Sketch of an Ontology Underlying the Way We Talk about the World

Jerry R. Hobbs
Artificial Intelligence Center
SRI International
Menlo Park, California

1 Introduction

If we are going to have programs that understand language, we will have to encode what words mean. Since words refer to the world, their definitions will have to be in terms of some underlying theory of the world. We will therefore have to construct that theory, and do so in a way that reflects the ontology that is implicit in natural language.

There are wrong ways to go about this enterprise. For example, we could take our underlying theory 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 prepostions in terms of that.

In this paper, I propose a general structure for a different underlying conceptualization of the world—one that should be particularly well suited to language. It consists of a set of *core theories* of a very abstract character. In this paper I discuss some of the most important of these, in particular, the core theories that explicate the concepts of systems and the figure-ground relation, scales, change, causality, and goal-directed behavior. These theories are too abstract to impose many constraints on the entities and situations they are applied to. In fact, the reader may complain that they apply to anything. But the main purpose of the core theories is to provide the basis for a rich vocabulary for talking about entities and situations. The fact that the core theories apply so widely means that they provide a great many domains of discourse with a rich vocabulary.

The enterprise is therefore to axiomatize these core theories in as clean a fashion as possible, and then to define, or at least characterize, various words in terms of predicates supplied by these core theories. For example, a core theory of scales will provide axioms involving predicates such as *scale*, $<$, *subscale*, *top*, *bottom*, and *at*. Then, at the "lexical periphery" we will be able to define the rather complex word "range" by an axiom such as the following:

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

$$\wedge (\forall\, u \in x)(\exists\, v \in s_1) at(u, v)$$

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

In this paper, I sketch the core theories and mention some of the words that would be in the lexical periphery of the core theories.

Two methodological principles should be mentioned first. Above, I said "define, or at least characterize, various words". 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 sort

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

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.

The second methodological point is that we need to be careful how we use an argument from the "naturalness" of an expression. Not all expressions that will be allowed by our core theories will sound natural. Our knowledge of language consists of thousands of very specific conventions, each of which has a rationale in terms of core theories. But not everything that has a rationale has been conventionalized. Conventional expressions sound natural. Other expressions with a rationale are interpretable, but may not sound natural. For example, it is conventional to say "at work" and "in progress", and recently in corporate America, the expression "on travel" has become conventional. There is no particular reason that these expressions are better than "on work", "on progress", and "at travel". It just happens that the latter did not become conventional. The account of lexical meaning given here is intended to provide a rationale for expressions, but not to explain why one version rather than another has been conventionalized.

2 Granularity

A road can be viewed as a line, a surface, or a volume. When we are planning a trip, we view it as a line. When we are driving on it, we have to worry about our placement on it to the right or left, so we think of it as a surface. When we hit a pothole, it becomes a volume to us.

This shifting of granularity is a general property of cognition. We are very good at adopting small, on-the-spot theories of situations that include just the aspects relevant to our current concerns. Notions of granularity will have to pervade the knowledge base we build. Many concepts are inherently granularity-dependent, and many other concepts provide us with means for imposing granularities on situations.

A granularity is defined by an indistinguishability relation $\sim$, or equivalently, a set covering. If the set covering is a partition, the indistinguishability relation is transitive. An example is when we are concerned only with the country a location is in and not any finer discrimination. Any two locations in, say, Italy would be indistinguishable under this relation. If the set covering is comprised of overlapping sets, the indistinguishability relation is not transitive. An example is when we do not distinguish any two points lying within 1 cm of each other. When we view a road as a line, we are not distinguishing between two points that are at the same place along its length, even though they are, for example, in different lanes.

3 Systems and the Figure-Ground Relation

A *system* is a set of entities, their properties, and the relations among them. The concept of system 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 system. For this reason, a vocabulary for talking about systems will be broadly applicable.

The elements of a system can themselves be viewed as systems, 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 in a system is its relations to the other elements of the system, 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 system 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, and so on. In spelling out the structure-function articulations, we might say something about how its content determines its place in the library.

A system can serve as the *ground* against which some external *figure* can be located or can move. A primitive predicate *at* expresses this relation. In

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

s is a system, y is an element in the system, and x is an entity not in the system. It says that the figure x is at a point y in the system 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

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

Such axioms constitute the source of spatial terminology and spatial metaphors. Some 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.

In computer science, a variable at a value in a range of values:
I goes from 1 to 100.

A person's salary at a particular point on the money scale:
John's salary reached $75,000 this year.

A particularly important example of an *at* relation is predication itself. We can view a set of predicates as constituting a system, where the relations among the elements are the implication and mutual exclusivity relations. Axioms of the form

$$(\forall\, p, x, s) p \in s \wedge p(x) \supset at(x, p, s)$$

say that for a predicate in a system of predicates to be true of an entity is for the entity to be at that predicate in the system. This makes the rich vocabulary of spatial relationships available for predication.

