The Theory of Cognitive Prism–Recognizing Variable Spatial Environments

Tiansi Dong

Cognitive Systems Group, Department of Mathematics and Informatics
University of Bremen, 28334, Bremen, Germany
tiansi@informatik.uni-bremen.de

Abstract

This paper outlines a symbolic computational theory for recognizing variable spatial environments-The Theory of Cognitive Prism, (Dong 2005). This theory defines distance and orientation relations between extended objects using the connectedness relation. The metaphor of cognitive prism is proposed to understand the representation of a snapshot view of an environment such that a cognitive system neglects part of the objects in a snapshot view and subjectively re-arranges selected objects, just like an optical prism which reflects some of the incoming light, re-arranges the incoming part, and forms a spectrum. The recognizing process is interpreted as a judgement of the compatibility between two representations that decides to which qualitative degree they belong to the same environment.

Introduction

When you enter your office in the morning, you may find the location of your chair is different from the place when you left yesterday. However, you can still recognize your office. The aim of this article is to introduce a computational theory which computationally simulates recognizing changed spatial layouts.

From a view of your office, you can recognize objects in it. That is, you can recognize objects from their partial images, e.g. (Buelthoff & Edelman 1992), (Tarr & Buelthoff 1998). Research in cognitive psychology also shows that recognizing furniture means categorizing furniture, e.g. (Liter & Buelthoff 1996). The most preferable level of category is called the basic level category in (Rosch *et al.* 1976). That is, if the category of your office chair has 100 members, recognizing your chair means that it is yours with the probability of 1%. However, when you enter your office, you recognize your office chair confidently, rather than that it is yours with 1% probability. How come?

It is your office chair means that it belongs to the same category as your office chair and that it is located in your office. The question is: what exists in your office that makes it to be your office? It is obvious that not everything inside helps you to recognize your office. For example, the air in it–suppose that you do not recognize your office by smelling

the air inside. (Wilson, Baddeley, & Young 1999) reported a brain impaired artist, LE. She can only see contours of objects, and even cannot recognize her husband, however she can locate objects, and amazingly she can recognize her home[1]. Research in cognitive neuroscience shows that the visual system consists of at least two subsystems: the "what" cortical system and the "where" cortical system, e.g. (Ungerleider & Mishkin 1982), (Rueckl, Cave, & Kosslyn 1989), (Creem & Proffitt 2001). That is, LE's "what" cortical system has some problems, while her "where" system is normal. This does not prevent her from recognizing her home. The conclusion is that recognizing spatial environments requires category information of single objects and that spatial relations among objects play an important role.

The starting point is that objects are categorized and spatial relations among them are known based on observation. The research questions are as follow: How shall we represent spatial relations among extended objects (the foundation)? How shall we represent spatial environments (the representation)? How shall we compare two representations (the reasoning)?

The foundation: Spatial relations between extended objects

Spatial relations include topological relations, distance relations, and orientation relations, e.g. (Stock 1997). In the field of qualitative spatial representation, topological relations[2] are based either on the connectedness relation–the RCC8 theory, e.g. (Randell, Cui, & Cohn 1992), (Cohn 1993), or on the intersection relations, e.g. (Egenhofer 1991), (Egenhofer 1993); distance relations between "solids" were represented informally by the connectedness relation in (de Laguna 1922); orientation relations between extended objects were represented with the occupation area in oriented grids, e.g. (Goyal 2000). In the Theory of Cognitive Prism, distance and orientation relations are represented by the connectedness relation.

