

6. Abstraction and Analogy

6.1 Introduction

Abstraction and analogy are two methods of understanding something new in terms of previous experience. Figure 26 shows a basic relationship between abstraction and analogy. Whenever A is analogous to B, one can form a description C which characterizes them both. For example, the oscillation of a weight connected to a spring is analogous to the oscillation of current in an electrical circuit composed of a capacitor, inductor, and resistor in series. Both can be described by the same first order differential equation.

This relationship between analogy and abstraction makes it more difficult to know how to deal with many expressions. For example, should we say that a *dog house* is analogous to a *house* (for people) or should we say there is one abstract sense of house which covers both of these. It is possible that people have all the information: concepts of house, dog house, an abstract notion of house, and an awareness of and ability to play on the analogy between a house and a dog house. (I remember when I first took my son to the zoo after he started to talk. We were watching the camels when the food truck arrived. I said *Look, the camels are going into their house for supper*. His response, *Do they eat supper in the kitchen or the dining room?*)

In linguistics it is sometimes tacitly assumed that if people have a general rule, they will not also have all the specific cases. But consider, for example, *crackage*. People know what this means, and they also know that it is not a word. They know it fits with *breakage* and *damage*, but that these are used and *crackage* isn't. Thus, at least in some cases people know both the general rule and all of its normally occurring cases.

Why have both the rule and the examples? The rule facilitates understanding new cases. Consider *transmote*, most people don't feel they know what this means until they learn that President Johnson wanted to get rid of some government officials which he couldn't promote or demote, so he decided to *transmote* them. People have the examples *promote* and *demote* needed to understand *transmote*, but they don't already have a rule which these and *transmote* will fit.

Fig. 26. Abstraction and Analogy

< = <fig5-1.press<

It could be that individual cases are remembered for efficiency reasons, and it could be because we usually know some idiosyncratic information about each word sense. At any rate, it would be false to argue that abstractions don't exist because the instances of them are idiosyncratic in some way.

In summary, it seems likely that a computer system should have rote memory, analogy, and abstraction.

6.2 Three Criteria for Filling a Slot

An important use of abstraction is in determining what can fill a given role in a description (slot in a frame). It is possible to distinguish three levels of acceptability.

- (74a) Conceivable e.g., *poison is edible*.
- (74b) True of the world e.g., *grass is edible*.
- (74c) Typical of the world e.g., *lettuce is edible*.

We see the progression in

- (75a) John ate the movie. *inconceivable*
- (75b) John ate the rubber ball.
- (75c) John ate the leaves.
- (75d) John ate the peas. *typical*

The notion of conceivability has been developed by Keil [45], who in turn drew his inspiration from papers by Sommers [98]. Keil suggests that if something is inconceivable, then its negation is also inconceivable, while if something is merely false its negation is perfectly acceptable. For example, *that bachelor is married* is conceivable but false. Instead, *that bachelor is not married* sounds fine. By contrast, *the idea is green* is inconceivable, and *the idea is not green* is no better, just as inconceivable.

The split between conceivability, truth, and typicality could be modeled in a computer program. For each frame slot, we would have a structural description of what could fill that slot. Whenever the slot was filled, the filler would have to match that description. For example, the sentences

- (76a) I sliced the ice cream.
- (76b) I mopped up the ice cream.

show how since sliced takes a solid and mopped up a liquid, one must pick the appropriate phase of ice cream to fill the slot. As a second example, note that *on* takes a surface and *in* an enclosure, so that

- (77a) I looked out the window and saw it on the grass. 2-D

evokes a two dimensional grass image while

- (77b) He got down on his knees and hunted for it in the grass. 3-D

evokes a three dimensional grasp image. The structural description of what can fill a slot would presumably be quite abstract, allowing many things to be viewed in such a way that they could fill the slot. If there was no way of viewing a particular thing, x , as fitting the slot's description, then x filling the slot would be inconceivable.

Whether something is true in the world could be tested by comparison with known examples. Knowledge of how the world is would be much more articulated, much less abstract, than knowledge about what can fill a slot.

Finally, the question of being typical of the world would be addressed as suggested by Winograd [109], by prototypical descriptions of what can fill a slot.

In summary, knowledge of what can fill a slot may involve abstractions both on what is conceivable, and on what is typical. It may also involve a wide knowledge of what actually does fill slots.

6.3 Predicability

Working in the tradition of logic, Keil [45] saw the question of what could fill a slot as the question of what descriptions would accept a given predicate. He worked mainly with one place predicates. Figure 27 shows what happened when a group of college students were given a list of objects and a list of predicates and where asked to indicate which predicates could be applied to which objects. Keil found that some predicates, like IS INTERESTING, could be applied to all of the objects. Other predicates like IS HONEST, were said to apply only to a few objects, here men and girl. The surprising result Keil reports is that if a predicate P_1 applies to an object applied to by a second predicate P_2 , then P_1 applies to every object applied to by P_2 . This makes it possible to organize the predicates into a tree as shown in Figure 27.

