Spatial Representation and Reasoning

Jerry R. Hobbs and Srini Narayanan
Artificial Intelligence Center
SRI International

1 Introduction

A robot moving through an environment, an interface to a geographical database, and a natural language program for giving directions all need means for representing and reasoning about spatial properties and relations. These include shape, size, distance, orientation, and relative location. The most precise and highly developed system of spatial representation is the mathematics of Euclidean space. But very often this is too precise for the purposes at hand. We may not have exact information and we may not need precise answers. For example, if we are telling someone how to get to the post office, we don't need to be more precise than the streets to follow. For this reason, much work in artificial intelligence has focused on *qualitative* spatial representation and reasoning.

Indeed, studies of human problem solving and language understanding [Talmy 1983, Herskovits 1986, Langacker, 1987, Lakoff, 1987, Tversky 1993, Landau & Jackendoff 1993] indicate that we draw fairly subtle and important qualitative spatial distinctions. It is accepted that a topological description of the environment is central to building a cognitive map which is developmentally prior to a metrical description. Infants have been shown [Landau et al., 1992] to be sensitive to shape distinctions and invariants very early in childhood, suggesting that some qualitative representation of space precedes, or at least coexists, with a more metrical one.

Research in primate vision [Ungerleider & Mishkin, 1982, Milner & Goodale, 1995] further suggests the existence of two separate pathways for visual information processing; a dorsal pathway which projects from the visual cortex to the posterior parietal cortex and a ventral pathway which projects into the inferotemporal cortex. The dorsal pathway is referred to as the "where" system and is suggested as primarily computing task related spatial information (such as location in egocentric coordinates or hand pre-shaping for a grasping task) while the ventral or "what" pathway is suggested as computing primarily object related characteristics (such as shape or other visual features). While there seems to be robust evidence (including double dissociation evidence) of these two different pathways, it is also clear that there are complex interactions through multiple cortico-cortical connections between them. Based on these findings, [Landau & Jackendoff 1993] suggest the existence of related subsystems in language; a "where" system corresponding to spatial predications (such as prepositions) and a "what" system corresponding to object naming and nouns.

From a computational perspective, the most general and most challenging problem in spatial representation and reasoning is route planning: What paths can an object of a given shape, size and location follow through an environment with obstacles of particular shapes, sizes, locations, and trajectories? Most researchers have studied restricted versions of this problem. For example, the problem may be merely to place an object within an environment and not to move it. Or the size and shape of an object may be viewed as negligible as it moves through the environment.

Thus, the first broad problem that must be addressed is to devise ways for representing and reasoning about relations and measures between objects at a range of granularities, from the purely topological to fully metric. The second problem is how to represent and reason about shape in two and three dimensions, again at a range of granularities. Finally, there is the problem of how to represent and reason about objects of particular shapes when they are moving.

2 Levels of Structure

Spatial and other quantities can be represented at a range of granularities. The coarsest level in common use is what is known as the "sign algebra" ([Forbus, 1984] among others); the real line is divided into the negative numbers, zero, and the positive numbers, and what is known about a quantity is only which of the three regions it lies in. This has proven very useful in qualitative physics for reasoning about increases and decreases in quantities and direction of flow. A more refined structure can be imposed by dividing the real line into further intervals, with ordering relations between them. For example, for a pot of water sitting on a stove, we might want to distinguish the intervals between the freezing point, the boiling point, and the temperature of the stove, as well as the transition points themselves. Moreover, "landmarks" can be set during the reasoning process. For example, to determine whether an oscillation is damping, we should compare successive maxima; these would be the landmarks [Kuipers, 1986].

Some work has been done on providing logarithmic structure to scales. This is known as order-of-magnitude reasoning [Raiman 1991]. We may not know whether Los Angeles is closer to San Diego or Santa Barbara, but we certainly know that it is closer to both than it is to Chicago, because these distances are of different orders of magnitude. Most work in order-of-magnitude reasoning has sought to exploit the fact that quantities at lower orders of magnitude can be ignored in operations involving higher orders of magnitude. Thus, we know that adding a stamp to a letter will not increase its weight enough that more postage will be required.

A very fine-grained representation of space is one that places ϵ-neighborhoods around points and considers points indistinguishable if they are within the same ϵ-neighborhood [Roberts, 1973].