---

[1]Personal communication with Allan Baddeley

[2]For the difference between topology in math and topological relations in qualitative spatial representation, please see (Smith 1994).
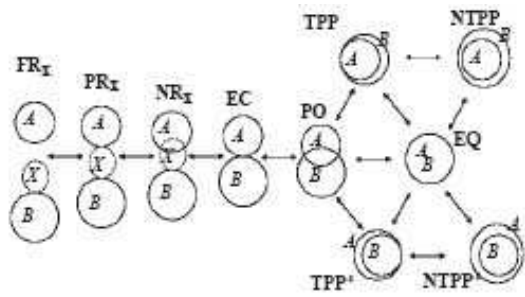
Figure 1: The conceptual neighborhoods network of RCC10. The disconnected (DC) relation in RCC8 is specified by three qualitative distance relations: far (FR_X), penumbra-far-or-near (PR_X) and near (NR_X)

RCC10: Integrating distance relations into RCC8

In physics, people would say that the Andromeda Galaxy is 2.3 million light-years away, which means that it will take the light 2.3 million years to reach the Andromeda Galaxy from the earth. In everyday life, people would say that the window is 8 feet away from the door, which means that it takes 8 feet (each has the same size as the British imperial foot) to reach the window with the condition that the 8 feet are connected one after another. Distances, either qualitative or quantitative, can be understood by the reachability between two objects. In RCC8, distance relations are covered by the DC (disconnected) relation. Therefore, by specifying the DC relation, we can integrate distance relations into RCC8 as follows: We categorize objects by their sizes such that objects of the same size are in one category. Let X be such an object category, e.g. light-year, foot, then for two disconnected object A and B, "A is near B" can be defined as that there is an X in X such that X overlaps B with the condition that X connects A; "A is far away from B" can be defined as that X connects neither B nor A; "A is penumbra-far-or-near[3] from B" can be defined as that X externally connects B with the condition that X connects A. The conceptual neighborhoods network is shown in Figure 1.

Integrating orientation relations into RCC10

When people say "the bike is in front of the house", they agree that the bike is nearer to the front side of the house than its other sides. Orientation relations can be interpreted as the distance comparison between the location object and different sides of the reference object. Let F_A, L_A, R_A, and B_A be the four sides (*front, left, right,* and *behind*, respectively) of an object A[4], then "object B is in front of A" can be interpreted as B is nearer to F_A than the other three sides of A. If the location object is nearer to two (or more than two) sides of the reference object, then the orientation relation is undetermined, shown in Figure 2. By specifying different sides of the reference object, we can integrate orientation relations into the RCC theory, shown in Figure 3. FR_X, PR_X,

---

[3]The notion of 'penumbra' is taken form (Freksa 1982).

[4]A general understanding of orientation relations between extended objects is presented in (Dong 2005, p. 54).
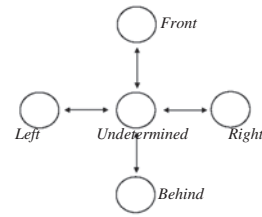

Figure 2: The conceptual neighborhoods network of the qualitative orientation relations. Each transition from one qualitative orientation to another should pass the undetermined relation
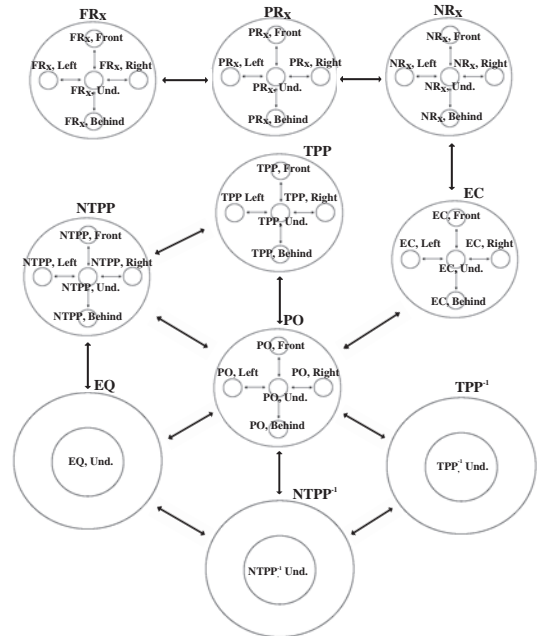

Figure 3: The conceptual neighborhoods network of RCC-10 which nests qualitative orientation relations, "Und." stands for "Undetermined$_4$"

NR_X, EC, PO, TPP and NTPP are specified into five nodes neighborhood networks. EQ, TPP^{-1} and $NTPP^{-1}$ are specified into one node, because under these three relations the location object covers all sides of the reference object, therefore, the orientation relation between them is undetermined.