Each node in the tree in Figure 27 defines an abstract class of objects which are accepted by all the predicates at or above that node. In Figure 5.3, Keil has attempted to give names to these classes.

That these classes would form a tree is an unexpected result and certainly deserves closer examination. First note, that if we were describing truth instead of predicability, there would be no chance of getting a tree. For example, Figure 29 shows that taxonomies organized by truth-values are not Hierarchical. For example, while both EXPENSIVE and EDIBLE are true of caviar, neither dominates the other as each is true of something the other isn't. Thus, it is only because we have relaxed the test from truth to predicability that the possibility of a tree exists.

Some apparent counterexamples to a tree can be explained as a confusion of word senses. For example, Figure 30 shows that IS RATIONAL applies to both man and 6, but IS HONEST applies to man and not 6 while IS DIVISIBLE BY 5 applies to 6 and not man. This configuration defeats formation of a tree, but it is eliminated by realizing that two different senses of IS RATIONAL are in question.

Not only is it necessary to recognize different senses of a predicate in order to get a tree, but it is also necessary to assume that some predicates apply only to certain descriptions of an entity. Figure 31 illustrates this. Flag poles and people, but not companies, can be tall, while people and companies, but not flag poles can owe money. A solution here is to distinguish between a person's physical description, which can be tall

Fig. 27. Keil's Taxonomy
< == <fig5-2.press<

Fig. 28. Leif's Taxonomy with Names
< = <fig5-3.press<

Fig. 29. Truth-Value Taxonomies are Non-Hierarchical

<= <fig5-4.press<

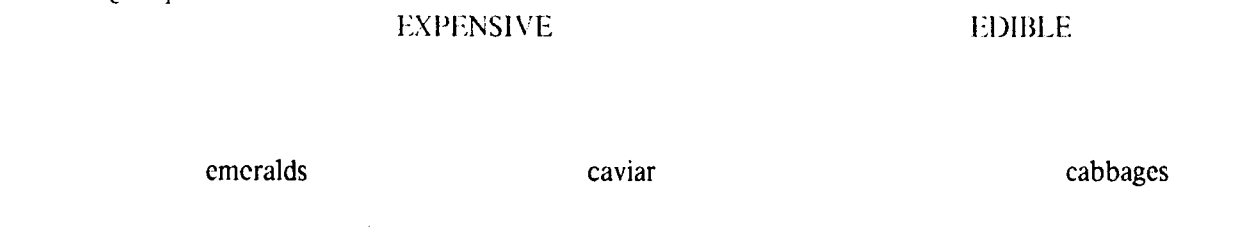


Fig. 30. A Confusion of Word Senses

<= <fig5-5.press<

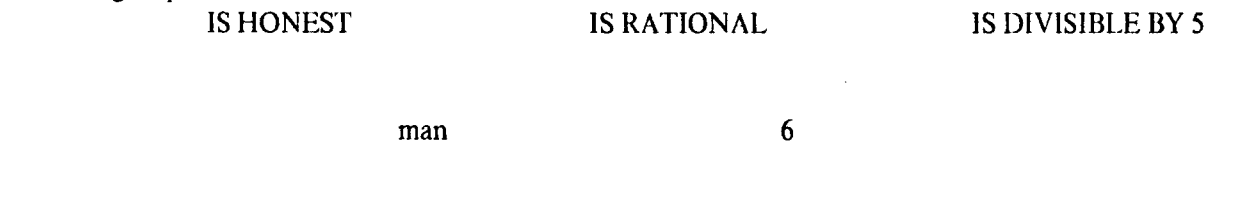
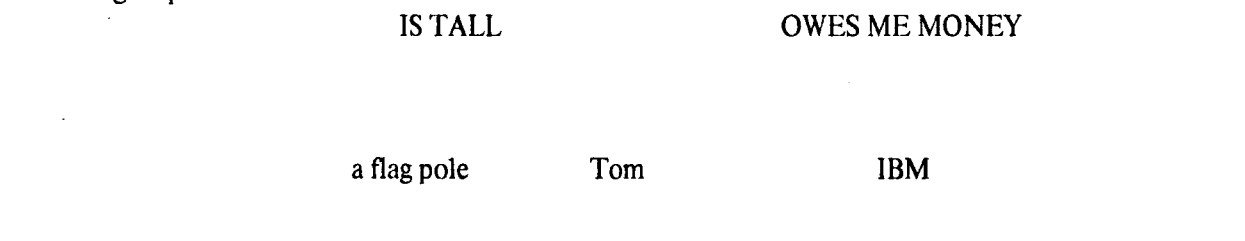


Fig. 31. A Confusion Between Descriptions

<= <fig5-6.press<



and his description as a social entity which can owe money. Support for this approach will come from the discussion near the end of this chapter. Other examples requiring this trick are

- (78a) ?Italy is sunny and democratic.
- (78b) ?The fire was red and an hour long.
- (78c) ?The song was about birds and an hour long.
- (78d) ?The book was about birds and weighed two pounds.

which sound odd because they conjoint two incompatible descriptions. Keil suggests that when one is dealing with multiple descriptions of the same entity it is usually possible to find paraphrases which make these descriptions explicit. For example, the climate in Italy is sunny and its society is democratic.

