COGNITIVE SCIENCE, 2, 129-153 (1978)

Modeling Spatial Knowledge*

BENJAMIN KUIPERS

Massachusetts Institute of Technology

A person's cognitive map, or knowledge of large-scale space, is built up from observations gathered as he travels through the environment. It acts as a problem solver to find routes and relative positions, as well as describing the current location. The TOUR model captures the multiple representations that make up the cognitive map, the problem-solving strategies it uses, and the mechanisms for assimilating new information. The representations have rich collections of states of partial knowledge, which support many of the performance characteristics of common-sense knowledge.

1. INTRODUCTION

Common-sense knowledge of space is knowledge about the physical environment that is acquired and used, generally without concentrated effort, to find and follow routes from one place to another, and to store and use the relative positions of places. Among other things, this knowledge allows me to follow the familiar route between my home and MIT; to think up a new and shorter route to the shopping center; to elaborate my "mental map" when given a guided tour; to point toward places I cannot see; and to face North. This body of common-sense knowledge is often called the "cognitive map." The research described in this paper proposes representations and inference mechanisms for that knowledge in the cognitive map that deals with large-scale spatial relations.

Large-scale space is space whose structure cannot be observed from a single viewpoint. This generally includes street networks observed by traveling through them, and excludes visual recognition of particular places, maps, or aerial photographs, although these can properly be considered part of the "cognitive map" in the wider sense. Naturally, this definition depends on the observer, so a city might not be large-scale when viewed from an airplane, while a map might be large-scale when viewed through a small hole.

We are concerned here with everyday activities: learning and problem solving

*This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Project Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.

Requests for reprints should be sent to the author at Division for Study and Research in Education, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

in a large-scale urban environment without the use of map or compass. The observations available are the sequence of places and paths encountered on a route, the magnitudes of turns and distances traveled (to some low accuracy), and the observed positions of remote landmarks. The manifest behaviors produced by the cognitive map are to solve route-finding and relative-position problems. However, the central object of interest is the cognitive map itself: the physically unobservable structure of information that represents spatial knowledge. Learning is assimilation of observations into that structure; problem solving is extracting the answers to particular questions from it.

Although the model presented here (called the TOUR model) deals only with the observations and behaviors described above, there are other sources of evidence about the structure of the cognitive map. Under the assumption that the same or similar representations are involved, we can learn from people's behavior at other spatial tasks. Lynch (1960) interviewed residents of three cities and categorized the spatial elements mentioned as landmark, node, path, edge, and district according to the roles they played in the cognitive map. Among other observations, he points out that errors in cognitive maps are most frequently metrical, and rarely topological. Appleyard (1970) analyzed sketch maps, identifying several variant types, and observing the frequent occurrence of highly structured areas loosely connected. Beck and Wood (1976) provided students with a formal mapping language as they explored London, and analyzed the features and distortions included in the sketch maps they produced. Golledge (1976) has also studied distortions in sketch maps, assuming the cognitive map to be a metrically distorted two-dimensional space whose mapping function can be determined from sketch maps. McCleary and Westbrook (1974) studied the effects of different kinds of printed maps on travel patterns of visitors to a historical village. Linde and Labov (1975) studied verbal apartment descriptions and showed that most descriptions followed an underlying tour route whose structure could be predicted.

The basis for a great deal of work on the development of spatial concepts in children is the research of Piaget and Inhelder (1967) which, though not restricted to large-scale space, provides an important classification of spatial knowledge into the categories of topological, projective, and Euclidean. Hart and Moore (1973) survey this large body of literature and propose a modified version of the Piaget and Inhelder model. Siegel and White (1975) concentrate their review on large-scale space, and show a strong parallel between a child's acquisition of spatial competence and an adult's acquisition of the spatial structure of a new environment. They further speculate that both kinds of learning occur through the "Now Print!" mechanism that takes a "snapshot" of the state of the nervous system at critical moments. As will be seen below, the TOUR model is a very different, accretionary mechanism for the assimilation of spatial information. The TOUR model takes no position on the mechanisms of development.

The vocabulary of spatial concepts in the TOUR model is derived in large part from these observations, particularly those of Lynch (1960) and Piaget and Inhelder (1967). As a further source of evidence, I conducted approximately 50 interviews with adult subjects, asking them to solve route-finding problems, describe cities, and draw sketch maps. These interviews and sketch maps suggested parts of the conceptual vocabulary that is used to represent spatial knowledge.

Yet another source of constraints on the representations for common-sense knowledge is the performance requirements they function under. Kuipers's (1978) performance requirements are that learning should be easy and performance should degrade gracefully under resource limitations, and he derives several properties that are desirable in the special-purpose representations for common-sense knowledge: (1) The amount of processing time and working memory required to answer common types of requests should be small. (2) A subset of a meaningful state of the representation should be meaningful. (3) The amount of processing required to change the representation from one state to another should be small. (4) Very few observations should be discarded because there is no reachable state of the representation that can incorporate them. These are, in a sense, design goals, and they cannot all simultaneously be satisfied. All of these design goals can be illustrated by comparing the use of a partial order to a total order as a representation for the states of partial knowledge encountered as a one-dimensional order is being observed and learned. These design goals are important because the special-purpose representations required to satisfy (1) make certain kinds of operations easy at the expense of others, and may make it impossible to represent certain states of knowledge at all. Thus the set of meaningful states of knowledge is an important consideration about a special-purpose representation.

Thus the TOUR model is presented as a psychological model of human common-sense knowledge of large-scale space, but one which has been constructed with primary attention to scope and general agreement with the observations available in the literature and my own interviews. The point of this paper is to describe the representations and operations that make up the TOUR model, referring anecdotally to the knowledge being represented. The clearest picture of the variety of knowledge in the cognitive map is painted by Lynch's (1960) delightful book *The Image of the City*. This paper does not include a detailed comparison between an empirical description of human behavior and the TOUR model's computational description of spatial knowledge. That comparison will have to wait on further empirical research now under way. Neither does it include a comparison of alternate computational models for the same phenomena, for alternate models are only recently being developed. This comparison also awaits further research.

The TOUR model is simulated by a computer program (written in LISP on the PDP-10) that takes as input simulated observations, assimilates them into its cognitive map representation, and solves route-finding and relative-position problems. The model and the computer simulation are described in more detail in (Kuipers, 1977). Although, in principle, I suppose the TOUR model could have been created without a working computer simulation, the program has been invaluable for testing the consistency of the representations and inference rules, and for debugging them into better forms.

