

PAUL THAGARD

FRAMES, KNOWLEDGE, AND INFERENCE*

The problem of knowledge representation has been central to recent work in artificial intelligence and cognitive psychology. In AI, researchers have been concerned to find a way of representing knowledge¹ for efficient use by computers. Psychologists have been struck by the difficulty of accounting for the ability of human beings to store and retrieve large amounts of information. In both these fields, emphasis has been placed on new kinds of "knowledge structures," which are more complex than propositions and are intended to provide a more effective means of representing and processing information. The most commonly used terms for such knowledge structures are "frame," "schema," and "script."

This paper is intended to acquaint philosophers with the new approach to knowledge representation, and to provide a preliminary assessment of the epistemological significance of frames and other processing notions. In section 1, I argue for a framework in which, contrary to the views of Frege and Popper, psychological investigations can be highly relevant to epistemology. This clears the way for exposition in 2 of the notion of a frame and its place in AI research. Then in 3, I describe the allied psychological notion of a schema and review the empirical applications which psychologists have found for it. Sections 4 and 5 contain two epistemological applications of the frame/schema notion: the first concerns methodological conservatism arising from the procedural character of frame systems, and the second concerns the significance of default reasoning and the relative unimportance of consistency as a property of knowledge systems.

1.

Epistemology, as I see it, is the theory of the structure and growth of knowledge. Its two central questions concern how knowledge is organized and how it can be increased. The standard philosophical answer to the first question is that knowledge consists of propositions. This goes hand in hand with the typical answer to the second question,

which is that knowledge primarily grows by increasing the number of propositions known through *inference*, wherein a new proposition is inferred from a set of old ones.²

If the cognitive sciences offer a revision of standard views of the structure and growth of knowledge, one would expect those views to have immediate epistemological significance. But consider the following argument, which is akin to ones offered by Frege (1893) and Popper (1972). Epistemology, the argument runs, is as unconcerned with psychology as is logic. Psychology describes what inferences people do make, but logic is concerned with what inferences people *should* make. Similarly, epistemology is the normative theory of objective knowledge, and need not take into account what psychology determines to be the nature of the belief systems in individuals. Propositions, or rather sentences expressing them, can be conclusions of arguments and can be written down in books for public scrutiny. To examine the structure of individual belief systems would be only to encourage a kind of subjectivism which abandons the traditional noble concerns of epistemology – justification and truth – for a vapid relativism.

However, a concern with psychology need not engender epistemological skepticism. Haack (1978) recommends a *weak psychologism*, according to which logic is prescriptive of mental processes. This position is distinguished from both *anti-psychologism*, which is the Frege/Popper view that logic has nothing to do with mental processes, and *strong psychologism*, the view that logic is descriptive as well as prescriptive of mental processes. Weak psychologism uses empirical psychology as a starting point, since it presupposes an empirical account of what mental processes to be prescriptive about, but goes beyond mere description of actual mental processes to consider what sorts of inferential practices are normatively correct (Thagard, 1982a, Goldman 1978). Hence weak psychologism can escape the charge of subjectivism and relativism which is the chief motivation for resistance to admitting the relevance of psychology to epistemology.³

Knowledge should be both private and public, inhabiting the brains of particular thinkers, but also subject to inter-subjective communication and assessment. Weak psychologism aims to capture both these aspects. The real test between weak psychologism and anti-psychologism consists in seeing which framework can develop a comprehensive and rich account of human knowledge. Sections 4 and

5 of this paper are attempts to describe some possible results of the weak psychologicistic research program.⁴

2.

Thus, whether the cognitive sciences have anything to contribute to epistemology is a question to be determined by seeing whether empirical questions help to generate a more powerful epistemological theory. I shall argue that work on knowledge representation in artificial intelligence and cognitive psychology can indeed make such a contribution, by suggesting new issues as well as new approaches to old issues. My discussion of frames and schemas makes no claim to bibliographical completeness; I shall merely refer to those aspects of the voluminous literature in each field which seem most useful in developing the epistemological questions which are the focus of later sections.

Marvin Minsky's (1975) paper on frames has had a major impact on work in artificial intelligence. The strongest constraint on theorizing in AI is the danger of a combinatorial explosion: a procedure which requires exponentially increasing amounts of processing capacity will quickly exhaust the resources of the largest available computers. In the 1960's, researchers in AI often used forms of representation based on first order predicate calculus. Inferences could be made from an initial data base by mechanically deriving theorems from an original set of sentences. The problem, however, is to produce out of the infinite set of possible theorems just those theorems which contain the information needed. A program which required the deduction of all consequences of a set of formulas would encounter a combinatorial explosion. Even without such an unrealistic requirement, enormous difficulties arise in determining just what deductions are to be made. More control and organization of knowledge is needed than is provided by a logistic approach.⁵

Minsky's frame notion was a vague but rich suggestion concerning how knowledge might be organized. He described a frame as data-structure consisting of a network of nodes and relations, used for representing a stereotypical situation. Attached to the frame is information about how to use the frame, what to expect to happen, and what other frames it might be appropriate to move to in certain

circumstances. The frame contains a large body of richly interconnected information about a single topic, organized around typical observations and procedures. Minsky proposes, for example, that when one enters a room, one retrieves from memory a frame which represents a typical room, with walls at right angles, a ceiling, and a floor. Some aspects of the frame are fixed, but others contain *terminals*, which are slots that must be filled in by specific instances of data. These terminals are initially filled with "default" assignments containing information which is supposed until new information displaces them. If a retrieved frame fails to fit reality, i.e., we cannot find terminal assignments which closely match our frame-based expectations, then a new frame is elicited by means of an interframe procedure. For example, if one walks into a hexagonal room, an unusually complex frame must supplant the frame for a rectangular room which we most frequently employ.

But what, the philosopher will naturally demand, *are* frames? Are they really anything different from sets of propositions? Is the frame notion really more than a vague metaphor for organized knowledge? I want to answer yes to both questions, and will defend the answer at two levels. First, it is possible to give a precise account of a frame as a structure in a programming language such as LISP. Second, it is possible to describe empirical consequences of the supposition that human knowledge is organized into frame-like structures (see section 3).

At the first level, we can characterize a frame as a data structure in LISP, the list processing language which is the programming language most commonly used in work in artificial intelligence.⁶ The basic data structure in LISP is a *list*, for example, (FIRST SECOND). More elaborate lists can also be constructed containing lists: ((FIRST (SECOND THIRD))) is a list containing two lists, with one and two elements, respectively. Programs in LISP consist of series of definitions of functions which perform operations on lists. An *association list* is a list of pairs; for example, we could construct an association list of names of people and their ages: ((PAUL 33) (NANCY 9) (ED 35)). It is possible, however, to have the second element in the pair be a list, too, for example, if the association list pairs names of people with names of their children: ((FRED (BILL ALICE)) (JANE (NANCY CARL))). Such a structure is called a *nested* association list. A *frame* is a particular kind of nested association list.