Unfortunately, not all problems in the construction of a tree are naturally resolved by an appeal to multiple predicate senses or multiple descriptions. Figure 32 shows two such examples. It seems less reasonable to split the description of a tree up into a part accepting dimension and a part accepting weight than to split a person into social and physical characterizations. However, this example suggests a way to save the tree scheme. Predicates like tall and heavy apply to anything which has a single attribute, here height, or weight. It may be possible to either place a predicate in the tree or have it test for one or more particular attributes. This should be explored.

Fig. 32. Incomparable Attributes

< = <fig5-7.press<

TALL

HEAVY

shadow

tree

milk

The importance of Keil's observations to computational linguistics is that it opens up the possibility of making a test for what can fill a slot based either on a) being a member of an abstract class, b) possessing a particular feature. Both of these tests can be efficiently implemented as long as the number of classes and features is not too large. One can see from the abstract nature of Keil's classes that there will not be a very large number of classes. There may be more features, and some sort of inheritance might be needed to efficiently attribute all the needed features to each description.

6.4 Prototypes

Predicability constraints can serve as an initial filter on what can fill a slot. At the opposite extreme comes detailed knowledge of the world. A proposed slot filler can be checked with actual experience.

Another use of past experience is in filling slots with default values. It has been suggested that this can be facilitated by the construction of prototypical individuals. For example, imagine a ball, when asked to do this most people will imagine one of their own personal set of prototypical balls. Sometimes its an actual ball they once had and sometimes it isn't. Usually, though, it is representative of balls, it isn't made of ice cream, isn't the world's largest ball etc. Rosch [87] has done experiments to determine what features a prototype of a category will have. The results are as one would expect. A prototype

- (79a) Should have features in common with many members of the category,
- (79b) should not have features which other members of the category don't have,
- (79c) Shouldn't be a good prototype of some other category,
- (79d) Should be commonly occurring.

The use of prototypes is not very well understood. It isn't known how abstract they should be, for example. Suppose someone hears the sentence

- (80) Harry walked over to the bench and picked up one of the new tools.

How much detail should be filled in on that tool. At its most abstract, the physical description of a prototypical tool might consist only of a handle and a business-end. If more detail is needed, a switch is made to some prototypical type of tool which fits the situation, say a screwdriver. This has a handle and a shaft, but perhaps we don't know if it is a Phillips head screwdriver or not. It seems possible that at each level of abstraction, e.g. tool, we have a list of typical handles, typical blades, etc. and for each of these we also have

classes of less abstract tools, e.g. screwdriver which would have a handle of that sort. This allows us to move down in abstraction to whatever level is necessary to get understanding, hypothesizing part from whole and whole from part.

6.5 Is There a Basic Level of Abstraction

Objects in very abstract categories, e.g. animal, have few properties in common. Consequently, to say that something is an animal is not to say very much about it. At the other extreme, objects in very particular categories, e.g. orange tabby Persian, have much in common with other categories. That level of differentiation is rarely called for. Roesch [88] argues that there must be some intermediate level, she calls it the base level, which affords the proper degree of discrimination.

Objects in very abstract categories differ so much there isn't much one can learn which applies to them all. Consequently, one can't tie much of his knowledge to a level like animal. At the same time it is redundant to store all of one's knowledge at a level like orange tabby Persian because most of this could be moved up to Persian, even to cat.

In considering this suggestion the reader might find it helpful to look at Figure 33 and mark what he would consider to be the basic categories. As yet, no one has made effective use of basic categories in computational linguistics. It has been argued that they could be used for matching by setting a common level at which matching could be done.

6.6 Types of Abstractions

One frequently possessed both structural and functional knowledge about a concept. For example both structural and functional knowledge are used in accessing category membership. Subjects assessing whether each of the objects in Figure 34 is a cup or a bowl are influenced in their judgment by whether it contains coffee or mashed potatoes.

Even spatial concepts may have a functional interpretation. To explore this, consider the assessment of when a situation may be described as *x is near y*. Three criteria are proposed:

- i) *x* and *y* are separated by less than 1/8 (say) the maximum distance evoked by the context, e.g. in a 16 x 16 room a desk 2 ft. from the window is near the window.
- ii) The distance between them is small compared with the size of the objects — measured in any way you feel appropriate.
- iii) The situation is recognized as fitting a frame in which there is a threshold distance required for some process — near is under this threshold.

The reader might want to test these in the following sentences.