Returning to the cognitive map itself, let us consider some of its aspects via the venerable story of the Blind Men and the Elephant. Imagine, however, that each man had his own elephant, and that one actually was shaped like a snake, while another genuinely resembled a tree, another a wall, and so on. Not only does the cognitive map appear different depending on how you approach it, but it actually *is* different in different people. Thus the metaphors we will look at must do double duty, both for different aspects of the cognitive map and for individual variation.

First, the cognitive map is like a map in the head. More accurately, it is like many maps in the head, loosely related, for the cognitive map certainly lacks the global consistency of a single printed map. It can be used to solve spatial problems, and some people claim to "see" a map when they answer spatial questions. However, the "map in the head" is only part of the answer, because people have many kinds of knowledge that do not correspond to any partially drawn map.

Second, the cognitive map is like a network. It is made up (in part) of streets and intersections, and the exact shapes and lengths of the links are often unimportant. Route-finding is a search for a set of links in the network leading from one place to the other. Spatial errors often correspond to distortions of the space preserving the network structure.

Third, the cognitive map is like a catalog of routes. Each route is a procedure for getting from one place to another, and they are essentially independent. A place may not be recognized as the same from two directions, and to attempt a shortcut is to court disaster. A particular route procedure might be an ordered sequence of actions to be followed, or it might be an unordered collection of actions, each triggered by the appropriate feature of the environment.

Each of these metaphors captures an important aspect of the cognitive map. Each may characterize some individuals' cognitive maps completely. The TOUR model attempts to fit these different aspects into a common framework, giving each a precise computational description, showing how most cognitive maps include all three aspects, and demonstrating how a set of simple inference rules accomplishes the interactions between them.

2. THE TOUR MODEL

The TOUR model divides spatial knowledge into five categories and contains five corresponding representations for knowledge about particular environments. Each representation provides a particular range of states of partial knowledge, communicates with certain other representations, and is able to solve a certain range of problems.

1. A route is represented as a sequence of actions taking the traveler from one place to another. Spatial knowledge is assimilated into the rest of the cognitive map from relations implicit in the actions that make up particular routes. A route description represents knowledge from three sources: observations of the environment (simulated in the TOUR model), recalled versions of previously traveled routes, and intermediate states of the route-planning process. These different kinds of knowledge are treated similarly in many ways, an important and somewhat surprising result.

2. The topological structure of a street network is represented by descriptions of streets and places that include partial knowledge of the order of places on a street and of the local geometry of the intersection of two streets. A street description defines a one-dimensional orientation on the street. Information in the topological description is assimilated from route descriptions.

3. The relative position of two places is defined as a vector with respect to a coordinate frame with a limited domain of applicability. This defines a two-dimensional orientation with respect to which a heading can be defined and attributed to streets, and which can tie together the local geometries of several intersections.

4. Dividing boundaries, defining regions to either side, provide a qualitative nonvector partial knowledge of position. This kind of knowledge is particularly easy to acquire, is particularly useful in route-finding, and lends itself to the definition of larger structures such as bundles of parallel streets and rectangular grids.

5. Regions, related by containment, provide useful levels of abstraction for stating relations among their elements. These levels of abstraction make it possible to state general principles that can be used to find routes or relative positions in a large number of particular cases.

To represent this knowledge and its uses in the cognitive map, the TOUR model has three classes of representations: (1) representations for knowledge about a particular environment; (2) a description of the current position (the "You Are Here" pointer); and (3) representations for inference rules which manipulate knowledge of the other two kinds.

Knowledge about particular environments is divided into the five categories presented above. This knowledge is encoded in descriptions of route instructions, places, paths, regions, and coordinate frames. Such a description is made up of a number of properties and their values (implemented in LISP as the property list of a generated atom and a collection of associated access functions). The detailed structure of these descriptions will be presented in the sections discussing the different categories of knowledge. See Kuipers (1977) for the complete descriptions.

The cognitive map solves problems by applying inference rules to information represented in these environmental descriptions. Most of the computation in the cognitive map takes place on assimilation, so that problem solving is usually quite simple. Assimilation of new information takes place by transferring small pieces of information from one description to another, often via the "You Are Here" pointer.

The current position of the traveler is represented by a small working memory called the "You Are Here" pointer, which describes the current position in terms of place, path, one-dimensional orientation on that path, current coordinate frame, and two-dimensional orientation with respect to that coordinate frame.

YOU ARE HERE:
  PLACE: ⟨place description⟩
  PATH: ⟨path description⟩
  DIRECTION: ⟨1-D orientation: +1 or −1⟩
  ORIENT: ⟨coordinate-frame description⟩
  HEADING: ⟨2-D orientation: 0 to 360⟩

(I will present individual descriptions in this format to emphasize their coherence as descriptions, and to allow the individual elements to be complex data structures (sets or partial orders) if necessary. This contrasts with the familiar circles-and-arrows semantic network diagrams which I feel are frequently cluttered, confusing, and deceptive.)

Most manipulations of knowledge in the TOUR model take place through an interaction between the environmental descriptions and the "You Are Here" pointer. Furthermore, the only environmental descriptions that are accessed are typically the ones referred to by the "You Are Here" pointer and the current route instruction. These amount to a focus of attention for the inference rules that manipulate the descriptions. This lack of search makes most operations quite efficient.

Both the "You Are Here" pointer and the environmental descriptions may be incompletely specified. In most cases the TOUR model will function with incompletely specified descriptions, although with degraded performance. For example, if only a partial route description can be retrieved from memory, it may still be possible to follow it from one end to the other, but perhaps not to perform additional assimilation into place or path descriptions, or to rehearse the route in the absence of its physical environment.

The inference rules that manipulate knowledge embedded in these various representations are represented as productions: simple modules that wait for a certain set of conditions to be true and then perform some action. They are organized around a process that follows a route description as a sequence of instructions to move the "You Are Here" pointer through the environmental description. The source of these instructions may be direct observations of the environment, a route description recalled from memory, or the route-planning process. Since this process resembles a computer executing a computer program, the collection of inference rules is known as the "TOUR machine." The inference rules also fall into several categories according to the common ways the environmental descriptions are used.

