THE TOUR MODEL: A THEORETICAL DEFINITION

BENJAMIN J. KUIPERS

APRIL 1988       AI88-78
The TOUR Model: A Theoretical Definition*

Benjamin Kuipers†

April 29, 1988

This is a section extracted from a longer article entitled "Navigation and Mapping in Large-Scale Space" by Benjamin Kuipers and Tod Levitt, to appear in the AI Magazine.

1 The TOUR Model

The TOUR model was primarily developed between 1974 and 1979 [Kuipers 1977, 1978, 1979b]. Its goal is to account for the robust performance and wide variety of incomplete knowledge exhibited by the human cognitive map. The cognitive map has a number of attractive properties as a research domain for the investigation of states of incomplete knowledge:

- Knowledge is acquired over significant periods of time, since observations are constrained by the speed of physical travel.

- Cognitive development takes place over a period of a decade or so, and there has been extensive study of the developmental stages of the cognitive map in children [Piaget and Inhelder, 1967; Siegel and White, 1975].

- There is significant individual variation among cognitive maps in adults [Lynch, 1960], leading one to suspect that there are distinct pathways for the development of a cognitive style.

- There is significant variation in cognitive maps as a function of the physical structure of the environment and type of experience with it [Lynch, 1960; Chase, 1982], making it easy to test knowledge representation hypotheses with *gedanken* experiments.

The TOUR model was designed to factor the issue of image processing of sensory input from the assimilation of observations into a description of the large-scale environment. Accordingly, it treats the sensory apparatus as a "black box," making only the simplest assumptions about its internal

---

*Copyright 1988 Benjamin Kuipers
†Computer Sciences Dept, University of Texas at Austin, Austin TX 78712
structure. The spatial environment is treated as something that is only indirectly perceived, whose description must be constructed from a person's egocentric sensory input.

The TOUR model has been implemented in Lisp, several times over the past decade, and tested on a variety of small simulated environments, up to a few dozen places and paths and perhaps a hundred or so views. Locality of access to the knowledge base ensure that the program runs very quickly, and its computational costs are not particularly sensitive to the overall size of the knowledge base. Elliot and Lesk [1982] implemented a route-finder that generated an internal map structure from USGS data tapes of a large region of northern New Jersey, and demonstrated that cost-effective route-finding could be done using a version of the boundary heuristic discussed below.

The elements of the TOUR model representation are presented here in a declarative, in some cases axiomatic, form, following Marr [1982] and Hayes [1979] who argue for defining the mathematical structure of the knowledge prior to its computational and procedural implementation. However, a little reflection shows that in almost all cases, the declarative form translates straight-forwardly into an incremental assimilation rule for acquiring the knowledge from observations.

1.1 Sensorimotor Interaction With The Environment

The sensorimotor world of an agent, in this case a traveller in a fixed environment, is a purely egocentric description of sensory input and motor output, and contains no references to fixed features of the external environment. Models of spatial learning in which sensory perception provides direct information about places or objects in the world are, at best, glossing over an important and non-trivial inference step, and more likely, committing a serious category error.

The TOUR model [Kuipers, 1977, 1978, 1979b], assumes a sensorimotor world consisting of views and actions.

- A view represents the traveller's sensory input at a given instant. While recognizing that this sensory description may be arbitrarily complex, from the point of view of the TOUR model, it is treated as an opaque object which can be matched against other views, and used as an index for associative retrieval of other structures.

- An action represents motion that the traveller can take within the environment, and typically changes the current view. An action is modeled as a change of state, with the current view defined immediately before and after an action, but not changing continuously during it.

Views are not necessarily visual: for a blind traveller, the sensory input at an instant may consist of a collection of aural, tactile, and olfactory stimuli, which are assimilated into a cognitive map much as visual views are by a sighted traveller. The principal requirement is that views be distinctive enough to allow assimilation of the structure of the environment. Open ocean, and the neophyte's view of desert or forest, may fail to satisfy this requirement.
The TOUR Model

Views and actions are regarded as the internal descriptions provided by the sensorimotor system in response to (recent) past sensations and actions. They are assumed to be correct descriptions of the actual experience, though possibly quite abstracted and incomplete.