Fig. 33. Taxonomy of Basic Categories

PHYSICAL-OBJECT
TOOL
 HAMMER
 BALL-PEN-HAMMER
 CLAW-HAMMER
 SAW
 HACK-SAW
 CROSS-CUT-SAW
TREE
 MAPLE
 SILVER-MAPLE
 SUGAR-MAPLE
 OAK
 WHITE-OAK
 RED-OAK
CLOTHING
 PANTS
 LEVIS
 DOUBLE-KNIT-PANTS
 SHIRT
 DRESS-SHIRT
 T-SHIRT
FISH
 BASS
 SEA-BASS
 STRIPED-BASS
 TROUT
 RAINBOW-TROUT
 STEELHEAD-TROUT

Fig. 34. The Cup/Bowl Continuum

< = <fig5-9.press<

-
- (81a) His office is on the 8th floor, near Minsky's.
 - (81b) Don't stand near the tracks.
 - (81c) The dog was near the telephone pole.
 - (81d) The bug was near the telephone pole.
 - (81e) The dog was near the fire hydrant.
 - (81f) The two pages were near each other in that 400 page book.
 - (81g) The mark was near the end of the yardstick.
 - (81h) Apartment for rent, near MBTA.

Some concepts seem to be abstractions primarily along structural lines, e.g., square, some along functional lines, e.g. container, and some along both, e.g., mammal.

6.7 Analogy and Transfer frames

Winston [110] suggests that understanding an expression like *Robbie is like a fox* involves a transfer of structure from fox to Robbie as illustrated in Figure 35. Close relatives of the source and destination frames are used to help guess the intended analogy. Not only are properties transferred, but a transfer frame is constructed. This can be used for similar analogies in the future. Winston's algorithm is shown in Figure 36.

The transfer frame is a key idea, because it allows us to record, as data, how one frame is to be related to another.

6.8 Word Sense Extension

When something new comes along which doesn't match any of our concepts too well, we may extend or specialize the sense of a given concept. Suppose that the prototypical HOUSE is one for humans. It is misleading to refer to a house for dogs as a HOUSE, because it differs from the HOUSE prototypes in too many ways. What English lets us do is to indicate by DOG HOUSE that the HOUSE concept is to be modified in some way so as to make it appropriate to DOG. As Marchand [59] points out, this process of word formation is not limited to nouns, e.g., *color blind*, and *hog tie*. We postulate that when someone hears a term like *dog house*, he can use some mix of the following three strategies in understanding it.

Fig. 35. Winston's Transfer Frames

<= = <fig5-10.press<

The basic idea behind the theory of learning presented in this paper. The teacher specifies a source and a destination and possibly the slots that are relevant. The student analyzes the source, the destination, and other aspects of the situation to discover and use a transfer frame.

Fig. 36. Winston's Algorithm

< = <fig5-11.press<

Overall organization of the hypothesizing and filtering methods. Hypothesizing methods are tried until one produces one or more slots that are not filled in the destination. After grouping into transfer frames, all filtering methods are used in an effort to reduce the number of surviving transfer frames. Filters have an effect only if they recommend dropping some, but not all of the transfer frames that they see.

- a) He may have a concept for which he knows this to be an unanalyzed name, e.g., *bull dog* or *skid row*. How many people know that a bull dog is a dog for fighting bulls or that skid row was a row of shanties along a log skid in early day Seattle, Washington. We know what horse shoes and alligator shoes are.
- b) He may have a concept for which this would be an appropriate name. For example, although people don't have a name for the little plastic cylinder on the end of a shoestring [116] they know of it and they can understand the sentence *my shoestring end came off*. This presumably because END could be extended in the context of SHOESTRING to mean this cylinder, and this can come off.
- c) He may create a concept by principled analogy. For example, by analogy with *dog house* and *bird house* one could understand *grasshopper house*. We contend that this process of analogy is always at work. We understand *crackage* by analogy with *breakage*, *pilferage*, *damage*. We understand *inchage* by analogy with *mileage* and *yardage*.

6.9 Naming Concepts

It is interesting to ask whether this process of word formation might be formalized as a way of getting a systematic method of naming concepts.

McDermott [67] has suggested that concepts should be given names in some invented language, lest we confuse our understanding of the name with an understanding of the concept. It seems unrealistic, however, to ask someone to learn several thousand names in order to use or evaluate a system. McDermott's suggestion seems practically limited to small or one person systems. For others, it will be necessary to name concepts using terms the user is familiar with, either directly or by supplying a translation. Since a natural language, such as English, is the only widely known form of expression for non-mathematical concepts, it is the only real choice for concept names. Most all semantic networks do, in fact, use English in creating their node and link names. The issue is to what extent this can be done in a principled way.

In a semantic network concepts can be named and they can also be individuated (or defined) by their relationships to other concepts in the network. When a person is studying one concept in the network it is extremely helpful if he can get at least a general idea of what the concepts are that it is related to in the network just from their names. Otherwise mapping the semantic networks onto his knowledge of the world becomes a combinatorial puzzle. Fortunately, this can be made possible by capitalizing on a natural skill of every English speaker, the ability to extend a sense of a word for use in a particular context.

