An examination of the nature of memory reveals that the representation of knowledge cannot be separated from the uses of knowledge. The answering of questions is not a simple retrieval and response of stored information; rather, the process is embedded in a general structural framework containing knowledge of the questioner, the question, and the world around. The teaching of knowledge requires an interactive process based on the knowledge the other person holds or lacks. A general formal structure for representing semantic information is proposed with examples of network structures for encoding general and specific knowledge. The structure is being tested by simulation on a digital computer. The result of these investigations is the realization that there is much more to the memory process than has heretofore been described in our theories. (Author/WH)
MEMORY, KNOWLEDGE, AND THE ANSWERING OF QUESTIONS

Donald A. Norman

CENTER FOR HUMAN INFORMATION PROCESSING

University of California, San Diego
La Jolla, California 92037

Chip 25
May, 1972
Memory, Knowledge, and the Answering of Questions

Donald A. Norman

University of California, San Diego

Paper presented at the
Loyola Symposium on Cognitive Psychology

(Chicago, 1972)
Memory, Knowledge, and the Answering of Questions

Donald A. Norman
University of California, San Diego

The Answering of Questions

How do people answer questions? At first, the process appears reasonably simple. A person is asked a question, he retrieves the relevant information from his memory, and then he responds with the appropriate answer. According to this notion, the traditional psychological studies of memory should tell us something about the way that knowledge is used to answer questions. This is not so; there is much more to answering questions than simply retrieving something from memory. For one thing, the question may be phrased in ways that differ from the format of the storage used for the relevant information. For another the answer may have to be derived from the logical consequences of the information that is known about the question-ed topic combined with the general knowledge that has been acquired about the world and its physical and logical laws of operation. And then, even once the relevant information has been collected together, the answer still cannot be given directly, for it is also necessary to consider why the question was asked to determine what the answer should be. As a result, it often happens that the best way to answer a question is to respond by asking a question. And finally, suppose the information relevant to the ques-
tion is not available, what kind of an answer should be given then? Thus, the study of the way that people answer questions leads to a study of memory, thought and understanding. A wide variety of cognitive capabilities are involved in dealing with this deceptively simple part of our everyday activities.

In this paper, I concentrate my discussion upon the answering of simple questions, but even this restriction will not greatly reduce the range of topics that must be covered. Memory, thought, and understanding seem all intertwined in that vast morass known as "cognitive processing." To study even one small aspect of human thought appears not to be possible, for to study one topic leads immediately to a study of all the others. This exercise has value in that it forces the examination of a large series of important issues of human performance. What starts off as a simple question about one aspect of human performance leads to the development of a theory of cognition, including a theory for the representation of knowledge in human memory, a theory about the use of cognitive structures, a consideration of the knowledge one person must have of the knowledge and motivations of another, and even a discussion of learning and teaching --- the first steps towards a theory of instruction.

The simplest type of question that can be asked of a person is to ask whether he is familiar with a fact. This question is related to the tests in the traditional psychological studies of human memory: a list of items is presented to a subject and then, at some later time, he is asked what he can recall or recognize. These studies have yielded much information about the overall structure of the information processing system: about short-term and long-term memory, about phonemic and semantic encoding, about organization, and about simple decision strategies. But it is not with this level of question that I will be concerned. These are not the types of questions
that characterize our daily, normal interactions.

The paradigm question, the one that started my interest is this:

What is the telephone number of Charles Dickens, the novelist? (1)

According to what is usually assumed about the retrieval of information, this question should be answered by searching for the information in memory about Charles Dickens and telephone numbers -- it is a simple paired-associate task. Because no such association will exist, the subject should immediately respond, "I don't know." No one says "I don't know" to this question. Rather, people claim not even to bother to search for the answer; they simply reject the question as illegitimate. In fact, by examining a set of questions of this nature, it becomes clear that considerable pre-processing precedes question answering. Similarly, there are other stages of processing, including search, deduction, and then, various levels of postprocessing. In order to know how people attempt to answer questions, we need to know how people store information and how they combine general information about the world with information about the question to derive an appropriate answer -- in short, how people think. Moreover, we will see that in order to answer questions, people must use:

* simple inference
* knowledge of causality
* their understanding of physical laws
* general knowledge of the world
* their understanding of what the person asking the question already knows.

Question Answering by Simple Retrieval

Just how do questions get answered? Let us start our studies by examining the difficulties encountered with a strategy of simple retrieval.
The Telephone Number Problem

Consider the following set of questions:

What is X's telephone number? Where X is:

John Happenstance
Charles Dickens
The President
A local restaurant
A friend
You

If these questions could be answered by simple retrieval from memory, the algorithm would look something like this:

1. Search memory for the structure equivalent to "the telephone number of A is B."
2. If successful, then B is the number.
3. If not, then the answer is "I don't know."

This algorithm has one immediate difficulty, that caused by what has come to be known as the paraphrase problem. That is, the information may actually exist in memory, but in a different format than that of the question. We might have the telephone number of X's home, or of his spouse or roommate, or we may not encode the number with the relation the-telephone-number-of but rather with phone-number, or extension, or simply as his-number-is. All these variations on the information require expansion of the simple algorithm: if the first search fails, it is necessary to consider variations on the question.