More generally, the level of structure we want to view space at will depend on functionality. If we are planning a land trip and are only concerned with getting the right visas, then we can view countries as nodes in a graph, regardless of their size, where each country is connected by an arc to each of its neighbors. When we are traveling in the country, we need a finer-grained view.

3 Dimensionality

The most important feature that distinguishes space from other quantitative domains is the fact that it has more than one dimension. The minimal condition that is required before a notion of dimensionality makes sense is that there must be two or more scales where the order of elements on one scale cannot be determined by their order on the other scales. We cannot infer from the fact that Chicago is east of Denver which of the two is farther north. The scales need not be fully numeric. A spatially represented bar graph may have a numeric vertical dimension while its horizontal dimension is a discrete space of alphabetically ordered named entities. The dimensionality of an entity is dependent on perspective. A road, for example, may be viewed as one-dimensional when we are planning a trip, two-dimensional when we are driving on it, and three-dimensional when we hit a pothole.

When one admits several dimensions, problems arise involving objects of different dimensions. For example, can we have an object consisting of a volume and a line segment? Does the boundary of an object have a lower dimension, and is it part of the object? These problems are related to more general problems concerning open and closed sets. [Galton, 1996] has developed an interesting solution to these problems. In his system, an object cannot be *part* of an object of a different dimension, but objects are *bounded* by objects of lower dimension.

Where there are dimensions, there must be frames of reference for describing locations in each of the dimensions. In human cognition, as evidenced by language, there are a number of frames of reference characterized by how they are anchored [Carlson & Irwin 1993]. There are the self-anchored frame of reference, right-left-front-back, and the world-anchored frame of reference, north-south-east-west. There can be frames of reference anchored on the vehicle one is in, port-starboard-bow-stern, or one's geographical region—the Hawaiian language has prepositions for "toward the center of the island" and "toward the ocean". There can be frames of reference determined by the forces that are acting on one, such as windward-leeward and upriver-downriver. Neurobiologists have also shown the existence of mulitple reference frames including ones anchored on specific joints [Rizzolatti et al., 1997].

4 Regions

In recent years, there has been a good deal of research in purely qualitative representations of space [Randall et al., 1992, Gotts et al., 1996, Gotts 1996, Cohn, 1997, Lemon & Pratt, 1997, Cohn & Hazarika, 2001]. Much of this work attempts to build axiomatic theories of space that are predominantly topological in nature and based on taking *regions* rather than *points* as primitives. Topological relationships between regions in two-dimensional space have been formalized, with transitivity inferences similar to those used in temporal reasoning, identified for vocabularies of relations [Randall et al., 1992, Cohn, 1997]. The basic primitive is a notion of two regions x and y being connected if the closures of x and y share at least one point.

This primitive *connect(x, y)* has been shown to be extremely powerful and has led to the development of a rich calculus of spatial predicates and relations referred to as the RCC calculus. [Gotts et al., 1996, Gotts 1996] show how the RCC calculus can describe and distinguish between complicated objects such as loops, Figure-8's and doughnuts with degenerate holes. However, this expressiveness is costly and reasoning with the general RCC calculus is undecidable [Cohn, 1997, Lemon & Pratt, 1997]. There has been some work on tractable subclasses. The best known is based on identifying a pairwise disjoint, jointly exhaustive set of eight spatial relations called the RCC8 calculus. The RCC8 calculus is able to describe the relations *inside (touching the boundary or not)*, *outside*, *touching*, *overlapping* for objects describable as *regular sets*; regular sets cannot have holes. The RCC8 subset also has a well-founded semantics (unlike the RCC calculus) based on Euclidean Geometry [Lemon & Pratt, 1997] and has been shown to be complete under this interpretation.

5 Orientation

In Euclidean geometry, the representation of orientation is straightforward. There is a fixed set of axes, and a direction corresponds to a vector, normalized to length 1, from the origin to some point in the space. That vector can be specified by projecting its endpoint onto the axes.

Qualitative representations of orientation can be inherited from qualitative representations on the axes. For example, if the structure of each of the two axes in a two-dimensional system is $\{+, 0, -\}$, then the corresponding representation of orientation maps the direction into one of eight values—along one of the four axes or in one of the four quadrants.