The problem with using expressions like *dog house* as the names of concepts is that they can be ambiguous. For example, *snake poison* can mean either poison from snakes or poison for snakes. A *steel drill* can be either made from steel or for drilling steel. A *river bank* could be a financial institution.

The ambiguity is bad for two reasons. First the user does not know which concept is meant. While this is bad, the user can always study the semantic network in order to see how the concept is used and thus what is meant. A worse problem is that the naming scheme does not generate enough distinct names; it proposes the same name for two different concepts. We wish a scheme which will allow us to generate a distinct name for each concept.

The ambiguity in *river bank* arises from the ambiguity in *bank*. It seems reasonable to resolve this by saying that names will be constructed from senses of words instead of words. Thus, RIVER BANK and RIVER FINANCIAL-BANK. The ambiguity in *snake poison*, however, does not seem to arise from any ambiguity in *snake* or *poison*. Instead it arises from two different relationships between these terms. We have already suggested in Section 6.8 that a person uses his knowledge of the world in order to determine what relationship is meant. It has long been recognized that there are certain regularities to this process. For example, corresponding to any substance such as steel, a *steel x* can mean an x composed of steel.

Rhyme [90] discusses attempts by various authors to capture these regularities in the case of compound nouns. These may be divided into attempts to classify the semantic relationship between the two constituents of the compound and methods for deriving the compound from a relative clause or other expression.

The difficulty with expressing the relation between the elements of a compound by a linguistic expression may be seen by considering *widget store*. Everyone knows what stores are for and so one meaning of *widget store* is clear. But how should this be expressed by a relative clause. Is it a store where widgets are sold, a store which carries widgets, a place to go to buy widgets, etc.

We postulate that somewhere in the semantic network there is a concept which corresponds to the store's wares. A *widget store* is a store where this wares concept may be characterized as widgets. There may be many different relative clauses which can express the relationship between the store concept and its wares concept, since in general they will be linked by a network of different relationships.

Loosely speaking, our solution to this problem will be to take names as triples. The second element of a triple will indicate how the first and third are related. In an example like *widget store* the second element will indicate how WIDGET and STORE are related by providing access to the store's WARES concept mentioned above.

6.10 A Model of the World

Whorf [103] has written a number of essays supporting the view that the language of a culture determines (or demonstrates) their model of the world. He believes that different languages show that quite a wide variety of models are in fact workable for common sense reasoning. Describing the Hopi language, he says:

After long and careful study and analysis, the Hopi language is seen to contain no words, grammatical forms, constructions or expressions that refer directly to what we call 'time,' or to past, present, or future, or to enduring or lasting, or to motion as kinematic rather than dynamic (i.e. as a continuous translation in space and time rather than as an exhibition of dynamic effort in a certain process), or that even refer to space in such a way as to exclude that element of extension or existence that we call 'time' and so by implication leave a residue that could be referred to as 'time.' Hence, the Hopi's language contains no reference to 'time' either explicit or implicit.

The metaphysics underlying our own language, thinking, and modern culture (I speak not of the recent and quite different relativistic metaphysics of modern science) imposes upon the universe two grand COSMIC FORMS, static three-dimensional infinite space, and kinetic one-dimensional uniformly and perpetually flowing time — two utterly separate and unconnected aspects of reality (according to this familiar way of thinking). The flowing realm of time is, in turn, the subject of a threefold division: past, present, and future.

The Hopi metaphysics also has its cosmic forms comparable to these in scale and scope. What are they? It imposes upon the universe two grand cosmic forms, which as a first approximation in terminology we may call MANIFESTED and MANIFESTING (or UNMANIFEST) or, comprises all that is or again OBJECTIVE and SUBJECTIVE. The objective or manifested comprises all that is or has been accessible to the senses, the historical physical universe, in fact, with no attempt to distinguish between present and past, but excluding all that we call future. The subjective or manifesting comprises all that we call future, BUT NOT MERELY THIS; it includes equally and indistinguishably all that we call mental — everything that appears or exists in the mind, or, as the Hopi would prefer to say, in the HEART, not only the heart of man, but the heart of animals, plants, and things, and behind and within all the forms and appearances of nature in the heart of nature, and by implication and extension which has been felt by more than one anthropologist, yet would hardly ever be spoken of by a Hopi himself, so charged is the idea with religious and magical awesomeness, in the very heart of the Cosmos, itself. The subjective realm (subjective from our viewpoint, but intensely real and quivering with life, power, and potency to the Hopi) embraces not only our FUTURE, much of which the Hopi regards as more or less predestined in essence if not in exact form, but also all mentality, intellection, and emotion, the essence and typical form of which is the striving of purposeful desire, intelligent in character, toward manifestation — a manifestation which is much resisted and delayed, but in some form or other is inevitable. It is the realm of expectancy, of desire and purpose, of vitalizing life, of efficient causes, of thought thinking itself out from an inner realm (the Hopian HEART) into manifestation. It is in a dynamic state, yet not a state of motion — it is not advancing toward us out of a future, but ALREADY WITH US in vital and mental forms, and its dynamism is at work in the field of eventuating or manifesting, i.e. evolving without motion from the subjective by degrees to a result which is the objective. In translating into English, the Hopi will say that these entities in process of causation 'will come' or that they — the Hopi — 'will come to' them, but, in their own language, there are no words corresponding to our 'come' and 'go' that mean simple abstract motion, our purely kinematic concept. The words in this case translated 'come' refer to the process of eventuating without calling it motion — they are 'eventuates to here' (pew'i) or 'eventuates from it' (angqo) or 'arrived' (pitu, pl. oki) which refers only to the terminal manifestation, the actual arrival at a given point, not to any motion preceding it.