The sensorimotor input for the traveller, from which the TOUR model constructs a cognitive map, is modeled as an alternating sequence of views and actions:

\[ V_0, A_0, V_1, A_1, V_2, \ldots V_{n-1}, A_{n-1}, V_n. \]

This discrete simplification of the continuous spatial world makes the research more tractable, of course, but it is also justified by observations that people frequently recall only those places at which decisions are made, and not the continuous paths between them [Lynch, 1960].

The TOUR model includes only two actions, each of which may include optional descriptions of quantitative parameters:

- (Turn \( \alpha \)), where \( \alpha \) describes the amount of rotation;
- (Travel \( \delta \)), where \( \delta \) describes the distance of travel.

The extent to which these descriptions of quantities are used, and whether they are treated as qualitative or quantitative descriptions, varies with the level of the semantic hierarchy.

1.2 The Procedural Map

The purpose of the procedural map is to represent a description of sensorimotor experience sufficient to allow the traveller to follow a previously experienced route through the environment. This is non-trivial if the representation must be robust in the face of resource limitations when information sensed, stored, or retrieved may be incomplete.

The basic element of the procedural map is a production-like schema:

A sensorimotor schema is a 4-tuple \( \langle \text{goal}, \text{situation}, \text{action}, \text{result} \rangle \), where action is an action description, and goal, situation, and result are all views.

The procedural interpretation of a sensorimotor schema is: "When attempting to reach goal, if the current view is situation, do action and expect result." The declarative interpretation is: "If the current view is situation, then doing action will produce result."

A procedural description of a route is a collection of sensorimotor schemas describing the fragments of the route. For example, suppose that the traveller, on his way to goal \( V_n \), has the sensorimotor experience,

\[ V_0, A_0, V_1, A_1, V_2, \ldots V_{n-1}, A_{n-1}, V_n. \]

A complete procedural route description would be the following set of sensorimotor schemas:

\[ \{(V_n, V_i, A_i, V_{i+1}) \mid i = 0, \ldots n - 1\}. \]
### The TOUR Model

<table>
<thead>
<tr>
<th>purpose</th>
<th>sensorimotor schemas</th>
<th>retrieval use</th>
</tr>
</thead>
<tbody>
<tr>
<td>describe or follow route</td>
<td>(goal, situation, action, result)</td>
<td>((G, V) \rightarrow (A, V'))</td>
</tr>
<tr>
<td>travel within the environment</td>
<td>(goal, situation, action, -)</td>
<td>((G, V) \rightarrow A)</td>
</tr>
<tr>
<td>record of experience</td>
<td>([-, situation, action, result])</td>
<td>((V, A) \rightarrow V')</td>
</tr>
<tr>
<td>recognition of familiar place</td>
<td>([-, situation, -, -])</td>
<td>(V \rightarrow \text{true</td>
</tr>
</tbody>
</table>

Table 1: The performance of a collection of sensorimotor schemas degrades gracefully as the schemas are less complete.

A route description can be followed as a procedure for navigating in the environment, depending on two factors:

1. There must be a complete set of schemas at least at the \((goal, situation, action, -)\) level.

2. The *action* descriptions must be sufficiently well specified, given the environment, to initiate the appropriate action. (Termination may be environment-driven, by activating the next schema, with obvious risk of overshoot.)

Even without being complete enough to allow the route-following procedure to be successfully executed, the route description in the procedural map is useful for recognition of landmarks and as a basis for subsequent assimilation. The procedural map is robust because many purposes can be effectively served by collections of incompletely specified sensorimotor schemas (Table 1).