The following expressions, for example, tap into a system of predicates about human activities and states of consciousness:

at work, at play, on travel, asleep, awake, on drugs, ...

4 Scales

A very common and very useful kind of system is one in which the relations among the entities are an indistinguishability relation $\sim$ and a partial ordering $<$. We can call this a *scale*.

A core theory of scales will provide definitions for such concepts as a subscale, a total ordering, a scale being dense, the top and bottom of a scale, and the reverse of a scale. Allen's relations among time intervals (Allen and Kautz, 1985) are in fact relations among subscales and are straightforward to define. If we have a primitive notion of points on a scale being adjacent, we can define connectedness in terms of it. A scale is a system, so the Figure-Ground relation applies to it. We can talk about an external entity being *at* a point on a scale.

An obvious example of a scale is the Number scale. Integers can be defined in the standard way using the successor function, i.e., by counting. The ordering can be defined recursively in the standard way:

$$(\forall\, n)n < n + 1$$

$$(\forall\, n_1, n_2, n_3)n_1 < n_2 \wedge n_2 < n_3 \supset n_1 < n_3$$

The indistinguishability relation is just equality.

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 number scale:

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

This again 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.

Just as we can have systems of predicates, we can have scales of predicates. The simplest such scale, for any predicate p, is the scale whose elements are p and $\neg p$, whose indistinguishability relation is equality, and whose ordering specifies that $\neg p$ is less than p.

We can also have more complex scales of predicates, such as

$$cold < cool < warm < hot$$

$$none < few < many < all$$

It is very useful to be able to isolate the high and low regions of a scale. We will do this with operators we can call 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 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 fairly 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. The Hi and Lo regions must be related to common distributions of objects on the scale, so that if something is significantly above average for the relevant set, then it is in the Hi region. They 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.

It is useful to be able to state the relationship between a scale and the absolute form of an adjective directly. For this, we will use the predicate *scale-for*. We cannot define it precisely, but it has the following property:

$$(\forall\, s, p) \mathit{scale\text{-}for}(s, p) \supset (\forall\, x)[p(x) \equiv (\exists\, y) at(x, y, s) \wedge y \in Hi(s)]$$

That is, if s is the scale for the predicate p, then p is true of some entity x if and only if x is at some point y in the Hi region of the scale s. The Height Scale is the scale for the predicate *tall*.

The notion of "scale for" is used in defining the comparative and superlative forms of adjectives. For x to be more p than y is for x's location on the scale for p to be greater than y's location.

It is possible to define composite scales. If a scale s is a composite of scales s_1 and s_2, then its elements are the ordered pairs $< x, y >$ where x is in s_1 and y is in s_2. An external entity is *at* a point $< x, y >$ in the composite scale s if and only if it is at x in component scale s_1 and at y in component scale 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.

Composite scales have two or more dimensions. In general, we could create two-dimensional structures in two different manners. The first is to follow the lead of Cartesian space and take two-dimensional space to be simply a composite scale, that is, a set of ordered pairs. Graphs, representing functions from one scale to another scale, are sets of points in a space defined in this way. Bar graphs are possible when one of the scales has a finite number of elements. A limiting case is when one of the scales has an empty ordering, thus reducing to a mere system. For example, when we are graphing people's incomes, the component "scales" are the set of people and the Money Scale. There is an natural ordering on the Money Scale, but no obvious, natural ordering on the set of people.

The second manner is to take two-dimensional space to be a set of elements with two independent ordering relations. The minimal two-dimensional space under this definition consists of three points A, B, and C, such that

$$A \leq_1 B,\ B <_1 C,$$
$$A \leq_2 C,\ C <_2 B$$

These two definitions are not equivalent. Under the first definition, it is not possible to have a three-point space. Under the second definition, neither of the two orderings can be empty. In our previous work, we have taken the second definition to be the basis for our axiomatization of space (Hobbs et al., 1987).

If we have a notion of adjacency in the two component orderings, it is straightforward to define the notion of adjacency for the two-dimensional space, or more generally, we can take the latter to be primitive. Given that, we can define the notions of connectedness, density, region, boundary, contact, and so on, in a straightforward manner.

Three-dimensional space can be defined in an analogous fashion.

In order to model a notion of orientation, and consequently in order to model a notion of shape, we require more structure in two- or three-dimensional space.

Material can be characterized in terms of extension and cohesion. Extension can be axiomatized in a straightforward manner by associating bits of material with the regions they occupy at a given time. Cohesion can be axiomatized in a way that parallels the axiomatization of connectedness in space.