1. Rules which compare the current route instruction, the "You Are Here" pointer, and the topological descriptions of the environment. They can act to fill gaps in each representation with information from the others. In particular, this is how the topological description is originally created from information in the route description.
2. Rules for maintaining the current heading, or two-dimensional orientation, with respect to the current coordinate frame. They operate with the relation between the one- and two-dimensional orientations represented in the "You Are Here" pointer and in the current place and path descriptions.
3. Rules which detect special structural features of a part of the environment, such as paths which act as dividing boundaries separating places. Dividing boundaries can then be combined into larger structures, such as bundles of parallel streets, and rectangular grids. These rules act within the focus of attention provided by the current route instruction and the "You Are Here" pointer.
4. Rules which solve route-finding and relative-position problems using knowledge in the hierarchy of regions and in the descriptions of coordinate frames, boundaries, grids, and street networks.

3. ROUTE AND TOPOLOGICAL REPRESENTATIONS

A route description is a sequence of actions, obtained initially as simulated observations of the environment, which have the effect of moving the "You Are Here" pointer from one place to another. At the same time, information from these observations is assimilated into the topological representation. The topological representation consists primarily of partial orders of places on the same street, and local geometries of intersections. The more global descriptions of the environment are built on the foundation of this topological description. (An exception to this is a strategy for exploring unknown territory which is discussed in Section 8 below.)

Two kinds of spatial relationships can be observed while traveling along a route through an environment. The two observations are immediately represented by filling in the observable parts of the TURN and GO-TO descriptions. Thus an external observation is immediately converted to a memory representation which can be modified, stored, recalled and rehearsed, or forgotten completely.

> TURN—the selection of a path to follow from a set of alternatives available at a given place (e.g., making a turn at the intersection of two streets). The selection specifies the path chosen and the amount of the turn taken. One-dimensional orientation (to be defined below) is not

physically observable, but refers to a relationship with other descriptions in long-term memory.

TURN:

PLACE: ⟨place description⟩
PATH1: ⟨path description of starting path⟩
DIR1: ⟨1-D orientation on PATH1⟩
AMT: ⟨amount of turn⟩
PATH2: ⟨path description of resulting path⟩
DIR2: ⟨1-D orientation on PATH2⟩

GO-TO—the ordered pair of places encountered along a path without an intervening decision (e.g., the starting and ending point of a route segment along a single street).

GO-TO:

FROM: ⟨starting place description⟩
TO: ⟨ending place description⟩
PATH: ⟨path description⟩
DIR: ⟨1-D orientation of travel on PATH⟩
DIST: ⟨distance traveled⟩

A route description is a sequence of TURN and GO-TO descriptions leading from an initial to a final place. Since the world is continuous, the route description provided by physical observation must contain no gaps; i.e., a TURN must be followed by a GO-TO on the selected path, and a GO-TO must be followed by a TURN at its endpoint. However, a route description as stored and recalled may contain many kinds of gaps.

There are clearly other kinds of spatial information that we are ignoring, including the sensory impressions of places encountered. These sensory impressions are necessary for recognizing the same place when it is revisited, but we will consider this process primitive and opaque.

The topological representation consists of PLACE and PATH descriptions. The PATH description includes a partial order of PLACEs which are on that path. This partial order represents states of partial knowledge about the total order which places actually have on a path. A PATH has a one-dimensional orientation with respect to this order: +1 represents facing in the direction of the order, and −1 represents facing against the order. The partial order data structure is a list of sequences whose transitive closure is the desired partial order. The external behavior of the partial order is to incorporate additional fragments of order, and to answer questions about the relative order of two given places with "+1," "−1," or "don't know."

PATH:

NAME: ⟨name⟩
ROW: ⟨partial order data structure⟩

A PLACE description includes a description of the local geometry of paths which intersect at that place. This local geometry describes the relations among paths, their one-dimensional orientations, and their radial headings in the local coordinate frame of this intersection. It will accept new information about that relationship from a TURN observation, or it will attempt to deduce either the amount or destination of a turn, given the other.

PLACE:
    NAME: ⟨name⟩
    ON: ⟨list of PATHs⟩
    STAR: ⟨local geometry data structure⟩

For the purposes of the topological representation, the "You Are Here" pointer describes three aspects of the current position: the current place, the current path, and the current one-dimensional orientation on that path. Some of these may be left unspecified.

YOU ARE HERE:
    PLACE: ⟨place description⟩
    PATH: ⟨path description⟩
    DIRECTION: ⟨ +1 or −1 ⟩

When following a route description, the TOUR machine initializes the "You Are Here" pointer to the beginning of the route. Each observation acts as an instruction with the effect of changing the "You Are Here" pointer to its destination. Meanwhile the inference rules that make up the TOUR machine take fragments of information from one description and put it into parts of the others that have been left unspecified. These inferences are of several kinds:

1. Inference rules that take information about the current position in the "You Are Here" pointer and use it to fill unspecified parts of the current instruction. For example, the "You Are Here" pointer may have a DIRECTION component provided by previous inferences, and it can be used to supply the missing DIRECTION component of the current GO-TO instruction.

2. Inference rules that take information from the current instruction and add it to the description of the current place or path (specified by the "You Are Here" pointer). For example, if a GO-TO instruction states that two places are related by a given direction on a given path, this information can be added to the partial order of places which is part of that PATH description. These rules may extend the topological description by adding a new PLACE or PATH description to the cognitive map when required by the current observation.

3. Inference rules that take information from the current place or path description (specified by the "You Are Here" pointer) to provide missing information for the current GO-TO or TURN instruction. For example, the current PATH

description may be able to supply missing information about the direction component of a GO-TO instruction. The updated route description may be stored in memory for later retrieval.

4. Inference rules that fill gaps in a route description by posing them as problems for the problem-solving component of the TOUR model. The solution to the problem can itself be an incomplete route description, requiring further calls on the problem solver. This allows skeletal route descriptions to be used as intermediate states of the route-planning process, to be repeatedly refined by the TOUR machine until complete. It also allows incompletely recalled route descriptions to be filled in.

The states of partial knowledge in the route and topological representations result from the partial order and local geometry descriptions, and from the ability of the TOUR machine to tolerate underspecified descriptions of route and environment. Fully specified descriptions are very useful for filling in missing parts of new descriptions; incompletely specified descriptions are usually adequate for driving the "You Are Here" pointer.

These states of partial knowledge make it possible for the individual inference rules to be very simple, so that a single pass through a route description can assimilate useful amounts of information into the topological representations at low cost. Usually, several passes through a particular route description are necessary before all the useful information is extracted, so it is important that route descriptions be available in a compatible form both from observations and from memory.

