THE REPRESENTATION OF KNOWLEDGE AND THE USE OF KNOWLEDGE IN DISCOURSE COMPREHENSION

Walter Kintsch

University of Colorado

SCIENCE FORUM 1986: LANGUAGE PROCESSING IN SOCIAL CONTEXT

Heidelberg, 1. - 5. September, 1986

Technical Report No. 152, Institute of Cognitive Science,

University of Colorado, Boulder CO 80309

June, 1986

THE REPRESENTATION OF KNOWLEDGE

AND

THE USE OF KNOWLEDGE IN DISCOURSE COMPREHENSION

Walter Kintsch

1. Introduction

In this talk I shall explore some ideas about knowledge representation, and about how knowledge is used in discourse comprehension. I shall first discuss some approaches to the problem of knowledge representation that have been used in psychology and AI. From these, I attempt to formulate principles that would seem to be desirable for an adequate system of knowledge representation. I shall also discuss a few experimental results, some of them from my laboratory, which appear to be of direct relevance to the issues under consideration. Thus equipped with guidelines and data, I shall broach the problem of constructing a knowledge system and show how such a system might operate in discourse comprehension. What I present here is far from a worked out theory; all I can do is raise some issues about knowledge representation in the context of discourse comprehension. I believe that the problem of knowledge representation and knowledge use is a focal one at the current stage of development in AI, linguistics, and psychology and that further progress in these disciplines will depend on finding solutions which are more adequate than what has been tried so far.

Let me give a simple example, just to specify more precisely the nature of the problem I am concerned with. Consider what sort of knowledge use is involved in understanding the following three sentences:

After an unusually heavy thunderstorm,

The water overflowed the bank of the river.

It was heavy work to clear the mud from the streets.

I am first of all interested in how just the right knowledge about word meanings is activated during comprehension. The principal dictionary meaning of *heavy* is something like *hard to lift because of its weight*, but a *heavy thunderstorm* is something quite else, and *heavy work* is something else again; nevertheless, we unhesitatingly and quite unconsciously arrive at the correct interpretation of these phrases. We know where *the water* comes from, and otherwise strong associations of *water* such as *glass, drink, ocean,* and *liquid* are unlikely to come to mind in this context, while *flood* probably would. When we read *bank,* we don't think of money and building at all.

How is it possible that exactly the right knowledge about a word is activated in the discourse context, and that everything else that we know about it doesn't intrude? The "meaning" of a word seems to be constructed appropriately for each context, and is therefore always a little different. How does this happen? More precisely, there seem to be two questions involved here: what is the *knowledge organization* that permits this astonishing degree of context sensitivity, and what do we know about the *process of knowledge use?*

2. Approaches to the representation problem

2.1 Associationism

Associationism was psychology's inheritance from philosophy. It has dominated psychological theory for a long time. It never was, of course, a monolithic system, but subsumed some rather distinct schools of thought.

An associative network is a structure in which the nodes are unanalyzed concepts, and the links are unlabeled, but vary in strength. An example of such a network is shown in Figure 1. Operationally, such networks are obtained in free-association experiments: the stream of thought brings forth what is similar, opposite, or spatio-temporally related - as Aristotle said it (after Strube, 1984).

Figure 1

Knowledge is thus represented as a network of ideas, with interconnections determined by the laws of association, i.e. "resemblance, contiguity, and cause-and effect" (this is Hume's list - a variety of related proposals have been made). Note that this is a messy, perceptual-based system, not a clean, neat, logical-conceptual structure.

Not everybody, not even every associationist, felt that such a network provided a sufficient basis for human cognition. Even Locke viewed association by contiguity merely as an element of randomness supplementing rational thought. Two developments which took place within psychology during the early part of this century, and which are very prominent in current AI, seem to me crucial for understanding these rational aspects of human thought.

2.2 Control Structures

In a classical associative network, the strongest association wins; The German psychologists Ach (1910) and Lewin (1917) showed that this was not so with real people: not necessarily the strongest association occurs, but that which corresponds to our train of thought guided by the "determination" (Ach); what happens depends both on the nature of the associative network and the person's "action readiness" (Lewin). Ach and Lewin clearly recognized the importance of control processes in knowledge activation.

Recent developments in cognitive science and AI have underscored the importance of these considerations. Classical associative nets (as well as semantic nets and frame-based systems) are usually thought of as being passive: spreading activation (or marker passing, in the terminology of AI) provides the basic processing mechanism; whatever connections exist determine the pattern of activation, once the exact rules by means of which activation spreads have been specified. We know, from the psychological work of Ach and Lewin that this is not enough to account for the orderly progression of human thought. It may also be not enough for efficient knowledge retrieval in AI. Recently, Kolodner (1983) has strongly emphasized the active control processes that characterize memory retrieval. Within a frame-like representation system, Kolodner developed a very active retrieval process: retrieval often requires searching for something other than what was requested, and sophisticated executive strategies are necessary to control this process. Psychological data on knowledge retrieval (Walker & Kintsch, 1985) have revealed the existence of a passive retrieval mechanism, which determines what is retrieved once a memory probe has been formed, and of control strategies, which are needed to put together an appropriate probe. Thus, we have good reason to believe that an adequate knowledge representation must be an *active* system.

2.3 Semantic nets

While associative nets are messy, perceptually based, semantic nets are orderly and conceptually based. Indeed, their originators wanted to represent the objective part of word meanings for use in human-like systems (Quillian, 1968; Collins & Quillian, 1969). They did this by designing a network the nodes of which were word concepts, linked by labelled relations, such as the "ISA"-relation in the hierarchy studied by Collins & Quillian. These links defined the meaning of a word concept, much as in a dictionary definition: the nodes to which the word concept is linked to form a "plane", which can be considered its definition. The inferential capacity of the system were of prime importance: for instance, in Collins & Quillian's taxonomic hierarchy, properties of higher order nodes were inherited by the lower nodes, as shown in Figure 2. Thus, SHARK was defined merely by the properties CAN BITE and IS DANGEROUS, but since it was linked to FISH via an ISA-link, it inherited such FISH properties as CAN SWIM, HAS FINS, etc. FISH in turn, was linked to ANIMAL, which was assigned attributes common to all animals (HAS SKIN, etc.), which then through an appropriate chain of inferences, could be inherited by all of its subordinates.

Figure 2

Semantic nets can be used for many different purposes, and a great deal of confusion has occurred in the literature because of a failure to recognize the existence and justification of different types of semantic nets. Brachman (1979) listed the following types of semantic nets, starting with the most abstract one (I am omitting his category of implementational nets, which are mere data structures without epistemological implications):

1. Logical nets Links represent logical relations and nodes are predicates and propositions

2. Epistomological nets The links are inheritance and structuring relations and the nodes are concept types (rather than particular concepts - this is the level of abstraction Brachman himself argued for);

3. Conceptual nets The links are semantic relations (cases) and the nodes are primitive objects or actions (this is the most common type of structure, as in Schank, 1972; Norman & Rumelhart, 1975);

4. Linguistic nets The primitives are language dependent and all meaning

derives from context, and changes as the network grows.

There is little to choose among these four types of networks a priori: which is best depends on the task at hand, and the specific way the network is designed. Psychologists will presumably be most interested in last two types of networks, while logicians or formal semanticists have different requirements.

A canonical form is implicit in the first three types of nets, with logical primitives, knowledge structuring primitives, and semantic primitives as the units, respectively. The various arguments (e.g. Kintsch, 1974, Fodor, 1983) made against the notion that concepts must always be decomposed into some set of primitives when used in either comprehension or production are, therefore, also arguments, against representation types 1-3 above. However, since it is clearly the case that people can decompose semantically complex concepts into simpler ones (not necessarily into a finite set of primitives, though), any psychologically plausible system of representation must permit such decomposition, though it should not require it.