A finer structure on the axes yields a finer structure on orientations. For example, if each axis has regions corresponding to "slightly positive" and "slightly negative", then it will be possible to define the notion of angles being "slightly acute" [Liu, 1998].

6 Shape

Shape is one of the most complex phenomena that must be dealt with in the qualitative representation of space. There is a range of possibilities. In topology, a circle, a square, and an amoeba-shaped blob are all equivalent, as long as they don't have holes in them. In Euclidean geometry, a large square and the same square with a slight nick in one side are different. Shape is very important in commonsense reasoning because very often the shape of an object is functional. The shapes of objects and obstacles determine possible paths. It is important for a door and a doorframe to be the same shape, and the fact that they are the same shape is the reason doors must be open for things to pass through them.

One form of representation for shape is the use of *occupancy arrays* that encode the location of an object in a 2-D or 3-D grid [Funt 1980, Glasgow et al., 1995]. These representations have a number of attractive properties including ease of visualization and a natural parallel implementation of operations such as intersection, translation, and computing spatial relations between objects. However, occupancy arrays are inflexible and cannot express abstractions and partial knowledge.

Another way to characterize shapes is by the shapes of their boundaries.

In one approach [Hoffman & Richards 1982], the sides of a complex figure are classified as straight or curving in or out, and the vertices as angling in or out. Boundary information can be combined with qualitative length information so that we can distinguish between "thin" and "fat" rectangles, for example, or between rectangles that are wider than they are tall and ones that are taller than they are wide. Boundary information can also be combined with theories of orientation, so that we can get qualitative measures of angles between sides.

Another represenation of shape is in terms of its "bounding box", or smallest containing quadrangle, or in terms of its convex hull. These representations are useful in moving objects through an environment. To determine whether a car will fit through a garage doorway, you need to take the side mirrors into account, but not the cavities due to open windows.

A commonly used representation for complex three-dimensional objects is cylinders [Davis, 1990]. This representation is subject to variations in granularity. For example, a human being can be represented at a very coarse grain as a cylinder. At a finer grain, the person is resolved into a cylinder for the torso, cylinders for each arm and leg, and a cylinder for the neck and head. At even finer grains, there are cylinders for the upper and lower arms or for individual fingers.

7 Motion

For any quantity or quality that can be represented at an instant in time, we can also imagine it changing across time. Topological relations between entities can change as the entities move. The distance between two objects, the orienatations of two lines, the shape of two objects can all change with time.

In general, given any qualitative theory of a spatial feature and any qualitative theory of time, we can develop a qualitative theory of how the spatial feature changes with time. In most instances, this will involve transforming what is a continuous motion in Euclidean space into a motion of discrete jumps in the qualitative space, as when an entity suddenly crosses the boundary between region A and region B of a plane. Various qualitative theories impose various constraints on the possible motions, or possible changes of state. Thus, we cannot go from positive to negative without passing through zero. Representing the behavior as a transition graph (fully qualitative or with metric information attached to the nodes), allows for algorithms that generate and recognize behaviors using a variety of analytic and simulation techniques. Modeling spatially distributed motion is more complex and some efforts have attempted to use spatial aggregation based on motion field invariants and other methods from computer vision to represent qualitative structures of fluid motion [Yip 1995] and geometric features in phase space [Zhao 1994].

8 Imagistic and Propositional Representations

There is a long-standing debate among researchers in spatial representation and reasoning about whether representations of spatial relationships should be imagistic or propositional, that is, more like pictures or more like linguistic descriptions [Pylyshyn, 1981, Shepard & Cooper, 1982, Kosslyn 1994]. Proponents of propositional representations argue that imagistic representations reduce to propositional ones, but impose constraints on what situations can be represented that make them of very limited use [Lemon & Pratt, 1997].