4. TOPOLOGICAL ASSIMILATION EXAMPLE

This example shows how the route description and the environmental descriptions interact to fill unspecified parts in each other. The scenario takes place near Central Square in Cambridge, whose simplified map is illustrated in Fig. 1. We

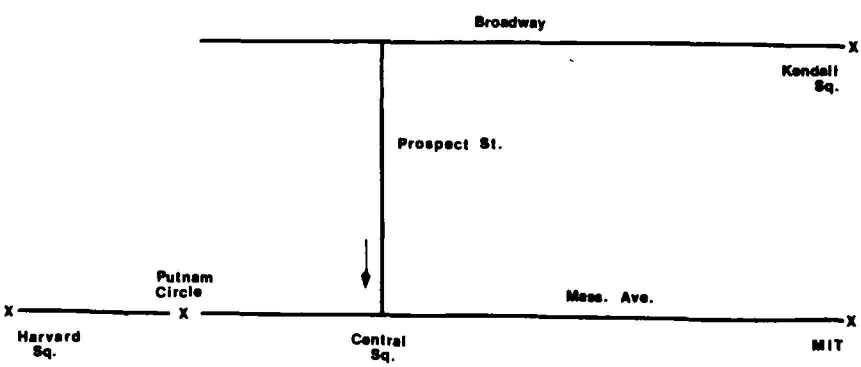

FIG. 1 A map of the area in Cambridge to be explored.

begin this scenario at Central Square, on Prospect Street, having come from Broadway. We will turn right and proceed to Putnam Circle. The current position is in the "You Are Here" pointer:

```
YOU ARE HERE:
            PLACE:  [PLACE2: Central Square]
            PATH:  [PATH3: Prospect Street]
            DIR:  +1
```

The relevant parts of the cognitive map are the descriptions of Central Square, Mass. Ave., and Prospect Street.

```
PLACE2:
            NAME:  Central Square
            ON:  [PATH1: Mass Ave]
                 [PATH3: Prospect Street]
            STAR:  (0. PATH1 -1)
                   (90. PATH3 -1)
                   (180. PATH1 +1)
PATH1:
           NAME:  Mass Ave
           ROW:  ([PLACE1: Harvard Square]
                  [PLACE2: Central Square]
                  [PLACE3: MIT])
PATH3:
           NAME:  Prospect Street
           ROW:  ([PLACE4: Broadway & Prospect Street]
                  [PLACE2: Central Square])
```

The first action in the route is to turn right. The observation of this action is a TURN description with six elements: the place, path, and direction preceding the turn, the amount of the turn, and the path and direction resulting from the turn. In this case, most of the elements are left unspecified.

```
TURN:
          AT:
          ST1:
          DIR1:
          AMT:  90.
          ST2:
          DIR2:
```

The "You Are Here" pointer provides the current context to fill in the first three missing elements of the TURN observation, corresponding to the starting position.

TURN:

AT: [PLACE 2: Central Square]
ST1: [PATH 3: Prospect Street]
DIR1: +1
AMT: 90.
ST2:
DIR2:

At this point, the local geometry description in the STAR property of PLACE2 can be used to describe the result of the turn. The local geometry associates a heading with certain (PATH DIRECTION) pairs as they radiate from a place, and the amount of the turn specifies the new heading which can, perhaps, specify a new (PATH DIRECTION) pair. The absolute values of these headings are meaningless, and can only be used to compute such differences. In this case, the result is to specify the TURN instruction completely:

TURN:

AT: [PLACE2: Central Square]
ST1: [PATH3: Prospect Street]
DIR1: +1
AMT: 90.
ST2: [PATH1: Mass Ave]
DIR2: −1

The final operation is to update the "You Are Here" pointer to reflect the result of the TURN instruction. Note that, in conjunction with the description PATH1, this implies "facing Harvard Square."

YOU ARE HERE:

PLACE: [PLACE2: Central Square]
PATH: [PATH1: Mass Ave]
DIR: −1

The second action of our brief tour takes us to Putnam Circle, which is shown on the map above, but which is completely new to the cognitive map, so it must create a new PLACE description. The observation corresponding to this action is:

GO[TO:

FROM:
TO: [PLACE5: Putnam Circle]
PATH: [PATH1: Mass Ave]
DIR: −1
DIST:

The "You Are Here" pointer again provides the current context, including the direction along Mass. Ave. that we are traveling, so we can fill in missing parts of the GO-TO instruction:

GO-TO:

FROM: [PLACE2: Central Square]
TO: [PLACE5: Putnam Circle]
PATH: [PATH1: Mass Ave]
DIR: −1
DIST:

When processing the previous TURN instruction, the PLACE description was used to add information to the instruction. Here the GO-TO instruction will be able to add information to the description of the environment. First, we may add to PLACE5 the fact that Putnam Circle is on Mass. Ave.

PLACE5:

NAME: Putnam Circle
ON: [PATH1: Mass Ave]
STAR:

Second, since the GO-TO instruction gives an order relation (−1) between Central Square and Putnam Circle, we can add this information to the partial order in PATH1. Notice that we do not know where Putnam Circle is with respect to Harvard Square, but we do know that both are on the same side of Central Square.

PATH1:

NAME: Mass Ave
ROW: (PLACE1 PLACE2 PLACE3)
(PLACE5 PLACE2)

It is illuminating to consider the effect of a partially specified "You Are Here" pointer. If there had been no local geometry information in PLACE2 about Central Square, for example, the direction of travel would have been unspecified in the GO-TO instruction, and the partial order in PATH1 showing the position of Putnam Circle would then have been:

PATH1:

NAME: Mass Ave
ROW: (PLACE1 PLACE2 PLACE3)
(PLACE5)

The corresponding route description, if stored in memory, would still contain the information that a right turn from Prospect Street at Central Square points you toward Putnam Circle, but this fact would not be represented in the more generally accessible PLACE and PATH descriptions. Once other observations had provided more useful information either to the local geometry of PLACE2 or the partial order of PATH1, a subsequent rehearsal of the route description would extract the remaining information it contained.

This example shows how the topological properties of places and paths are represented in their descriptions, and how information is assimilated from the relatively inaccessible route instructions into the more globally useful topological descriptions. The assimilation process takes place through very simple, and computationally inexpensive, interactions between the current instruction, the environmental descriptions, and the "You Are Here" pointer. Notice that, since the only environmental descriptions accessed are those referred to by the current instruction and the "You Are Here" pointer, processing time is independent of the total amount of information in the cognitive map.