While acquiring a route description during actual experience travelling in the environment, working memory holds the current view and very likely the current action. Once an action has been initiated, an interval of time occurs before the result is observed as the new current view. During this time, distractions may occur, losing some or all of the contents of working memory. It is relatively easy to create a sensorimotor schema and fill its *situation* and *action* components. However, the *goal* and *result* components depend on information persisting in working memory over a significant time-interval, so they may frequently be left empty.

The first two uses of sensorimotor schemas in Table 1 above illustrate the graceful degradation of performance that occurs when resource limitations prevent the schema from being filled. A set of complete schemas allows the route to be followed or described, using the *result* component as the forward pointer in a linked list, to the schema describing the next action and its result. If the *result* component is missing, the route can still be followed, but only within the physical environment, where the environment will produce the result of the action, and allow retrieval of the next schema. This level of performance appears to account for a common state of incomplete knowledge of a route, frequently described, “I could take you there, but I can’t tell you how.”
Preservation of one's goal in working memory during a sequence of actions has obvious survival value. Nonetheless, it is not uncommon to experience "capture errors" [Norman, 1981], when inattention leads a familiar procedure to "capture" the execution of a less familiar procedure. For example, I may be driving to do an errand on a Saturday morning, and inattentively discover myself arriving at work. This phenomenon has a plausible explanation in terms of the interaction between sensorimotor schemas and working memory. If the current goal is dropped from working memory, a given situation may evoke several (?,?,situation,action,??) schemas. With no reason to select among them, the most common is selected and executed.

Within the procedural map, route-finding can take place either by retrieving known routes, or by searching for concatenations of known routes. While this is adequate for many purposes, it does not provide any capability to search for novel routes, find short-cuts to optimize familiar routes, or cope without guidance with unknown environments.

1.3 The Topological Map

Where the procedural map is defined in terms of the egocentric sensorimotor experience of the traveller, the topological map is defined in terms of the fixed features of the environment — places, paths, and regions — and topological relations among them, such as connectivity, order, and containment.

Once the procedural map is sufficiently complete — i.e. there is a stream of sensorimotor schemas of the form (?,?,situation,action,result) — it becomes possible to assimilate the information in sensorimotor schemas into the topological map. Since it doesn't matter whether such a schema comes from a goal-indexed route description, or is simply a record of experience while wandering around, we omit the goal component from our schema notation.

The different topological relations in the map are defined in terms of sensorimotor experience, in such a way that the computational structures representing knowledge of those relations can be accumulated incrementally from the stream of observations.

There are two levels of the topological map in the TOUR model: (1) a topological network of places and paths, and (2) the containment and boundary relations of places and paths with regions.

1.3.1 The Network of Places and Paths

To build a topological map consisting of a network of places and paths, we need only consider actions to be described as Turn or Travel, ignoring any associated description of magnitude.

1. A Turn action leaves the traveller at the same place, even though it may change the current view. A place is identified with the set of views obtainable at it. The relation at(view,place) indicates that view is associated with place.

\[\langle V_1, \text{Turn}, V_2 \rangle \equiv \exists P[\text{at}(V_1, P) \land \text{at}(V_2, P)]\]
Figure 1: Places and paths make up a network of topological relations.

Since the place associated with a given view is unique, we may define a function \( \text{place}(V_1) = \text{place}(V_2) = P \). (In the real world, similar or identical views may be associated with different places. A generalization to handle such a situation is not difficult, but is beyond the scope of this discussion.)

2. Similarly, a Travel action does not leave the traveller at the same place, but to leave him on the same path. A path is a topologically one-dimensional subset of the environment. If a view lies on a path, we say \( \text{path} = \text{path}(\text{view}) \). When a view is on a path, we define a function \( \text{direction}(\text{view}, \text{path}) \rightarrow \{+1, -1\} \), which discriminates between the two directions one might be facing along a path. The connectivity between a place and a path is expressed by the relation \( \text{on}(\text{place}, \text{path}) \).