A frame (in Winston and Horn, 1981) has the basic structure:

```

(FRAME-NAME
(SLOT-NAME      (FACET-NAME      (values ...))
                 (FACET-NAME      (values ...))
                 (more facets ...  ))
(SLOT-NAME      (FACET-NAME      (values ...))
                 (more facets ...  ))
(more slots ... ))

```

Each *slot* is an associated pair consisting of a slot-name and an associated list of lists, each of which is an associated pair consisting of a facet-name and a list of values of that facet in that slot. A frame minimally has one slot, one facet and one value.

To be more concrete, let me describe a modest frame system called EAT which I have implemented on the University of Michigan Amdahl. EAT makes inferences and answers questions about Ann Arbor restaurants. Among its frames are ones which contain information about specific restaurants, for example:

```

(MACDONALDS
(A-KIND-OF      (VALUE (RESTAURANT)))
(LOCATION        (VALUE (ANN-ARBOR))))
(SERVING-STYLE (VALUE (FAST-FOOD)))
(FOOD-STYLE    (VALUE (AMERICAN)))
(COST          (VALUE (CHEAP)))
(DECOR         (VALUE (MINIMAL)))
(RATING        (VALUE (FAIR))) )

```

Here the slots contain information about what kind of thing Macdonalds is, its location, and so on. The slots include only one facet, which gives an actual VALUE for the restaurant.

More elaborate frames are required to represent what we know about the general concept of a restaurant:

```

(RESTAURANT
(A-KIND-OF      (VALUE (BUSINESS))))
(LOCATION        (RANGE (ANN-ARBOR YPSILANTI
                       DETROIT))
                (DEFAULT (ANN-ARBOR)))
(SERVING-STYLE (RANGE (FULL-SERVICE
                       FAST-FOOD CAFETERIA))
                (DEFAULT (FULL-SERVICE)))
(FOOD-STYLE    (RANGE (AMERICAN CHINESE)

```

	FRENCH OTHER))
	(DEFAULT (AMERICAN)))
(COST	(RANGE (EXPENSIVE REASONABLE CHEAP))
	(IF-NEEDED (GUESS-COST))
(DECOR	(DEFAULT (REASONABLE)))
	(RANGE (ATTRACTIVE PLUSH MINIMAL))
	(DEFAULT (MINIMAL)))
(SCRIPT	(DEFAULT ((GET 'RESTAURANT' SCRIPT))))
(RATING	(RANGE (EXCELLENT GOOD FAIR POOR))
	(DEFAULT (FAIR)))

Here there are four possible facet-names with four different sorts of value. The A-KIND-OF slot has an *actual* VALUE as did the MACDONALDS frame, but the location slot has two different facets, for RANGE and DEFAULT. The RANGE facet contains a list of values representing the range of possible locations for a restaurant, and the DEFAULT facet has as its value the value which one would suppose to be the location of a given restaurant unless information were provided to the contrary. The other facet-name is IF-NEEDED, in the cost slot, and will be explained below.

The RESTAURANT frame is by no means intended to be a *definition* of the concept of a restaurant. It does not give necessary and sufficient conditions for something being a restaurant. The function of its various slots and facets is to say what is *typical* of restaurants, not what is universally true of them. The DEFAULT facets in particular should not be interpreted as giving universal properties of restaurants; defaults function rather to provide a quick and easy value when precise information is not available. Section 4 contains a discussion of the epistemology of default reasoning.⁷

The point of all this structure is not simply to contain a lot of information, which could be done just as well in a set of sentences, but to enable the information to be economically processed. EAT is able to use the structured information in frames to make inferences about restaurants for which only partial information is given. The A-KIND-OF slot links things and concepts into a hierarchy which makes possible a search procedure which provides answers using a process analogous to deductive inference.⁸ Suppose you ask whether Macdonalds pays

taxes. EAT fails to find the answer in the MACDONALDS frame, so it moves up the hierarchy using the A-KIND-OF slot to the RESTAURANT frame. Again failing to find an answer, it moves to the BUSINESS frame, whose relevant slot is:

(BUSINESS
(TAX-STATUS (VALUE (PAYS-TAXES))))

EAT then returns the answer that Macdonalds pays taxes. A logician would see this as a deductive inference from the propositions that all restaurants are businesses, and all businesses pay taxes, and Macdonalds is a restaurant, but procedurally, EAT uses frames and slots to generate the appropriate answer.

A second procedural resource of frames derives from the presence of *default* values. The slot FOOD-STYLE admits a range of different styles of food, but it contains as a default the value AMERICAN. In the processing system, the effect of this value is that unless there is evidence to the contrary, it will be assumed that a given restaurant serves American style food, producing very economically the answer "American" to a question concerning what kind of food a restaurant serves. Although this may seem inferentially precipitous, it establishes great gains in speed of information processing, and disaster is avoided by procedures which supercede the default value if there is any positive information about what style of food is to be found in a given restaurant.

Defaults and inherited values are relatively immediate and inexpensive ways of filling in slots, requiring little complex processing. In some cases, we need a special procedure to fill in a slot, and the incorporation of *attached procedures* is the third salient feature of the restaurant frame. The slot for COST of restaurant contains not only a range of values and a default value, it includes a procedure for determining the value of the slot if one is needed. The IF-NEEDED value is a specific function, GUESS-COST, which in EAT contains the rules that if the FOOD-STYLE of a particular restaurant is FRENCH, then the value for COST is EXPENSIVE, and if the value for SERVING-STYLE is FAST-FOOD, then the value for COST is CHEAP. Thus, given only the following information about Burger-King, we could nevertheless use EAT to infer that it is cheap:

```
(BURGER-KING
(A-KIND-OF      (VALUE  (RESTAURANT)))
(SERVING-STYLE (VALUE  (FAST-FOOD))))
```

The program uses the A-KIND-OF value to consult the RESTAURANT frame, where in the IF-NEEDED facet of the COST slot, it finds the function GUESS-COST, which calculates using the value for SERVING-STYLE of BURGER-KING that its COST is CHEAP. As with the use of inherited values, the use of the IF-NEEDED slot could be translated into a straight-forward deductive inference, but only at the expense of ignoring the actual procedural nature of the inference. In addition to IF-NEEDED facets, other sorts of procedural attachments are sometimes used, but I shall neglect their operation here.

The use of inherited, default and if-needed values and the sorts of inferences described will be standard to any frame system, but in EAT I have also used the resources of the system to perform other sorts of inferences, more akin to what are usually classed as inductive inferences. The implementations I will now describe are not intended to be optimal ways of performing analogical and other inferences, but are suggestive of how such inferences might operate in a frame system.