5. ORIENTATION

A "sense of direction" is the ability to define one's current heading (or two-dimensional orientation) and the relative positions of remote places with respect to the same coordinate frame. By having many different coordinate frames, a person may represent the positions of many places without the requirement that they fit into a single consistent framework. A "sense of direction" is therefore not a sense at all, but knowledge that is dependent on a coordinate frame for the current position and a particular set of remote places. The constraints on newly added position information can be relatively weak, so the position representation has many states of partial knowledge, and learning is relatively easy. Knowledge about the relation between two coordinate frames can also be represented as part of the coordinate frame descriptions.

COORDINATE-FRAME:
    TYPE: ⟨"local" or "regional"⟩
    DOMAIN: ⟨place or region⟩
    OTHERS: ⟨other coordinate frames with relative headings⟩
PLACE:
    NAME: ⟨name⟩
    ON: ⟨list of PATHs⟩
    STAR: ⟨local geometry data structure⟩
    ORIENT: ⟨coordinate frame⟩
    VIEW: ⟨set of triples: (place heading distance)⟩

The ability to support multiple coordinate frames for positions allows the TOUR model to include the "multiple map" metaphor for the cognitive map. This independence of coordinate frames is required because it is quite common for a person to be well oriented within each of two different regions, but have very little notion of the relation between them. Knowledge about the relation between two coordinate frames can be learned or forgotten separately from position information within each one.

The topological representation supports only a one-dimensional orientation with respect to the order on a particular path. Thus, we must augment the TOUR model to include descriptions of coordinate frames, and knowledge in the PATH and PLACE descriptions about the relation between one- and two-dimensional orientations. The "You Are Here" pointer must be augmented to include the current coordinate frame and the current heading with respect to that coordinate frame.

YOU ARE HERE:
PLACE: ⟨place description⟩
PATH: ⟨path description⟩
DIRECTION: ⟨1-D orientation: +1 or −1⟩
ORIENT: ⟨coordinate-frame description⟩
HEADING: ⟨2-D orientation: 0 to 360⟩

The current implementation allows the heading to be specified as an integer from 0 to 360 degrees. In fact, only the eight headings at 45 degree intervals are actually used, referring to a range of actual headings within about 30 degrees of the given value. It seems very likely that partial knowledge of heading in people includes more imprecisely specified headings. When a theory of partial metrical knowledge has been worked out, numerical values for heading and distance can be replaced by more realistic descriptions without changing the overall structure of the TOUR model. This is an area of current research interest.

PLACE and PATH descriptions may contain information about the relationships between one- and two-dimensional orientations. The local geometry of a PLACE can be defined so that its headings are consistent with those of a particular coordinate frame. Unlike in the topological representation, this means that the absolute values of the headings in the local geometry data structure have meaning: they can be compared with headings in other PLACE descriptions with the same coordinate frame. A PATH description may include the heading of its +1 direction, which can have multiple values associated with different coordinate frames.

As the TOUR machine drives the "You Are Here" pointer along a route, its problem is to maintain the current HEADING and to transfer orientation information between the "You Are Here" pointer and the PLACE and PATH descriptions. There are three kinds of inference rules to accomplish this:

1. Inference rules that update the current heading for a TURN whose amount is known, and that check to see that the PATH followed by a GO-TO is straight before allowing the heading to remain fixed.
2. Inference rules that set HEADING or DIRECTION in the "You Are Here" pointer by examining the current PLACE and PATH descriptions.
3. Inference rules that add information to the current PLACE and PATH descriptions about the interaction between the current HEADING and DIRECTION as they appear in the "You Are Here" pointer.

In addition to representing knowledge from visual observation or verbal report, knowledge of heading can be used to implement a "dead reckoning" technique for computing the relative positions of the source and destination of a given route. Dead reckoning requires the "You Are Here" pointer to be expanded to hold X and Y values for the current position in rectangular coordinates with respect to the current coordinate frame. The distance traveled by a GO-TO on a known heading can be converted to those rectangular coordinates, and the result at the end of the route converted back to polar coordinates. More than any other part of the TOUR model, this process will be affected by a change from a numerical representation of metrical information to a more cognitively realistic one.

6. TWO-DIMENSIONAL ORIENTATION EXAMPLE

Information about the current heading is maintained and updated in much the same way as the topological information in the previous example. For example, in the "Turn right" instruction at Central Square, the enlarged "You Are Here" pointer would acquire a heading and coordinate frame from the local geometry description in PLACE2. Notice that PLACE2 must explicitly include the coordinate frame description ORIENT3, because ORIENT3 may be shared with other places.

```
PLACE2:
          NAME:  Central Square
          ON:  [PATH1: Mass Ave]
               [PATH3: Prospect Street]
          STAR:  (0. PATH1 -1)
                 (90. PATH3 -1)
                 (180. PATH1 +1)
          ORIENT:  ORIENT3
YOU ARE HERE:
           PLACE:  [PLACE2: Central Square]
           PATH:  [PATH1: Mass Ave]
           DIR:  -1
           ORIENT:  ORIENT3
           HEADING:  0
```

The domain of a given coordinate frame can propagate along lines of frequent travel. Assume that we arrive at Putnam Circle and make a turn, so that information must be added to its local geometry description. Rather than creating a new, local coordinate frame for just PLACE5, its local geometry would be defined with respect to ORIENT3, and would thus be closely related to the local geometry of Central Square.

Two coordinate frames *collide* when one defines the local geometry of the current place, while the other defines the heading in the "You Are Here" pointer. A collision can have two outcomes. If one of them is local to its particular place, it may be replaced by the coordinate frame with the larger domain, which therefore continues to propagate. If both have substantial domains, the relationship between them is stored in the coordinate frame descriptions. In either case, the heading in the "You Are Here" pointer can be maintained and updated along a route that travels quite far from where its coordinate frame was originally defined.

At the same time, the HEADING property of a PATH description can hold the relationship between the one-dimensional and the two-dimensional orientation of the path. As seen below, the heading of the path's +1 direction is paired with its coordinate frame and stored.

PATH1:
NAME: Mass Ave
ROW: (PLACE1 PLACE2 PLACE3)
(PLACE5 PLACE2)
HEADING: (ORIENT3 180)

If more than one heading is put into the same PATH description, we again have a collision of coordinate frames.

