Representing the Local Space Qualitatively in a Cognitive Map

Margaret Jefferies (jefferies_me@cs.otago.ac.nz)
Artificial Intelligence Laboratory
Computer Science Department
University of Otago, Dunedin, New Zealand

Wai-Kiang Yeap (yeap_wk@cs.otago.ac.nz)
Artificial Intelligence Laboratory
Computer Science Department
University of Otago, Dunedin, New Zealand

Abstract

The cognitive maps that humans compute as representations of the spatial environment they have visited are rarely even close approximations to what was actually experienced. When we experience the environment we seem to see it all so perfectly, yet rarely are we able to reproduce from memory an exact description of the places visited. Yet these vague, muddled descriptions of the places visited are adequate for many spatial reasoning tasks. But how is such an impoverished representation computed from what is initially delivered by one's senses? And what effect does this representation have on the construction of the cognitive map? We present one method for computing a vague description of each local space visited. It is derived from the initially accurate description needed for the actions the viewer might perform within the local space. We show the effect of this representation on the structure of the cognitive map.

Introduction

Urban planners, geographers and environmental psychologists have long been interested in the human's perception of the environment (e.g. Downs & Stea (1973) Lynch (1960)), and Artificial Intelligence researchers have in the past attempted to develop computational models of the underlying processes (e.g. Davis (1986), Engelson (1994), Kortenkamp (1993), Kuipers (1996), Yeap (1988)). The term cognitive map has been widely used to describe the representations which result from these processes. Since the term was first coined by Tolman (1948), there have been numerous theories proposed to describe these representations (e.g. Downs & Stea (1973) Gallistel (1989)).

In our own laboratory a computational theory of cognitive maps has been developed which seeks to explain what information is made explicit at each step in the cognitive mapping process starting at the bottom with what is delivered by vision (Yeap, 1988; Yeap, Holmes, & Jefferies, 1994; Yeap & Jefferies, 1997; Yeap, Naylor, & Jefferies, 1990). We are interested in how spatial knowledge develops from the most primitive representations computed directly from the senses and on through various stages to ever more sophisticated representations.

The predominant theory regarding the development of spatial knowledge is that of Siegel and White (1975) – it suggests that the progression of spatial knowledge in a cognitive map is from landmark to route to survey map. The essence of this theory is that landmarks are remembered first and this is followed by an initial topological network and then a much expanded one and finally euclidean information becomes available. The main argument in favor of this theory is the experimental evidence which shows that as one's knowledge of the spatial environment progresses landmark knowledge is far more accurate than route knowledge and likewise route knowledge is more accurate than survey knowledge (Cousins, Siegel, & Maxwell, 1983; Lee & Schmidt, 1988; Moeser, 1988). Given the complexity of computing survey knowledge as compared to landmark knowledge such a finding is not surprising. However it does not necessarily follow that route and survey knowledge cannot be computed simultaneously alongside landmark knowledge (see Yeap & Jefferies (1997) for a detailed discussion). Recently, Montello (1993) criticized the landmark/route/survey hypothesis arguing that pure landmark or route knowledge always coexist with metric knowledge about distance and direction, however vague they may be. Metric knowledge begins to be acquired the first time one encounters an environment and like all spatial knowledge the quantity and quality of the information stored improves with repeated exposure to its source. Indeed if one considers the information which is delivered by the senses as for example in vision, then metric knowledge is as computable from this input as landmarks or routes and is abundantly available when compared with landmark knowledge.

Humans "see" the places they visit with much precision. Even though we can easily determine where objects are when we are looking directly at them, rarely are we able to reproduce from memory an exact description of the places we have visited. Not often do we remember exactly where objects are when we cannot physically see them, and nor do we easily remember their exact size and distance to other objects. Yet we are able to make good use of the vague and imprecise memories we have for our environment. The decisions we make on how to get from one place to another are often based on rather sketchy memories for the places we have visited along the way. Continuing on with our bottom-up approach to computing a cognitive map our concern is with how representations of such poor quality are computed from one's seemingly rich and detailed experience of the environment.