Suppose you have only the following information about Burger-King:

```
(BURGER-KING
(FOOD-STYLE    (VALUE  (AMERICAN)))
(DECOR         (VALUE  (MINIMAL)))) )
```

A reasonable way of inferring the cost of this restaurant would be to use an analogical inference, concluding that the cost of Burger-King is the same as the cost of the restaurant most similar to it in other respects. EAT uses a pattern-matching procedure to find which frame of the restaurants in the data base has most in common with the frame for BURGER-KING. In this case, it gets the best match between the slots of BURGER-KING and the slots of MACDONALDS, therefore picks MACDONALDS as the analog of BURGER-KING, and then concludes that, like Macdonalds, Burger-King is cheap.⁹

A similar pattern matching procedure is used by EAT to classify

objects. If the A-KIND-OF slot in an object's frame is not filled, EAT can classify the object by finding the general frame whose slots best match those of the object's frame. For example, the BURGER-KING frame in the last paragraph can be matched against the frames for RESTAURANT, GROCERY-STORE, and BUSINESS. EAT finds that the best match between the slot-names of BURGER-KING and the slot-names of the three general frames is with RESTAURANT, and therefore concludes that Burger-King is a kind of restaurant.¹⁰

After an inference has been made, it would seem natural that an inferrer would store the new information so that if needed in the future it can be simply retrieved rather than have to be inferred again. EAT has this capability: once an inference is made about something using its frame, the frame is altered to include the new information. Thus, the BURGER-KING frame discussed in the past two examples is altered by the addition of new slots and values following the inference.

Because EAT deals only with a very limited domain, it does not have an important general feature which Minsky proposed for frame systems. For most efficient use, a computer's (or a human's) frame system might be organized into a network of sub-systems. The frame system EAT would be a component of a much larger system; such sub-systems of frames might be called *macrostructures* (cf. van Dijk, 1980). A large frame system will need a control structure for dealing with different macrostructures. The existence of macrostructures can have epistemological consequences which are discussed in section 4. Even within a macrostructure such as the one in EAT, there must be procedural devices for determining what sorts of inferences are made when. In its current form, EAT makes inferences in response to specific questions put to it, but a more realistic implementation would allow it to make inferences more autonomously and automatically in response to less specific prompting.

These general procedural capabilities, in addition to the particular ones we saw above (inheritance, defaults, and procedural attachments), should suffice to convince the reader that frames are more than thinly disguised sets of propositions in predicate calculus. Granted, as Hayes (1979) argues, the *content* of a frame can easily be translated into predicate calculus. However, this observation is irrelevant to assessing the significance of frames, since their potential values lie in organizing knowledge for the sake of more effective *processing*. Predicate calculus is not in itself a processing system at all, and the sort of processing

system most naturally attached to it, based on the usual rules of deduction, seems to many researchers in artificial intelligence to be too indiscriminating to serve effectively.

The point requires emphasis. The information in a frame might instead be represented in a set of sentences or a set of formulas in predicate calculus. (The only difficulty would come in representing default values, which would require something like a modal operator “typically *p*”.) But the possible translation does not in the least diminish the significance of frame systems, which is largely *procedural*. The key claim is that information can be processed much more effectively by a frame system than by a system based on predicate calculus. Consider, for example, the inference, described above, that Macdonalds pays taxes. In form more familiar to logicians, this inference could be represented:

All businesses pay taxes
 All restaurants are businesses.
 Therefore, all restaurants pay taxes.

All restaurants pay taxes.
 Macdonalds is a restaurant.
 Therefore, Macdonalds pays taxes.

If the sentences in these syllogisms were formalized, it would be easy to prove using your favorite axiomatic or natural deduction system that the inferences are valid. But that is not the question we are asking, which concerns how the information that Macdonalds pays taxes is to be derived from a data base of facts about restaurants and businesses. If the data base were deductively closed, it would already contain the information that Macdonalds pays taxes, but it is computationally absurd to have a deductively closed and hence infinite data base.

When a question such as “Does Macdonalds pay taxes?” is asked, we need some means for generating an answer. One means might be to take basic facts such as that Macdonalds is a restaurant and match them on the basis of their logical form with deductive rules such as universal instantiation:

All *A* are *B*.
x is *A*.
 Therefore, *x* is *B*.

Undoubtedly such a system *could* be designed. In the 1960's, especially, ingenious systems were designed which used deductive theorem proving based on the method of resolution, wherein a theorem is proved by constructing derivations from the negation of the theorem expressed in conjunctive normal form (see, e.g., Nilsson 1971). Unfortunately, resolution theorem proving becomes impossibly slow when the number of facts in a database becomes more than a few (Barr and Feigenbaum 1981, p. 168). A combinatorial explosion occurs because there is no satisfactory way to prune the number of ways to combine facts to make inferences.

Frame systems avoid this problem by not doing the usual sort of deduction at all. EAT derives the conclusion that Macdonalds pays taxes without explicitly using any deductive rule such as universal instantiation or modus ponens. Instead, it uses the LISP function already mentioned which directly searches the hierarchy of frames established by the A-KIND-OF slots. This has the same result as deductive inference of the traditional sort, but is considerably more economical in respect to processing. Similarly, great gains of efficiency accrue through IF-NEEDED procedural attachments and frame-based mechanisms for analogical inferences.

These general procedural properties of frame systems show that frame systems really are importantly different from the logistic systems more familiar to philosophers. For computational purposes, knowledge organization of the sort found in frame systems seems to be needed. I shall now review some psychological studies which suggest that the human information processing system also employs some of the principles of organization found in frame systems.

3.

If frame systems were only a device to make computers process information more efficiently, they would not necessarily be of any philosophical interest at all: why should an epistemologist be any more concerned with frames than with FORTRAN? The answer is that parallel with the developments about frames in AI have been empirical studies in cognitive psychology which suggest that human thinking employs frame-like structures.¹¹

In cognitive psychology, frames are more commonly referred to as "schemas" or "schemata". Although intense discussion of schemas has

occurred only recently, the notion of a schema can be found in Bartlett (1932), and related notions occur in Head (1926) and Kant (1787). The psychological literature on schemas has become voluminous and diverse, and I shall not attempt to survey it.¹² I will try, however, to give a succinct characterization of schemas, their frame-like properties, and the role they are claimed to play in human cognition.

Whereas frames can be given a relatively exact characterization as structures in LISP, schemas are theoretical psychological entities postulated to explain a variety of observed phenomena of human cognition, so no exact definition is to be expected. Roughly, a *schema* is a large, complex unit of knowledge expressing what is typical of a group of instances. Whereas we had frames for both particular objects and general concepts, schemas, by the very nature of the term, involve some sort of abstraction and generalization. The RESTAURANT frame described above fits this definition, as it is more complex than a proposition and expresses in a nondefinitional way what is typical of restaurants. Schemas are posited to provide the same advantages to human cognition which frames provide to a computer. A set of schemas serves to generate a set of expectations, so that the thinker need not confront external information passively. Incoming information is processed by matching it with existing schemas, which immediately makes possible the utilization of information already acquired. For example, recognizing something as a restaurant produces a set of expectations and possible inferences using procedural mechanisms such as inheritance and defaults. Schemas, like frames, are presumably organized into macrostructures.