The representation: The cognitive spectrum of spatial environments

Spatial knowledge can be explored through spatial linguistic descriptions, e.g. (Foos 1980), (Ullmer-Ehrich 1982), etc. The knowledge representation of snapshot views is based on the Schematization Similarity Conjecture that *to the extent that space is schematized similarly in language and cognition, language will be successful in conveying space*, (Tversky & Lee 1999).
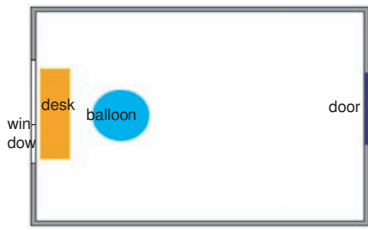
Figure 4: A simple spatial environment

Figure 5: The asymmetric relations of objects in the simple layout

The selective-ness

(Talmy 1983) discussed how language is effective for conveying spatial information. He proposed that language schematizes space by selecting certain aspects from a referent scene to represent the whole, while disregarding others. This suggests that the knowledge representation of spatial environments only include part of the objects inside. To recognize your office, you pay attention to the room, windows, doors, desks, couches, rather than apples, books, pens, etc.

The commonsense knowledge of relative stability

We say "the book is on the table" not "the table is under the book"; "the picture is on the wall" not "the wall is behind the picture"; "you are in the room" not "the room contains you", because objects should be anchored to more stable objects. If a pilot in a plane has lost his current location information, he expects something like "you are above the South Pole" rather than "you are in your plane". People must have commonsense knowledge of stabilities of objects in spatial environments, which results in the asymmetric relations between location objects and reference objects in their linguistic descriptions, (Dong 2005, p. 56).

The cognitive prism metaphor

When a beam of white light reaches an optical prism, part of the light will be reflected and the part that passes through will be re-arranged forming a spectrum: red, orange, yellow, green, blue, and violet. The cognitive system works, therefore, like the optical prism that selects certain objects from the spatial environment while neglecting others, and that rearranges these selected objects based on the commonsense knowledge of stability. The knowledge representation of spatial environments through a cognitive prism is called a cognitive spectrum, (Dong 2005, p. 45).

An example

For the simple spatial environment in Figure 4, people would give following spatial descriptions as they recall the configuration: *the door and the window are opposite to each other*; *the desk is close to the window*; *the balloon is in front of the desk*.

We first identify the relative stability relations by checking location-reference relations in the sentences. The door and the window are referenced each other, therefore, they are of the same stability; the desk is referenced to the window, therefore, the desk is less stable than the window; the
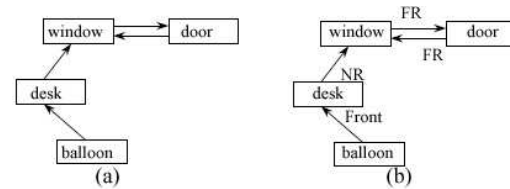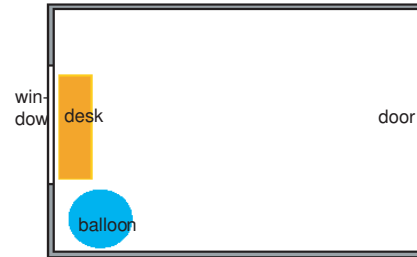

Figure 6: The current perceived spatial environment

balloon is referenced to the desk, therefore, it is less stable than the desk. The diagrammatic representation is shown in Figure 5 (a).

If we interpret the proposition 'opposite to' as the 'FR_X' (or 'FR' for short) relation, 'close to' as the 'NR_X' (or 'NR' for short) relation, and 'in front of' as the 'Front' relation, then the spatial relations among the objects can be formalized as shown in Figure 5(b), which is called a cognitive spectrum of the spatial layout.

