Title
Where am I? Similarity Judgement and Expert Localization

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 13(0)

Authors
Smith, Kip
Heinrichs, Marian
Pick, Herbert

Publication Date
1991

Peer reviewed

Where am I? Similarity Judgment and Expert Localization

Kip Smith, Marian Heinrichs, & Herbert Pick, Jr.

Center for Research in Learning, Perception, and Cognition
University of Minnesota
205 Elliott Hall
75 East River Road
Minneapolis, Minnesota 55455
kip@umnsom.bitnet

Abstract

How do skilled map-readers use topographic maps to figure out where in the world they are? Our research addresses this question by studying the problem solving of experienced map-readers as they solve localization Where am I? - problems. Localization relies upon judgments of similarity and difference between the contour information of the map and the topographic information in the terrain. In this paper we discuss experiments that focus on how map-readers use attributes and structural relations to support judgments of similarity and difference. In our field and laboratory experiments, experienced map-readers implicitly define attributes to be detailed descriptors of individual topographic features. They use structural relations that link two or more topographic features as predicates. The time-course of their problem solving suggests that attributes and relations are psychologically distinct. Attributes like slope, e.g., "steep (hill)", support only initial judgments of difference. Relations like "(this hill) falls steeply down into (a valley)" are more powerful, supporting both judgments of difference and judgments of similarity. Judgments based on relations are used to test hypotheses about location. Experienced map readers exploit the distinction between attributes and relations as they solve localization problems efficiently.

Localization - the 'Where am I?' problem in navigation[1]

Localization is the familiar task of finding the point on a map that represents your viewpoint in the world. Anyone who has ever been lost knows that localization can pose a difficult problem. It is a fundamental component of all navigation in large-scale space. Diverse professions (e.g., geology and airborne infantry) require individuals to become skilled at localization. The work reported here elucidates the roles of topographic features, attributes of features, and relations among features in the judgments that establish correspondence between map and terrain.

Maps are representations that preserve with fidelity a selected subset of the information available in a section of the world. The information contained in a map provides a context for the map-reader: localization judgments based upon a map can only be made with reference to the type of information it makes available. We restrict our study to topographic maps because they provide a clear, familiar, and pragmatically useful context for constructing a theory of localization that will assist the design of intelligent systems to control vision-based robot navigation.

When using a topographic map to solve a localization problem, the map-reader must find the location among the contours that matches the *viewpoint* in the terrain. The viewpoint is the location in the world where one happens to be standing. It determines what can be seen and what is occluded. The viewpoint dependence of terrain information tightly constrains problem solving. It determines the topographic features and relations among those features that can be used to generate and test hypotheses about location. Map-readers necessarily relate terrain information to their viewpoint.

The constraint of viewpoint dependance and the context provided by topographic maps transform localization into the task of finding the contours on the map that characterize a layout similar to that seen from the viewpoint. Determining the correspondence between map and terrain relies on judgments of similarity and difference between the contour information of the map and the topographic information in the terrain.

There are frequently many locations on a topographic map that appear similar to the viewpoint. Each may be entertained as a hypothesis. Selecting the hypothesis that provides the best match to the terrain relies on judgments that discriminate among competing hypotheses. Thus, there are two basic steps to localization problem solving: (1) generating hypotheses that relate map and terrain information and (2) testing these hypotheses by identifying the best match. Similarity judgment is essential to both.

Localization and similarity judgment

Gentner (1983) and Medin, Goldstone, and Gentner (1990) emphasize the role of relations among objects in judgments of perceptual similarity. Their notion of structure-mapping holds that the relations among objects constrain judgments of similarity and are, in fact, more central to the process of judgment than are the individual objects themselves. This emphasis on the structure that relations impose on their constituent objects is intuitively consistent with the correspondences that map-readers must make to compare a map to the terrain they see.

Tversky (1977) introduces the notion that judgments of similarity depend on the context of the task in which they are embedded. He specifies a rule for calculating similarity. He implies that application of the rule is dependant on the task context but does not indicate specifically how. Medin et al. (1990) seize on this insight and suggest that the

[1]This work was supported by NSF Grant IRI-8901888 (funded in part by DARPA) and AFOSR Grant 88-0187.

relational structure among objects provides the context missing in Tversky's (1977) model. This too matches the demands of the localization task. Topography provides not only a context but also an intrinsic structure within which a map-reader views both the terrain and the map.