Semantic nets are very popular in AI (see the discussions in Brachman & Levesque, 1985). Indeed, most representation schemes used in work on natural language processing are some sort of semantic net, often elaborated to incorporate frame structures (to be discussed below). As models for the knowledge structures people use, however, semantic network models have been a failure (for a critical review see Kintsch, 1980). The very features which make semantic nets so attractive computationally - their clear conceptual structure - clashes with the openness and flexibility that characterizes human knowledge use. The wide use that AI makes of these techniques is born more of necessity and a lack of alternatives than a satisfaction with the status quo. Ideally, knowledge systems in AI should be just as flexible, context-sensitive as human memory; really large systems probably have to be in order to be workable.

2.4 Frames, Scripts, and Schemata

The major departure in AI from pure semantic nets has not been towards greater flexibility, however, but into the opposite direction: towards a more structured knowledge representation in the form of frames, scripts, and schemata. Knowledge is often used in well-structured chunks: A fixed skeleton of knowledge can hold together information of a certain kind and assign it a global meaning which each piece alone would not have. The term "schema" was used for such structures, first in England by the neurologist Head (1920) and the psychologist Bartlett (1932).

Frames were introduced into modern cognitive science by Minsky (1975) in the context of visual perception. Scripts were popularized primarily by Schank & Abelson (1977) in their work on natural language processing, for much the same reasons which motivated Bartlett (1932) and Selz (1922) in psychology several decades before. What was needed was a method for organizing knowledge representations to facilitate inferencing and retrieval on the one hand, and to form expectations that could focus and guide attention during processing on the other.

Frames, or to use the more general term, schemata, consist of a heading and various slots. A well-known example is illustrated in Figure 3. The slot name specifies the relation of the information contained in the slot to the schema as a whole (e.g., it is the LOCATION of of the object named in the heading), and the slot specification constrains the nature of the information that can be assigned to that slot (e.g., it must be an ADDRESS). Schemata can get very complicated, with slots and subslots, and embedding of other schemata. Thus, the RESTAURANT frame, everybody's favorite example, has a slot EVENT-SEQUENCE which contains a script, EAT-AT-RESTAURANT. It, in turn, has such slots as PARTICIPANTS, PROPS, etc., as well as another EVENT-SEQUENCE specifying the usual sequence of actions involved in eating at a restaurant. Note that frames combine both declarative and procedural knowledge, e.g. about various types of restaurants, and what to do in each particular case.

Figure 3

The computational power of frames arises in part from their inferencing capabilities: each slot has its default value which can be activated if no other information is available. Thus, as soon as a frame is invoked, a great deal of well-organized knowledge becomes available, without the need for elaborate computations.

Frames, however, also play a role in the processes of comprehension and perception themselves: they permit expectations to be formed, which can be used to guide these processes.

While there is considerable evidence for frame-based inferences, and while humans can be shown to use script-like knowledge structures on occasion (e.g. Bower, Black, & Turner, 1979), scripts and frames cannot be the ultimate answer to the problem of knowledge organization. It has become more and more clear (Schank, 1982; van Dijk & Kintsch, 1983) that such fixed structures are much too inflexible, both to simulate human knowledge use and to support more sophisticated AI systems. What is required in using knowledge is a system that structures knowledge in a way appropriate for the specific context in which the knowledge is to be used. A "generic" frame is insufficient - each situation seems to require its own, context-specific frame. There appears to be no way one can foresee all posibilities in a situation, or make room for all contingencies in precompiled knowledge structures such as frames or scripts. Rather, we need to find a way to generate such structures in just the right way for the particular context at hand from a flexible knowledge structure that contains frames and schemata only as possibilities to be realized on demand.

2.5 Production systems

Production systems (Newell, 1973) are a form of knowledge representation which provides for a good deal of flexibility. Productions are condition-action pairs, much as the stimulus-response pairs of behavioristic psychology, but without the observability constraints of the latter. The set of productions by itself is quite unstructured. To make it work it needs two kind of control processes. One is a short-term memory buffer: only the data currently held in that buffer can activate the condition of a production;

thus the flow of data in and out of short-term memory determines in part what productions are executed. However, since it will frequently be the case that more than one production condition matches the data in the short-term memory buffer, some kind of conflict resolution procedure is required. A fraction of a production system for performing addition is shown in Figure 4.

Figure 4

For the purpose of knowledge representation, production systems have some very attractive properties. We don't have to worry about their psychological plausibility: for decades psychologists knew nothing else but S-R connections. Of course, behaviorism is dead today, but that does not mean that the idea of condition-action pairs is a bad one, only that the constraints the behaviorists put on their S's and their R's are unacceptable. From the standpoint of AI, it is clearly an advantage that all knowledge is represented in the same way in production systems. Furthermore, since the system itself is not structured, it is quite modular, so that, for instance, it is relatively easy to add or delete productions without affecting the remainder of the system.

Production systems have their disadvantages, too. It is not easy to understand what actually happens in a large production system. Also, such systems may not be as modifiable as one would suppose: interactions between productions can have surprising outcomes. Even an expert system that peforms quite well, such as MYCIN, must be supplemented by more structured, declarative knowledge in order to make it capable of learning from its experience (instead of just having new productions added to it by an outside agency) and explaining itself (Clancey, 1984). An appropriate guiding structure is missed also in another way in large production systems: such a structure could greatly improve the efficiency of the system. Great computational effort is wasted by matching the conditions of numerous productions which a well-organized system would never even consider in certain contexts. Thus, while frames and semantic nets gave us too much structure (or, rather, too inflexible a structure), we may get too little from production systems.

2.6 Associative nets

An old idea in philosophy and psychology, associative nets have only recently been investigated in AI as systems for knowledge representation (McClellamd & Rumelhart, 1985; Waltz & Pollack, 1985). Here we find the ultimate lack of structure. Knowlegde is represented simply as a large, highly interconnected set of nodes, with the connections varying in strength and nature (facilitatory versus inhibitory connections). Or, alternatively, we could talk about a set of neurons and their interconnections much as we find in the brain.* No rules, production or otherwise, are built into such a system, but rule-like behavior can nevertheless arise from it.

---

* While the brain analogy is striking, at least superficially, it is in no way essential to the use of associative nets as knowledge representation systems.

Modern associative nets, or connectionist system as they are called, are graphs with weighted nodes and links, and an iterative operation which recomputes the activation level of each node. The links are either excitatory or inhibitory in nature. The activation of a node is a function of its current value and the inputs, excitatatory as well as inhibitory, which it receives from the other nodes it is connected with. Thus, if some node is activated, say because the word IRON has been read, excitation from that node spreads to its neighbors, in proportion to the strengths of their links with IRON. STEEL, perhaps, might be the most strongly excited neighboring node at this point. As other nodes in the system become activated, however, inhibitory effects develop. For instance, if IRON is read in the context of "ironing clothes", the original activation of STEEL will decay within a sufficient number of cycles, because the meaning of IRON as METALL will be suppressed as incompatible with the contextually dominant meaning.

In a simple example worked out in Waltz & Pollack (1985) 50 iterations were needed to arrive at a stable activation pattern. In more complex systems, many more cycles might be required. It can be shown, however, that for certain assumptions, such systems eventually do stabilize. Figure 5 shows the ambiguous input sentence "John shot some bucks" with a fraction of the associated network; it is assumed that the context node "HUNT" is activated. In Figure 6 the activation pattern that develops in this system is illustrated: the contextually appropriate word sense of "shot" is now activated and the inappropriate associates of "shot" are suppressed.