The reasoning: Recognizing spatial environments

If we understand that recognizing a snapshot view of an environment is the judgment of whether the perceived snapshot participates into the 4-dimensional target environment, we will be in trouble, because generally speaking, a 4-dimensional environment can not be completely known by the person who wants to identify it. The exact locations of books and cups in your office cannot be known exactly at the next minute.

The theory of cognitive prism takes the alternative approach (inspired by (Grenon & Smith 2004)) as follows:
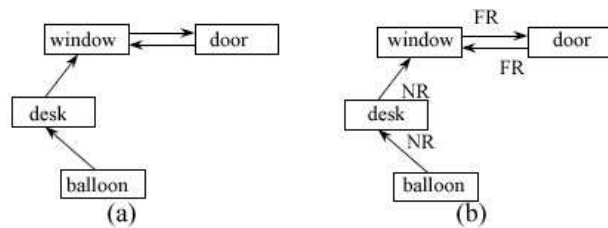

Figure 7: The cognitive spectrum of the perceived environment

recognizing spatial environments is a judgment on whether the perceived snapshot view of an environment is compatible with the remembered snapshot view of the target environment. Two snapshots are compatible means that they participate into the same 4-dimensional environment at different temporal points: one participated some time ago, the other participates at the moment and that the perceived one is transformed from the remembered one. The ease of transformation from the remembered snapshot view to the perceived one determines the degree of the compatibility. For example, two snapshots are compatible, if only the locations of cups in them are different (moving a cup is very easy); two snapshots might be compatible, if the locations of desks are different; two snapshots are hardly compatible, if the locations of the windows are different (moving windows is not an easy job).

The process of inspecting cognitive spectrums

The condition of recognizing a target environment is that the target environment is remembered and used as the criteria to inspect the current perceived snapshot. When people enter a room, they first pay attention to the shape of the room, then windows and doors, then big furniture, like desks and shelves, etc. (Minsky 1975), (Brewer & Treyens 1981). This implies the top-down sequence of inspecting the perceived cognitive spectrum. To continue the above example, we suppose the current perceived spatial environment is shown in Figure 6. Only the location description of the balloon is changed as follows: *the balloon is near the desk*. The reference relations remain the same as before, shown in Figure 7(a). The spatial relation between the balloon and the desk is changed into 'NR', if near is interpreted as 'NR', shown in Figure 7(b). Let Figure 5(b) and Figure 7(b) be representations of the remembered snapshot and the perceived snapshot, respectively. The top-down process works as follows: it first recalls what are located at the top of the hierarchy (here are the window and the door) and what are their spatial relations (here are two FRs); then it checks the top level of the perceived one–there is also a window and a door and they are also FR relations with one another. Therefore, the window in Figure 5(b) and the window in Figure 7(b) are mapped, which means that they are believed to be the same. So are the door in Figure 5(b) and the door in Figure 7(b). The second step is to inspect location objects which are anchored to the mapped objects in both of the representations. They are the desk that is anchored to the window in Figure 5(b) with the spatial relation NR and the desk that is anchored to the window in Figure 7(b) with the spatial relation NR. Then the two desks are mapped and believed to be the same. The balloon in Figure 5(b) is anchored to the desk with the spatial relation Front, however, there is no balloons that are anchored to the desk in Figure 7(b) with the relation Front. This is called qualitative destruction, following (Grenon & Smith 2004). The balloon in Figure 7(b) is anchored to the desk with the spatial relation NR, however, there is no balloons that are anchored to the desk in Figure 5(b) with the relation NR. This is called qualitative creation, following (Grenon & Smith 2004). Therefore, the two balloons are not mapped.

The degree of the compatibility

Spatial environments are dynamic. This results in the difference between the current perceived snapshot and the remembered one. The question is raised: should the observation of these differences lead to believe that the current perceived snapshot view does not belong to the target environment?