\[
(V_1,\text{Travel},V_2) \equiv \text{place}(V_1) \neq \text{place}(V_2) \\
\exists!\text{path} [\text{path} = \text{path}(V_1) = \text{path}(V_2)] \\
\text{on} (\text{place}(V_1), \text{path}(V_1)) \land \text{on} (\text{place}(V_2), \text{path}(V_2)) \\
\text{direction}(V_1, \text{path}(V_1)) = \text{direction}(V_2, \text{path}(V_2))
\]
3. We can induce a (partial) order relation on the places lying on each path, by defining a function 

\[ \text{order}(\text{place1}, \text{place2}, \text{path}) \rightarrow \{+1, -1, \text{nil}\}, \]

\[ \langle V_1, \text{Travel}, V_2 \rangle \equiv \text{order(place}(V_1), \text{place}(V_2), \text{path}(V_1)) = \text{direction}(V_1, \text{path}(V_1)). \]

These relations can be naturally translated to rules that respond to a stream of \( \langle V, A, V' \rangle \) schemas, and incrementally build up a collection of places and paths, asserting connectivity relations (at, on, place, and path) and order relations (direction and order) among them (figure 1).

This gives us a topological network of places and paths, which can be used to find novel routes among places using the usual graph-search algorithms. The collection of known \( \langle V, A, V' \rangle \) schemas could also be treated as a graph and searched, but at a finer granularity and restricted to schemas describing experienced actions, resulting in both less efficiency and less success at finding routes.

1.3.2 Regions, Boundaries, and Containment

As we have seen, a path defines an order relation over the places on the path. It also defines a binary classification over the places not on the path: those to the right and those to the left of the path (figure 2a).

A boundary is a sequence of one or more directed paths. Here, we will restrict our attention to a single path directed in its +1 direction, but the concept generalizes to sequences of segments, including simple closed curves. A region is a set of places; a boundary region is the set of places defined to be on one side of a boundary. A path is associated with two boundary regions: right(path) and left(path).

To extract boundary region information from sensorimotor experience and the procedural map, we need a slightly more detailed description of the Turn action, including a qualitative description of magnitude, to distinguish among:

- **Turn 0** — do nothing.
- **Turn Around** — ... to face the opposite direction along the same path.
- **Turn Right** — ... but not as far as a Turn Around.
- **Turn Left** — ... but not as far as a Turn Around.

Now, suppose we observe a sequence of sensorimotor schemas matching the following pattern:

\[ \langle V_1, \text{Travel}, V_2 \rangle \]
\[ \langle V_2, \text{Turn Right}, V_3 \rangle \]
\[ \langle V_3, \text{Travel}, V_4 \rangle \]
• (a) Each path divides space into its right and left boundary regions.

• (b) Boundary relations between individual places and paths can be acquired incrementally during travel.

Figure 2: Boundary regions
The rules for the place-path network would define the following places and their relations with the observed views (figure 2b):

\[
\begin{align*}
P_1 &= \text{place}(V_1) \\
P_2 &= \text{place}(V_2) = \text{place}(V_3) \\
P_3 &= \text{place}(V_4) \\
p_1 &= \text{path}(V_1) = \text{path}(V_2) \\
p_2 &= \text{path}(V_3) = \text{path}(V_4) \\
d(V_1, p_1) &= d(V_2, p_1) \\
d(V_3, p_2) &= d(V_4, p_2)
\end{align*}
\]

Membership in the boundary regions is then defined by:

\[
\begin{align*}
d(V_2, p_1) &= +1 \Rightarrow P_3 \in \text{right}(p_1) \\
d(V_2, p_1) &= -1 \Rightarrow P_3 \in \text{left}(p_1)
\end{align*}
\]

A similar rule can place \( P_1 \) to one side or the other of \( p_2 \).

Clearly, for each place and each path, the place lies either on the path, or in one of its boundary regions. The actual exploration history of the traveller determines which boundary relations are actually observed and assimilated. Once sufficiently many boundary relations have been accumulated, they provide a useful topological route-finding heuristic, one instance of which is:

**Boundary Heuristic:** To find a route from A to B, if \( \exists \text{ path} \) such that \( A \in \text{right}(\text{path}) \) and \( B \in \text{left}(\text{path}) \), look for routes from A to \( \text{path} \) and from \( \text{path} \) to B.

