *The Perception of Speech: From Sound to Meaning*, *Philosophical Transactions of the Royal Society B: Biological Sciences* 363:1105-1122.

Keywords: *lexical semantics*, *world knowledge*, *lexical decomposition*