But even the most casual thought about these questions indicates that it is the basic philosophy of the algorithm that is wrong. For some of the questions, we do not even bother checking in memory for the relevant information. With the request for the telephone number of John Happenstance, the
typical response is something like "Who on earth is he?" With the requests about the numbers of Charles Dickens and of the President, it is clear that we determine that the phone number was never learned (or could not even have existed), so again the response is either explanation or an incredulous return question, not a simple "I don't know." In the situation where the correct answer is actually known, such as when asked for the number of an infrequently called telephone, then the answer is likely to take a long time to be retrieved. This is in fact just the opposite of the prediction from the simple search algorithm: the algorithm should produce the fastest response time when the number is actually known.

Stages of Question Answering

Several important conclusions result from these simple examples. First, question answering proceeds by stages. There appears to be a preliminary rapid, cursory search of the information presented to determine if anything at all is known about the query. If this rapid search fails, then the reason for the failure determines the type of response made to the question. Thus, to the question about Happenstance, the response is to get more information about Happenstance himself.

The need for an initial, rapid search shows up in most of the psychological experiments that have been performed on search strategies. For example, Collins and Quillian (1969) found that it took less time for a person to answer that the statement a canary is a fish is false than to answer that a canary has skin is true. Similarly, Meyer (1970) found that it took less time for his subjects to answer false to the statement some chairs are people than true to the statement some pilots are men. For both these examples, the simple algorithm stated earlier would predict longer search times for the false answers. In these cases it looks like people make a rapid ini-
tial search to find first whatever relation does exist in memory among the items in the question. Then, the results of that search are examined to see whether they are sensible or not. If not, the search is abandoned. If so, then a more detailed analysis and, perhaps, more search is required.

In Meyer's case, the initial search is used to see whether there is any overlap in the meaning of the two terms to be compared: if not, answer false; if so, more analysis is needed.4

Memory and General Knowledge

The Floorplan Problem

External knowledge about the world offers constraints on the possible interpretations of information retrieved from memory. Consider how someone draws a floorplan from memory. Buildings have walls and supporting structures. If a staircase occurs on one floor, then it must appear in a corresponding location on the next adjoining floor. Toilets are often located one above the other, especially in public buildings. There must be passageways for people to get from one room to another, and all rooms must have entries. In remembering a floorplan, all these facts combine with whatever is actually remembered. In drawing floorplans, the role of inferred knowledge is made clear because the constraints of building construction are well known. In most memory retrieval tasks, external knowledge is also used, but usually its role is not so easy to discover.

We get clues about the way different types of knowledge interact by examining errors. Consider the following example. In the married students' housing at the University of California, San Diego, all apartments have a balcony that is entered from the living room. Figure 1 shows the architect's floorplan of the apartment.
Residents of these apartments have some surprising ideas about their own apartments. Figure 2 shows a typical floorplan drawn by a resident of one of the apartments: note the way he drew the balcony.

People who had lived in the apartments for long periods of time (measured in years in some instances) thought the balcony was constructed flush with the exterior of the house, whereas, in fact, it extended beyond (in the normal way). The reason for the confusion is clear from the real floorplan in Figure 1. This balcony design is quite unusual in that there is a solid wall on both sides of the balcony; the wall on the right is assumed to be the outer wall of the adjoining bedroom. Normally, this would be an excellent assumption, but in this case it is false. This particular building has a peculiar balcony design, but one that is immediately obvious simply by looking at the building from the outside. The error reveals the constructive nature of the retrieval process. Forty-seven percent of the people who have been tested made this same error in drawing the balcony, even though some of them sketched the plan while they sat within the living room itself. Another 20% of those tested have had difficulty with the bal-
cony, correcting their drawings several times before satisfying themselves. (A total of 15 people have been tested.) The error and the difficulties reveal the potent influence of knowledge of the world in reconstructing knowledge from memory.

Knowledge of the World

Here is another example to illustrate how a detailed knowledge of the world gets used even in the most unexpected places. I draw the example from a discussion by Minsky (1968, p. 22).

Peter put the package on the table.

Because it wasn't level, it slid off.

The pair of sentences is easy to understand. The problem is to determine just how the pronouns get matched up properly with their appropriate referent: the first it in sentence 3 (it₁) could refer to either the table or the package, the second it (it₂) must refer to the box. Syntactical knowledge alone cannot identify with which referent each it belongs. Consider the pair of sentences

Peter put the package on the table.

Because it was round, it rolled off.

In sentence 4, it₁ must refer to the package, not the table.

To decipher sentences 3 and 4 properly requires some knowledge about objects, about causes of movement, about the things that can roll and slide, and, in general, about some of the laws of physics. Moreover, some deduction is added freely, because if now asked

Where is the package now?

the answer would probably be something like "I'm not sure, but probably on the floor."
When concepts are represented within memory they must fit within the framework provided by the knowledge of the world. This general world knowledge is likely to be extremely extensive, containing as it must all the learned information that we have come to take for granted. In fact, it is exactly the material that we take for granted that is most important to understand: any concept we have to exert any conscious effort to use is probably not an important one for everyday understanding. Unfortunately, it is the latter concept that is easier to study: the former the more difficult.