5 Change

A primitive predicate of central importance is the predicate *change*. This is a relation between situations, or conditions, or predications, and indicates a change of state. In this paper, to avoid an overgrowth of notation, I will write

$$change(p(x), q(x))$$

where, strictly speaking, I should, in the ontologically promiscuous notation of Hobbs (1985), write

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

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

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

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

The predicate *change* possesses a limited transitivity. There was a change from Reagan being an actor to Reagan being President, because they are two parts of the same ongoing process, even though he was governor in between. But we probably do not want to say there was a change from Reagan being an actor to Margaret Thatcher being Prime Minister. They are not part of the same process.

Any given process, that is, any sequence of events linked by *change* relations, is a scale whose partial ordering is induced by the predicate *change*.

The Time Line could be taken as primitive, with the *before* relation as its ordering and an *at-time* relation relating states and events to points and intervals on the Time Line. The *at-time* relation would be an *at* relation, giving us the common spatial metaphors for time. Such an ontology seems to be justified by the clock and calendar terms in modern languages. In this ontology, we could define the predicate *change* to be true when different properties are true of an entity at different times.

It seems to me, however, that the notion of *change* is more basic. It is built into the more "primitive" parts of language, such as the event verbs. Even the words "before" and "after", which might seem to relate directly to the Time Line, carry a whiff of causality. The sentence,

> The French Revolution broke out after George Washington was elected president.

seems to convey some causality or suggest that somehow the two events are part of the same process.

If we take *change* to be the basic notion, 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 Bureau of Standards—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 calling TIME.

In any case, there is no need to decide between these two ontologies, since they are inter-definable in a straightforward fashion (Hobbs et al., 1987).

6 Causality

The next primitive predicate of central importance is *cause*. As with *at* and *change*, it has no definition. There is no axiom of the form

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

but the knowledge base is rife with axioms of the form

$$cause(p(x), q(x))$$ [1]

[1] This should also be seen as an abbreviation for an ontologically promiscuous, first-order representation.

expressing causal connections among states and events. We don't know precisely what causality is, but we know lots and lots of examples of things that cause other things.

There is a question as to what the arguments of *cause* can be. Some would urge that they 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 (the first argument) 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.

I have spoken loosely of states and events. We are now in a position to characterize more precisely the intuitive notions of state, event, action, and process. A state is an *at* relationship, $at(x, y, s)$, or more generally, a predication. 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:

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

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:

$$cause(a, change(at(x, y, s), at(x, z, s)))$$

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 tense in English. Thus, fluctuating can be viewed as a state.

The world is laced with threads of causal connection, and therefore our knowledge base must be rife with axioms encoding causal connections. In general, if two entities x and y are causally connected with respect to some behavior p of x, then 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.

$$cause(p(x), move(z, x, y)) \wedge cause(move(z, x, y), q(y)) \supset cause(p(x), q(y))$$

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.

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, the nail, which causes or should cause a change in the mental or emotional state of the Beneficiary, Bill.

$$\begin{aligned} Agent \text{ --}cause\text{-->} \;\; & change(at(Instrument, x), at(Instrument, Object)) \\ \text{--}cause\text{-->} \;\; & change(at(Object, y_1), at(Object, y_2)) \\ \text{--}cause\text{-->} \;\; & change(p_1(Beneficiary), p_2(Beneficiary)) \end{aligned}$$

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.

Croft (1991) has pointed out that the preposition used to label an argument of a verb is determined by the argument's place in this causal chain. Verbs pick out a particular entity as the Object. Then arguments that are upstream from the Object in the causal chain are signalled with "by" or "with", including the Agent, Instrument, and Comitative roles. Arguments that are downstream from the Object are signalled with "to" or "for", including the Goal and Beneficiary roles. The preposition use in the following two sentences results from the fact that causality runs from John to the hay to the wagon.

John loaded the wagon with hay.
John loaded hay onto the wagon.

Another important role for causality is in linking two scales. It often effects a monotonic, scale-to-scale function. The general pattern is this:

$$cause(change(at(x, y, s_1), at(x, z, s_1)), change(at(w, u, s_2), at(w, v, s_2)))$$

where if $y < z$ on s_1, then $u < v$ on s_2. That is, if there is a change from x being at y on s_1 to x being at a higher point z on s_1, then this causes there to be a change from w being at u on s_2 to w being at a higher point v on s_2. This is the basis of our many "The more ..., the more ..." rules, such as

The more you press on the accellerator, the faster you go.

A concept closely related to causality is enablement. It can be defined as follows:

$$(\forall e_1, e_2) enable(e_1, e_2) \equiv cause(not(e_1), not(e_2))$$

That is, e_1 enables e_2 if e_1 not happening will cause e_2 not to happen. Enablement is crucial in the core theory of goal-directed systems.