On the other hand, Imagistic representations are common in models of human spatial reasoning. In human communication, imagistic representations such as sketches, maps and figures are commonly used to communicate information about shape, size, routes, and spatial arrangements. Images and diagrams are also commonly used to communicate and comprehend abstract structures ranging from data-structure visualizations and process-flow diagrams to corporate hierarchies. Reasoning with imagistic representations has often been argued to be fundamental in human cognition [Kosslyn 1994, Farah 1995]. It has been shown that infants, very early, perhaps pre-attentively, do grouping based on similarity of shapes, orientations and sizes [Julesz 1984] as well as shape closure [Treisman 1982, Treisman, 1985]. There is evidence that even symbolic tasks such as language understanding may use imagistic representations [Langacker, 1987, Lakoff, 1987]. A computational model of the role of such representations in the acquisition of spatial prepositions can be found in [Regier 1992].

There is growing evidence that acquiring, storing and reasoning with spatial concepts requires the coordinated use of heterogeneous representation and inferential processes that involve both propositional and imagistic components (qualitative and metric) [Glasgow et al., 1995], and much of current research is exploring how this can be accomplished computationally.

References

[Carlson & Irwin 1993] Carlson-Radvansky, L. A., and D. E. Irwin. 1993. "Frames of Reference in Vision and Language: Where is Above?" Cognition, 46, 223-244.

[Cohn, 1997] Cohn, A. G. "Qualitative Spatial Representation and Reasoning Techniques", Proceedings of KI-97, edited by G. Brewka and C. Habel and B. Nebel, LNAI, 1303 , pp 1-30, Springer Verlag, (1997).

[Cohn & Hazarika, 2001] Cohn, A. G., and S. M. Hazarika. "Qualitative Spatial Representation and Reasoning: An Overview", *Fundamenta Informaticae*, 46 (1-2), pp 1-29, (2001).

[Davis, 1990] Davis, E., 1990. *Representations of Commonsense Knowledge*, Morgan-Kaufmann, 1990

[Farah 1995] Farah, M., 1995. "The Neural Bases of Mental Imagery", In M.S. Gazzaniga (Ed), *The Cognitive Neurosciences*, (963-975). Cambridge, MA: MIT Press.

[Forbus, 1984] Forbus, K., 1984. "Qualitative Process Theory.", in *Artificial Intelligence*, 24 pp. 85-168.

[Funt 1980] Funt, B., 1980. "Problem Solving with Diagrammatic Representations.", in Glasgow, J., B. Karan, and N. Narayanan, (Eds.), 1995. *Diagrammatic Reasoning*, Cambridge Mass.: AAAI Press/The MIT Press, pp. 33-38.

[Galton, 1996] Galton, Anthony, 1996. "Taking Dimension Seriously in Qualitative Spatial Reasoning", in *Proceedings*, 12th European Conference on Artificial Intelligence, Budapest, Hungary, August, 1996, pp. 501-505.

[Glasgow et al., 1995] Glasgow, J., B. Karan, and N. Narayanan, (Eds.), 1995. *Diagrammatic Reasoning*, Cambridge Mass.: AAAI Press/The MIT Press.

[Gotts et al., 1996] Gotts, N. M., J. M. Gooday, and A. G. Cohn, 1996. "A Connection Based Approach to Common-sense Topological Description and Reasoning", *The Monist*, 79 (1), pp 51-75, (1996).

[Gotts 1996] Gotts, N. M. 1996."Toplogy from a Single Primitive Relation: Defining Topological Properties and Relations in Terms of Connection", *Report* 96.23, School of Computer Studies, University of Leeds, (1996).

[Herskovits 1986] Herskovits, A. 1986. *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English.* Cambridge, England: Cambridge University Press.

[Hoffman & Richards 1982] Hoffman, D. D., and W. A. Richards, 1982. "Representing Smooth Plane Curves for Recognition: Implications for Figure-Ground Reversal", *Proceedings*, National Conference on Artificial Intelligence, Pittsburgh, PA, August 1982, pp. 5-8.

[Julesz 1984] Julesz, B. 1984. "A Brief Outline of the Texton Theory of Human Vision", *Trends in Neurosciences*, 7, 41-45.

[Kosslyn 1994] Kosslyn, S.M., 1994, *Image and Brain: The Resolution of the Imagery Debate*, Cambridge, MA: MIT Press.

[Kuipers, 1986] Kuipers, B., 1986. "Qualitative Simulation.", in *Artificial Intelligence*, 29 pp. 289-338.

[Landau et al., 1992] Landau, B., L. Smith, and S. Jones, (1992). "Syntactic Context and the Shape Bias in Children's and Adults' Lexical Learning", *Journal of Memory and Language*, 31.