A *script* (Schank and Abelson, 1977) is a frequently discussed kind of schema which describes a typical sequence of events. People generally expect events in a restaurant to occur in an order much like that described in the following structure, a frame in EAT:

```
(RESTAURANT-SCRIPT
(FIRST (VALUE (ENTER)))
(THEN 1 (DEFAULT (BE-GREETED)))
(THEN 2 (DEFAULT (BE-SEATED)))
(THEN 3 (DEFAULT (ORDER-DRINKS)))
(THEN 4 (DEFAULT (ORDER-FOOD)))
(THEN 5 (DEFAULT (EAT)))
(THEN 6 (DEFAULT (PAY-BILL)))
(LAST (VALUE (LEAVE))) )
```

You would be very surprised if food were brought to you before you had ordered it, or if you were in a standard full-service restaurant and were asked to pay immediately after ordering. The restaurant script has other scripts embedded within it, and is itself embedded in the general frame or schema for restaurant.

Other examples of areas in which schemas have been thought to be psychologically important include category terms such as *bird*, and social stereotypes such as *Irish*. A schema for bird would differ from a conceptual analysis providing necessary and sufficient conditions: the point of a schema is not to define what properties belong to all and only birds, but to list properties typical of birds. Thus, a schema for bird would likely contain the default information that birds fly. The schema for the stereotypical Irishman might include a default for being Catholic. Obviously, there are potential costs as well as benefits in the use of schemas in reasoning, since error can result from careless use of default values.

The existence of such errors in human reasoning is one of the sorts of evidence which researchers have used to support claims for the psychological reality of schemas. In Brewer and Treyens (1981), subjects were asked to recall what items were in a university office in which they had been kept waiting. They often made the mistake of "recalling" that there were books in the office, even though there were none. Having a schema for university office which included a default for the presence of books would explain their error.

Schemas also affect the speed with which items can be recalled from memory. It should be easier to recall events which instantiate schemas and thereby are more easily accessed. There is a mounting list of results which support the view that the human processing system employs something like schemas to facilitate encoding and recall of information. Schemata have been postulated to have important functions in perception, discourse understanding, learning, remembering, and problem solving.¹³ Although schema theory is by no means universally accepted, the wealth of empirical applications which have been found for it provides at least a presumption that the human information processing system uses framelike structures.

Because schemas are theoretical entities, the relation between schemas and propositions is more difficult to discuss than the relation between frames and sentences in predicate calculus. Schemas are

psychological entities, and there are grave problems in saying just what their structure is, how many of them there are, and how they are used. Of course, there are also great difficulties, both philosophical and psychological, in saying what a proposition is (see, for example, Gale, 1967). In the foreseeable course of empirical inquiry, there does not seem to be an experimental way of determining whether schemas have propositions as constituents. Nevertheless, the schema notion clearly has significance independent of the issue of whether schemas are constituted of propositions, since even if they are, they nevertheless have important emergent procedural properties. The schema, as a whole, plays a role in information processing which cannot be ascribed to unconnected, unorganized sets of propositions. Hence, despite the vagueness of the notion of a schema, there are empirical and conceptual reasons for supposing that schemas go beyond the propositional knowledge structures usually discussed by philosophers.

Although the question of whether schemas have propositional components is not at present empirically answerable, it is not without theoretical interest. We should at least leave open the possibility that the human "language of thought" is organized rather differently from a natural language. After all, other animals besides humans, as well as human infants, perform complex tasks without the benefit of knowledge of natural language (Churchland 1979). Moreover, the basic structure of the frames discussed above is apparently not propositional, even though the slots could be translated into propositions with subject and predicate.

My claim about the possible nonpropositional nature of schemas will be more plausible to psychologists if it is made clear that I am using "proposition" in a sense closer to that common in philosophy than to that common in psychology. The rough meaning of the term in philosophy is that which is expressed by a declarative sentence. Since frames do not contain declarative sentences, slots being verbless, it is tempting to say that they are nonpropositional. In contrast, the term in psychological investigations such as those of John Anderson (1981) usually means a relational structure of nodes in a processing system. The psychological notion of proposition is already a procedural one, whereas the philosophical notion only becomes procedural when some processing system is adduced.

Another argument for not considering schemas to be mere sets of

propositions is based on the nature of propositions. Propositions are often said to be constructed out of *concepts*, but what is a concept? The proposition that restaurants serve food is constructed, perhaps, out of the concepts of a restaurant, serving, and food. But what, in psychological terms, is the concept of a restaurant if not the restaurant *schema*? Thus it seems at least as legitimate to say that the components of propositions are schemas as it does to say that the components of schemas are propositions.

Problems remain, particularly concerning the actual size and structure of schemas. In an artificial frame system, we can have a large number of discreet, complex frames, identifiable by reading the code of the program defining them. In contrast, speculation about the human schema system is open to all sorts of construals of what goes into a schema. Original work on scripts (Schank and Abelson, 1977) assumed that scripts such as the one for eating in a restaurant were fairly rigid structures, playing a unitary role in processing. More recently, empirical and theoretical discussions have suggested that such structures do not exist in memory precompiled, but must be constructed from smaller knowledge structures (Bower, Black & Turner, 1979; Schank, 1979, 1980). Schemas therefore appear to be more fluid and flexible than frames. They can nevertheless play the important roles in cognition described above.¹⁴

Before proceeding to epistemology, I should make a terminological note. In the cognitive science literature, "frame" and "schema" are sometimes used interchangeably. I have adopted the policy of using "schema" for the rather vague psychological notion, reserving "frame" for the much more precise structure in LISP described above. For epistemological analysis, the precise frame notion is the appropriate one, although it must be kept in mind that the epistemological significance of frames depends on their being idealizations of schemas.

4.

If frames do play an important role in cognition they are potentially of epistemological significance. This follows from the position of weak psychologism defended above in Section 1: if logic is to be prescriptive about mental processes, we must at least know what mental processes to be prescriptive about. Construing our knowledge as structured in frame systems has many interesting epistemological implications, but only two

kinds of consequences will be discussed. In this section I shall argue that the construal of knowledge as structured in procedurally oriented frame systems suggests an alternative view of inference and alteration of knowledge which has methodological conservatism as an ineluctable result. Then in section 4 I argue that default reasoning introduces novel epistemological issues, and that potential problems of inconsistency are easily handled in a frame system by procedural means.

According to the familiar philosophical view sketched at the beginning of section 1, knowledge (or, if you prefer, belief) is structured as a set of propositions, which changes by addition to and deletion from the set by means of inference based on other members of the set. In contrast, frame theory suggests that the growth of knowledge in an individual must be understood as involving the incorporation of a complex cognitive system, not simply a set of propositions. Experts and novices may have the same set of explicit beliefs about a domain, i.e., they would assent to the same sentences, yet have very different abilities to handle information in that domain. Chi, Feltovich and Glaser (1981) report important differences in the ways in which physics problems are represented by experienced and beginning physicists. Whereas novices categorize problems by surface structure, experts are able to map a problem into a richer set of problem schemas with attached procedures. Chi et al. found that whereas novices tend to categorize physics problems in terms of types of physical objects, experts tend to classify problems in terms of the major physics principle governing the solution of each problem. Moreover, experts associate these principles with procedural knowledge about their applicability. In frame terminology, we would say that the experts had acquired a frame system which included frames for kinds of problems as well as associate procedures for solving those problems. To know an area of physics, then, is to have incorporated a very complex structure over which one has little or no conscious control.