Our recent work on cognitive mapping has been concerned with deriving a representation for the viewer's local space from visual input (Yeap, et al., 1990; Yeap, et al., 1994; Yeap & Jefferies, 1997). This is a fundamental step in the construction of a cognitive map and we have shown how this could be done. Our current algorithm emphasizes the importance of detecting exits in view from the surfaces perceived, and from these exits, a boundary of the local space is computed. Each local space is computed using a cartesian coordinate reference frame and this provides an adequate representation for the task of determining where things are when the viewer is looking directly at them. The representation is called an Absolute Space Representation (ASR), a term which emphasizes the independent, local nature of each local space visited. Once computed the individual ASRs can be connected together in the way they are experienced to form a cognitive map. Computing the local space in this way is a necessary first step but once one moves out of the current space, it is evidently clear that, at least for humans, one does not remember the exact details of its shape. How then would the representation devolve into what the viewer actually remembers, i.e. a vague representation of what was initially computed and if such a representation is computed, what effect does this have on the construction of a cognitive map? In this paper we describe one method we have devised for devolving the initial description of the local space into a much simplified representation and we show the effect of using it to build a cognitive map. We have written computer programs to test our ideas. The results from two experiments which simulate a viewer with a 150^0 view moving through a complex 2D environment are presented.

In recent years the qualitative spatial reasoning (QSR) community has proposed many methodologies for representing and reasoning with this vague and uncertain spatial knowledge (Clementini, Di Felice, & Hernandez, 1995; Cohn, Randell, & Cui, 1995; Egenhofer & Khaled, 1992; Hernandez, 1993; Rohrig, 1994). In the next section, on related work, we examine some of their methods and briefly mention one cognitive mapping implementation which makes use of qualitative representations and reasoning techniques (Kuipers, 1996).

Related Work

Cohn & Gotts (1996b) state that: "The challenge of QSR is to provide calculi which allow a machine to represent and reason with spatial entities of higher dimension, without resorting to the traditional quantitative techniques prevalent in, for example, the computer graphics or computer vision communities". Much of the motivation for QSR comes from the fact that accurate metric representations grossly over-specify the accuracy which is achievable. For a robot the sensors employed are incapable of delivering such accuracy. Humans on the other hand are able to compensate for this with high level conceptual knowledge about how their environment should appear to them, however the problem for them is that the representation seems to quickly degrade once there is no longer immediate feedback from the environment. A common theme is apparent across many of the QSR models – spatial relations are defined in such a way that the various relationships between objects can be distinguished at an appropriate but usually coarse granularity. Given some facts about pairs of objects the transitivity of the spatial relations is exploited to infer further facts. For example, if A is in front of B and B is in front of C then it can be inferred that A is also in front of C. The possible inferences are usually stored in composition tables.

For Hernandez, Clementini and Di Felice (Hernandez, 1993; Hernandez, Clementini, & DiFelice, 1995), the key to defining qualitative spatial relationships is to make explicit just those spatial relations necessary for a particular context, thus eliminating unnecessary detail. Within the local space orientation is expressed using relations such as *left, right, front,* and *back* (for a level with four distinctions); *left-back, right-back, left-front,* and *right-back,* would be added for a level with eight distinctions. In large-scale geographic space the absolute orientation relations *north, south, east, west,* etc would be used. Eight topological relations are defined –*disjoint, tangent, overlaps, contains-at-border, included-at-border, contains, included,* and *equal.* Hernandez (Hernandez, 1993) defines neighboring structures to simplify the process of calculating the composition of pairs of relations. In Hernandez et al. (1995) a methodology for representing and reasoning with qualitative distance as it pertains to large-scale geographic space is developed. The problem in representing distance relations is that our notions of vague terms such as *far* and *near* are defined by the context in which they are used. To define a distance relation Hernandez et al. (1995) specifies a *primary object,* a *reference object,* and a *frame of reference.* Distance relations can be specified at various levels of granularity, for example a level with four distinctions would comprise *very close, close, far* and *very far.* Thus a system of distance relations is defined along with a set of structure relations which define how the distance relations relate to each other. The composition of spatial relations as in A *far* B and B *close* C results in a relation between A and C over some range of distances specified by a lower and upper bound.