A key component of the structure mapping hypothesis is a fundamental distinction among objects, relations, and attributes. We embrace this distinction. In this paper, we call the topographic objects that capture a map-reader's attention *features*, e.g., "I see a valley". An *attribute* is a property, like gradient, that embellishes the description of a feature, e.g., "I see a steep valley". A *relation* is a connective property that cannot be hung on any one feature; relations span two or more features, e.g., "When I look southeast, I see the ground falls abruptly into a valley". In this example, the relation is a predicate that links the map-reader's viewpoint to a distant feature.

Since localization is a veridical task, individuals who have developed this skill are readily identifiable. They include professionals who make their living finding their way around the world using topographic maps (e.g., geologists and wilderness guides) and serious recreationists (e.g., orienteers and outfitters). By investigating the problem solving of experienced map-readers as they solve localization problems, we gain insight into the methods used in efficient localization problem solving.

These considerations lead us to believe that studies of localization problem solving can shed light on three current issues in similarity judgment: (1) the claim that the structure of relations among features is more vital to these judgments than features taken independently, (2) the utility of making a distinction between relations and attributes, and (3) the processes by which these judgments are made.

Experiment 1: Field studies

The goal of Experiment 1 was to address these issues using as data the thinking-aloud reports (protocols) of experienced map-readers solving a localization problem (Thompson, Pick, Bennett, Heinrichs, Savitt, & Smith 1990).

Method

Subjects. A total of 29 experienced map-readers including professional geologists, champion orienteers, and wilderness guides participated in Experiment 1.

Procedure. Individual subjects were blindfolded and driven approximately 30 miles to a road access about one-quarter mile from the station point: the point in the terrain to be found on the map. They were led across a level field and up a hill to the station point. Once there, the blindfold was removed and they were given a topographic map attached to a clipboard. The map is a cropped U.S.G.S. topographic map from which all non-topographic information (culture) has been deleted. The map contains only contour information about elevation.

The station point is a roughly circular hill that is the westward extension of a larger highland. A distinctive attribute is its steep slope to the southwest. A second round hill with a similar orientation is selected as an alternative hypothesis by all subjects. This hill forms a garden path hypothesis (Johnson, Moen, & Thompson 1988): its similarity to the correct solution and its position in the center of the map lead many subjects to consider it early in their problem solving. Both have a pond to the north. Other alternatives are also considered by most subjects. Some subjects consider as many as eight different alternatives. Selection of the correct solution does not appear to depend on the number of alternatives considered.

The subjects' task was to find their viewpoint on the map. During the drive to the site, they had been briefed on the procedure and instructed to think aloud and to point to what they were talking about as they addressed the task. Subjects spent an average of 45 minutes on the task.

Subjects' verbal reports were recorded as they thought aloud. Simultaneously, their behavior was videotaped to provide information about where they were looking (and pointing) while thinking aloud. The verbal reports were transcribed and the composite audiovisual protocols coordinated and scored. Scoring focused on two aspects of problem solving: on the type of information attended to and how that information was used.

The scoring procedure identifies the source of information, the map or the terrain, and three categories of information - features, relations, and attributes. We define *features* as individual topographic objects that our subjects identify with a familiar count noun, e.g., hill, valley, pond. Each subject's lexicon is small and consistent. The composite lexicon across subjects provides a taxonomy of useful topographic terms. *Attributes* are properties that modify individual features. Subjects tended to use bipolar, qualitative attributes to differentiate among similar features, e.g., narrow or wide, steep or shallow. *Relations* are connectives that conjoin two or more features into a single structural unit we call a *configuration* of features. Some relations are purely topologic connectives, e.g., behind, below. Most configurations are expressed by qualitative predicates, e.g., "and then it (feature 1) gets steep down into (feature 2)". Use of quantitative relations, e.g., higher than, a mile apart, is less common.

Configurations constrain problem solving more effectively than do individual features. For example, there are fewer matches to "a high spot going down steeply to some lakes" than to an individual hill or pond. Distinguishing among features, attributes, and relations is consistent with the arguments made by Gentner (1983) and Medin et al. (1990).

Analysis of the protocols also identifies components of the problem solving process. Three of these processes involve judgments of similarity and/or difference. Localization problem solving is initiated by an extended period of reconnaissance. *Reconnaissance* identifies features, attributes, and relations for subsequent processing. Subjects return to reconnaissance to gather additional information. *Matching* is a form of argument that marshalls evidence that the features or configurations seen in the terrain correspond to those seen in the map, or vice versa. *Hypothesis generation* is an explicit statement about a particular location on the map that may represent the viewpoint. Localization concludes with wholesale acceptance of a hypothesis.