If Whorf's analysis is correct, it shows that there is more than one workable world model. Hence, we may regard the world model that has evolved for use in English as somewhat arbitrary. Suppose we construct a formal interpretation of the world model of English speakers. This will be somewhat arbitrary, though we might hope that that arbitrariness is comparable to the arbitrariness of interpretation of a typical competent and intelligent English speaker. Such a world model may, nevertheless, be quite useful, particularly in a heuristic rather than a scientific sense. As with Hopi, we may find it alien to our thinking rather than ineffective for the intended use. This is because we are not used to a computer using animistic and metaphoric rather than scientific concepts. On the other hand we will argue that the speaker of English lives in a metaphorically rich environment. It is only because he is trained to regard his models as "idioms" or non-scientific ways of thinking that the Hopi model seems so different.

We feel that English resembles Hopi in that it too uses a change model as one of its most basic notions. The form of the change model is very simple and has been noticed by Schank [95], among others. Something goes from an initial state along a trajectory to a final state. In its most general form, this *going* is not in physical space, but in an abstract state space, that is, the object going along the trajectory is undergoing a change of state, represented by a change in one or more of its properties. Only if its LOCATION is changing is an object *going* in physical space.

Looking in the dictionary at the definition of go we find such sentences as

(82a) The car is going to my house.

(82b) The car is going to pieces.

(82c) The car is going to run.

(82d) The firm is going to the dogs.

(82c) The motor is going.

(82f) Tom and Sue are going steady.

The preposition to typically signals the destination. In the first four sentences the destination shows us what type of state change is involved. As shown the change could be a location, a fragmentation, a change in activity, or in a more abstract state. The notion *go* is that of a predictable directed change. A motor that is *going* is not just moving, but is moving in a systematic predictable way. If a couple is going steady, they are acting in a systematic predictable way with respect to one another or combining to form a functional unit.

A hat and dress can *go together*. This meaning of *go*, which takes a plural object, seems to mean combine to form a functional unit. Here, the final state appears to be the combination of the two objects. This meaning is, however, derivative of the one above. In fact, *go* is not suitable as the word for our notion of change. For example, it contains a default notion of *away* (as opposed to *come*). We will introduce the concept TRANS of which GO is a specialization.

An object can be accompanied on the trajectory. This can be indicated by *with*. Also, there can be an AGENT which is responsible for the change.

Although one can think of the change model as underlying every change described in English, the typical sentence describes only certain aspects of the change. Since the OBJECT, SOURCE, TRAJECTORY, and DESTINATION are slots in a TRANS frame, one may optionally include all of them in a sentence using *go*:

(83) John went (from the house) (through the gate) (to the store).

Here, the SOURCE, TRAJECTORY, and DESTINATION are used to form the initial state, location of John at house, trajectory, and final state, location of John at store, of TRANS. Almost all verbs, however, require the user to be less general and complete. They

- a) Focus on only part of the change.
- b) Specialize the change so that it can apply to only particular kinds of states, e.g., LOCATION.

For example, the verb *put* concentrates on the final state. If I say that I put the pencil in the desk then I mean I caused the final state. Note that *put* doesn't even allow us to mention the SOURCE or TRAJECTORY. One cannot use *put* without implying that the final state has been reached. On the other hand, one can *send* an arrow into the air with no implication of a destination. Send means *cause to go, without going with*.

When someone dies (goes to his reward) then he has experienced a particular change of state, and he has completed it. (The destination state is known as death. We are *close to death*.) We don't have to say *He finished dying*. This notion of completion of state change is an important one. When one begins eating dinner, how much does he have to eat before he has eaten dinner? From the point of view of the eater, we ask if he is finished eating. From the point of view of the dinner, we ask if he ate the dinner up. Notice the difference between:

- (84a) He ran to the tree.
 (84b) He ran for the tree.

This use of for marks a destination of the state change, as does to; but, unlike to, for does not imply the reaching of that destination.

At the most primitive (or highest) level of abstraction, let us postulate two models: the state change model and the life history of a thing. An object can come to exist, but if it is not also available, and operational it is of reduced importance. There is a notion of being in the collective center of things. Since conservation of matter holds, objects are not caused to exist out of nothing. Rather, they are either a) brought into the center of things, b) noticed for the first time even though they have already been there, or c) formed from something else. For example an object can appear and then disappear. It can be produced (L. *producere*: pro- forward + *ducere*, to lead, draw) or disbanded. It may be born into the world or carried out. Something goes from raw material to a finished product, or something goes out of one form into another.