This has a clear consequence for the question of conservatism in science. Consider the Kuhn-Popper debate over the attitudes of scientists toward their theories or paradigms (Kuhn, 1970a, 1970b; Popper, 1970). Popper maintains that it is a mark of scientific rationality that scientists be readily prepared to abandon their conjectures when they encounter experimental disconfirmation. Kuhn presents an image of a scientist as tenaciously retaining a familiar paradigm in the face of difficulties which are treated as mere puzzles to be solved using the

paradigm. A paradigm is abandoned only when a more powerful paradigm becomes available. As a description of the history of science, Kuhn's account can easily be shown to be superior to Popper's, but what of the normative issue: should scientists be as conservative as Kuhn describes?

If a theory is a complex frame system, then it would seem that we have little choice in the matter. Unlike a propositional conjecture, a frame system is not easily abandoned. If learning a theory requires adoption of a whole network of frames for solving problems and handling the flow of information, then there is very great utility attached to its retention, over which, in any case, we do not have voluntary control. It becomes fully understandable why a scientist would not abandon a fertile way of thinking in the face of a few anomalies, and why it is a prerequisite for abandonment that a new paradigm (frame system) becomes available. Prescription about what revisions a person ought to make in his or her belief system therefore must take into account the difference between accepting or rejecting a proposition, on the one hand, and acquiring or supplanting a frame system on the other.

Accepting a proposition is analogous to adding a slot and value to a frame, and rejecting a proposition corresponds to deleting a slot and value. But adding an entire frame is *not* the same as accepting a set of propositions. A new frame must be integrated with existing knowledge by establishing its place in the processing system: new information is useless until procedural connections with existing frames are in place. Once a complex frame system such as that needed for solving physics problems is functioning as a whole, piecemeal revision becomes problematic.

Sklar (1975, p. 378) has discussed the following principle of methodological conservatism:

If you believe some proposition, on the basis of whatever positive warrant may accrue to it from the evidence, a priori plausibility, and so forth, it is unreasonable to cease to believe the proposition to be true merely because of the existence of, or knowledge of, alternative incompatible hypotheses whose positive warrant is no greater than that of the proposition already believed.

An analogous principle for frame systems might be something like:

If you have a frame system for a domain, it is unreasonable to give it up merely because there is an available plausible alternative frame system.

This principle derives its support from two considerations. First, because *ought* implies *can*, we cannot be enjoined to abandon a frame system if it is not possible for us to do so. We cannot reprogram ourselves: the construction and alteration of frame systems is not within our conscious control. We have no direct access to the structures and processes which guide our thinking. Hence it is not reasonable to expect someone to suspend use of a processing system just because an alternative exists. Second, even if it were possible, it appears that there are empirical grounds for viewing the abandonment of a functioning frame system as undesirable. For by and large it is better to have *some* elaborate system for processing information than none at all. To abandon a frame system for a domain is to be left with no categories for approaching problems in the domain, so that thought grinds to a halt. The rational strategy to adopt when one learns of the existence of a plausible alternative conceptual system to one's own is to attempt to learn the system in much the same way as its proponents have done. This is difficult, because of interference from the frames already in place. However, if one does succeed in developing the alternative frame system in parallel with one's original one, then rational choice of the second over the first becomes possible. Criteria can establish one frame system as a better explanation of the facts (Thagard 1978, forthcoming-ST). A new, whole frame system can then assume the procedural role of the previous system. But it is neither possible nor desirable to abandon an old way of understanding a domain until an alternative is in place.

I am not saying that the two competing frame systems need be so incommensurable that no overlap exists. My point is just that for a processing system to be effective it must be systematically interconnected by procedural mechanisms, in this case for interrelating frames. A modest holism, holding that piecemeal revision of isolated frames is difficult, immediately follows. Although it might be possible to revise a knowledge system by altering particular slots, significant revisions, affecting whole frames, will necessarily affect other frames because of essential procedural interconnections. To take a trivial example, if you altered the frame system EAT by removing the RESTAURANT frame, the hierarchical connection between MACDONALDS and BUSINESS would be broken, and important inferences would be blocked.

The conservatism and holism which follow from the frame theoretic

approach can be compared to similar views of W. V. O. Quine. Quine (1953, 1960; Quine and Ullian 1970) has suggested several powerful metaphors for the structure of knowledge. Science is a web of belief, a connected fabric of sentences which face the tribunal of sense experience collectively, all susceptible to revision and adjustment like the planks of a boat at sea. Part of the power of these metaphors is that they help lead us away from the standard view that knowledge consists in discrete sentences, confirmed and contentful in isolation. Rather, knowledge forms an interconnected whole. But Quine's metaphors do not take us far enough. How is a set of sentences connected by more than deductive relations? How do we juggle an interconnected set of sentences so as to come up with the best total account? Frame systems offer a much richer description of the interconnections of what we know, a description which follows the basic insight of Quine's holism that we do not infer sentences in isolation, but which moreover organizes our knowledge topically and dynamically in a psychologically realistic way. Richness comes through adoption of a procedural viewpoint and abandonment of the fabric of sentences in favor of more complex cognitive structures and operations. The web of belief does not consist of beliefs, but of frames and procedures for using them.

5.

Let us now look at those procedures in finer detail. In this section I want to discuss the epistemology of default reasoning and a procedural approach to the inconsistencies which can arise from default reasoning. In such reasoning, we infer that an object has a certain property in the absence of any information to the contrary. EAT, for example, infers that a restaurant is in Ann Arbor unless it is specifically told that the restaurant is located elsewhere. From a normative standpoint, the use of default values seems at first glance outrageous, since it will not infrequently lead the inferrer to a false conclusion, as when subjects in the office experiment described above reported the presence of non-existent books. Yet from a procedural standpoint, default values have such utility that their proneness to error does not warrant their proscription. Many processing tasks are done under time pressure and it can often be desirable to have a quick, approximate answer to a question than to wait a long time for a more reliable one. Accepting

default values also leaves one time to attend to more important matters.

Designing an inference system such as EAT raises normative questions quite unlike those usually discussed by epistemologists, since one has to decide what sort of priority is to be given to different sorts of inference. For example, the system can be set up to employ default values at an early stage in generating an answer to a question, or to defer the use of default reasoning until all other kinds of reasoning, using inheritance, procedural attachment, and analogy, have been performed. It is an interesting empirical question at what point in the process of inference human beings actually employ default values; and it is a challenging normative question at what levels of processing the different sorts of reasoning *should* be applied. To answer this question, we need more information about what revisions in order of processing we are capable of, and about the efficacy of different processing strategies in different circumstances.

Default reasoning is superseded once an actual value for a slot is provided. Accordingly, such reasoning is said to be *nonmonotonic*, meaning that addition of new information does not simply add to the total of existing conclusions, but requires the revision of previous conclusions. There is much discussion of nonmonotonic reasoning in AI today,¹⁶ but the nonmonotonicity of inductive reasoning, as opposed to theorem derivation in deductive systems, is no news to philosophers. To take just one well-known case, a theory may be judged to be acceptable on the basis of existing data, but the addition of new negative data may necessitate the rejection of the theory.