Condition 1. The 17 subjects in the first condition were instructed to remain at the station point as they attempted to solve the problem. They were permitted to move a few feet in turning around. As this task proved extremely difficult, a second condition was introduced.

Condition 2. In the second condition, the 12 subjects were free to walk about and to explore the terrain.

Results

Solution. Of the 17 stationary subjects, only one arrived at the correct solution. Six of the 12 exploring subjects arrived at the correct solution. This difference in performance is significant, $\chi^2 = 5.60$, $p < 0.05$, df = 1.

Judgments of similarity and difference. All subjects begin by identifying salient features from the terrain and the map. They may begin with the terrain and move to the map, or begin with the map and move to the terrain. Subjects may identify a large number of features or key on a few salient features. The subject highlighted in Table 1 begins by describing his own position in the terrain as a relatively high area and identifying similarly high areas in the map (lines 4-5). Based on the few features he extracts in the first 25 seconds of reconnaissance, he generates a pair of hypotheses (lines 10-11). One of these hypotheses is the correct location on the map. The second is the garden path hypothesis.

Reconnaissance followed by hypothesis generation is typical of highly proficient subjects. Identification of features appears to be sufficient to generate informed hypotheses. Many subjects spend considerable time identifying features and assembling configurations of features before generating hypotheses. Subjects then proceed to focus on relations and judgments of similarity and difference to evaluate those hypotheses.

Single attributes often provide sufficient information to judge that a map feature cannot stand for a terrain feature. That is, difference judgments are often based on single features. An example of a judgment of difference based upon an attribute is shown in Table 1 (lines 87-90).

As shown in Table 1, the subject follows his generation of hypotheses with the assembly of several configurations, one of which is contained in lines 15 to 19. He conjoins his description of his viewpoint, "a knob", to the "stream valley below" with the predicate "gets pretty steep down into". The steep descent from his knob to the stream valley becomes a structural constraint on similarity judgment.

After assembling several other configurations both in the terrain and the map, he proceeds to attempt to match them. This matching necessarily entails judgments of similarity. One such match is shown in Table 1 (lines 38-43). He begins by reiterating a configuration extracted from the terrain (line 38). He turns his attention to the map to match features constrained by the same relation (lines 41-43). He then judges the two configurations to be sufficiently similar to support the hypothesis.

TABLE 1 SIMILARITY JUDGMENT IN LOCALIZATION

key: 4 - line number; M - map information; T - terrain information

1 Identify features including viewpoint, relations, and attributes. Match features to guide assembly of configurations.	4	T	All right, well I noticed I'm at one of the higher points within this area, so that's important.
	5	M	So I'm first looking on the map, for some higher points on the map.
2 Assemble configurations - descriptions of the topographic layout of relations among features including the viewpoint	15	T	but what I was actually looking at is how steeply the hill drops off.
	16	T	And it's kind of a knob right here we're standing on,
	17	T	and then generally not very steep
	18	T	and then it looks like it gets pretty steep down into a valley to the east.
	19	M	Ok, so I'm looking for the same types of things on the map.
3 Generate viewpoint hypotheses	10	M	Um, for example say, somewhere here (HYPO 1 - CORRECT),
	11	M	or on a hill here (HYPO 2 - GARDEN PATH).
4 Eliminating alternatives using attributes to make judgments of difference	87	M	And, see I'm kind of looking up here (HYPO 3)
	88	M	'cause this also has a hill
	89	T	but um, ... Now straight to the north it should be quite steep
	90	M	So, that doesn't seem likely. (REJECT HYPO 3)
5 Matching configurations using relations among features to make judgments of similarity	38	T	I'm at a high point. Directly north is a fairly flat area and north of that it get's steep and then there's the lake, ok.
	39	T	and I'm trying to match those features with what I see here on the map.
	40	M	Ok, ah, for instance, again. Let's go back to this place (HYPO 1) , ok, so here's a higher area.
	41	M	Here's a generally flat area.
	42	M	Then it goes down steeper here and it looks like there's valley coming through here.
	43	M	And then possibly some lakes or ponds here.
6 Comparing hypotheses using relations to make judgments of difference	47	M	Ok, I still kind of like this area (HYPO 1),
	48	M	But then I was looking up on the map. I also have a high here, (HYPO 2)
	49	M	and with (a pond?), .. that's not very far at all.
	50	M	It just doesn't seem to work well
	51	M	because there's a fairly steep and long gradient here before you get to a flat part
	52	T	and I don't see that where we're standing.

This is a consistent pattern in our protocols. Judgments of similarity are used to support hypotheses. They are also used to compare hypotheses. One such comparison is shown in lines 47-52. In this passage, a map configuration relating a high area and a pond is compared with the terrain configuration stated in line 38. This relation is a suitably strong constraint to reject this alternative.