The Structural Framework

One view of the role of world knowledge is to consider it as a structural framework upon which newly acquired information must be fastened. This skeletal or schematic representation then guides both the interpretation of information and also the search for new information to fill the gaps left in the structure. This notion is especially clear for the analysis of language. Verbs, for example, imply a particular structure of concepts and other events. Thus, upon hearing the phrase

...put the package...

the verb put carries with it a framework that must be filled for complete understanding to occur. In particular, a verb like put requires an object (the package, in this case), a location, and an agent, someone who instigates the action. Moreover, the object must be movable as in "put the package on the chair" or creatable, as in "put the answer in the square," the location must be capable of holding the object, a prior location of the object is assumed to exist (although it need not be stated), and the agent must be one capable of moving the object.
In a different manner, suppose we are told that

Chris is the husband of Pat
Pat is Tom's mother
Sally is Tom's sister

We now know the sexes of all the characters, the fact that Pat is the wife of Chris and the mother of both Tom and Sally, that Chris is the father of both Tom and Sally, and even that Chris and Pat are older than Tom and Sally. Moreover, we would not be surprised to learn that they all live in the same house. In short, each simple familial relationship carries with it a complex structural framework. This structure tells us how to elaborate upon what is given. Moreover, it tells us what information is missing and what of the information we have been given does not fit. When does the elaboration occur? Is it at the time the sentences are heard or when a question about the information is answered? It is easy to show that at least some simple elaboration occurs very early. If we are told the following

Sam is Henry's father
Sam is Tom's wife

we know immediately that either there are two people named Sam or that something is wrong.

The Representation of Information in Memory

Let us now try and see just what a memory representation must contain in order that it be capable of dealing with the several issues just discussed. I have suggested that the retrieval of information is preceded by a quick stage of checking, searching the memory for any relationship that exists among the items being asked about. Then there follows a more careful examination of the information stored in memory. In addition, the structural framework of information about the world supports the specific knowledge of gen-
eral events. We know that Charles Dickens had no telephone even though no one had ever told us that fact before. We know (incorrectly) that the balcony is flush with the rest of the house because we can "see" the exterior wall of the bedroom running up to the outer edge of the balcony.

The job now is to devise a formal, particular representation that allows us to test our ideas of memory, of the interactions of specific information and of the supporting structures of world knowledge, and of specific retrieval schemes. My colleagues and I have devised one such possible representation. It seems quite effective, but it is yet too early to tell (see Rumelhart, Lindsay and Norman, 1972). Formal methods of encoding information in a large data base are reasonably common, and our procedure is an amalgamation of ideas borrowed from many sources, with a few original ideas scattered here and there. We represent knowledge by means of a network representation, with the nodes standing for concepts or events and the directed, labeled relations that connect the nodes providing the meaning structure. We started off with a scheme heavily dependent upon linguistic consideration, especially upon Fillmore's (1968, 1969) case grammar, but this dependence has been much reduced as we have gone on to deeper, more conceptually based representations. Most important, we have learned that the network representation should be active, with some of the nodes standing for programs that operate upon the network itself. The static nature of the network owes much to the inspiration and insights given to us from our study of Fillmore (1968, 1969) and Quillian (1968, 1969), plus discussions with Kintsch (1972). The need to consider causal relations and to add a deeper conceptual level of knowledge to the network came primarily from Schank (In press), and the importance of dynamic representation, with the memory being an active
collection of processes rather than a simple passive network has been considerably aided by the works of Reitman (1965) and Winograd (1972). These latter two points will be emphasized in the remaining part of this paper. Collins and Quillian (1972) have discussed a number of the points made in this paper.

The Representation of Actions and Concepts

The representation to be described here is presented in more detail and with more justification in the paper by Rumelhart, Lindsay and Norman (1972). Basically, we represent concepts as directed, labeled graph structures. A concept is represented by a node in the space: it is defined by its relationships to other concepts. Nodes are connected to other nodes by labeled, directed relations. Because each relation also has an inverse, the network is bidirectional, although obviously the meaning of a relation usually differs from the meaning of its inverse.

Events are specified in a similar way, except that actions require certain arguments. Thus, the node that represents an action may have obligatory relations leading from it specifying such things as the agent, location, and object of that action.

Most actions and concepts in the network have a single primary node (or type node) that encodes its definition, and numerous secondary nodes (or token nodes) that represent specific instances of the primary one. Almost all encodings of specific scenes are done by means of secondary nodes.

Although the structural network is presented here and illustrated in the figures as if it were a graph, there is a simple formal equivalence
between such a graph structure and both list structures and functional notation. All three notations are closely related—networks, lists, and functions—and which is used at any one time becomes a matter of convenience and personal preference.

The basic unit in the memory space is the scenario: an action that consists of events, actors, locations, and objects. Later, we will see that scenarios also imply causal factors and results.

Let us illustrate the representation system by analyzing sentences 2 and 3.

Peter put the package on the table. (2)
Because it wasn't level, it slid off. (3)

Figure 3

Figure 3A shows a possible simple encoding for sentence 2 which includes some of the underlying structures of the action. (Notice that the method by which the location was caused to change is left unspecified.) Figure 3B shows a similar encoding for sentence 3. These two figures show how the frameworks associated with the verb and concepts must define a more complete representation of the events depicted by the sentence. The final picture is derived from a consideration of much more information than is in the two sentences themselves:

A. There is no raised edge on the table (else the package would not have rolled off).
A: PETER PUT THE PACKAGE ON THE TABLE.
B: Because it wasn't level, it slid off.
B. When the package slid off the table, it was left unsupported.