Figure 5 & 6

2.7 Issues

A number of general issues have emerged from this discussion of current knowledge representation schemes. I have already discussed the need to distinguish between different levels of abstraction: what is good for formal semantics may not necessarily be good for an expert system, and a simulation of human knowledge organization may require something else, still. It has also been noted that knowledge representations must be built so that they permit the operation of control strategies and problem solving mechanisms. Several further issues appear worth comment.

Some of the systems I have discussed tend to be more perceptual and chaotic in character, while others are more conceptual and logical.

How much, and what sort of structure "is" there in a knowledge system, versus how much structure is generated in a context-dependent way whenever a task requires it, and how is it generated?

The traditional view is in terms of fixed knowledge structures, both in psychology and AI. Associative nets (Anderson & Bower, 1972), semantic nets (Collins & Quillian, 1969), or schemata (Rumelhart & Ortony, 1976), are thought of as the structure of the mind; frames (Minsky, 1975), structural inheritance nets (Brachman, 1979), scripts (Schank & Abelson, 1977), and the like make up the data structures in the best known AI programs. It has, of course, not escaped the attention of the designers of these systems that the relatively inflexible nature of such pre-existing knowledge structures makes for a lot of problems. Schank & Abelson, for instance, tried to give their scripts the required flexibility by introducing different tracks. But solutions of that nature eventually have to be abandoned: in the long run, it turns out that almost every time a script is applied, a new "track" would be needed. Its almost limitlesss flexibility and sensitivity to context is perhaps the most salient characteristic of human knowledge use, and it is becoming more and more clear that a similar flexibility will be required for AI systems, if they are to approach human-like performance levels. The problem is by now widely appreciated; the question is how to conceive of flexible, context-sensitive knowledge systems that nevertheless provide the right kind of organization and structure when it is needed.

3. Some experimental results

How does context influence the process of word identification? Recent data from psychological experiments have forced a reassessment of the widely held belief that context-based expectations directly facilitate or interfere with the way in which the perception of a word makes contact with the knowledge about that word. Instead, the context becomes effective only after this initial knowledge activation phase, when information from many different sources is integrated.

3.1 Priming effets in lexical decision experiments

Consider the following experimental paradigm. A subject listens to a brief text, such as my old example

After an unusally heavy thunderstorm,

the water overflowed the bank of the river. At unpredictable intervals, while the subject is listening to the text, letter strings appear on a screen in front of the subject, and the subject is asked to decide as rapidly as possible (by pressing a response key) whether the string is an English word or not. Thus, the subject performs two tasks concurrently: a listening comprehension task, and a lexical decision task. In the lexical decision task, priming effects occur: e.g., if the word "river" is presented immediately after the spoken word "bank", the reaction time to identify this word is reduced in comparison to unrelated control words: the associative/semantic relations between "bank" and "river" facilitate access to the second word, once the first has been activated. Such priming effects are well known, in discourse contexts (e.g. Swinney, 1979), as well as in list contexts (e.g. Meyer & Schvaneveldt, 1971).

The discourse context is irrelevant to this priming effect: if, in the example above, the context appropriate associate "river" is replaced with the context inappropriate associate "money", an equally strong priming effect is obsered (e.g. Swinney, 1979; Kintsch & Mross, 1985; Till, Mross, & Kintsch, 1986; similar results were obtained with a naming task by Seidenberg, Tanenhaus, Leiman, & Bienkowsky, 1982). What matters seems to be the relatively fixed lexical context of "bank" - not the momentary discourse context in which this word is used! The lexical connection between "bank" and its associates "river" and "money" are both activated, irrespective of the sense in which "bank" is used in the discourse. This activation is, however, only a brief one: if "money" or "river" are not presented immediately after "bank", but are delayed for 500 msec, only the context appropriate "river" will be primed. The fleeting nature of this priming effect undoubtedly accounts for the fact that the context-inappropriate meaning of "bank" does not rise to the level of consciousness under normal reading conditions.

Not only does the discourse context not suppress the inappropriate meaning of "bank", it also does not facilitate the identification of words that are highly context appropriate, but which are not associatively or semantically related to "bank": if "flood" is presented as the target word in the lexical decision task immediately after "bank", it is identified no faster than context-irrelevant control words (Kintsch & Mross, 1985; Till et al., 1986; Seidenberg et al., 1982), in spite of the fact that "flood" is a highly probable inference at this point. However, if one waits a second to give the subject a chance to actually make that inference, "flood" will be identified significantly faster than context-unrelated control words (Till et al., 1986; for word recognition, McKoon & Ratcliff, 1986).

The only "context" that seems to have immediate effects on word identification consists of the associative and semantic relations of a word in the subjective lexicon and/or general knowledge base. The discourse context, on the other hand, becomes effective only after the inital knowledge activation has occurred.

3.2 The time course of word identification

Thus, word identification seems to be a more complex process than we have heretofore supposed, with a strictly botttom-up, data-driven initial phase in which the discourse context plays no role, followed by an integration phase in which the discourse context shapes the incoherent bits and pieces of knowledge that have been activated into an integrated whole. Figure 7 is an idealized summary of the experimental data which have been discussed here.

Figure 7

The very first stage of word identification is perceptual - a process of feature detection which makes contact with an appropriate lexical node; this process appears to be fairly far advanced 50 msec after a word is seen. Once a lexical node is activated, activation spreads to its neighbors in the lexical network in a context independent fashion: all word senses contacted by the perceptual analysis are activated at this point. This stage of *Sense Activation* lasted at least until 333 msec in the Till et al. study. By 500 msec, however, *Sense Selection* had occurred in that experiment: only the context appropriate associate continued to be primed at this point. Activating the right lexical sense of a word is, however, far short of establishing its full contextual meaning: A process of contextual enrichment or *Sense Elaboration* is necessary to generate the latter. In the Till et al. experiment, this process was sufficiently far advanced 1000 msec after a word was seen, so that priming effects for inference words were obtained; in less clear-cut situations, sense elaboration may require considerably more time (e.g. when subjects have to infer the meaning of a novel word from the context alone, as in Clark & Gerrig, 1983), or may remain incomplete, especially when subjects must operate under time pressure.

Figure 7 illustrates the time course of word identification processes, as inferred from the studies above. While many questions remain open, and further experimental results will undoubtedly modify our present understanding, the data are sufficiently suggestive to serve as a basis for a model of how people use their lexical knowledge in discourse comprehension.

4. Knowledge use in Word Identification

The model of word identification in discourse which is proposed here is based on two assumptions: (1) knowlegde is represented as an associative network the nodes of which are propositions, and (2) the time course of context effects in word identification is approximately as described in Figure 7.

4.1 The knolwedge net

The knowledge representation considered here is an associative net with both atomistic and holistic properties (Mudersbach, 1983). The net is atomistic because its nodes are propositions which have an internal structure and meaning of their own, and it is holistic because the full meaning of a node must be generated from its position in the whole network and from the momentary context.

The propositions which are the nodes of the net are predicate-argument structures, as they are used in much current work in this area.

Links among nodes vary in strength from facilitatory to inhibitory, with values between 1 and -1. One can think of the net as a huge matrix with the nodes of the net as the rows and columns and the entries indicating the strength of the connection between any two nodes.