With the observation of the location changes of books, cups, pieces of paper, we do not believe that the perceived snapshot view belongs to a new environment; with the observation of the location changes of windows or doors, we would believe that the perceived snapshot belongs to a different environment; with the observation of big furniture, such as the writing-desk, couches, shelves, we would doubt about it and ask for the reason if we believe that it does not belong to a different environment.

As a 4-dimensional entity, a spatial environment has its temporal parts and it unfolds itself phase by phase, (Grenon & Smith 2004). Object locations can be changed in different phases. In a particular environment, objects inside have different frequencies in their location changes. The frequency of location changes of windows and doors is almost zero; the frequency of location changes of writing-desks, couches and shelves are much smaller than that of cups and pens. These frequencies are known by observers, which are called the commonsense knowledge of stability, and further lead to the asymmetric relations of location objects and reference objects in their spatial linguistic descriptions.

The compatibility refers to the believed ease of the transformation from the remembered snapshot to the perceived one. The degree of the compatibility is determined by the commonsense knowledge of stabilities of un-mapped objects. In the above example, the un-mapped objects are two balloons. Balloons are believed to be movable, as they are located at the lowest level in the cognitive spectrums. Therefore, the two snapshots are compatible. The two layouts in Figure 4 and Figure 6 are believed to belong to the same spatial environment.

An implementation: the LIVE model

A symbolic model, the LIVE model (Lisp representations of Indoor Vista spatial Environments), is implemented in LispWorks4.2 both on the Linux Susie 7.3 platform and on the Windows XP professional platform.

The general architecture of the LIVE model is shown in Figure 8. It has five sub-models: the furniture system, the configuration files, the viewing system, the drawing system and the comparison system.

The furniture system stores all the furniture information that the LIVE model represents. The configuration files store all the symbolic representations of different configurations of Mr. Bertel's room. The viewing system provides a graphical interface for the symbolic representation of configurations. The drawing system provides a graphical interface for creating or modifying configurations. The comparison system compares the compatibility of two selected configurations.

The furniture system is the basic system. It provides class information for the viewing system, the drawing system, and
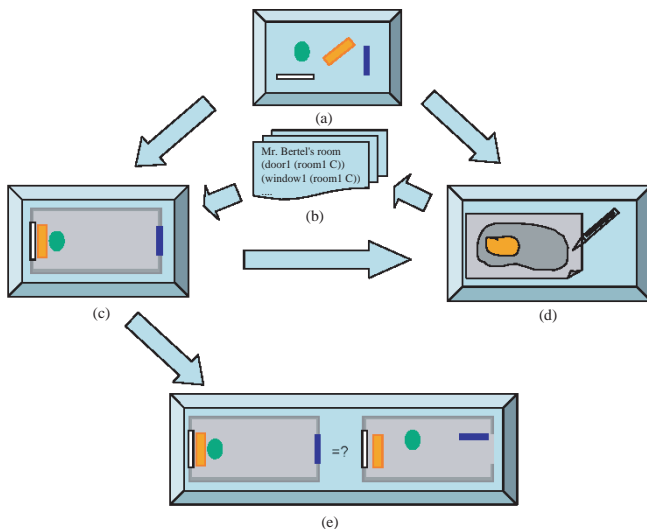
Figure 8: The general architecture of the LIVE model: (a) the furniture system, (b) configuration files, (c) the viewing system, (d) the drawing system, (e) the comparison system. Arrows in the picture represent information flow

the comparison system. The viewing system provides an easy way for the drawing system to create new configurations by modifying already existing one.