Cohn et al.'s (1995) RCC-theory defines some basic relations to specify the connectivity of a pair of spatial regions. RCC-5 has the set of five relations, *partially overlapping, proper part, equal, proper part inverse,* and *distinct regions.* A finer granularity of relationships is achieved by splitting individual members in the set into two or more disjoint relations. For example, in RCC-8, the set of eight relations, *partially overlapping* is split into *tangential proper part* and *non-tangential proper part.* In (Cohn & Gotts, 1996a) the RCC-theory is extended to encompass regions with indeterminate boundaries. While Cohn's group has mostly been concerned with qualitative representational issues, Bennett (Bennett, 1994) has proposed a method for automatically generating the entries in the composition table of RCC relations using propositional logic.

One implementation that does compute qualitative representations from sensory information is that of Kuipers (1996). His Spatial Semantic Hierarchy comprises five layers from sensorimotor to metric level with the topological level being immediately prior to the metric level. Assimilation of knowledge proceeds from layer to layer with each layer providing the properties that the next one depends on. The topological layer consists of places, paths and regions along with connectivity and containment relations. Following on from the topological layer, the metric layer adds metric attributes so that places, paths and regions are linked by metric relations, such as relative and absolute angles and distances, according to some framework.

Computing a Qualitative Local Space Representation

Determining the ongoing nature of spatial memories when they are no longer receiving immediate feedback from the environment is not easy. Studies which examine this problem are mostly concerned with the manner in which the representations are distorted and their significance altered once they are merged into the wider "picture in the head". Variables such as size, distance and location are often systematically distorted by containment relations and the significance of an object as compared with others it is related to (Hirtle & Jonides, 1985; Holyoak & Mah, 1982; Sadalla, Burroughs, & Staplin, 1980; Tversky, 1992). But these modifications result from some top-down processing, i.e. the input to the process isn't only what has been computed bottom-up from the senses but also includes the results of earlier computations, often higher-level representations which are conceptually more sophisticated.

Our concern at this stage is only with what can be computed bottom-up from the senses. It is our contention that the initial representation computed for a local space is computed for the viewer's immediate needs, to provide a locus for the objects surrounding the viewer, and the activities which involve these objects. But while much of the detail is forgotten or goes unnoticed one can still remain cognizant of the local space for a long time after it was occupied. To encapsulate a local space in this way would only require representing its extent in very rough terms, but the resulting representation could still provide a useful framework for reasoning about the local space and could then become more elaborate as the viewer's familiarity with the environment increases. Such a representation is "qualitative" in the sense of QSR; the representation may never be isomorphic to the actual environment and the "quality" of the information represented would be extremely variable.

We compute such a representation by devolving the initial representation computed into a rectangle which roughly approximates its extent. A straightforward algorithm is used - points on the surfaces forming the boundary of the ASR are sampled to firstly find a good length for the rectangle and then the length itself is sampled to find a good width. We call this representation a fuzzy ASR. Figure 1 (a) shows an initial ASR computed, its surfaces are labelled s1 – s5 and its exits e1 – e4 (for a detailed description of this algorithm, see Yeap & Jefferies (1997)). The fuzzy ASR computed from this ASR

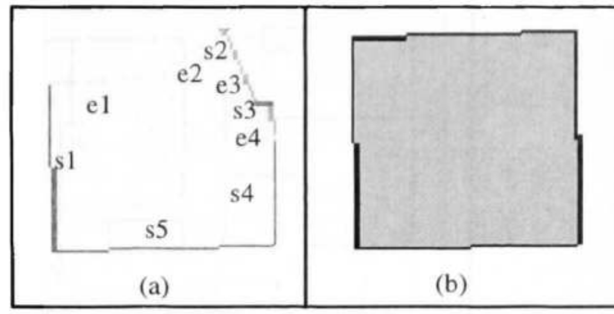

Figure 1. (a) An exact ASR computed while the viewer occupied the local space. (b) the fuzzy ASR description which devolves from the ASR in (a)

is shown in Figure 1 (b). No claims are made as to the plausibility of our method. In reality many processes would be operating to modify the original ASR and we cannot claim to fully understand these. This is but one method for producing a fuzzy ASR. There will be many, many more.