C. All objects heavier than air must be supported or else fall.

D. Therefore, the package fell.

E. The usual supporting surface for a table is a floor.

F. Hence, the package fell to the floor.

G. When an object falls from a height, it is likely to break.

Expanding the Meaning-structure

The representation of the scenario shown in Figure 3 is not complete. There are two types of expansions that should be performed. First, there is expansion to reach the underlying meaning of each to the actions and concepts. Second, there is conceptual expansion to consider all the implications of the scenario.

Expansion of the individual entries to get to the underlying meaning is necessary not only to enable that meaning to be derived, but also to solve the problem of paraphrase. Consider sentence 2,

Peter put the package on the table.

(2)

What we mean is that the person we know as Peter caused a particular package to be moved from its former location to some unspecified spot on the top surface of a particular table. Moreover, although this need not necessarily hold, it is strongly implied that Peter carried the package in his hands and then, once the package was on the table, he let go of it. No matter what words are used to describe the action we wish to get the same underlying interpretation: We must be able to recognize that a sentence like (2) is
completely equivalent to the dialogue of (12) and (13):

That package on the table--how did it get there?

It was Peter, he did it. (12)

Actually such analysis is not too difficult. Many actions reduce to basic concepts. In both (2) and (12-13), the important concept is that Peter caused the package to change location from its former to its final position, the tabletop. If we had a basic underlying concept of a change-in-location concept, and if all actions that reflect changes in location were translated into this underlying meaning, then we would have the solution to several problems, including the one of paraphrase. To be able to do this, we must have a lexicon that describes the method by which actions are defined in terms of more fundamental units. A possible start towards this has been shown in Figure 3.

Conceptual Representation

When a movable object is placed on a fixed, tilted surface, the object is likely to slide along the surface. When an object starts to move it continues to do so, either as long as the causal factor of the movement is operating, or until it reaches some barrier to its movement. In this case, the causal factor is that of gravity acting upon a movable object on a tilted surface. When the surface terminates, then, if there is no barrier (an edge on the table), the object must slide off the surface. But an object that is not upon a surface will fall until it reaches one: all heavier-than-air objects must be supported. The table must also be supported, usually by a large surface called the floor. Thus, the package will fall from the tabletop to the floor.
All of this must be represented: the major question is how. This is not the place to attempt more formal analyses for these items, except to say that the change in location of the package is the easiest concept of the entire scenario: all the complications reside in the explication of the movement. Moreover, we have barely begun: each action in the method part of the scenario can be expanded. Is there any end?

The one major conceptual difficulty here is that the complexity of the analysis is in conflict with our impressions of how we ourselves understand action sequences. Suppose I had actually watched Peter move the package, what would my memory representation be? Intuitively, my picture is something like that of Figure 4: a cognitive knowledge of the fact that

---

Figure 4
---

... a transition in the package location has taken place, but with the method of the movement filled by something like a sensory image of the acts. If this is so, then actions should probably not be defined in terms of more elementary actions indefinitely, but rather as programs or plans for performing or thinking about events.

The Representation of Memories as Plans

The difficulty with the static, formal network representation of memory just described is that it is too abstract: there is no contact with physical reality. With the human cognitive system, all input and output information comes through the sensory and motor systems. Thus, our image of a concept such as a table or of an action such as to walk is probably
Figure 1
not specified in terms of more and more detailed refinements on the semantic level, but rather as a sensorimotor image of our experiences. We could think of walk as a form of motion in which the actor and object are the same and in which the method is by means of a certain set of motions made by the feet of the actor. We could think of stroll, saunter, and amble as more specific types of walking. But these definitions seem unwieldy: we know what these types of walking are through our sensorimotor routines for performing them.

One way of cutting the tangled definitional web is to determine the place for sensorimotor images. Fortunately, one method is readily available: it has already been partially illustrated in Figure 4. We let the primitive definition of actions be the sensorimotor instructions for performing those actions. This by itself is not enough. If the representation of a sentence such as

Move the object to the table.

involves the sensorimotor commands for movement, then thinking of the sequence also is likely to lead to the movement. This is not what we want. We need several different ways of using the same sensorimotor representation. We should be able to do at least three different things with a given sensorimotor plan:

A. The sensorimotor system can be examined, as data.

B. The sensorimotor system can be activated, causing the action.

C. The sensorimotor system can be simulated, causing an internal representation of the action sequence.

To illustrate the properties of such a representation, consider what the structure of a program might look like. Consider some information stored
in memory that can also be used as instructions to act upon the memory itself. Suppose you are asked to state what is in common between two concepts, say dogs and cats—how would you do it? Presumably, one does this by going to the representation for the two concepts and comparing the relation concept pairs that each are associated with: the pairs that are the same for both are the things in common, the others can be ignored. This description of how one might go about comparing two concepts is an example of using the sensorimotor representation as data (an example of use A, above). Actually performing the comparison between the concepts would be an example of activating the representation (an example of use B, above).

A formal description of the operation for finding the factors common between two nodes called A and B and putting them on a new node called \textit{in common} might read like this:

Define the following as an operator: \textit{common factors between} A and B with results at \textit{in common}.

\textbf{Paragraph 1.}

Connect \textit{in common} to the token of \textit{common factors} with result.

\textbf{Paragraph 2.}

Find a new node-relation-pair associated with A.

If the pair doesn't exist, then stop.

If that relation connects B to that node, then connect \textit{in common} to the node with that relation.