Symbolic representations

In the LIVE model, a piece of furniture is represented by an instance of a class. The class has categorical knowledge about the object, such as the name of this category, default qualitative values of stabilities, sides, etc. The location information are represented by two structures: object-object relation and object-face relation. The distance and the classic topological relations are represented by the object-object relation list, whose first element is the name of the location object. Each element of the tail is a two-element list–the first element is the reference object, the second element is one of RCC10 relations. For example, `(|desk1|((|window1| NR)))`, which stands that `|desk1|` is near (NR) `|window1|`. The location of an object is further specified by orientation relations between the location object and reference objects. For example, the location of the desk can be further specified by saying that the desk is in front of the window, which is represented by the object-face relation, whose first element is the name of the location object. Each element of the tail represents an orientation relation. An orientation relation is represented by a list whose first element is the name the reference object, whose second element is the name of its side and whose last element is the nearer relation NRR, e.g., the above orientation relation is represented `(desk1 ((|window1| WINDOW_FACE1 NRR)))`, where `WINDOW_FACE1` stands for the front side of the window. Given a set of symbolic representations of a spatial layout, the viewing system illustrates the hierarchical structure of stabilities, shown in Figure 9.
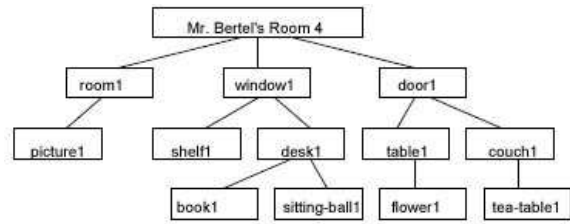

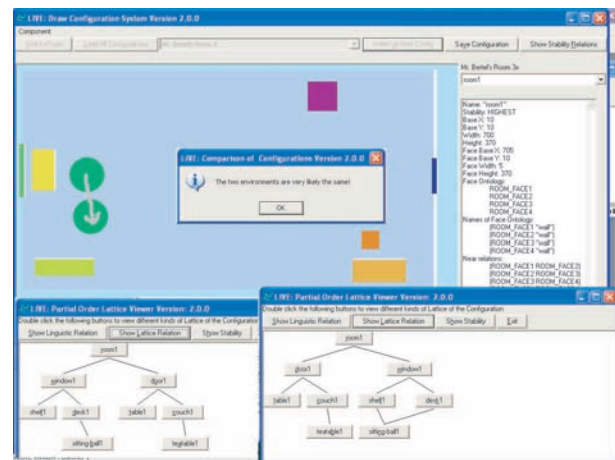Figure 9: The hierarchy of the stability a spatial environment in the LIVE model

Figure 10: An example of the LIVE model: the sitting-ball is moved a little bit, the comparison process makes a positive judgement: the changed layout is compatible with the old one

An example

In Figure 10, the balloon is located differently in two views; and the location description is changed from "the balloon is in front of the desk" to "the balloon is between the desk and the book-shelf". Therefore, its reference relation changes accordingly. However, in both spatial layouts two balloons are located in the lowest level of the hierarchy of the stability, therefore, the recognition process reported that "They are QUITE POSSIBLE the same", which means that the two views are compatible.

Conclusions and outlooks

This article introduced a new symbolic computational theory-the Theory of Cognitive Prism-to simulate recognizing variable spatial environments. It is composed of three parts: the topological definitions of orientation and distance relations (the foundation), a representation of an indoor spatial environment, and the spatial reasoning for the compatibility between two representations. Recognizing a changed spatial environment is interpreted as the compatibility between the current perceived environment and the remembered one, which can be acquired from spatial linguistic descriptions.

The current representation of indoor spatial environments

is based on spatial linguistic descriptions. For future work, we shall include sensor perception and movement of the observer, so that we can find the co-relations among language, perception and action. The expected result shall be a cognitive theory which integrates language understanding and generation, perception, and action. The theory shall be implemented on a cognitive robot which can perceive, move, and make dialogs with other cognitive agents.

It is also an interesting topic for the future to think about spatial cognition based on different sensors. In nature, some animals use eyes, like people, some animals can rely on noses, like dogs, some can rely on ultra-sonic, like blind bats, etc. Their perceived world may be totally different from each other, however, they are capable of recognizing their homes, navigating in dynamic environments. This leads to the research topic of cognitive ontologies. The results can guide to build robots using different kinds of sensors, like laser scanners.