The real significance of the fuzzy ASR for our computational theory is the manner in which the representation is able to be used to structure the cognitive map, however poorly. The fuzzy ASR does not comprise actual surfaces or exits, it merely represents a portion of space once occupied by the viewer. But one would expect the viewer to remember some of the connections to neighboring spaces, confused though they may be. Thus we retain the connections to neighboring "ASRs", but only in the loosest sense. We conducted two experiments with the program by varying the amount of knowledge the viewer retained for the connections between ASRs.We thus showed how a fuzzy cognitive map might be structured and how useful such a map might be.

In the first experiment the viewer remembers how many exits there are in an ASR but no locational information is retained for them. For the fuzzy ASR in Figure 1 (b), for example, the viewer remembers just that there are four exits, e1, e2, e3 and e4. When the ASR is exited a connection is made to the ASR just entered but the viewer does not remember which exit was used. Our viewer has a very poor memory indeed! The outcome of this is a scenario often faced by humans "I know I've been here before so which doorway did I use to get to..." Thus the information made explicit in a fuzzy-ASR comprises the rough extent of the ASR, the number of exits in the ASR and which neighboring ASRs have been experienced as connected to this one. The results of the experiment are displayed in Figure 2. Figure 2 (a) shows the portion of the environment traversed and Figure 2 (b) a cognitive map constructed from the "exact" ASRs computed for each local space visited. Note that although for display purposes the ASRs are laid out as if there is one global coordinate system, in reality this is not the case. Each ASR is independent of all others with its own local coordinate system, and the only links to other ASRs are through the exits used to traverse them. Thus the viewer knows exactly where each surface and exit in the ASR is located, and exactly which exits are used to connect to neighboring ASRs. The actual structure of the cognitive map is mostly route-like, except where previously visited ASRs are able to be recognized (see Yeap, Jefferies, & Naylor (1991)). These parts of