What would happen if default reasoning gave one answer to a question while other kinds of reasoning produced a contradictory answer? That is, what happens if the frame system is inconsistent? The question is of general interest, since it has been argued (Campbell 1981, Routley 1979) that inconsistency in our knowledge system may be natural and unavoidable. I shall now argue that a procedural perspective implies that we need not be terribly worried about inconsistencies.

On the traditional philosophical view, logical contradictions are disastrous because the presence of both *p* and *not-p* in a set of propositions would entail that the set contained every proposition, since any proposition can be derived deductively from a contradiction. To deal with contradictions logistically, we need to take a logically radical step such as that recommended by Routley (1979), who advocates a nonstandard "dialectical" logic based on relevance logic, which does

not permit the derivation of just any proposition from a contradiction. Such radical revision of our logical principles is unnecessary, since contradictions can be dealt with procedurally. Instead of restricting what *may* be inferred, which is what revision of basic logic does, we instead make procedural restrictions on what as a matter of fact does get inferred. Such restrictions follow naturally from the point discussed above that the danger of combinatorial explosion precludes deductive closure.

Any real processing system with limited resources must be judicious in what inferences it makes, and it is easy to see how a frame system could handle the presence of a contradiction. First, the effects of a contradiction could be localized within a particular macrostructure. The modular character of cognitive processes allows you to have a serious contradiction in your political beliefs, for example, while having no difficulty in dealing with grocery shopping. Your RESTAURANT macrostructure might contain the information that all fast-food restaurants are cheap, that Saco-Taco is a fast-food restaurant, but that Saco-Taco is expensive, without this contradictory information having an adverse effect on the system in general. Inferences need not resonate throughout the processing system, and indeed it would be computationally disastrous if they did so. Second, even within a particular macrostructure, what is believed need not be deductively closed. Rules of inference tell us only what we *may* infer, not what we *must* infer (Thagard, 1979). A frame based processing system could easily have a mechanism for suspending the production of new inferences when a contradiction occurs. Thus, the presence of a contradiction need not be epistemologically disastrous. That is, with appropriate procedural constraints in place, it is quite possible to have a contradictory set of beliefs without destroying the general efficacy of the system.¹⁷

This is not to say that contradictions are especially valuable. The significance of inconsistencies is dynamic:¹⁸ discovering that we believe both *p* and *not-p* is not an occasion for going on an inferring spree or for concluding that we are irrational, but it does signal that critical revision is in order. By critical revision, I do not mean merely adjustment in our set of propositions believed and not believed, but the more fundamental task of frame system adjustment. As we saw above, this is not a voluntary act: we cannot intentionally adjust a frame system any more than we voluntarily decide what to believe (Goldman 1978). But we can pursue information and investigation which will have the

desired effect of revision of the system. Revision may be as simple as replacement of a single slot in a single frame, or, as in the case of theories, it may involve reconstruction of entire macrostructures.

6.

To many epistemologists, it might seem that the present emphasis on procedural matters has abandoned traditional concerns with justification and truth. It might be charged that frames are relevant only to systems design, not to the theory of knowledge. I would reply that if epistemology is construed, as I urged earlier, as the theory of the structure and *growth* of human knowledge, psychological and procedural questions are unavoidable. In Thagard (1982a) I develop a methodology for determining the relevance of descriptive, psychological matters to normative, logical matters. This methodology is an alternative both to a priori, foundationalist approaches to logic, and to approaches which see the justification of logical principles to derive from some kind of reflective equilibrium between inferential principles and inferential practice (Goodman, 1965; Stich and Nisbett 1980). A logical principle is justified, on my account, if it is part of the best available inferential system, where an inferential system includes in addition to the set of normative principles, a set of descriptions of general inferential practice, a set of inferential goals, and a set of background psychological theories about the capacities of the inferer. One of the criteria for selecting the best inference system is how successfully a set of principles enables an inferer to meet his or her inferential goals given his or her capacities. If, *ex hypothesi*, the human information processing system is akin to a frame system, then the capacities and limitations which accompany such systems must be taken into account when judging what inferential principles to adopt. Questions of justification therefore *can* be raised from a procedural perspective, such as that provided by frame theory.

This paper is only a brief introduction to the epistemology of frame systems (see further Thagard forthcoming-ST). I have tried to characterize the notions of a frame and a frame system as clearly as possible, and I have described implications of frame theory for two important epistemological questions. Of course, it could turn out as an empirical matter that the human information processing system is not a frame system. At present, however, frame-like knowledge representations

provide arguably the most attractive program for work in cognitive psychology and artificial intelligence. I have tried to show that philosophers concerned with epistemology and logic may also gain from concern with knowledge structures.

NOTES

* For valuable comments and suggestions, I am grateful to Nancy Cantor, Mary Gick, Mark Golden, Alvin Goldman, Daniel Hausman, John Holland, Keith Holyoak, John McCumber, William Merrill, Marvin Minsky, Richard Nisbett, Michael Shafto, Luc Steels, Stephen White, and Jan Zwicky. This research has been aided by grants from the National Science Foundation (SES 81-07556), the University of Michigan Center for Cognitive Science, the University of Michigan Rackham School of Graduate Studies, and the University of Michigan-Dearborn.

¹ This AI use of "knowledge" will seem odd to philosophers for whom questions of knowledge are intimately connected with truth and justification. It would be closer to standard philosophical usage to say that AI is concerned with the representation of *belief*. However, this is also misleading since beliefs are usually taken to be atomistic and propositional, and it is a matter of debate whether our representational scheme should be. I will therefore stick with the AI terms "knowledge structure" and "knowledge representation", noting that these terms do not assume that the representations in question are veridical or justified.

² This is obviously a caricature of what I see as the dominant view, and does not take into account such conflicting views as those of Quineans who eschew talk of propositions in favor of sentences, or Bayesians and Popperians who resist the suggestion of the *acceptance* of propositions.

³ For further arguments, see Haack (1975).

⁴ Kindred attempts include the naturalistic epistemology of Quine (1968), the evolutionary epistemology of Campbell (1974), and the epistemics of Goldman (1978).

⁵ But there are dissenters to this claim. See for example Nilsson (1983).

⁶ My account of frames is based on Winston and Horn (1981), chap. 22, which in turn is based on the frame representation language FRL: see Roberts and Goldstein (1977) and Goldstein and Roberts (1979). Other useful discussions include Charniak (1977), Bobrow and Winograd (1977, 1979), various essays in Bobrow and Collins (1975), Schank and Abelson (1977), Moore and Newell (1973), and, of course, Minsky (1975). It should be noted that philosophical antecedents of the frame notion can be found in such thinkers as Kant, Hegel, Husserl, Heidegger, Wittgenstein, and Kuhn.

⁷ The presence of defaults makes possible a kind of concept formation different from the standard methods of definition and abstraction from experience. For a frame-based account of conceptual combination, see Thagard (1984).