4.2 The Process

As a sentence is read, each word contacts its lexical node, thereby activating the information available at that node. The information thus activated provides merely a sketch of the intended meaning. This core meaning is enriched by a process of random sampling of related propositions: a certain number of neighbors of the core proposition are selected at random, with probabilities proportional to the strengths of their connection to the core proposition. Thus, if a whole sentence is read, a subnetwork of interrelated propositions is generated, consisting of

1. The text propositions themselves which represent the meaning of the sentence; these are all positively connected, in proportion to their distance from each other in the textbase, and

2. The knowledge nodes (propositions) activated by the text

propositions.

The propositions in the sub-network thus generated are richly interconnected, both positively and negatively. We now let the activation spread in this net: the net undergoes repeated cycles of stabilization, during which the activation which originally was concentrated on the actual text propositions, spreads to the other nodes in the system, in accordance with the positive and negative interconnections in the system. If all goes well, a stable pattern of activation will be achieved. (In the examples below this happens in from 10 to 30 cycles). The propositions that are still highly activated at the end of this process - which may be some or all of the actual text propositions, plus whatever bits of knowledge survived the integration process - constitute the end result of comprehension: a knowledge-enriched textbase.

If the pattern of interconnections in the mini-network is such that the activation does not stabilize within some reasonable number of cycles, "immediate" or "automatic" comprehension fails, and recourse must be taken to strategic processes, e.g. new nodes can be sampled from the network, in the hope that the added interconnections will permit the system to find a stable solution, or other, more sophisticated problem solving heuristics may be required to infer missing links in the textbase, or to reinteprete the data. Thus, comprehension in this view is both direct, immediate, automatic, perception-like, in other words, and if this fails, deliberate, strategic, conscious, that is, like problem solving. Here, I am concerned only with the first phase.

4.3 Examples: The literal meaning hypothesis

The view of knowledge use in discourse sketched here has implications for the long-standing controversy about literal meaning (e.g., Gibbs, 1984). What I would propose is to regard the propositional core meaning, with only minimal knowledge elaboration, as the literal meaning of a sentence. Since this semantic sketch is very superficial, very rough, and may be incoherent (as when both meanings of a homophone are activated), this is probably not at all what proponents of the literal meaning hypothesis have in mind. Indeed, since this initial semantic representation is not even conscious, it is far away from a well-defined, sentence meaning. But it does represent a common core of meaning, upon which all further contextual elaboration and interpretation is based.

Consider

(2) *The cat sits on the mat.*

According to this hypothesis, the initial semantic sketch of that sentence would be the proposition SIT[CAT,ON MAT], together with some minimal elaboration of SIT, CAT, and MAT. E.g., for CAT this might be *cats are pets, and I love them,* with associations such as *cats purr, my cat is black, tigers are a kind of cat, etc.* The literal meaning of this sentence is this set of immediately activated semantic nodes, forming shells around the three lexical nodes. The full meaning of the sentence, out of context, is probably not much richer than that for most readers of that sentence. It is true, one could construct an image, work out presuppositions (the mat flies through the air like a magic carpet, the mat is on the floor) - but there is little incentive to do so, other than for the purpose of philosophical discussion. In context, the situation might be quite different: imagine a story about a young couple whose favorite cat was lost, and they searched for it all over town; when they come home, the woman utters "The cat sits on the mat". A great deal of contextual elaboration can now take place - a particular cat and a particular mat are introduced, and the sentence becomes the resolution to a whole story.

Idioms and metaphors are at the heart of the literal meaning controversy. In the present view, following Ortony (1979), there is no processing difference between the metaphorical meaning of a sentence and its non-metaphorical meaning. Both start out with the same semantic sketch, which then needs to be contextually elaborated. There are no reasons why one kind of elaboration should always be more complex than the other. A familiar idion like

(3) *He let the cat out of the bag.*

has a semantic skectch consisting of the proposition LET[HE,CAT,OUT-OF-BAG], where HE, out of context, is merely a placeholder, CAT is elaborated more or less as in the previous example, and BAG is similarly specified; the *the* are potential pointers that point nowhere. In a context where *he* refers to a politician with sinister plans, the elaboration of this primitive semantic sketch will procede along the idiomatic meaning of the phrase. In the context where *he* is a thief who has just stolen a valuable Siamese, the elaboration will take a very different course. There is no reason why one should be more difficult than the other. No wonder, as numerous experiments have shown (e.g.Glucksberg, Gildea, & Bookian, 1982), it takes people about equally much time one the average to come up with either kind of elaboration.

As a final example, consider

(4) *In the beginning was the word.*

The initial semantic sketch here is probably even less elaborate than for "The cat sits on the mat": *beginning* and *word* are readily identified, but the resulting elaborations will probably turn out rather incoherently, with various mutually incompatible associations; once again, the *the's* point to a void and are not very helpful. We have no problem understanding, but we don't understand much. The contextual process of elaboration, in this case, has gone on at the cultural rather than the personal level for over 2,000 years. Different knowledge, biases, beliefs, and goals have yielded many different meanings - often very deep and elaborate ones. But out of context "In the beginning was the word" is about as trivial as "The cat sat on the mat", only more vague.

4.4 Calculations

Finally, I present a worked-out numerical example, to illustrate my discussions of knowledge representation and knowledge use in discourse. This example combines connectionist ideas about knolwedge representation and activation (Waltz & Pollack, 1985; McClelland & Rumelhart, 1986) with the model of discourse processing I have been working on for some time (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). In presenting this example, I have chosen an extremely informal notation in order to make a complex story understandable - even at the risk of some lack of precision.

I shall be concerned with the following mini-discourse:

(5) *John was thirsty.*

(6) *John took a glass of water.*

The propositional representation of this discourse is

(7) JOHN(X)

THIRSTY(X)

GLASS(Y)

WATER(Z)

TAKE(X,Y)

IN(Y,Z).

Each of these propositions is connected to many other nodes in the reader's knowledge net. I have to make specific assumptions about these nodes and their interconnections. Since *John* is merely a dummy in this context, I assume that JOHN(X) samples only two knowledge nodes - that JOHN is the NAME of a PERSON, which is MALE. For THIRSTY and WATER, existing associative norms can give an idea what other nodes are closely related to them. Thus, I have assumed that THIRTSY samples the nodes DRINK, DRY, HUNGRY, SUMMER, and COLD WATER from the many it is connected to in the knowledge net; WATER is assumed to sample OCEAN, LAKE, WET, DRINK, AND LIQUID from its neighbors in the net; similarly for the other propositions.

All these propositions, whether derived from the text or from the knowledge base, are interconnected, and the connections may be either positive or negative. For each interconnection I have asigned a value between 1 and -1 on the basis of my intuitions and my knowledge of the world and language. The particular values I have chosen are reasonable, but by no means compelling. For instance, I connected DRINK and WATER positively with a strength of .5, but DRINK and HUNGRY negatively with a strength of -.5. Thus, a connectivity matrix was obtained. The exact numerical values in this matrix are not very important for my example. This connectivity matrix was then repeatedly multiplied with an activation vector to update the activation values of each node in the net, until the pattern of activation stabilizes. This mathematical operation simulates the spread of activation in a neural system.

Figure 8 depicts the resulting pattern of activation graphically: out of context, the sentence "Johnis thirsty" is not overly meaningful; the model picks up mostly on the name *John* because of some strong facilitatory inter-connections, while *thirsty* is deemphasized: the semantic material it activates is inconsistent, and the nodes inhibit each other, thereby taking away activation even from the parent node. sp. 1 The activation pattern for "John took a glass of water" was calculated in the same way, and is also shown in Figure 8. The text propositions end up most strongly activated, plus a few moderately strong inferences, including JOHN-IS-A-NAME, as in the previous example. (But note that this inference is now less strong because of the increased competition in this richer context).