Summary

The field protocols reveal the critical role of similarity and difference judgments in localization problem solving. Of the six components itemized in Table 1, the final three involve similarity and difference judgments. Topographic relations that link features (including the viewpoint) support both similarity and difference judgments. Attributes of features support difference judgments.

This difference in power between relations and attributes may explain the difference in performance between the stationary and exploring conditions. Subjects who were allowed to explore the terrain had better access to information in general and better information about their viewpoint in particular. They used this information to assemble richer configurations that included the viewpoint. The experimental manipulation cannot distinguish whether the information from the viewpoint or about the terrain at large is the more valuable for successful localization.

Experiment 2: Laboratory studies

Experiment 2 is a laboratory simulation of the localization task in which the amount of map information available to the subjects was manipulated. In an earlier study, maps were masked so as to obscure a portion of the map (Heinrichs, Montello, Nusslé, & Smith 1989). There were three masking conditions in that study: the mask covered either the inner one-third, the outer two-thirds, or the outer one-third of the map. The control condition used an unmasked map. Subjects were presented with one of the four conditions. Five pairs of masked and control conditions were selected for the present experiment.

The goal of experiment 2 was to determine whether map information around the viewpoint is generally more informative than other regions of the map. A second aim was to elucidate better the different roles of relations and attributes. The third aim was to examine a variety of locations in order to generalize beyond the single location used in the field experiment.

Method

Subjects. Ten subjects from the same pool of experienced map-readers participated in Experiment 2.

Apparatus. Subjects were presented with topographic maps of five locations, two in Minnesota, two in New Mexico, and one in Arizona. As in Experiment 1, the maps are enlarged and cropped copies of U.S.G.S. topographic maps with all culture removed. The maps were marked with a single point at the center of the map that identified the location from which color slides were taken. The slides were taken with a camera mounted on a tripod at eye level. The line of sight was horizontal. A complete set of twelve pictures covered the whole 360^{o} panoramic view. For the experiment, either one view or three non-overlapping views were selected for presentation to the subject. Slides were presented on a rear projection screen in a darkened laboratory room. The subject sat 120 cm away from the screen with a reading lamp illuminating the map from behind. A remote control allowed the subject to advance or reverse the slides at will.

Procedure. Subjects solved two types of localization problems. In the first, one arrow was drawn on the map leading from the center point. Subjects used a remote control to view the three slides. Their task was to select the slide that corresponded to the terrain that would be seen looking in the direction of the arrow on the map. In the second type of problem, three arrows separated by 120^{o} were drawn on the map leading from the center point. Subjects were shown only one slide. Their task was to select which of the arrows on the map corresponded to the view of the terrain in the slide. These procedures presented options that subjects could entertain as hypotheses.

As in the field study, subjects were asked to think aloud while solving the problems. Collection of concurrent verbal reports, videotaping, and scoring of the resulting protocols followed the procedures of Experiment 1.

Condition 1. To investigate whether information about the viewpoint is favored over information from more distant areas, map information was selectively masked in the first condition. In four of the five trials (one trial for each set of maps and slides) a black circle masked the inner 1/3 of the map. In the fifth (Arizona) a black annulus masked the outer 2/3 of the map.

Condition 2. As a control condition, subjects also solved the same set of five tasks in a 'full map' condition in which the masks were removed from the maps. Each subject solved the problems twice, first in the masked condition and in the full map condition. This manipulation allowed within-subjects and within-location comparisons.

Results

Solution. Accuracy was significantly better in the full map condition (66%) than in the masked condition(44%), $t(9) = 3.16$, $p < 0.05$. In the full map condition, performance is significantly different from chance, $t(9) = 6.27$, $p < 0.001$, but not in the masked condition.

Judgments of similarity and difference. In this section we compare judgments within subjects and across conditions for three of the five tasks.

In the first task, subjects viewed one slide of rolling terrain typical of southeastern Minnesota. The major discriminating feature in the slide is a prominent hill. Many subjects find the hill so salient that they base their judgments on a match to this feature. They were given a map with three arrows to choose among. In the masked condition, the mask covers the inner 1/3 of the map and totally obscures the prominent hill. One of the three arrows crosses a hill in the unmasked region of the map. Subjects who spend a disproportionate amount of time on the slide select this (incorrect) arrow. The salient hill leads them down a garden path. They ignore information about the relation of distance between the viewpoint and the feature and are led to an incorrect solution.