Repeat from Paragraph 2.\footnote{This set of instructions is a program that readily translates into Figure 5, the format for representing events in the data base. It can be...}
used in exactly the three ways we need. First we can examine it as something that has been learned and stored in memory. Thus we can answer questions about how the program works. Second, we can actually initiate the program. Thus, if we were asked what was in common between dogs and cats, presumably we would do so by doing something analogous to the program for commonfactors, using dog and cat as input nodes (in place of A and B). Third, we could simulate the routine by examining each statement in turn and seeing what it accomplished. (This would be more appropriate to physical actions such as typing a paper than to routines that are performed internally, such as comparing two concepts.)

Note that the routines are defined in terms of primitive actions—the sensorimotor instructions. In the computer realization of this idea (which is now in operation), the primitives are machine instructions capable of being executed. If the instructions refer to internal operations (such as to modify the node structure for a concept), then the primitives are defined as operations on the network itself in the internal programming language used by the computer. If the instructions refer to external operations, the "sensorimotor" system of the computer, then the result is to control the external devices of the computer to type messages or read cards or tapes. These primitives could just as easily move an arm or receive input from a television "eye," if our system were implemented on one of those half-dozen machines that have them.
COMMONFACTORS

RESULT

ACTION-CONNECT

NEW-RELATION

NEW-NODE

LOGICAL-CONNECT

ACTION-CONNECT

STOP

Figure 5
This multiple representation has one more virtue. If in the human, the end referent of a node is a sensorimotor experience, then we see how someone can:

A. Learn by doing—trial and error construction of a network;
B. Learn by being instructed—building a network that then refers to appropriate (and already existing) sensorimotor schemata;
C. Learn by observation.

Because the program is both executable instructions and data, it can be changed or modified in the same way as other information stored in memory can be modified. Thus, just as we can learn to modify our knowledge of concepts and of events, so too can we learn to modify our actions.

Causal Factors and Memory

A major feature of the way that a person views the events of the world is in terms of their causal factors. That is, we tend to disbelieve that an event could simply happen by itself; rather we tend to believe that an event must have a cause. The tendency to give causal reasons for events probably is beneficial in the long run, but such a tendency also gives rise to many discrepant, superstitious beliefs.

We find evidence for the distortions introduced by the apparent necessity to find causal factors of events in much anthropological and sociological literature (see D'Andrade, Quinn, Nerlove and Romney, 1972) and even in the distortions of stories reported by Bartlett (1932). This tendency is especially marked in the reproductions of Bartlett's story, "The Son Who Tried to Outwit His Father" (pp. 129-146). Here we find persistent attempts to add reasons for the peculiar acts depicted in the story. Bartlett calls this phenomenon "rationalisation." According to Bartlett, "They were all introduced
unwittingly, although in one or two cases a subject said, after giving her
version, 'I like to have reasons for things, and there aren't many here'" (p. 136).

Now, what does all this mean for the structure of memory? A new exam-
ple illustrates what has to be done.

The Three Drugstores Problem

The basic problem before us was elegantly posed by Abelson and Reich
(1969). I paraphrase their version of the problem in this way:

Suppose an individual says a sentence such as,

"I went to three drugstores." (15)

A response based on syntax only might be,

"How did you go to three drugstores?" (16)

A response based on some semantics might be,

"What useful things did you buy in three drugstores?" (17)

But the most natural response ought to be,

"How come the first two drugstores didn't have what you wanted?" (18)

Let us look at the specific knowledge that must exist in order to provide
each of these possible answers. First, however, there is even one level of
behavior below the one described by Abelson and Reich. Most memory systems
that have been formulated today would respond,

"Thank you. I have now stored that information." (19)

This is the level of response implied by most current theories of human mem-
ory. But we expect humans to respond with something like sentence 18, not
that of 19, so let us work our way up in complexity, and see what extra in-
formation or knowledge is requested to respond with the increasing levels
of sophistication shown by the sequence of (19), (16), (17), and finally,
the most reasonable one (18).

To respond as in sentence 19 requires an efficient storage system as well as
a good understanding of English. These are not easy tasks to master, but they are only the first steps. If we simply stored sentence 15, then in terms of our network representation, the input sentence would be represented somewhat in the format of Figure 6.

Figure 6

Now, consider what knowledge is required to respond as in (16). Here, we must realize that go is a specialized form of motion (location change), with a structural framework that has slots for a prior location, a final location, and a means of travel. Thus, to say, "I am going to the drugstore" implies that I will travel by some unspecified means from my present location (which is not that of the drugstore) to the location of the drugstore. With this sentence, several pieces of information that go with the verb framework are unspecified—locations, instruments, times, etc. (see Figure 7).

Figure 7

If we know this much, then to attempt to fill in these missing slots is only natural: it represents an obvious strategy to complete the structure by asking:
Figure 6

Figure 7
"How did you go to three drugstores?"

"From where did you start?"

"Where did you go after the last drugstore?"

A Semantically-based System.

What useful things did you buy in three drugstores? (20)

Here, motives and purposes must get added to the system. To do this is to add more than semantics—some understanding of human behavior is also added. In specific, it requires the knowledge that most actions are actually part of a deeper schemata, that human behavior is goal-oriented.