- (85a) I made a dress from cloth.
 (85b) I made a dress out of cloth.
 (85c) I made the cloth into a dress.

Once something exists, it is generally not “un-made”. Rather, the end of its useful life is remembered by the state change which moved it out of the center of things or transformed it so it is no longer suitable for its old role. If John eats a piece of cake and someone asks *where is that piece of cake?* the answer is usually *John ate it*, not *inside John*. The cake still exists in a sense, but it is *gone*. A typical answer to *Where is that jar?* might be *I threw it out*. Out where? Out of the center of things. Another answer might be *I made a bird house out of it*.

6.11 Micro-Worlds

In a way, life is quite simple. Agents use information to form plans and carry them out. To describe this, English relies on analogies to the physical world! It is as if we only have one set of conceptual machinery, this being suitable for describing configurations of objects and their motion in a three dimensional setting. The more abstract concepts must then be fit to the same machinery. Our representation of them winds up as a vast complex of analogies to the physical world.

We define the physical world to be made of PHYSICAL-OBJECT's and their physical attributes. Basically, a physical object can be characterized as:

- (86a) something made of matter, or
 (86b) a “picture producer” (Schank's terminology), or
 (86c) something made of the “four elements”: earth, air, fire, and water, or
 (86d) something “concrete” as opposed to “abstract”.

By physical aspects we mean those which are directly observable, like size and color. Time advances in the physical world, and the activities which are aspects of physical objects take place. At any point in time, the physical world is completely described by listing all of the objects in it with their physical aspects.

As a thinking being, it is the goal of an English speaker to predict what will happen next in the physical world. This may be an explicit prediction, or it may be only implicit, in the sense of appropriate reaction to previous events. It is fair to judge a world model, be it in a computer or an English speaker, primarily in terms of its predictive capabilities. For this reason, the notion of AGENT of a change is of utmost usefulness. Even if we do not model the motivations of the AGENT, observed sequences of acts by the same AGENT give us predictive power. In the sentence, *The dog hurt the man*, we have an abstraction of the dog as an agent of hurt, but what are we to do with *The government hurt the man*? In the physical world there is no *government*. However, the set of physical events that *hurt* the man become more understandable if the abstract concept of *government* is specified as agent. Formalization of this leads us to the notion of the system world, a world of entities which serve as agents. Some entities in this world and their aspects are shown in Figure 37.

There is a hierarchy of agents which may be discovered using the pattern "x was given y". For example, *the dog was given the bone* and *the bone was given the dog* both mean the same thing because dogs come above bones in the anthropomorphic hierarchy. An entity in the system world can be a system abstraction of an entity in the physical world, or it can be unique to the system world. For example, in the sentences

- (87a) The man carried the item.
- (87b) The store carried the item.

we see that in the first sentence carry is ambiguous between the physical meaning of carry and the system world meaning of carry appearing in the second sentence. Store selects a meaning of carry in the system world, and man tends to select the meaning in the physical world, though the system abstraction of man allows the system world interpretation. Note that there are also physical world stores (a building set up in a certain way) but they cannot carry anything. Note that there is a certain analogy between the two meanings of carry. In fact, there is quite a strong analogy between the physical and system worlds. The most richly described system in the system world is the social system and its various subsystems. "Locations" in the social system are:

- (88a) on the social scene
- (88b) in an organization
- (88c) in a social class
- (88d) at the top of an organization
- (88e) under a bad boss

The physical location prepositions are still used, but their social abstractions have less meaning. One can use these locations in analogies with the physical world like

Fig. 37. Entities and Properties in the System World
<= =<fig5-12.press<

- (89a) He came from the lower class.
- (89b) he dropped out of school.
- (89c) He got out of trouble.
- (89d) He followed a career path in the organization.

Some aspects in the systems world are FAMILY-RELATIVE, CUSTOMER, BUSINESS, OCCUPATION, RESPONSIBILITY, and AUTHORITY. Corresponding to PART in the physical world, one can be a PART of an organization. A kind of a PART is a MEMBER. An organization can have a top, a bottom, arms, and a head. In addition, the systems world is the locus of "possession". In the physical world we observe the following relations between "have" and "location".

- (90a) There is a cookie in the box.
- (90b) The box has cookie in it.
- (90c) The box contains a cookie.
- (91a) There is a cookie on the plate.
- (91b) The plate as a cookie on it.
- (92a) There is a tree on the mountain.
- (92b) The mountain has a tree.
- (93a) *There is a cookie at the ice cream.
- (93b) There is a cookie with the ice cream.
- (94a) There is a dog at the house.
- (94b) *There is a dog at Jim.
- (94c) There is a dog with Jim.
- (94d) Jim has a dog with him.

| <u>Action</u> | <u>Result</u> |
|-----------------------------|--------------------|
| I take the dog to the tree. | He is at the tree. |
| I take the dog to Jim. | Jim has him. |
| I take the dog. | I have him. |

From this we argue that to "have" an item in the physical world means for it to be located there. If the location is *in*, then the have can be further specified to contain.