Chase [1982] and others have observed that many human cognitive maps are organized around a “skeleton map” of major streets within which most problem-solving takes place, with final links from the original start and goal points to the nearest point on the skeleton. The statistical frequency with which routes follow a minority of streets should give those streets a preponderance of the boundary relations, and thus determine which streets are treated as part of the “skeleton” for a given problem.

### 1.4 The Metrical Map

Metrical descriptions may be added to the map if we consider the quantitative components describing the magnitude of actions:

- (Turn \( \alpha \)), where \( \alpha \) describes the amount of rotation;
- (Travel \( \delta \)), where \( \delta \) describes the distance of travel.

For the metrical map, the TOUR treats these descriptions of magnitude as numbers or intervals.

The metrical map extracts this quantitative information acquired during travel, and assimilates it into two levels of metrical description:
The TOUR Model

1. Local Geometry — the locally-defined headings of paths radiating from a place, and the position of places within a locally-defined 1-D coordinate system of each path.

2. Orientation Frames — sets of places and paths whose local geometry has a known relation with a common coordinate frame of reference; supporting the use of two-dimensional vectors to define relative position and orientation.

1.4.1 Local Geometry

The conceptual prerequisite for the local geometry level of the metrical map is the network level of the topological map, not the topological map as a whole.

The directions radiating from a given place are associated with views. If we can determine the relative angles between views from the magnitudes of particular Turn actions, we can incrementally associate an absolute heading with each view, using coordinates locally meaningful at each place.

\[ (V_1, (\text{Turn } \alpha), V_2) \equiv \text{localheading}(V_2, \text{place}(V_2)) \]
\[ = \alpha + \text{localheading}(V_1, \text{place}(V_1)) \]

The assignment of local headings can take place incrementally if the first request for a local heading of a view at a place assigns that view the arbitrary heading of zero.

This association of local headings with views automatically induces an association of local headings with directed paths at each place.

\[ \text{localheading}(\text{path}(V), \text{direction}(V, \text{path}(V)), \text{place}(V)) = \text{localheading}(V, \text{place}(V)). \]

An analogous relation allows us to assimilate the relative distance between two places, and incrementally create a 1-D coordinate system associated with each path, giving an absolute 1-D coordinate to each place.

\[ (V_1, (\text{Travel } \delta), V_2) \equiv \text{localposition}(\text{place}(V_2), \text{path}(V_2)) \]
\[ = \text{localposition}(\text{place}(V_1), \text{path}(V_1)) + \delta \ast \text{direction}(V_1, \text{path}(V_1)) \]

(Recall that \( \text{path}(V_1) = \text{path}(V_2) \).)

From local 1-D coordinates, we can compute relative distances between places independently of having actually travelled between those two places:

\[ \text{pathdistance}(\text{place}1, \text{place}2, \text{path}) = |\text{localposition}(\text{place}1, \text{path}) - \text{localposition}(\text{place}2, \text{path})|. \]

Local geometry information is useful in problem-solving, making it possible to estimate the lengths of routes.
1.4.2 Orientation Frames

It is very useful for route-finding to be able to estimate the distance and direction from one place to another by integrating the distances and turns along the segments of a route leading from one to the other. This requires that relative distances and directions be in a common frame of reference, to allow vector addition. Unfortunately, in the local geometry, headings are strictly local to individual places, so comparisons are impossible.

However, a global orientation frame, such as the compass directions, can serve as a common frame of reference for multiple local coordinate systems at individual places. Globally meaningful headings $heading(V, OF)$ may be defined with respect to an orientation frame OF, with the constraint that the headings are compatible with observed turn information:

$$\langle V_1, (\text{Turn} \alpha), V_2 \rangle \equiv \forall OF[heading(V_2, OF) = \alpha + heading(V_1, OF)]$$

Note that orientation frames need not be genuinely global frames of reference such as the compass directions. Frequently, an orientation frame will be associated with a region of a city that has a regular rectangular pattern of streets, such as Back Bay in Boston, and ceases to apply at the boundaries of that region. In such cases, the regional orientation frame can be created by local propagation along paths, as neighboring places bring their local coordinate systems into correspondence.