7. BOUNDARIES

Boundaries, by specifying a division of the space into distinct regions, are very useful in describing the location of a place. For example, sitting in Technology Square, I can describe the location of the Cambridge Public Library (See Fig. 2) by saying that it is:

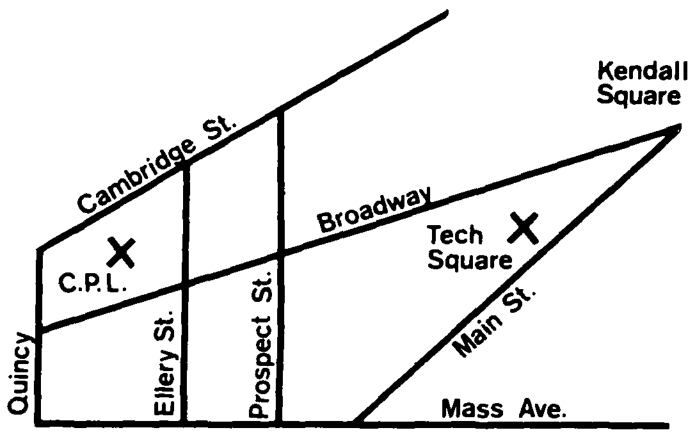

FIG. 2 Dividing boundaries provide partial specification of position.

on the other side of Prospect Street,
on this side of Mass. Ave.,
between Broadway and Cambridge Streets,
beyond Ellery Street,
before Quincy Street.

In each part of this description, I am using a street to draw a boundary, dividing the world (or at least a small part of it) into two regions, then specifying which of those regions contains the place I am describing. A boundary in this sense is not a barrier: it acts to define groups of positions, not to impede travel. Naturally, a barrier like the Charles River can also function as a boundary.

When a path acts as a boundary, the sides on the right and the left when facing the +1 direction are represented by REGIONs (sets of PLACEs) in the RIGHT and LEFT slots of the PATH description.

PATH:
    NAME: ⟨name⟩
    ROW: ⟨partial order of places⟩
    HEADING: ⟨list of pairs: (coordinate-frame heading)⟩
    RIGHT: ⟨region⟩
    LEFT: ⟨region⟩

Thus, specifying where a place lies with respect to a dividing boundary provides partial knowledge about its position. This knowledge is particularly easy to acquire, easy to combine with other similar pieces of knowledge, and easy to apply to route-finding problems.

A single turn in a route description can specify where a place lies with respect to a dividing boundary. If the route to the Cambridge Public Library involves a left turn from Mass. Ave. (in the +1 direction), then the Cambridge Public Library is on the left side of Mass. Ave., no matter how tortuous the rest of the route is (providing, of course, that it does not cross back over Mass. Ave.).

Although in principle the two regions defined by a boundary extend to infinity, in practice they include only places whose relationship to the boundary is represented in the cognitive map. Thus, if a street is involved in a large variety of different routes, it will have many boundary relationships with different and distant places. Otherwise the division it represents may apply only to the immediate neighborhood.

The description of the position of the Cambridge Public Library translates readily into useful constraints on a route from here to there. Prospect Street and Ellery Street are both potential intermediate subgoals, while Mass. Ave. and Quincy Street can act as barriers in case the route strays too far from the goal. In fact, any of the given streets can act as a skeleton from which to plan a route, since each has known relationships with both source and goal. If connections are found from source to skeleton and from skeleton to goal, then the skeleton street can join the pieces into a complete route.

If two boundaries are known to be parallel, then the regions on their sides have further useful relationships. For example, Prospect Street and Ellery Street are parallel and both lie between here and the Cambridge Public Library. By knowing the order of the two parallel boundaries, Prospect Street can be chosen as the first subgoal, followed by Ellery Street. A bundle of parallel streets allows subgoals to be ordered into a sequence of small steps that can be easily joined to construct a route.

Two streets can be found to be parallel by examining their headings. The two streets fall within the focus of attention of the TOUR machine when they are cross streets encountered on a GO-TO instruction. This allows "local parallel" relations to be found linking two streets. When a particular street has several local parallel relations to other streets, a gathering operation is initiated to follow the local parallel links and gather up and order a bundle of parallel streets. The data structure that represents such a bundle becomes part of the REGION description for that area.

A rectangular grid structure on an area amounts to two such bundles, perpendicular to each other. This makes route-finding even easier because the sequence of subgoals can be found within one bundle, while the connections lie in the other. Thus, the rectangular grid is a useful description of the geography because it allows certain very powerful route-finding strategies to be used. Since this route-finding power is relatively insensitive to irregularities in the geography, people are led to apply the grid description even when it is metrically incorrect. The incorrect description of an area as a rectangular grid is thus one of the most common, and the most pragmatically useful, of the "mistakes" to be found in people's cognitive maps.

8. EXPLORATION

Finally, we can discuss an interesting technique for exploring unknown territory. How does a person explore an unfamiliar area before he knows the topology of the street network? It is clear that accomplished explorers use their sense of direction to find the way back to familiar places while learning the new area. If an explorer can maintain his own heading with respect to a familiar street, and if he knows what side of that street he is on, he can always navigate back toward it when he wants to. What is the knowledge that permits him to do this?

A person in a new area can define a coordinate frame by the position of a prominent landmark, for example the John Hancock Tower in Boston. This coordinate frame can allow him to define the heading of a familiar street. Then, while exploring, he must maintain his current heading with respect to that coordinate frame, and remember what side of the familiar street he is on. He can maintain the heading by attending only to the amounts of turns and the curves of streets. If the unknown territory can be assumed to have a grid structure, the problem becomes much easier, because the heading must have one of four values. Then our explorer can always compute the direction toward the familiar street, and can always guide himself toward it, even with no knowledge of the street network. The only dynamically varying piece of information he needs to maintain is his own heading, and that need not be particularly accurate, since a street subtends a large sector of the space. An alternate strategy is to continually update the relative position of his starting point using the dead reckoning method described earlier, but this method places a greater processing load on the explorer, and the accuracy demands are greater.

Observations of people learning an area have revealed that while newcomers orient themselves with respect to conspicuous landmarks, long-time residents very seldom do (Lynch, 1960). Those with detailed cognitive maps of an area can orient themselves by local features of each place in the street network. Furthermore, they often have a sufficient stock of familiar routes that they need not maintain a two-dimensional orientation at all, but can just follow route descriptions.

9. REGIONS

Regions allow places to be grouped and referred to collectively. As such, they provide levels of abstraction for stating facts and answering questions. For example, I can give a route for getting from the West Coast to the East Coast and hope that it can be used to solve a variety of problems concerning particular places in the two regions.