We need a further scenario, one that states that people go to drugstores in order to purchase one or more of that special class of things sold in drugstores. Thus, the structural framework must now be expanded considerably to include purposes, as well as reasonably complete descriptions of the concepts and actions that are involved. The general structural representation for going to a drugstore starts to look something like that of Figure 8.

Figure 8

Here we have shown the encoding for the types of items contained within (American) drugstores and the purpose of going to drugstores. This scenario points out the need to get even more information about the drugstore trip: hence, two new possible questions are

"What useful things did you buy in three drugstores?"
Figure 8
or even, "Where did you get the money?"

A Conceptually-based System.

How come the first two drugstores didn't have what you wanted? (21)

For a system to generate this response, the underlying structural framework for an action must be further modified. The structure must take on the general characteristics of a rule. The basic conceptual structure might look something like that of Figure 9. It contains:

-----------------------------

Figure 9

-----------------------------

A. A conditional statement.
B. An action to be performed if the conditional is true.
C. A result of that action.

Now we can expand the picture given to us earlier. We go to a drugstore to purchase something. If we then go to a second drugstore, clearly the first trip must have failed in its purpose. Similarly, we need to visit a third store only if the trip to the second store was unsuccessful. Hence, the question we ask in response to the presentation of the original statement

"How come the first two stores didn't have what you wanted?"9

Finding the Suitable Answers to Questions

One last stage remains in our study of how questions get answered. Just as the acquisition and the search through the knowledge contained in memory is a more complex task than originally thought, so too is the way
Figure 9
that knowledge gets used when it is finally found more complex than is immediately obvious. The construction of a proper answer to a question is not performed simply by describing the relevant information that has been found. A number of different issues arise in the development of a suitable answer to a question: let me examine only one of these issues here.

The Empire State Building Problem

The most straightforward situation occurs when the question is found to have an answer. Except in the most simple psychological experiments on memory, however, we would not expect that the answer be describable by a single word or phrase. Usually we can expect that a large body of information is found relevant to the question, perhaps with several different possible interpretations. What then should the respondent do?

Suppose we consider a question with a reasonably unambiguous answer: 10

Where is the Empire State Building? (22)

But look, a part of my knowledge of the Empire State Building is shown in Figure 10. With which part should I respond? To answer this question properly requires that the respondent understand the requirements of the person who asked the question. If I were asked this question in Russia, I might well respond "In the United States." If I were asked by an adult in Europe, I would probably respond "In New York City." In the United States--especially in New York City, I would respond "On 34th Street." Finally, if
asked in the New York subway system I would not answer with a location, but rather with instructions on how to get there.

In order to answer a question appropriately, it is necessary to have a model of the knowledge of the listener, including knowledge of why the question was asked. Full exploration of this issue leads to the development of a major set of experimental and theoretical issues, so I will only touch on a few now. Basically, a person who is giving a serious answer to a question must consider the developing network of information owned by his listener and attempt to fill the gaps. To do this well requires reasonable depth of knowledge about the listener, or perhaps a sophisticated understanding of the reason that certain questions get asked. A humorous, yet insightful demonstration of the difficulties involved in answering questions without any appreciation for the knowledge base of the listener occurs at any cocktail party where strangers from a wide variety of occupations meet and ask "What do you do?" If the question is taken seriously, then it may require a good deal of conversation to establish sufficient common ground that it can be answered at an appropriate level.

Piaget has studied a highly related problem in the egocentric behavior of a young child. As is easily imagined, learning to consider another person's knowledge in answering a question is not a simple task. Piaget suggests that the child of seven years or less is unable to do this. The young child is egocentric in his behavior, describing everything from his own point of view. Thus, he frequently uses pronouns in his speech without any understanding that the referent may not be intelligible to the bewildered adult to whom he is speaking.
As far as I can determine, all of the existing computer-based information retrieval systems are egocentric: their designers are usually so pleased that they can sometimes find the information requested of them that they deluge the inquirer with more information than he can use, as well as information that he may already know. According to my interpretation of Piaget's stages of intellectual development, we would have to say that these systems had just barely started the acquisition of intelligent behavior.

To answer a question intelligently requires a large body of specific knowledge about the area being questioned, including general knowledge about the world and its causal and physical laws, and also an understanding of the knowledge and behavior of other people. Those of us trying to model the human use of memory have just barely gotten used to the fact that we must add knowledge of the world to the model, so it comes as a surprising and new challenge that we must also incorporate a person's understanding of other people into the model. A number of interesting philosophical issues are encountered in this new problem: I refer the reader to the insightful treatment of some of these issues in Minsky's chapter, "Matter, Mind, and Models" (1968).

The Acquisition of Knowledge

We have now examined a number of different features about the representation of information in memory. We have seen how an active representation based on a sensorimotor process seems to be necessary, with the previously acquired knowledge about the world providing a structural framework upon which to construct new knowledge. The pre-existing structure not only defines the way that new information will be represented, but it helps in organizing the information that is not yet known. Thus, the structure can maintain open
positions within its framework for the necessary causes and results, objects and actors, and even methods and actions that still remain unspecified. In this way, a person acquiring knowledge can be led to seek more information to fill the missing nodes in his developing memory representation.

**Learning**

Teaching and learning are the names we give to the activities of presenting and acquiring new information. There is no adequate theory of instruction to guide us in these activities: perhaps an examination of the structure of memory from the viewpoint represented by this paper might provide useful guidelines for the development of such a theory. Indeed, Macdonald-Ross (1972) has studied network relationships with just this goal in mind: to determine how to design an educational system for the Open University in England. Macdonald-Ross argues that structures of this nature do guide both students and faculty into an understanding of their subject matter.