This type of propagation is most likely to take place along straight streets. Incorporating a feature developed for the NX robot, we can generalize the quantitative information in a Travel action description to include a description of the net change in heading $\Delta \theta$ experienced during the action: $(\text{Travel} \delta \Delta \theta)$. For a straight street, of course, $\Delta \theta = 0$. This information imposes a second source of constraint on the headings asserted by an orientation frame.

$$\langle V_1, (\text{Travel} \delta \Delta \theta), V_2 \rangle \equiv \forall OF[heading(V_2, OF) = \Delta \theta + heading(V_1, OF)]$$

It is natural here to use an interval representation for incomplete knowledge of $\Delta \theta$. Travel along a straight street would yield $\Delta \theta = [0, 0]$, while a twisty mountain road might give $\Delta \theta = [-180, +180]$.

1.4.3 Vector Addition

It would be useful to be able to manipulate 2-D relative position vectors, as people clearly do in some cases. However, the extent of human capabilities along these lines is not at all clear, so the TOUR model takes no firm position on this topic, though several vector addition schemes have been implemented at various times. The problem is to define functions $distance(place1, place2)$ and $direction(place1, place2, OF)$ whose values represent the vector from place1 to place2 within the orientation frame OF, and to provide axioms corresponding to appropriate assimilation rules.
An assimilation process should yield 2-D distance and direction relations from experience such as

\[ (V_1, (\text{Travel } \delta \Delta \theta_1), V_2) \]
\[ (V_2, (\text{Turn } \alpha), V_3) \]
\[ (V_3, (\text{Travel } \delta_2 \Delta \theta_2), V_4) \]

Where \( \Delta \theta_1 = \Delta \theta_2 = 0 \), this is a straightforward vector addition. Where \( \Delta \theta_1, \Delta \theta_2 \neq 0 \), the quantitative description of Travel requires more detail, to capture the net straight-line distance covered by a curved path. It remains unclear (to BK) whether this is a realistic kind of knowledge for a person to accumulate, how it is inferred, how it is represented, and what its states of partial knowledge are.

Psychological studies of behavior related to vector addition have aroused considerable discussion about the existence of special-purpose spatial analog computational mechanisms in the brain, or at least in the mind. Part of this discussion was sparked by the dramatic results of Shepard and Metzler [1971] on mental rotation, leading to widespread interest in the properties of mental imagery (e.g. [Kosslyn, 1980]). Another part of this discussion arises from interest in the properties of the hippocampus, a portion of the brain with very regular anatomical structures, and closely tied with spatial reasoning and working memory, as shown by studies of cognitive deficits resulting from brain injuries [O'Keefe and Nadel, 1978].

1.4.4 Distorted Cognitive Maps

The values of heading\((V, \Omega l)\) are underconstrained by local observations in Turn and Travel actions. Not surprisingly, then, it is widely observed that human cognitive maps have "rubber sheet" distortions when compared with the actual environment, preserving topological and local metrical properties, but possibly grossly distorting global metrical distance and direction relations [Lynch, 1960].

As these relations are implemented with incremental assimilation rules, there are a number of heuristics, especially for dealing with quantitative observations, that can predictably lead to these distortions:

- In \((\text{Turn } \alpha)\), if \( \alpha \) is approximately equal to a right angle, assume it is exactly a right angle.
- In \((\text{Travel } \delta \Delta \theta)\), assume that \( \Delta \theta = 0 \).
- Given \((V_1, (\text{Travel } \delta \Delta \theta), V_2)\), assume that distance\((place(V_1), place(V_2)) = \delta\).
2 References