Regions are often defined in terms of legislative boundaries, visual texture, typical activities, ethnic composition, and other characteristics that are not strictly aspects of spatial cognition. Thus, unlike the other aspects of the cognitive map, I will not talk about how a region description is created, only how it is used. The problem is how the relationships among different region descriptions allow information to be stated at one level of abstraction and used at another.

For example, suppose I know a generally useful route for driving from Northern California to New England. How is this represented so that I can use it when my problem is getting from Stanford to MIT? The general idea comes from a proposal by Rumelhart (1974) to use a hierarchy of nested regions. (Rumelhart used the hierarchy to select the appropriate context for answering distance questions.) The method is to construct the sequences of containing regions about the two places, and find the smallest common containing region, in this case the continental United States. Then, proceeding downward in the two sequences, look for solutions to the problem indexed under pairs of disjoint containing regions. In this case we would look for possible routes from the West Coast to the East Coast, from California to the Northeast, from Northern California to New England, and from the Bay Area to Massachusetts. If several possibilities are found at different levels, use the most specific one.

The TOUR model could represent a hierarchy of nested regions quite easily by having a region description point to the next larger containing region. The set of nested regions about a given place forms a sequence, so it is easy to compare two such sequences to find the smallest common region. Unfortunately, it is difficult to add a new region to such a structure, since its relationship with all other regions must be known before it can be merged into the sequence. Furthermore, some useful regions overlap in a way that does not fit into a convenient hierarchy. For example, the Berkshires overlap Massachusetts, Connecticut, and New York, with no containment relations in either direction. Thus, the set of containing regions for the Tanglewood Music Center contains both Massachusetts and the Berkshires, but they cannot be ordered into a nested sequence. The TOUR model relaxes the nesting requirement to provide many more states of partial knowledge without great loss of performance.

In the TOUR model, a region description points to an unordered collection of containing regions. When a problem is posed concerning a particular place, the set of regions containing it is obtained by following these "upward pointers" and giving them a partial order. The longest totally ordered subset is taken from this partial order to be used for problem solving. This makes it possible to miss relationships which are actually represented, but only in cases where the containment order is partial. Thus, when solving a problem involving the Tanglewood Music Center, it could be regarded as belonging either to the Berkshires or to Massachusetts, but not to both simultaneously. Note that this limitation applies only to a particular problem-solving process; the memory representation includes both containment relations.

To solve a problem, the sequences of containing regions about the places of interest are compared from the "top down" to find the smallest region containing both places. The diverging parts of the two sequences are then examined to see if the desired relationship is found between disjoint regions contained in the common region. If several applicable relationships are found, the most specific is used. As well as being computationally efficient, this top-down access is consistent with the experimental results of Stevens (1976), who tested the relative difficulty people encountered in answering questions about relations among different sized geographical regions.

When an abstract relationship is found between regions containing two places of interest, it must still be matched to the concrete problem as originally stated. If my original problem was to get from Stanford to MIT, I might discover that a good route from Northern California to New England was to take I-80 to Cleveland, then I-90 to New England. To complete the details of the route, I must shift my focus of attention at the endpoints of the route to make them more specific. In other words, at the endpoints, there are (at least) two possible states of the "You Are Here" pointer that correspond to the same situation, and the problem is to get from the more abstract to the more concrete one.

This requires that a region description include information to permit a mapping from a general description of a place to a more specific one. The region Northern

California does not correspond to a particular more specific place or region, but Northern California *on I-80* does correspond to a specific place in San Francisco (actually, several places). Then I can pose the subproblem of getting from Stanford to that place. Similarly, I can get from I-90 in New England to MIT.

Thus, the structure of containing regions permits generally useful information to be stated at an abstract level and used at a more specific one. The containment relations permit a particular place to have a partially ordered set of containing regions, rather than simply a nested sequence. This places fewer constraints on the creation of new region descriptions. Only when a problem is posed are the local containment relations merged into a unified structure. It is also necessary to have a downward mapping, going from more abstract to more specific descriptions of places. This cannot be done by relating PLACE descriptions at two levels of abstraction, but becomes possible if the correspondence is between more and less abstract values of the whole "You Are Here" pointer.

10. CONCLUSIONS

I have presented the TOUR model as a model of spatial knowledge, consisting of a number of representations with large collections of states of partial knowledge. The knowledge in the cognitive map can be divided into five categories: route descriptions, topological street networks, coordinate frames for relative position, dividing boundaries and grids, and structures of containing regions. The representations for this knowledge can be classified as the environmental descriptions, the "You Are Here" pointer, and the inference rules that manipulate them. The TOUR model shows how the different kinds of knowledge are stored in the representations provided, and how new information is assimilated, changed from one representation to another, and used to solve problems.

The TOUR model, of course, addresses only part of the knowledge in the cognitive map. Consider, for example, the use of mental imagery to create a picture for the "mind's eye." This apparently makes some kinds of inference much easier than they would otherwise be, and many people report visualizing a "map in the head" as they explore a new area. The existence and properties of mental imagery is a matter of considerable controversy (Kosslyn & Pomerantz, 1976; Kosslyn & Shwartz, 1977; Marr & Nishihara, 1976; Pylyshyn, 1973; Shepard & Metzler, 1971), on which the TOUR model takes no stand.

The TOUR model also has little to say directly about map-reading, although maps are clearly an important source of common-sense spatial information. The process of reading a map involves a complex interaction between representations of visual space and large-scale space (Robinson & Petchenik, 1976). The TOUR model also omits any theory of visual place recognition or feature extraction, although this has been a favorite topic for urban planners (e.g., Appleyard, Lynch, & Myer, 1964). A computational theory of visual place recognition seems as difficult as the general computational theory of vision, so the TOUR model treats place recognition as a primitive "black box."

To be tested as a psychological theory, detailed predictions about human behavior must be drawn from the model, distinguishing it from other potential theories of spatial cognition. These predictions can be of several different kinds, each a major research area in its own right. The examples given below are extremely speculative and offered only for illustration.

1. Qualitative predictions about the kinds of distortions and geographical paradoxes that cognitive maps can and cannot contain.

A boundary relation, such as what side of a street a landmark is on, should seldom be recalled incorrectly. It would, however, be possible for the information not to be stored at all, as when a visual memory exists of the landmark on the street, but without enough information about the viewpoint to say what side of the street it is on. More commonly, many familiar distortions can be attributed to the false description of a street as straight or an intersection as orthogonal.

2. Qualitative predictions about the extent and nature of individual variation, based on possible alternate embodiments of the TOUR model.