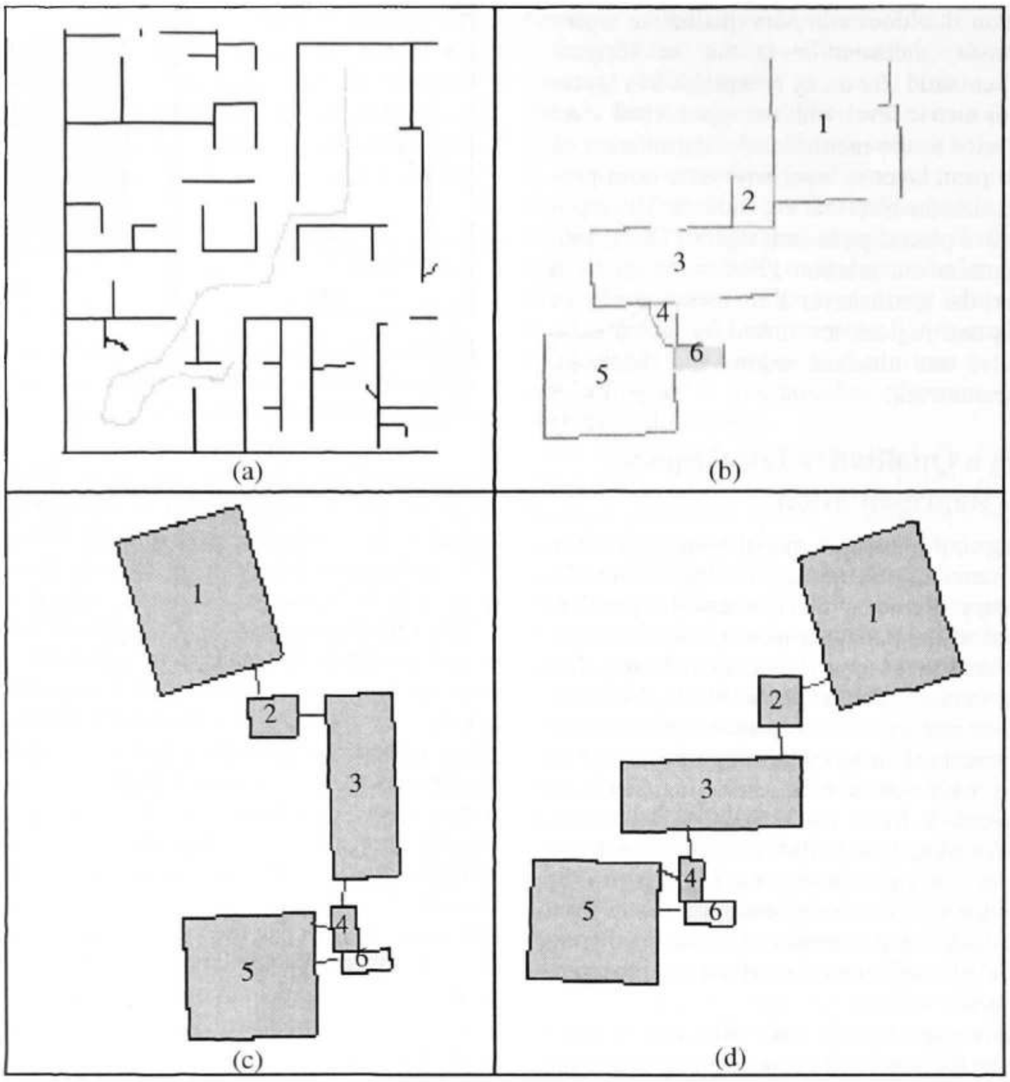

Figure 2. (a) The environment traversed, (b) a cognitive map computed from exact ASRs (c) the viewer's interpretation of a cognitive map computed from fuzzy ASRs where the viewer has no locational information for the exits. (d)the viewer's interpretation of a cognitive map constructed from fuzzy ASRs where the viewer knows on which side of the fuzzy ASR the exits are located but not their exact position.

the map exhibit a more integrated structure. Figures 2 (c) and (d) convey the underlying structure of the fuzzy cognitive map more realistically but this is only practical for a small number of ASRs. In all the figures the ASRs are numbered in the order in which they are visited.

The fuzzy cognitive map constructed for the path in Figure 2 (a) would comprise:

fuzzy-ASR1 with four exits, connected to ASR 2
fuzzy-ASR2 with two exits, connected to ASR 3, ASR 1
fuzzy-ASR3 with five exits, connected to ASR 4, ASR 2
fuzzy-ASR4 with three exits, connected to ASR 5, ASR 3
fuzzy-ASR5 with four exits, connected to ASR 6, ASR 4
fuzzy-ASR6 with three exits, connected to ASR 5

To demonstrate the usefulness of such a map, the viewer is told to repeat the journey from start to finish in its head. Figure 2 (c) demonstrates how confused a viewer making use of such a map could become. As the viewer imagines re-entering ASR 1, its knows from its fuzzy map that one of these exits leads into ASR 2 but not which one. The viewer randomly chooses an exit. The line emanating from the bottom of fuzzy ASR1, rather than its side, demonstrates that the viewer made an erroneous decision. It can be seen from the output from our computer simulations displayed in this figure that the errors made here result in rotation errors in the cognitive map and while they are not shown in this figure, translation errors are possible also.

In the second experiment we allowed the viewer to remember on which side of the fuzzy ASR the exits were located and thus on which side of a fuzzy ASR the connection to a particular ASR is located. Figure 2 (d) shows a viewer's attempt at using a fuzzy cognitive map constructed using this strategy. In ASR 1 the viewer recalls that ASR 1 connects to ASR 2 via an exit on the left side of ASR 1 and since there is only one such exit the correct choice is made. However, on the side of ASR 3 which connects to ASR 4 there are two exits. One leads directly into ASR 4 (see Figure 2 (b)) and one leads into an as yet unexplored region of the