By analogy, *to have an item* is specialized in the systems world to *to possess an item*. An item is located at a person or organization in the sense of possession. Possess can be further specialized to own or control. If someone controls you, you are under him. These specializations of have were pointed out by Charniak [13]. If my father gives me the car for the evening and I say, *I have the car*, I do not mean that it is physically *at me*, nor do I mean I possess it. I merely control it temporarily.

In this light we can associate give, get, and take with changes of having, recognizing the relationship between have and location and the fact that getting something to own can involve buying it as in

(95) I got this car for a song.

Agents take actions to get credit for them. If one does an action for somebody else, then the credit for that action goes to him. An agent hopes for success, not failure, in an organization.

In the system world, we describe all possible agents and the organizational constraints placed upon them. In order to better predict how these agents will act, we need the notion that some of these agents, namely animals and organizations, carry out actions in accordance with plans (possibly instinctive). Again, this is handled by analogy to the physical world. Thus we have the world of plans and actions. In addition to the plans and actions abstraction of agents, this world contains the entities: plan, intention, alternative, goal, resource, action and mistake.

Activities like maintain, insure, limit, and trick are abstract activities used only in the description of plans. Agents decide what alternative plan to follow. Wise agents make good decisions. Using resources, agents take actions in accordance with plans which they have. They have reasons for the decisions and actions. They carry the plan through to completion, they reach their goal, or they make a mistake. An action may be easy or difficult. An agent may encounter a difficulty. In this world an agent is located with respect to actions. He is *in the middle of dinner, in surgery, on trial, or under investigation*. A resource is *in use*. This analogy is much less complete than the systems world.

The mental world is even less complete. The mental world is the source of plans and reasons. It contains situations, problems, concepts, names, facts, ideas, views, and opinions. The mind is added as a location for these. The only agents in this world are the mind and the attention. Something can be located

- (96a) in mind
- (96b) on one's mind
- (96c) in the back (or front) of my mind
- (96d) on my tongue
- (96e) on the tip of my tongue
- (96f) in my head
- (96g) on a solid surface (He put the idea on paper).

The attention can be drawn to a problem or the problem can be brought to the attention. One thinks over a situation in order to discover a problem and solve it.

It is not clear how to classify the perception and communication of information, although it probably arises in the social world. This is done by the agents of the social world, but to understand their use of it we must move to the mental world. How should we analyze *She solved the problem with a suggestion from Bob?*

7. Procedures

7.1 Procedural Representation of Knowledge

It is useful to distinguish between procedural representation of knowledge and representation of knowledge about procedures. By procedural representation of knowledge we will mean casting that knowledge in the form of knowledge about the results of a procedure. For example, the knowledge that every boy wants a lion could be cast in the form: "if you check every boy and count those who want lions, then the count of those who want lions will be equal to the count of the boys checked."

Woods [115] introduced a FOR iteration construct for representing knowledge of quantified propositions procedurally. Examples of the use of this construct are:

- (97a) (FOR EVERY X / CLASS : (P X) ; (Q X))
Every X in CLASS that satisfies P also satisfies Q.
- (97b) (FOR SOME X / CLASS : (P X) ; (Q X))
Some X in CLASS that satisfies P also satisfies Q.
- (97c) (FOR GEN X / CLASS : (P X) ; (Q X))
A generic X in CLASS that satisfies P will also satisfy Q.
- (97d) (FOR THE X / CLASS : (P X) ; (Q X))
The single X in CLASS that satisfies P also satisfies Q.

Woods assumes that for each CLASS there exists an enumeration function which will produce all the members of the CLASS. The FOR statement specifies that the predicate P(x) be applied in turn to each element adduced by the enumeration function. The predicate Q(x) is then applied to those for which P(x) is true.

The first argument of the FOR divides the statements into those (EVERY, THE EQUAL GEN) which exhaust the enumeration function and those (SOME, ORDINAL, GREATER) which enumerate either until exhaustion or until a termination criterion is satisfied. The description of the termination conditions and the resulting states assumes that, in general, three counts are maintained.

- (98a) a count C of the elements enumerated so far.
- (98b) a count CP of the elements passing P(x) so far.
- (98c) a count CPQ of the elements passing P(x) and Q(x) so far.

For example, the (FOR EVERY ...) statement says that in the situation resulting from its execution $CP = CPQ$.

The FOR iteration construct allows knowledge about sets to be built up from knowledge about individuals contained in the sets. As was mentioned above, such a composition operation should not be used to the exclusion of decomposition. A decompositional approach allows predicates on sets of individuals rather than just predicates on individuals as is done here; predicates on individuals are inferred from predicates on the sets they are in. Both composition and decomposition are needed.