The tendency to confuse right and left should be closely associated with a very limited cognitive map, probably consisting only of descriptions of familiar routes. On the other hand, the ability to use partially specified positions, such as boundary relations, in route-finding, should be characteristic of the navigational expert.

3. Qualitative predictions about the order in which spatial abilities are acquired by children, based on their computational dependencies.

A child must have both the concept of one-dimensional orientation on a path (and hence, presumably some one-dimensional order of places on that path) and the concept of two-dimensional orientation at a place before being able to describe the two-dimensional heading of a path and use it to infer the relative positions of places on the path.

4. Qualitative predictions about the relative difficulties of different problem-solving tasks.

The difficulty of a "dead reckoning" task should vary according to the number of turns in the route, while the error rate and magnitude should vary according to the departure of the streets from straight and the turns from a right angle.

5. Qualitative predictions about the error rates to be expected with different problem-solving tasks.

Error rates at estimating relative position will be fairly low in an area organized as a rectangular grid, with perhaps a tendency to distort long, thin rectangles into shorter, fatter ones. In an area that is topologically a grid, with metrical distortions, route-finding should be almost as successful *as in a true grid, relative position errors will be of the obvious kind, and overall confidence in the correctness of solutions will be high. In an area which is not even topologically close to a grid, finding novel routes or relative positions will be characterized by high error rates, low confidence, and conservative strategies.*

It should be pointed out that some of these speculative predictions depend on the low-resolution representation of shape and magnitude. The precise details of such a theory do not affect the predictions, however.

How can this model of the cognitive map be applied to other problems in Artificial Intelligence or psychology? First, if an AI program is to use spatial information to solve problems, some or all of the TOUR model representations will be useful, depending on whether it is being asked to acquire information from local observations, supply routes from one place to another, understand stories by reference to regional similarities or differences, or supply relative position information. The individual representations are useful for their ability to answer a specific kind of request, with the most frequent kind of request being answered most efficiently while rarer cases require more inference. For example, the use of multiple coordinate frames allows relative position questions within a densely described area to be answered quickly, while the relation between places in different areas must be inferred from the relation between their coordinate frames.

Second, spatial metaphors are very common in our language for expressing many kinds of other relations: mental, social, musical, temporal, and so on. This suggests that the representations that we learn for spatial relations are also very useful in other domains. Thus we may find use for flexible representations of sequences of operations, one-dimensional order, multiple frames of reference, binary distinctions (intersecting and parallel), nested regions, and so on. The performance characteristics that make these representations generally useful are discussed in more detail in (Kuipers, 1978).

Third, there are many important psychological questions, both theoretical and practical, about how information gets in and out of the cognitive map: giving or understanding verbal directions, reading or drawing maps or diagrams, navigating if blind or brain-injured, the use of spatial mnemonics, etc. Each of these problems deals with a process that accesses the spatial knowledge stored in the cognitive map. In order to make computational theories of these processes, it is essential to have a theory of the representations in the cognitive map.

ACKNOWLEDGMENTS

I am grateful for graduate fellowships provided by the National Science Foundation and the Danforth Foundation. I would also like to thank the members of my thesis committee, Marvin Minsky, Seymour Papert, and Carl Hewitt, and the anonymous reviewers for their helpful comments.

REFERENCES

Appleyard, D. Styles and methods of structuring a city. *Environment and Behavior,* 1970, **2,** 100–117.

Appleyard, D., Lynch, K., & Myer, J. R. *The view from the road.* Cambridge, Mass. MIT Press, 1964.

Beck, R., & Wood, D. Comparative developmental analysis of individual and aggregated cognitive maps of London. In G. T. Moore & R. G. Golledge (Eds.), *Environmental knowing: Theories, research, and methods.* Stroudsburg, Pa.: Dowden, Hutchinson & Ross, 1976.

Golledge, R. G. Methods and methodological issues in environmental cognition research. In G. T. Moore & R. G. Golledge (Eds.). *Environmental knowing: Theories, research, and methods.* Stroudsburg, Pa.: Dowden, Hutchinson & Ross, 1976.

Hart, R. A., & Moore, G. T. The development of spatial cognition: A review. In R. M. Downs & D. Stea (Eds.), *Image and environment.* Chicago: Aldine, 1973.

Kossiyn, S. M., & Pomerantz, J. R. Imagery, propositions, and the form of internal representations. *Cognitive Psychology,* 1977, **9,** 52–76.

Kosslyn, S. M., & Shwartz, S. P. A simulation of visual imagery. *Cognitive Science,* 1977, **1,** 265–295.

Kuipers, B. Representing knowledge of large-scale space. Cambridge, Mass: MIT Artificial Intelligence Laboratory TR-418, July 1977. (Doctoral thesis, MIT Mathematics Department.)

Kuipers, B. On representing common-sense knowledge. In N. V. Findler (Ed.), *Associative networks —The representation and use of knowledge by computers.* New York: Academic Press, 1978.

Linde, C., & Labov, W. Spatial networks as a site for the study of language and thought. *Language,* 1975, **51,** 924–939.

Lynch, K. *The image of the city.* Cambridge, Mass.: MIT Press, 1960.

Marr, D., & Nishihara, K. H. Representation and recognition of the spatial organization of three-dimensional shapes. Cambridge, Mass.: MIT Artificial Intelligence Memo 377, 1976.

McCleary, G. F., & Westbrook, N. *Recreational and re-creational mapping.* Sturbridge, Mass.: Old Sturbridge Village, 1974.

Piaget, J., & Inhelder, B. *The child's conception of space.* (orig. publ. in French, 1948). New York: Norton, 1967.

Pylyshyn, Z. W. What the mind's eye tells the mind's brain. *Psychological Bulletin,* 1973, **80,** 1–24.

Robinson, A. H., & Petchenik, B. B. *The nature of maps.* Chicago: Chicago University Press, 1976.

Rumelhart, D. E. The Room Theory. Unpublished computer listing. La Jolla, Calif.: University of California, San Diego, 1974.

Shepard, R. N., & Metzler, J. Mental rotation of three-dimensional objects. *Science,* **171,** 701–703.

Siegel, A. W., & White, S. H. The development of spatial representations of large-scale environments. In H. W. Reese (ed.), *Advances in child development and behavior,* vol. 10. New York: Academic Press, 1975.

Stevens, A. L. The role of inference and internal structure in the representation of spatial information. Doctoral dissertation, Psychology Department, University of California at San Diego, La Jolla, Calif., 1976.