⁸ Frame systems share this feature with semantic networks; see, for example, Norman and Rumelhart, 1975, and Anderson and Bower, 1973.

⁹ For much more sophisticated discussion of the uses of frames in analogical reasoning, see Winston (1978, 1980) and Gick and Holyoak (1980).

¹⁰ EXPLORE, a program which uses a frame representation similar to that in EAT,

performs more sophisticated kinds of inference including inductive generalization (Thagard forthcoming-RV). Thagard and Nisbett (1982) argue that confirmation of generalizations by their instances is in part a function of background knowledge about the variability of the kinds of things concerned. EXPLORE uses A-KIND-OF and INSTANCES slots to make variability judgements which are critical for generalization.

¹¹ There are also analogous notions in linguistics: see Fillmore (1975) and Chafe (1976). Concern with knowledge structures is one of the most important strands uniting the new interdisciplinary field of "Cognitive Science". For a discussion of the relevance of artificial intelligence and psychology to philosophy of science see Thagard (1983).

¹² Useful surveys include Rumelhart (1980), Anderson (1980), chap. 5; and Hastie (1981).

¹³ In addition to references already cited, see: Bower, Black, and Turner (1979); Thorndyke and Hayes-Roth (1979); Cantor and Mischel (1979); Lichtenstein and Brewer (1980); Abelson (1981); Chi et al. (1981). This list makes no pretense of comprehensiveness.

¹⁴ For a notion of schema which combines the advantages of frames with the flexibility found in rule-based systems, see Holland, Holyoak, Nisbett, and Thagard (forthcoming).

¹⁵ See also Thagard (forthcoming-ST).

¹⁶ See the articles in the special issue on nonmonotonic reasoning in *Artificial Intelligence*, **13**, 1980.

¹⁷ For arguments that we should sometimes judge people to be systematically inconsistent, see Thagard and Nisbett (1983).

¹⁸ Similarly, Hegel sometimes explicitly embraced logical contradictions, but his primary concern was the process by which knowledge develops. Routley's "dialectical" logic fails to capture Hegel's complex dynamic notion of negation. See Thagard (1982b).

REFERENCES

- Abelson, R.: 1981, 'The Psychological Status of the Script Concept', *American Psychologist* **36**, 715-729.
- Anderson, J. R.: 1980, *Cognitive Psychology and its Implications*, Freeman, San Francisco.
- Anderson, J. R.: 1981, 'Concepts, Propositions, and Schemata: What are the Cognitive Units?', in J. Flowers, (ed.), *Nebraska Symposium on Motivation*, University of Nebraska Press, Lincoln, vol. 28, pp. 121-162.
- Anderson, J. R. and Bower, G.: 1973, *Human Associative Memory*, V. H. Winston, Washington.
- Anderson, R. C., Spiro, R. J. and Montague, W. E. (eds.): 1977, *Schooling and the Acquisition of Knowledge*, Lawrence Erlbaum, Hillsdale.
- Barr, A. and Feigenbaum, E.: 1981, *The Handbook of Artificial Intelligence*, vol. I, William Kaufmann, Los Altos, CA.
- Bartlett, F. C.: 1932, *Remembering*, Cambridge University Press, Cambridge.
- Bobrow, D. G. and Collins, A. (eds.): 1975, *Representation and Understanding*, Academic Press, New York.
- Bobrow, D. G. and Winograd, T.: 1977, 'An Overview of KRL, a Knowledge Representation Language', *Cognitive Science* **1**, 3-46.

- Bobrow, D. G. and Winograd, T.: 1979, 'KRL: Another Perspective', *Cognitive Science* **3**, 2-42.
- Bower, G. Black, J. and Turner, T.: 1979, 'Scripts in Memory for Texts', *Cognitive Psychology* **11**, 177-220.
- Brachman, R. and Smith, B. (eds.): 1980, Special Issue on Knowledge Representation, *ACM SIGART Newsletter*, no. 170.
- Brewer, W. and Treyens, J.: 1981, 'Role of Schemata in Memory for Places', *Cognitive Psychology* **13**, 207-230.
- Campbell, D.: 1974, 'Evolutionary Epistemology', in P. Schillp (ed.), *The Philosophy of Karl Popper*, Open Court, La Salle, Ill., pp. 413-463.
- Campbell, R.: 1981, 'Can Inconsistency be reasonable?' *Canadian Journal of Philosophy* **11**, 245-270.
- Cantor, N. and Mischel W.: 1977, 'Traits as Prototypes: Effects on Recognition Memory,' *Journal of Personality and Social Psychology* **35**, 38-48.
- Chafe, W.: 1976, 'Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View', in C. N. Li (ed.), *Subject and Topic*, Academic Press, New York.
- Charniak, E.: 1977, 'A Framed Painting', *Cognitive Science* **1**, 355-394.
- Chi, M., Feltovich, P. and Glaser, R.: 1981, 'Categorization and Representation of Physics Problems by Experts and Novices', *Cognitive Science* **5**, 121-152.
- Churchland, Paul: 1979, *Scientific Realism and the Plasticity of Mind*, Cambridge University Press, Cambridge.
- Dresher, E. and Hornstein, N.: 1976, 'On Some Supposed Contributions to the Scientific Study of Language,' *Cognition* **4**, 321-398.
- Fillmore, C.: 1975, 'An Alternative to Checklist Theories of Meaning', *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, 1975, **1**, 123-131.
- Frege, G.: 1893, *The Basic Laws of Arithmetic*, M. Furth, (trans), Berkeley, University of California Press, 1964.
- Friedman, A.: 1979, 'Framing Pictures: The Role of Knowledge in Automated Encoding and Memory for Gist', *Journal of Experimental Psychology: General* **108**, 316-355.
- Gale, R.: 1967, 'Propositions, Judgments, Sentences, and Statements', *Encyclopedia of Philosophy*, vol. 6, Macmillan, New York, pp. 494-505.
- Gick, M. and Holyoak, K.: 1980, 'Analogical Problem Solving', *Cognitive Psychology* **12**, 306-355.
- Goldman, A. I.: 1978, 'Epistemics: The Regulative Theory of Cognition', *Journal of Philosophy* **75**, 509-523.
- Goldstein, I. P. and Roberts, R. B.: 1979, 'NUDGE, a Knowledge Based Scheduling Program', in D. Metzger (ed.), *Frame Conceptions and Text Understanding*, Walter de Gruyter, Berlin, pp. 26-45.
- Goodman, N.: 1965, *Fact, Fiction and Forecast*, 2nd edition, Bobbs-Merrill, Indianapolis.
- Haack, S.: 1975, 'The Relevance of Psychology to Epistemology', *Metaphilosophy* **6**, 161-176.
- Haack, S.: 1978, *Philosophy of Logics*, Cambridge University Press, Cambridge.
- Hastie, R.: 1981, 'Schematic Principles in Human Memory', in E. T. Higgins, C. P. Herman and M. P. Zanna (eds.), *Social Cognition: The Ontario Symposium on Personality and Social Psychology*. Lawrence Erlbaum Associates, Hillsdale, N. J..
- Hayes, P. J.: 1979, 'The Logic of Frames', in D. Metzger (ed.), *Frame Conceptions and Text*