Figures 8 & 9

In Figure 9 the same two sentences are analyzed, but now I assume that the two sentences are read together. Nothing changes for the first sentence, but in accordance with Kintsch & van Dijk (1978), we assume that when the second sentence is read, some propositions from the previous sentence are held over in a short-term buffer to establish the coherence between the two sentences. In Kintsch & van Dijk (1978) the propositions to be held over were selected on the basis of structural criteria; alternatively, the most strongly activated propositions can be retained in the buffer. In the calculations on which Figure 9 is based, I have retained the five most strongly activated propositions. This greatly changes the pattern of activation for the second sentence: The text propositions are still strongly activated, but the rather irrelevant inferences that *John had, and wanted, the glass* are much less prominent, while *John drinks the water* has become dominant! An inference, in this case, a plausible macroproposition, becomes more strongly activated than the actual input.

Obviously, such an example does not prove much: merely that it was possible to arrange things in such a way as to produce an intuitively appealing outcome. But it is a first step, and it is easy to see, how other examples could be generated in which, for instance, a script or frame is inferred, and then used as a basis for the organization of the textbase, as described in van Dijk & Kintsch (1983). Maybe the model which I have sketched here - the connectionist assumptions about knowledge representation, and the process model of knowledge use as it was inferred from psychological laboratory experiments - is a step in the direction of being able to operate with knowledge in truly flexible, context-sensitive, human-like ways.

I am confident that further research will teach us what kind of knowledge representation is most adequate to simulate human knowledge use in discourse comprehension. Whatever this turns out to be, will it also be the most useful knowledge representation in AI? It is quite likely that this would be so, for humans are impressively good at understanding natural language! However, it is by no means necessary that the best human simulation will also be the most powerful one in AI: computers can do certain things much better than humans (e.g. to search huge search spaces), and, as far as artificial intelligence is concerned, we need to exploit these special capacities of computers to obtain systems which in some respects might someday even exceede human abilities.

REFERENCES

Ach, N. (1910) *Über den Willensakt und das Temperament* Leipzig: Quelle und Meyer.

Anderson, J. M. (1983) *The architecture of cognition*, Cambridge, MA.: Harvard Unversity Press.

Bartlett, F. C. (1932) *Remembering*, Cambridge, England: Cambridge University Press.

Bower, G. H., Black, J. B., & Turner, T. J. (1979) Scripts in memory for text. *Cognitive Psychology*, 11, 177-220.

Brachman, R. J. (1979) On the epistomological status of semantic networks. In N. V. Findler (Ed.) *Associative networks: Representation and use of knowledge by computer.* New York: Academic Press.

Brachman, R. J. & Levesque, H. J. (1985) *Readings in knowledge representation.* Los Altos, CA.: Morgan Kaufman.

Clancey, W. J. (1984) Methodology for buidling intelligent tutoring systems. In Kintsch, W., Miller, J. R., & Polson, P. G. (Eds.) *Method and tactics in cognitive science.* Hillsdale, NJ: Erlbaum.

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

van Dijk, T. A. & Kintsch, W. (1983) *Strategies of discourse comprehension* . New York: Academic Press.

Fodor, J. A. (1983) *The modularity of mind.* Cambridge, MA: MIT Press.

Gibbs, R. W. (1984) Literal meaning and psychological theory. *Cognitive Science,* 8, 275-304.

Glucksberg, S., Gildea, P., & Bookian, H. D. (1982) On understanding non-literal speech: Can people ignore metaphors? *Journal of verbal Learning and verbal Behavior*, 21, 85-98.

Head, H. (1920) *Studies in neurology*. Oxford: Oxford University Press.

Kintsch, W. (1974) *The representation of meaning in memory.* Hillsdale, NJ: Erlbaum.

Kintsch, W. (1980) Semantic memory: A tutorial. In R. S. Nickerson

(Ed.) *Attention and performance VIII*. Hillsdale, NJ.: Erlbaum.

Kintsch, W. & van Dijk, T. A. (1978) Towards a model of text comprehension and production. *Psychological Review*, 85, 363-394.

Kintsch, W. & Mross, E. R. (1985) Context effects in word identification. *Journal of Memory and Language*, 24, 336-349.

Kolodner, J. L. (1983) *Retrieval and organizational strategies inconceptual memory: A computer model* Hillsdale, NJ: Erlbaum.

Lewin, K. (1917) Die psychologische Tätigkeit bei der Hemmung von Willensvorgängen und das Grundgesetz der Assoziation. *Zeitschrift für Psychologie*, 77, 212-247.

McKoon, G. & Ratcliff, R. (1986) Inferences about predictable events. *Journal of experimetnal Psychology: Learning, Memory, and Cognition*, 12, 82-91.

McClelland, J. L. & Rumelhart, D. E. (1985) An interactive activation model of context efects in letter perception: Part I. An acount of basic findings. *Psychological Review*, 88, 375-407.

Meyer D. E. & Schvaneveldt, R. W. (1971) Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of experimental Psychology*, 90, 227-234.

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

Newell, A. (1973) Production systems: Models of control structure. In W. Chase (Ed.) *Visual information processing*. New York: Academic Press.

Norman, D. A. & Rumelhart, D. E. (1975) *Explorations in cognition*. San Francisco: Freeman.

Ortony, A. (1979) Beyond literal similarity. *Psychological Review*, 86, 161-180.

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

Rumelhart, D. E. & Ortony, A. (1977) Representation of knowledge. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.) *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Erlbaum.

Schank, R. C. (1972) Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 3, 552-631.

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

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

Selz, O. (1922) *Zur Psychologie des produktiven Denkens und Irrtums*. Bonn: Cohen.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M. & Bienkowsky, M. (1982) Automatic access of the meaning of ambiguous words in context: Some limitations of knowledge based processing. *Cognitive Psychology*, 14, 489-537.

Strube, G. (1984) *Assoziation*. Berlin: Springer.

Swinney, D. A. (1979) Lexical access during sentence comprehension: (Re)Consideration of context effects. *Journal of verbal Learning and verbal Behavior*, 18, 523-534.

Till, R., Mross, E. R. & Kintsch, W. (1986) Priming in a discourse context. ICS Technical Report.

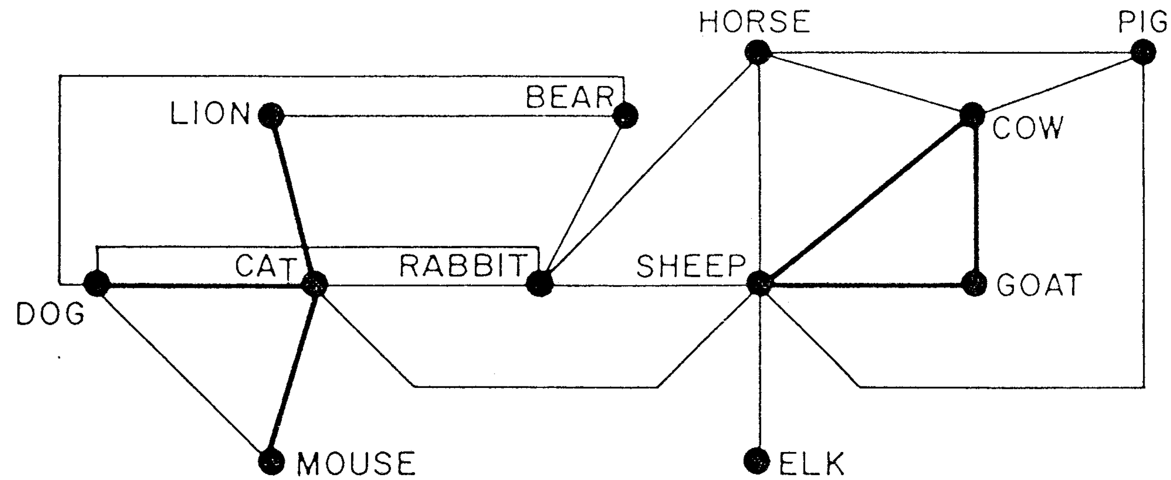
Walker, W. H. & Kintsch, W. (1985) Automatic and strategic aspects of knowledge retrieval. *Cognitive Science*, 9, 261-283.