7 Goals and Plans

The final primitive concept of central importance that I will discuss here is the concept of a *goal*. This again will not be defined, but axioms will link it in the right way with axiomatizations of belief and action, to make available intentional interpretations of human and other behavior. In particular, these core theories should insure that people's actions can be seen as attempts to achieve their goals, given their beliefs.

Among the most important facts about goals are those linking them with causality and enablement, for it is by manipulating the causal structure of the world that agents achieve their goals. The two primary axioms are as follows:

$$(\forall a, q, r) goal(a, q) \wedge enable(r, q) \supset goal(a, r)$$

$$(\forall a, p, q) goal(a, q) \wedge cause(p, q) \wedge choose(a, p, q) \supset goal(a, p)$$

The first axiom says that if an agent a has a goal q and r enables q, then a will have the goal r. This captures the prerequisites of the STRIPS operators of Fikes and Nilsson (1971).

The second axioms says that if an agent a has a goal q, where p causes q, and a chooses p as a way of achieving q, then a will have the goal p. I will not attempt to explicate *choose* here, but something like this is necessary to accomodate nondeterminism. There may be many things that will cause q, and the agent need pick only one of them. This axiom encodes the "body" of the STRIPS operators of Fikes and Nilsson (1971).

Thus, to achieve a goal, an agent must satisfy all the prerequisites, removing all the barriers to the goal, and then choose something that will cause the goal to come about. These two axioms allow us to construct hierarchical plans, decomposing goals into their subgoals. In the above axioms, q is the goal, r and p the subgoals. These subgoals can in turn be decomposed into further subgoals

The depth of decomposition in these plans is one of the prinicpal ways we impose a granularity on our view of behavior. It may be sufficient for our purposes to know that John drove his car to the airport, or it may be necessary to view it under a finer granularity that makes visible his actions of shifting the gears and turning the steering wheel.

A plan is essentially a representation of causal structure. It is therefore useful for explaining not just human behavior, but other phenomena as well. Artifacts and organizations can, for example, be viewed as plans made concrete.

Much of the knowledge we have about artifacts is best represented by the plan that it implements. Consider a very simple example. The function of coffe cup is to move coffee. We decompose this goal into two subgoals—containing the coffee in the cup and moving the cup. The subgoal of moving the cup is further decomposed into the subgoals of attaching the cup to the handle and moving the handle. It is a very common schema for artifacts that in order to do something to an object, we set up a causal connection, such as containment or attachment, to another object, and do something to that other object. We can continue to decompose in this fashion until we have specified the role or function of all the components of the artifact.

Similarly, organizations can be seen as having a goal and implementing a plan to achieve that goal, where the structure of the organization reflects the structure of the plan. Thus, the goal of an organization might be to provide people with cars. This decomposes into the subgoals of having one division of the organization manufacture cars and another division sell them to people. Each of these would decompose further. Eventually the plan would bottom out in sets of actions by single individuals. These sets of actions constitute the members' *roles* in the organization.

Any system that can be viewed as exhibiting functionality can be represented in terms of a plan that expresses the system's underlying causal structure. A tree, for example, can be viewed as a goal-directed system whose goal is to grow and reproduce.

8 Summary

A common way to encode knowledge for natural language and other AI programs is to proceed domain by domain. Many of the most common words is language, however, apply across many domains. What I have tried to do in this paper is to suggest some very abstract domains—systems, scales, change, causality, and goal-directed systems—that seem to underlie more specific domains. The more specific domains can be seen as instantiations of the abstract ones. Language provides us with a rich vocabulary for talking about the abstract domains. When we construct core theories of these domains, then we have a hope of being able to define, or at least characterize, the words in this vocabulary in terms provided by the core theories. When the core theory of an abstract domain is instantiated as a specific domain, then the vocabulary associated with the abstract domain is also instantiated, giving us a rich vocabulary for talking 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.

Acknowledgments

This paper is based in part on work that was done with William Croft, Todd Davies, Douglas Edwards, and Kenneth Laws. I have also profited from discussions with Annelise Bech, Bente Maegaard, and Costanza Navarretta.

References

[1] Allen, James F. and Henry A. Kautz. 1985. "A Model of Naive Temporal Reasoning." *Formal Theories of the Commonsense World*, ed. by Jerry R. Hobbs and Robert C. Moore, Ablex Publishing Corp., pp. 251-268.

[2] Croft, William, 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*, University of Chicago Press, Chicago, Illinois.

[3] Fikes, Richard, and Nils J. Nilsson, 1971. "STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving", *Artificial Intelligence*, Vol. 2, pp. 189-208.

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

[5] Hobbs, Jerry R., William Croft, Todd Davies, Douglas Edwards, and Kenneth Laws, 1987. "Commonsense Metaphysics and Lexical Semantics", *Computational Linguistics*, vol. 13, nos. 3-4, July-December 1987, pp. 241-250.