- Understanding*, Walter de Gruyter, Berlin, pp. 46–61.
- Head, H.: 1926, *Aphasia and Kindred Disorders of Speech*, Cambridge University Press, Cambridge.
- Holland, J., Holyoak, K., Nisbett, R., and Thagard, P. (forthcoming), *Induction: Processes of Inference, Learning & Discovery*.
- Kant, I.: 1787, *Critique of Pure Reason*, 2nd ed., N. Kemp Smith, (trans.), London, MacMillan, 1963.
- Kintsch, W. and Van Dijk, T. A.: 1978, 'Toward a Model of Text Comprehension and Production', *Psychological Review* **85**, 363–394.
- Kuhn, T.: 1970a, 'Logic of Discovery or Psychology of Research', in I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*, Cambridge University Press, Cambridge, pp. 1–24.
- Kuhn, T.: 1970b, *Structure of Scientific Revolutions*, 2nd ed., University of Chicago Press, Chicago.
- Lichtenstein, E. and Brewer, W.: 1980, 'Memory for Goal-Directed Events', *Cognitive Psychology* **12**, 412–445.
- Metzing, D. (ed.): 1980, *Frame Conceptions and Text Understanding*, Walter de Gruyter, Berlin.
- Minsky, M.: 1975, 'A Framework for Representing Knowledge', in P. H. Winston (ed.), *The Psychology of Computer Vision*, McGraw Hill, New York, pp. 211–277.
- Minsky, M.: 1979, 'The Society Theory of Thinking', in P. Winston and R. Brown (eds.), *Artificial Intelligence, an MIT Perspective*, MIT Press, Cambridge, pp. 421–450.
- Moore, J. and Newell, A.: 1973, 'How can MERLIN understand?' in L. W. Gregg (ed.), *Knowledge and Cognition*, Erlbaum, Potomac, MD.
- Nilsson, N. J.: 1971, *Problem Solving Methods in Artificial Intelligence*, McGraw-Hill, New York.
- Nilsson, N.: 1983, 'Artificial Intelligence Prepares for 2001', *The AI Magazine* **4**, 7–14.
- Norman, D. and Rumelhart, D.: 1975, *Explorations in Cognition*, Freeman, San Francisco.
- Popper, K.: 1970, 'Normal Science and its Dangers', in I. Lakatos and A. Musgrave, (eds.) *Criticism and the Growth of Knowledge*, Cambridge University Press, Cambridge, pp. 51–58.
- Popper, K.: 1972, *Objective Knowledge*, Oxford University Press, Oxford.
- Quine, W. V. O.: 1953, *From a Logical Point of View*, 2nd ed. Harper Torchbooks, New York.
- Quine, W. V. O.: 1960, *Word and Object*, MIT Press, Cambridge, Mass.
- Quine, W. V. O.: 1968, *Ontological Relativity and Other Essays*, Columbia University Press, New York.
- Quine, W. V. O., and Ullian, J.: 1970, *The Web of Belief*, Random House, New York.
- Roberts, R. B. and Goldstein, I. P.: 1977, 'The FRL Primer', MIT AI Laboratory Memo 408.
- Routley, R.: 1979, 'Dialectical Logic, Semantics, and Metamathematics', *Erkenntnis* **14**, 301–331.
- Rumelhart, D. E. and Ortony, A.: 1977, 'The Representation of Knowledge in Memory', in R. C. Anderson, R. J. Spiro and W. E. Montague (eds.), *Schooling and the Acquisition of Knowledge*, Erlbaum, pp. 99–135, NY.

- Rumelhart, D.: 1980, 'Schemata' The Building Blocks of Cognition', in R. Spiro, B. Bruce and W. Brewer (eds.), *Theoretical Issues in Reading Comprehension*, Erlbaum, Hillsdale, N.J.
- Schank, R.: 1979, 'Reminding and Memory Organization: An Introduction to MOPs', Yale Computer Science Report 170.
- Schank, R.: 1980, 'Language and Memory', *Cognitive Science* 4, 243-284.
- Schank, R. and Abelson, R.: 1977, *Scripts, Plans, Goals, and Understanding*, Lawrence Erlbaum, Hillsdale, N.J.
- Sklar, L.: 1975, 'Methodological Conservatism', *Philosophical Review* 84, 374-400.
- Stich, S. and Nisbett, R.: 1980, 'Justification and the Psychology of Reasoning', *Philosophy of Science* 47, 188-202.
- Thagard, P.: 1978, 'The Best Explanation: Criteria for Theory Choice', *Journal of Philosophy* 75, 76-92.
- Thagard, P.: 1979, 'In Defense of Deductive Inference', *Australasian Journal of Philosophy* 57, 274-279.
- Thagard, P.: 1982a, 'From the Descriptive to the Normative in Psychology and Logic', *Philosophy of Science* 49, 24-42.
- Thagard, P.: 1982b, 'Hegel, Science, and Set theory', *Erkenntnis* 18, 397-410.
- Thagard, P.: 1983, 'Artificial Intelligence, Psychology, and the Philosophy of Discovery', in P. Asquith and T. Nickles, (eds.), *PSA 1982*, vol. 2, Philosophy of Science Association, East Lansing, MI, pp. 166-175.
- Thagard, P.: 1984, 'Conceptual Combination and Scientific Discovery', in P. Asquith and P. Kitcher, (eds.), *PSA 1984*, vol. 1, Philosophy of Science Association, East Lansing, MI, pp. 3-12.
- Thagard, P. (forthcoming-RV), 'Representations of Variability'.
- Thagard, P. (forthcoming-ST), *Scientific Thinking*.
- Thagard, P. and Nisbett, R. E.: 1982, 'Variability and Confirmation', *Philosophical Studies* 42, 379-394.
- Thagard, P., and Nisbett, R.: 1983, 'Rationality and Charity', *Philosophy of Science* 50, 250-267.
- Thorndyke, P. and Hayes-Roth, B.: 1979, 'The Use of Schemata in the Acquisition and Transfer of Knowledge', *Cognitive Psychology* 11, 82-106.
- van Dijk, T.: 1980, *Macrostructures*, Erlbaum, Hillsdale, N.J.,
- Winograd, T.: 1977, 'A Framework for Understanding Discourse', in M. A. Just and P. A. Carpenter (eds.), *Cognitive Processes in Comprehension*, Erlbaum, Hillsdale, N.J., pp. 63-88.
- Winston, P.: 1978, 'Learning by Creating and Justifying Transfer Frames', *Artificial Intelligence* 10, 147-172.
- Winston, P.: 1980, 'Learning and Reasoning by Analogy', *Communications of the ACM* 23, 689-703.
- Winston, P. and Horn, B.: 1981, *LISP*, Addison-Wesley, Reading, Mass.

Program in Cognitive Science
 Perry Building
 University of Michigan - Ann Arbor
 Ann Arbor, Michigan 48104
 U.S.A.