Waltz, D. L. & Pollack, J. B. (1985) Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, 9, 51-74.

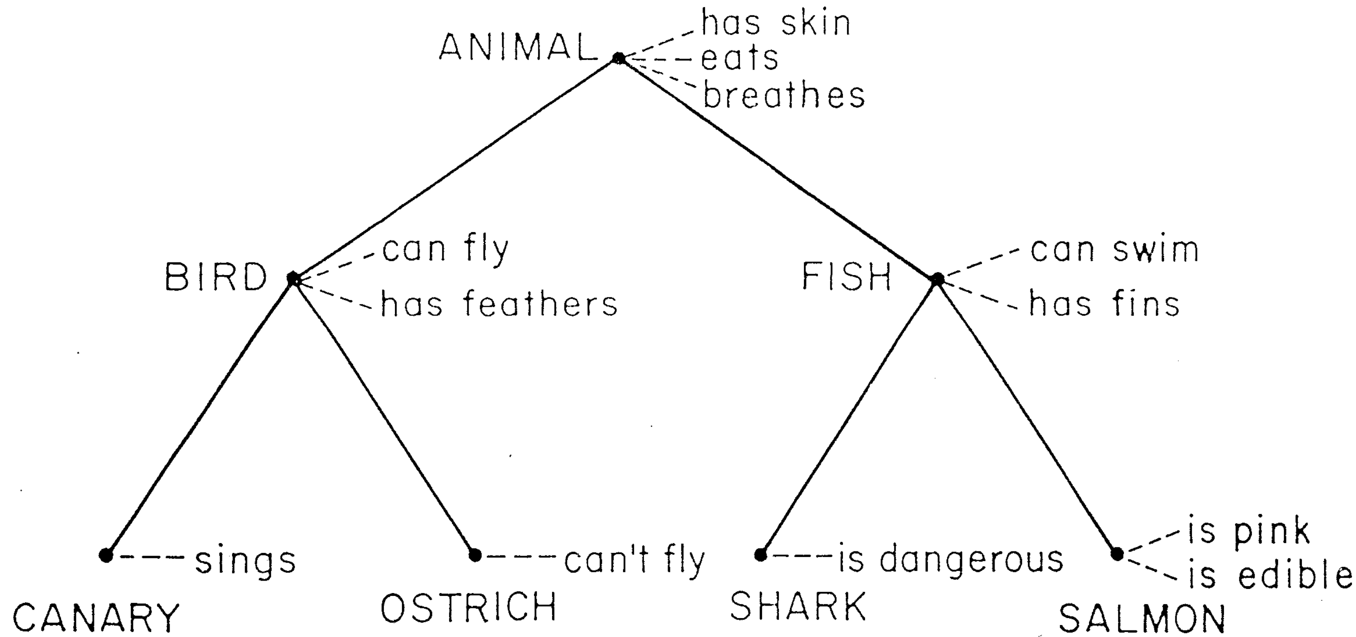
Footnote

This research was supported by National Institute of Mental Health Grant MH 15872.

List of Figures



ASSOCIATIVE NET:
12 MAMMALS
(after Strube '84)



SEMANTIC NET

(after Collins & Quillian, '69)

RESTAURANT FRAME:

Type: (Cafeteria, ...)

:

Location: an ADDRESS

:

Event-Sequence: Eat-at-Restaurant Script

Props: .....

Roles: .....

Event-Sequence:

first: Enter-Restaurant-Script

:

last: Leave-Restaurant-Script

Restaurant Frame
(after Schank & Abelson, '77)

A PRODUCTION SYSTEM FOR PERFORMING ADDITION

P1 . . .

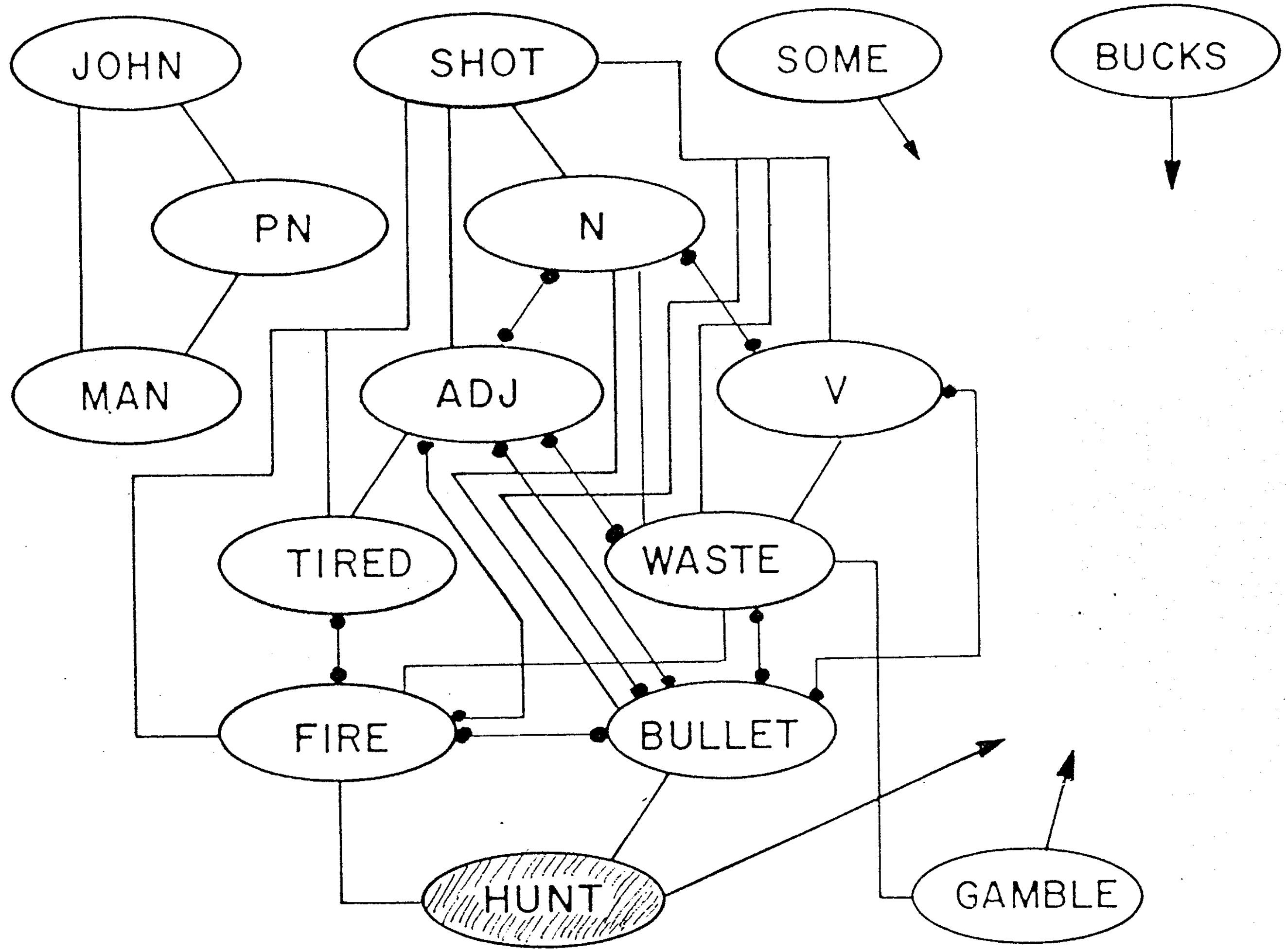
:

P4 IF the goal is to iterate through the columns
of an addition problem
and a column has just been processed
and there is a carry
THEN write out the carry
and POP the goal

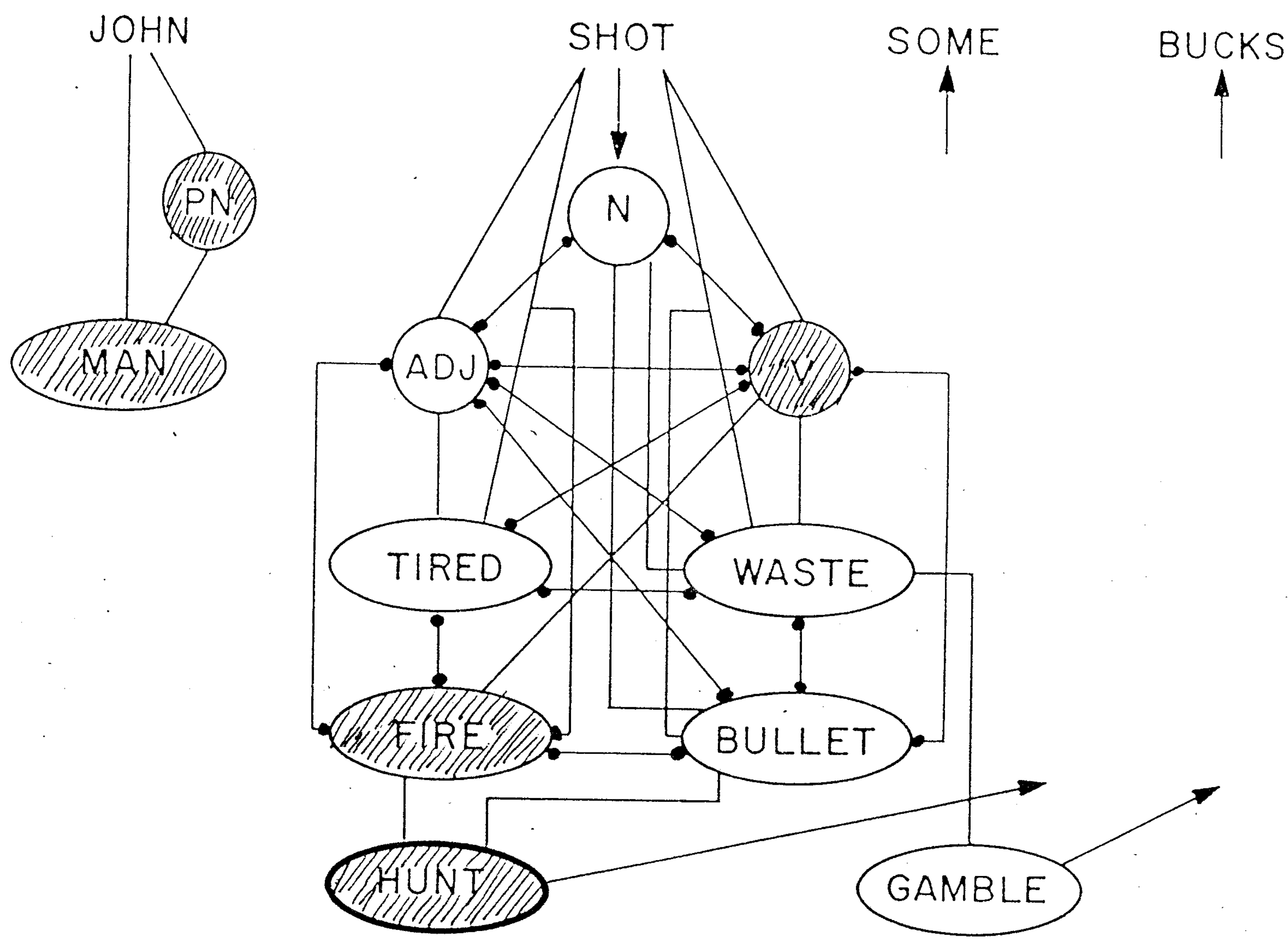
:

P11

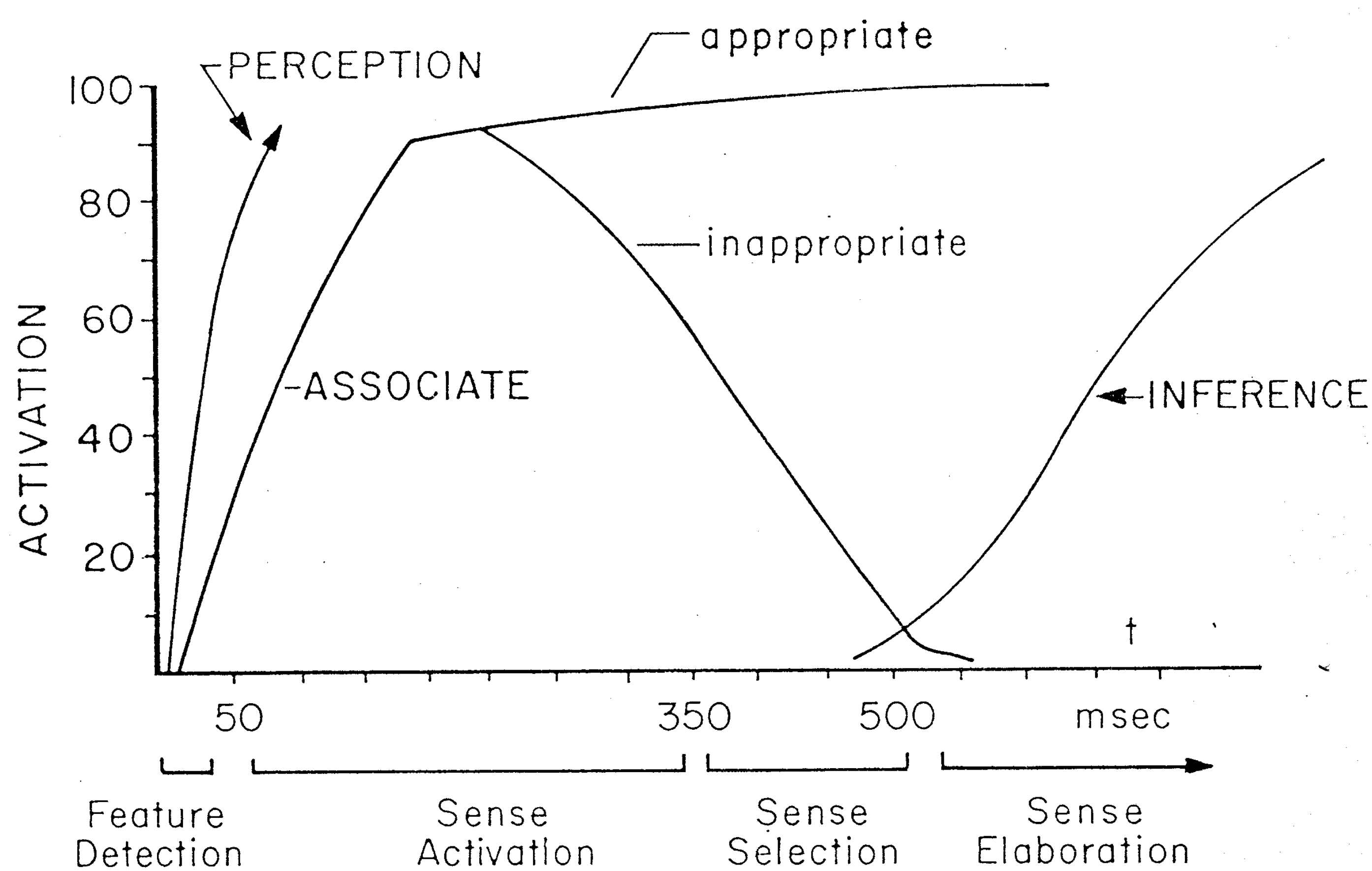
A PRODUCTION SYSTEM (after Anderson, '83)