[Landau & Jackendoff 1993] Landau, B. and R. Jackendoff. 1993. " 'What' and "Where' in Spatial Language and Spatial Cognition", *Behavioral and Brain Sciences*, 16, 217-265.

[Lakoff, 1987] Lakoff. G., 1987. *Women, Fire and Dangerous Things*, The University of Chicago Press: Chicago.

[Langacker, 1987] Langacker. R.W., 1987. "An Introduction to Cognitive Grammar", *Cognitive Science*, 10(1), 1-40.

[Lemon & Pratt, 1997] Lemon, O. and I. Pratt, 1997. "Spatial Logic and the Complexity of Diagrammatic Reasoning", *Machine Graphics and Vision*, 6(1): 89 - 109, 1997 (Special Issue on Diagrammatic Reasoning).

[Liu, 1998] Liu, Jiming, 1998. "A Method of Spatial Reasoning Based on Qualitative Trigonometry", Artificial Intelligence Journal 98 (1998), pp.137-168.

[Milner & Goodale, 1995] Milner D. A. and M. A. Goodale, 1995. *The Visual Brain in Action*, Oxford Psychology Series, No. 27, Oxford University Press.

[Pylyshyn, 1981] Pylyshyn, Z.W., 1981. "The Imagery Debate: Analogue Media Versus Tacit Knowledge", *Psychological Review*, (88) 16-45.

[Raiman 1991] Raiman, O., 1991. "Order of Magnitude Reasoning.", in *Artificial Intelligence*, 51, pp. 11-38.

[Randall et al., 1992] Randell D. A., Z. Cui, and A. G. Cohn, 1992. "A Spatial Logic Based on Regions and Connection" *Proceedings* 3rd International Conference on Knowledge Representation and Reasoning, pp. 165-176, Morgan Kaufmann, San Mateo, (1992).

[Regier 1992] Regier, T., 1992. "The Acquisition of Lexical Semantics for Spatial Terms: A Connectionist Model of Perceptual Categorization", Ph.D. thesis, Computer Science Department, University of California, Berkeley.

[Rizzolatti et al., 1997] Rizzolatti, Giacomo, Luciano Fadiga, Leonardo Fogassi, and Vittorio Gallese, 1997. "The Space Around Us", *Science*, Vol. 277, pp. 190-191, 11 July 1997.

[Roberts, 1973] Roberts, Fred S., 1973. "Tolerance Geometry", *Notre Dame Journal of Formal Logic*, vol. 14, no. 1, pp. 68-76.

[Shepard & Cooper, 1982] Shepard, R.N., and L. A. Cooper, 1982. *Mental Images and Their Transformations*, Cambridge, MA: MIT Press.

[Talmy 1983] Talmy, L., 1983. “How Language Structures Space”, In H. and L. Acredolo (Eds.), *Spatial Orientation: Theory, Research, and Application*, New York: Plenum Press.

[Treisman, 1985] Treisman, A., 1985. “Preattentive Processing in Vision”, *Computer Vision, Graphics, and Image Processing* 31:156-177.

[Treisman 1982] Treisman, A., 1982. “Perceptual Grouping and Attention in Visual Search for Features and Objects”, *Journal of Experimental Psychology: Human Perception and Performance*, 8 (2), 194-214.

[Tversky 1993] Tversky, B., 1993. “Spatial Mental Models” In G. H. Power (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, vol. 27, New York: Academic Press.

[Ungerleider & Mishkin, 1982] Ungerleider, L.G., and M. Mishkin, 1982. “Two Cortical Visual Systems”, In D.J. Ingle, M.A. Goodale and R.J.W. Mansfield (Eds.). *Analysis of Visual Behavior*, pp. 549-586, Cambridge, MA: MIT Press.

[Yip 1995] Yip, K., 1995. ”Reasoning about Fluid Motion I: Finding Structures”, *Proceedings* of IJCAI-95, pp. 1782-1788, AAAI Press.

[Zhao 1994] Zhao, F., 1994. ”Extracting and Representing Qualitative Behaviors of Complex Systems in Phase Space”, *Artificial Intelligence*, 69. pp. 51-92.