Figure 3.(a) The environment traversed, (b) a cognitive map computed from exact ASRs (c) the viewer's interpretation of a cognitive map computed from fuzzy ASRs where the viewer has no locational information for the exits. (d)the viewer's interpretation of a cognitive map constructed from fuzzy ASRs where the viewer knows on which side of the fuzzy ASR the exits are located but not their exact position.

environment – this exit can be seen as the lighter shaded gap in the boundary directly adjacent to the exit into ASR 4 in Figure 2 (b). To visit ASR 4 from ASR 3 the viewer must choose between these exits and does so correctly (this time). If the incorrect exit had been chosen a translation error would have occurred. This is the case in Figure 3 (d) when the viewer does make a wrong decision on which exit leads from ASR 3 into ASR 4. See the paragraph which follows for a more detailed explanation.

Figure 3 shows the results of applying the strategies of both experiments to a longer traversal of the environment. Again Figure 3 (a) shows the environment traversed, (b) the cognitive map constructed from exact ASRs, (c) the cognitive map constructed when the viewer has no locational information for the exits in a fuzzy ASR, and (d) the cognitive map constructed when the viewer remembers which side of the fuzzy ASR the exits are on. Note in Figures 3 (b), (c) and (d) that the viewer fails to recognize ASR 3 when the local space is re-entered from ASR 4 and a new ASR, ASR 8 is constructed. This is overlaid on top of ASR 3 only for display convenience. There is no such integration in the viewer's "head" and a one dimensional route-like structure is a better approximation of the actual structure of this part of the cognitive map. Figure 3 (c) has the expected rotation error. A translation error occurs at about fuzzy ASR 3 in both Figure 3 (c) and Figure 3 (d). This is most noticeable in the way in which fuzzy ASR 8 in particular, has shifted in relation to fuzzy ASR 3 in the display. In deciding which exit leads from fuzzy ASR 3 into fuzzy ASR 4 the viewer selects an erroneous one which is on the same side of the fuzzy ASR as the correct one. In Figure 3 (d) it is just possible to make out a corner of fuzzy ASR 4 underneath fuzzy ASR 5. Unfortunately in Figure 3 (c) fuzzy ASR 4 is completely hidden.

Conclusion

We have shown how a representation for the local space which is little more than a vague description for its extent could be computed from an initial accurate description of the local space. Also shown is the underlying structure of the cognitive map which emerges as the viewer explores its environment, computing these muddy descriptions, uncertain as to how they are connected. Such a map is not an unrealistic representation of a viewer's initial tentative exploration of the environment. However a viewer using one of these maps to navigate around its environment would soon become lost. How is such a map enriched as the viewer becomes more familiar with its environment, not in precise metric terms, but merely in terms of being able to work out roughly where places and objects are in relation to others? One would know exactly which exits in an ASR lead to which neighboring ASRs but still be no wiser as to their exact coordinate values in a cartesian frame of reference, for example. Is this the role of landmarks? The "fuzzy cognitive map" has given us a framework in which we can study these problems. The fuzzy ASR provides a structure in which the viewer's experience of the environment can be charted. Eventually important details will be recorded and significant events remembered, some will be remembered well, some poorly. The fuzzy ASR will continue to evolve to reflect the ever changing memories one has for the spatial environment.

References

Bennett, B. (1994). Spatial reasoning with propositional logics. *Proceedings of the Fourth International Conference on Knowledge Representation and Reasoning* (pp. 51-62).

Clementini, E., Di Felice, P., & Hernandez, D. (1995). *Qualitative representation of positional information* (Tech. Rep. FKI-208-95). Fakultat fur Informatik, Technische Universitat Munchen.

Cohn, A. G., & Gotts, N. M. (1996a). The 'Egg-Yolk' representation of regions with indeterminate boundaries. *Proceedings of the GisData Specialist Meeting on Spatial Objects with Undetermined Boundaries* (pp. 171-187).

Cohn, A. G., & Gotts, N. M. (1996b). Representing spatial vagueness: A mereological approach. *Proceedings of the Fifth International Conference on Principles of Knowledge Representation and Reasoning (KR96)* (pp. 230-241). San Francisco: Morgan Kaufmann.