Examine the structure of memory. It is an interconnected network, with new facts supported within the skeletal structure provided by the old. To acquire new information means to construct new nodes and the relations between them. Consider how that might get done. First, let Figure 11A represent a segment of knowledge in the memory. If we acquire two new concepts (C1, C2) and a relation between them (R) as shown in Figure 11B, there are no connec-
tions to the old network. Retrieval should be difficult, and perhaps even the acquisition is difficult. This is poorly learned information.

Now consider Figure 11C. Here, the two newly acquired nodes are well integrated. To use an obvious analogy drawn from the structure itself, we would say that the newly learned components of 11C are well supported; those of 11B lack support. Notice, too, that 11C is better integrated within the network, in part because it contains more information. If this view of things is correct, it explains why mnemonic techniques are so useful in causing arbitrary strings of items to be learned even though they add to the total amount of material that is to be acquired. They provide a firm, well integrated structural framework.

Teaching

Suppose we have a large body of knowledge to teach—what is the best way of doing it? Presumably, we need to interconnect the new information with the existing structure. One way to do this is to construct a supporting web structure first, and then fill in the details. To do the details first would not work, for without a supporting structure, the new material simply could not become integrated. In teaching, this means that an outline of the material to be learned should be acquired first, then a more detailed overview, and then progressively more and more detailed structures.

The network representation of knowledge can guide the process of instruction in two different ways. First, if we have a good representation of the knowledge we wish to teach, then we can organize it properly for efficient learning. Second, if we try to discover the network representation of the student, we can use this to guide our teaching. Knowing the knowledge structure of the student helps in devising the original level of organization of
the material. In addition, as the lessons progress, we can use our understanding of the development of the student's structures to guide us in teaching, telling us what old material has not been acquired and what new material might perhaps already be known. Thus, in theory it should be possible to tailor instruction to the knowledge base and competence of the student.

Whether the network representation makes the goal easier to attain remains to be seen. The major drawbacks have seemed to result from the tyranny of numbers: tutorial methods are too expensive to use in mass education, lectures, texts, and even are not flexible enough. Most teaching machine programs and computer-assisted instructional systems simply are not yet sophisticated enough in their implementation to attempt these goals (with the exception of the demonstration systems of Brown, Burton and Zdybel, 1972, and Carbonell, 1970). Nonetheless, the analysis is instructive even if the implementation remains in the future.

We can characterize two different strategies of presenting material: two different strategies of teaching. One is to present a cohesive organized structure to the student, carefully adding one new piece of information after another to the developing structure. This might be called linear teaching. It is the system that characterizes lectures, textbooks, and even the structure of this paper. The other method is to present a coarse web of information, outlining the topics to be discussed, then giving a general overview, more detailed overviews, and finally the detailed substructure. This procedure might be called web teaching. Web teaching is often prescribed, seldom done. It is difficult to perform well. But I wish to suggest that for the learning of complex topic matters, web teaching may at times be more efficient.
The distinction here between web and linear strategies are closely related to, yet different from, the distinction between holistic and serialistic strategies of learning (Pask, 1971; Pask and Scott, 1972). Pask and Scott gave students freedom to explore a network of knowledge about probability theory. Some students, the serialists, preferred to work on one node at a time, always completing that node before going on. Others, the holists, worked on many nodes at a time. These are interesting results that, hopefully, can be used to characterize styles of learning.

**Linear Teaching.** Linear teaching (and its complement, linear learning) might be characterized by the sequence shown in Figure 12. At first, there is a well developed structure of knowledge (Figure 12A). Then, we add a new set of nodes, say the linear string C1, C2, and C3, and their subnodes, as shown in Figure 12B. Finally, a longer linear string is acquired, nodes C1 through C5 in Figure 12C. Presumably, as learning increases, the older material becomes more thoroughly embedded into the knowledge structure, but the overall picture is still one of a systematic linear increase in the knowledge base. But this new knowledge is weakly supported. If one link in the chain is lost, either because it was poorly acquired or because the learner might have missed that part of the material, then the rest of the structure is weakened and, conceivably, may fail. This is a familiar experience to most instructors: a student who fails to learn the point made early in the course may thereafter be in difficulty, failing to understand all that has followed.
Figure 12
Web Teaching. Web teaching (and web learning) might be characterized by the sequence shown in Figure 13. Here, we use the first few nodes acquired (Figure 13B) to establish a coarse web, but one that is well integrated with the previous knowledge structure. Then we insert refinements within the structure created by the original web framework, as shown in Figure 13C. Finally, we can fill in the details, as shown in Figure 13D. In web learning, no single node is critical to the whole, so that poor acquisition or even the absence of a single set of concepts does not destroy the validity of the structure.

The structures described in these two diagrams (Figures 12 and 13) are at best weak analogies to the process of learning and teaching. As yet, they are far from being a formal, testable theory. But I find the analogies insightful ones, useful even at this early stage of development in the planning of course materials. Whether there is any more to the distinctions between web and linear teaching than simple analogies remains to be investigated.
Figure 13
Summary

When we examine in detail the nature of memory we find that we cannot separate the representation of knowledge from the uses to which we put that knowledge. It is possible to devise formal representation of knowledge within human memory. To do so requires some type of network, with distinctive means of representing concepts and actions. But even with such a representation, we find that there is much more to the process of memory than simply storing the information away.