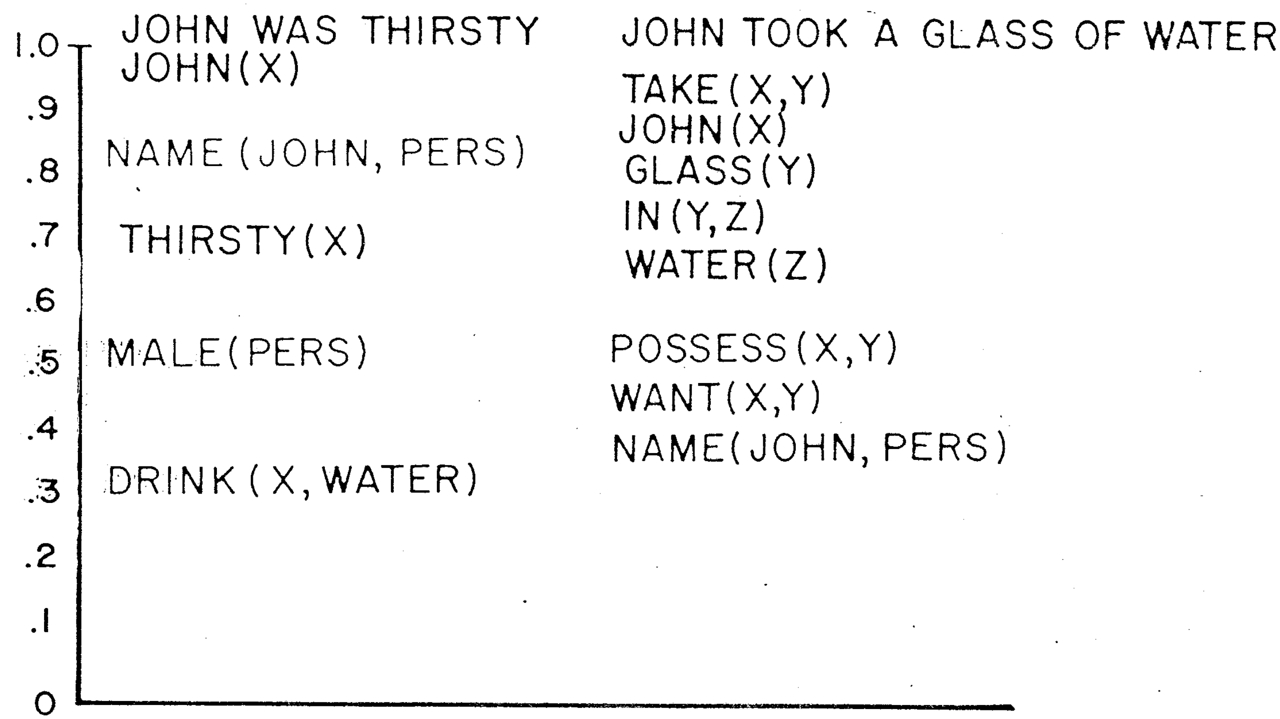


AN ACTIVATED ASSOCIATIVE NET
(after Waltz & Pollack '85)



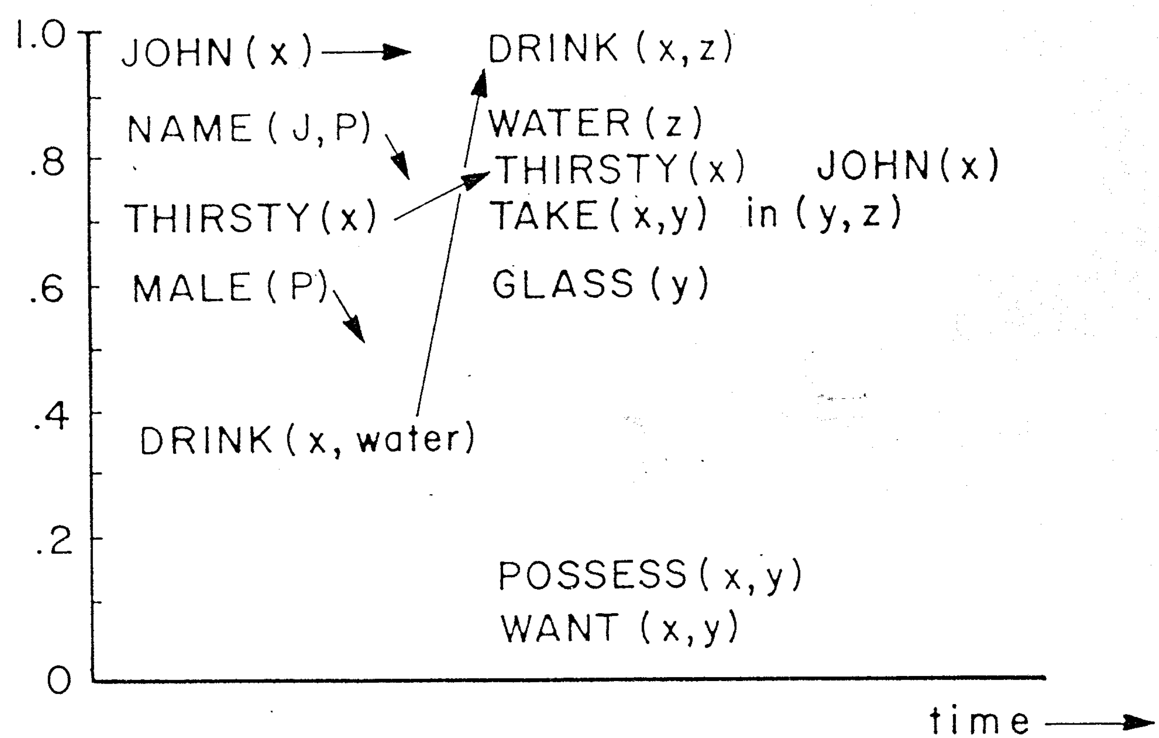
AN ASSOCIATIVE NET
(after Waltz & Pollack '85)


PERCEPTION
appropriate
100
80
60
40
20
ACTIVATION
inappropriate
ASSOCIATE
INFERENCE
t
50
350
500
msec
Feature
Detection
Sense
Activation
Sense
Selection
Sense
Elaboration



ACTIVATED PROPOSITIONS:
TWO SEPARATE SENTENCES

JOHN WAS THIRSTY. HE TOOK
A GLASS OF WATER.



ACTIVATED PROPOSITIONS:
TWO SENTENCES JOINED