Cohn, A. G., Randell, D. A., & Cui, Z. (1995). Taxonomies of logically defined qualitative spatial relations. *International Journal of Human-Computer Studies*, 43(5-6), 831-846.

Cousins, J. H., Siegel, A. W., & Maxwell, S. E. (1983). Wayfinding and cognitive mapping in large-scale environments: A test of a developmental model. *Journal of Experimental Child Psychology*, 35, 1-20.

Davis, E. (1986). *Representing and acquiring geographic knowledge*. Pitman.

Downs, R. M., & Stea, D. (1973). *Image and environment: cognitive mapping and spatial behavior.* Aldine Publishing Coy.

Egenhofer, M. J., & Khaled, K. A. (1992). Reasoning about gradual changes of topological relationships. In A. U. Frank, I. Campari, & U. Formentini (Eds.), *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. Springer Verlag.

Engelson, S. P. (1994). *Passive map learning and visual place recognition*. PhD, Yale University.

Gallistel, C. R. (1989). Animal Cognition: The representation of space, time and number. *Annual Review of Psychology*, 40, 155-189.

Hernandez, D. (1993). Reasoning with qualitative representations: exploiting the structure of space. *Proceedings of the III Imacs International Workshop on Qualitative Reasoning and Decision Technologies – QUARDET '93* (pp. 493-502).

Hernandez, D., Clementini, E., & DiFelice, P. (1995). Qualitative Distances. In A. U. Frank & W. Kuhn (Eds.), *Spatial Information Theory – A Theoretical Basis for GIS*. Springer-Verlag.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition*, 13(3), 208-217.

Holyoak, K. J., & Mah, W. A. (1982). Cognitive reference points in judgements of symbolic magnitude. *Cognitive Psychology*, 14, 328-352.

Kortenkamp, D. M. (1993). *Cognitive Maps for Mobile Robots: A Representation for Mapping and Navigation.* PhD, University of Michigan.

Kuipers, B. (1996). A hierarchy of qualitative representations for space. *Proceedings of the Tenth International Workshop on Qualitative Reasoning about Physical Systems (QR-96)*. Menlo Park, CA: AAAI Press.

Lee, Y., & Schmidt, C. G. (1988). Evolution of urban spatial cognition: patterns of change in Guangzhou, China. *Environment and Planning A*, 20, 339-351.

Lynch, K. (1960). *The image of the city.* MIT Press.

Moeser, S. D. (1988). Cognitive mapping in a complex building. *Environment & Behavior*, 20(1), 21-49.

Montello, D. R. (1993). A new framework for understanding the acquisition of spatial knowledge in large-scale environments. *Proceedings of the Specialist meeting of the NCGIA Research Initiative 10.* Lake Arrowhead, California.

Rohrig, R. (1994). A theory for qualitative spatial reasoning based on order relations. *Proceedings of the Twelfth National Conference on Artificial Intelligence AAAI-94* (pp. 1418-1423).

Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning & Memory*, 6(5), 516-528.

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

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4), 189-208.

Tversky, B. (1992). Distortions in cognitive maps. *Geoforum*, 23(2), 131-138.

Yeap, W. K. (1988). Towards a computational theory of cognitive maps. *Artificial Intelligence*, 34, 297-360.

Yeap, W. K., Holmes, K. L., & Jefferies, M. E. (1994). What is an exit? *Proceedings of the Third International Conference on Automation, Robotics and Computer Vision* (pp. 369-373). Nanyang Technological University, Singapore.

Yeap, W. K., Jefferies, M., & Naylor, P. (1991). An MFIS for computing a Raw Cognitive Map. *Proceedings of the 12th IJCAI* (pp 373-378). San Mateo, CA: Morgan Kaufmann.

Yeap, W. K., & Jefferies, M. E. (1997). *Computing a representation of the local environment* (Tech. Rep. AIM-35-97-1). Computer Science Dept, University of Otago.

Yeap, W. K., Naylor, P. S., & Jefferies, M. (1990). Computing a representation of the physical environment. *Proceedings of the First Pacific Rim International Conference on Artificial Intelligence* (pp. 847-852).