When people are asked questions, they do not simply go into their memory and respond with the appropriate information. Rather, they first investigate the question itself, determining whether it be sensible or not, or what its exact referents are. Even when some information is retrieved, it is likely to be deeply imbedded within a general structural framework determined by knowledge of the world itself, and this extra information can bias the type of memory responses that are given. Thus, it is possible to show how people might make errors in information that presumably they have learned very well, for example, even in recalling the floorplan of the apartment in which they are then living.

It is neither possible nor desirable to separate memory for action from the plans necessary to execute those actions. Thus, it should be possible to store within memory a sequence of instructions that can be used in different ways:
1. We can examine the instruction as something that has been learned, as data within the memory structure.

2. We can activate the instruction, causing the action thereby described to be performed.

That information stored in memory can be used in these two different ways is an important concept, for now we can see how learning can come about by actual performance, by instruction, by thought, or by observation.

When we teach someone else knowledge, we are trying to build within that person a data base comparable to that of our own for the particular subject matter of interest. But in order to do this we must know what the other person knows and what he lacks. What is needed is some sort of interactive process, in which we first question the other person to find out what is lacking, then teach, and then question again to find out how successful we have been.

In answering a question, it is important to be able to do more than simply combine information about the world with information that has been learned about the question. In order to derive the proper answer we must determine exactly why the question was asked, else we are likely to answer at the wrong level. This means that in addition to the knowledge of the subject being asked about, we must also have knowledge of the person who has asked the question. Consideration of these topics leads to a theory of instruction, a theory which states the importance of properly connecting new material into the framework provided by the old.

In this paper, a general formal structure for representing semantic information is proposed. Examples are given of network structures for encod-
ing both general knowledge about the world and also specific information learned about a particular concept or event. The structure is now being tested by simulating it on a large digital computer. The result of these investigations is the realization that there is much more to the memory process than has heretofore been described in our theories.
References


Footnotes

1. This research is supported by research grant GB 32235X from the National Science Foundation. Many of the concepts discussed in this paper have come from the activities of the "LNR Research Group," most especially the insights of David Rumelhart and Adele Abrahamson, Marc Eisenstadt, Dedre Gentner, Yaakov Kovarsky, Jim Levin, and Steve Palmer.

2. George Mandler is responsible for first demonstrating to me the importance of preprocessing by devising the question, "What is Professor Luria's telephone number?" The answer, "Why on earth would I know his phone number?" proved to be a central influence in developing the sets of queries and analyses presented in this paper.

3. In a network representation of memory, everything is eventually connected to everything else, so that if two items exist somewhere in memory, there must be some path connecting them, although that path may be non-meaningful.

4. The search probably takes place by some variation of a depth-first search, starting at all the nodes corresponding to the concepts of the input query. The intersections which result when the searches which start at different nodes overlap one another determine the paths to be examined. Many possible variations exist on this simple strategy for rapid, initial search, but the basic idea is to do an initial rapid preprocessing of the question in order to establish some relationship among the items as quickly as possible. Whether the search process is sequential or parallel, active or passive, is not known.

5. I thank Marc Eisenstadt and Yaakov Kovarsky for collecting some of the floorplans.

6. This type of expansion of a given lexical item into more primitive definitions has been taken from the work of the generative semanticists. I have been most influenced by Miller's analysis of verbs of motion (Miller, 1971) and Schank's (in press) structural analyses. The points that causality is an essential part of the
analysis of an action sequence, that the instrument of an action is, in fact, an entire action sequence, another event rather than simply a single concept, comes from Schank. (Schank calls his basic verb of change-of-location trans.) The particular expansions of Figure 3 are modifications of the analyses performed by the "Verb Group" subsection of our research group who actually do a more thorough analysis, breaking down a verb like put into the sequence of actions do, cause, location-change.

7. What I have in mind here is the fact that formal network or logic representations of events are unwieldy, perhaps fundamentally so. Analogical representation, however, is much richer in scope, for a single simple analog of a situation can be equivalent to a large complex of logical statements about the same situation, as well as containing information not easily (or even not possibly) represented in formal language systems. An analogical representation need not be the same as a picture in the head, for the representation could be much more abstract and stored in any one of numerous formats, including digitally. Much of my reasoning is captured by the recent discussion on analogical representation by Sloman (1971).

8. The program is written in the format of SOL II, a language developed and implemented by David Rumelhart. The parser that he has implemented along the lines of the recursive, augmented transition network parser of Woods (1970) transforms statements in SOL II into the data base format, as in Figure 5. Pronouns and anaphoric references across sentence boundaries are permitted, as in the example.

9. Most people on hearing this analysis object that it is too specific: there are lots of reasons for visiting several drugstores--casing them for a robbery, price shopping, selling newspaper advertisements. The objections simply serve to
strengthen the point, however, for they all assume a purpose and result for each visit. The point of the analysis is that people infer motives and causal factors, and that the structural framework for representing conceptual information must include causes and results.

10. The subtelties involved in answering questions of this sort and the Empire State Building example were pointed out to me by Marc Eisenstadt and Yaakov Kareev.


Bernstein, N. A. The role of the number of the number of the element relative phase of the component standards. March, 1970.