In the full map condition all subjects select the correct answer. The availability of information near the viewpoint pulls their attention to the distances between the viewpoint and the various hills along the arrows. As only one of these distances is similar to what is seen in the slide, the correct arrow is selected.

In the second task, subjects viewed one slide of mountainous Sonoran Desert terrain and were given three arrows on the map to choose among. In the masked condition, the mask covers the outer two-thirds of the map. This task is unique in that most subjects correctly answer both the masked and full-map conditions.

Subjects focus their attention on the orientation of a series of small ridges and valleys. It is clear that the relation of parallelism among these features (not including the viewpoint) is sufficiently diagnostic to raise only one of the offered choices to the level of a hypothesis.

The full map condition produces a second finding. It reveals a high hill in the distance to the left of one of the arrows. Subjects find that the hills in the slide are not as high and immediately eliminate that arrow from further consideration. This result supports the inference that attributes are sufficient to support judgments of difference.

In the third task, the inner one-third of the map is occluded and one arrow is shown on the map. Subjects viewed three slides of an area adjacent to a large river valley in eastern Minnesota. Their task is to select the slide that contains the terrain they would see looking in the direction indicated by the arrow on the map. The mask covering the viewpoint makes it appear as though the viewpoint is within a valley and the viewing direction is up at a hill. The viewpoint is actually on the crest of a small ridge that is completely obscured by the mask.

One of the slides contains a long gentle slope up to a distant hill. In the masked condition, many subjects make a reasonable assumption and incorrectly select this slide. By occluding the viewpoint, the masked condition eliminates vital information about the distribution and relations among features in the terrain and prevents correct solution. Correct solution of this localization problem clearly requires matching on the basis of relations of features that include the viewpoint.

Summary

The first task shows the superiority of a judgment of similarity based on a configuration over a judgment based solely on a salient feature. The second task also shows the superiority of a judgment of similarity based on a configuration over a judgment based solely on a feature. In addition, it reveals reliance on the attributes of a feature to justify a judgment of difference. The third suggests that successful localization often requires full knowledge of the relations that tie the viewpoint to nearby features.

Discussion

Experienced map-readers adopt a basic generate and test strategy to solve localization problems. They move in either direction, from map to terrain and from terrain to map, as they attempt to figure out where in the world they are. Judgments of similarity and difference inform both the generation and testing of hypotheses.

The two experiments reveal that structural relations of features play a key role in both the generation and testing of localization hypotheses. They also show a fundamental difference in the roles played by relations and attributes. The protocols reveal that experienced map-readers make this distinction. Attributes are used to make preliminary judgments about potential hypotheses (Table 1, Section 4) whereas relations are used to scrutinize hypotheses (Table 1, Sections 5 & 6). Features and relations guide the assembly of configurations. Attempts to match map and terrain configurations inform hypothesis testing. These tests rely on judgments of the similarity of relations. Relations are also used in judgments of difference to discriminate among competing hypotheses. In contrast, attributes are used only for judgments that either eliminate an alternative or raise it to the status of hypothesis.

Three questions remain: In judgments of topographic similarity and difference, are some *relations* more important than others? In the judgments of topographic difference, are some *attributes* more important than others? How do these vary with the nature of the terrain?

The roles played by relations and attributes in judgments of similarity and difference are part of a larger theory of localization problem solving. This theory is to be embodied in a system designed to control the navigation of vision-based mobile robots in dynamic environments.

References

Gentner, D. 1983. Structure-mapping: A theoretical framework for analogy. *Cognitive Science* 7:155-170.

Heinrichs, M.R, Montello, D.R., Nusslé, C.N., and Smith, K. 1989. Localization with topographic maps. In Proceedings of the AAAI Symposium on Robot Navigation, 29-32. Stanford, California.

Johnson, P.J., Moen, J. B., and Thompson, W.B. 1988. Garden path errors in diagnostic reasoning. In Bolc, L., and Coombs, M.J. eds. *Expert Systems Applications*. New York: Springer-Verlag.

Medin, D.L., Goldstone, R.L., and Gentner, D. 1990. Similarity involving attributes and relations: Judgments of similarity and difference are not inverses. *Psychological Science*. 1(1):64-69.

Thompson, W.B., Pick, H.L., Jr., Bennett, B.H., Heinrichs, M.R., Savitt, S.L., and Smith, K. 1990. Map-based localization: The "drop-off" problem. In Proceedings of the DARPA 1990 Image Understanding Workshop, 706-719. Washington D.C.

Tversky, A. 1977. Features of similarity. *Psychological Review* 84:327-352.