Acknowledgments

References

Brewer, W., and Treyens, J. C. 1981. Role of schemata in memory for places. *Cognitive Psychology* 13:207–230.

Buelthoff, H. H., and Edelman, S. 1992. Psychophysical support for a two-dimensional view interpolation theory of object recognition. In *Proceedings of National Academy of Science*. 60–64. USA.

Cohn, A. G. 1993. Modal and non modal qualitative spatial logics. In Anger, F. D.; Guesgen, H. M.; and van Benthem, J., eds., *Proceedings of the Workshop on Spatial and Temporal Reasoning*. IJCAI.

Creem, S. H., and Proffitt, D. R. 2001. Defining the cortical visual systems: "What", "Where", and "How". *Acta Psychologica* 107:43–68.

de Laguna, T. 1922. Point, line and surface as sets of solids. *The Journal of Philosophy* 19:449–461.

Dong, T. 2005. *Recognizing Variable Spatial Environments — The Theory of Cognitive Prism*. Ph.D. Dissertation, Department of Mathematics and Informatics, University of Bremen.

Egenhofer, M. 1991. Reasoning about binary topological relations. In Gunther, O., and Schek, H. J., eds., *Second Symposium on Large Spatial Databases*, volume 525 of *Lecture Notes in Computer Science*, 143–160. Zurich, Switzerland: Springer-Verlag.

Egenhofer, M. 1993. A Model for Detailed Binary Topological Relationships. *Geomatica* 47(3 & 4):261–273.

Foos, P. W. 1980. Constructing Cognitive Maps From Sentences. *Journal of Experimental Psychology: Human Learning and Memory* 6(1):25–38.

Freksa, C. 1982. Linguistic Description of Human Judgments In Expert Systems and In The 'Soft' Sciences. In Gupta, M., and Sanchez, E., eds., *Approximate Reasoning in Decision Analysis*. North-Holland Publishing Company.

Goyal, R. 2000. *Similarity Assessment for Cardinal Directions between Extended Spatial Objects*. Ph.D. Dissertation, Spatial Information Science and Engineering, University of Maine.

Grenon, P., and Smith, B. 2004. SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spatial Cognition and Computation* 4(1):69–103. Lawrence Erlbaum Associates, Inc.

Liter, J. C., and Buelthoff, H. H. 1996. An Introduction to Object Recognition. Technical report, Max-Planck-Institue fur biologische Kybernetik. Technical Report No. 43.

Minsky, M. 1975. A Framework for Representing Knowledge. In Winston, P. H., ed., *The Psychology of Computer Vision*. McGraw-Hill.

Randell, D.; Cui, Z.; and Cohn, A. 1992. A spatial logic based on regions and connection. In Nebel, B.; Swartout, W.; and Rich, C., eds., *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning*, 165–176. San Mateo: Morgan Kaufmann.

Rosch, E.; Mervis, C. B.; Gray, W.; Johnson, D.; and Boyes-Braem, P. 1976. Basic objects in natural categories. *Cognitive Psychology* 8:382–439.

Rueckl, J.; Cave, K.; and Kosslyn, S. 1989. Why are "What" and "Where" Processed by Separate Cortical Visual Systems? A Computational Investigation. *Journal of Cognitive Neuroscience* 1:171–186.

Smith, B. 1994. Topological Foundations of Cognitive Science. In Eschenbach, C.; Habel, C.; and Smith, B., eds., *Topological Foundations of Cognitive Science*. Workshop at the FISI-CS.

Stock, O., ed. 1997. *Spatial and Temporal Reasoning*. Kluwer Academic Publishers.

Talmy, L. 1983. How Language Structures Space. In Pick, H., and Acredolo, L., eds., *Spatial Orientation: Theory, Research and Application*. Plenum Press. 225–281.

Tarr, M. J., and Buelthoff, H. H. 1998. Image-based object recognition in man, monkey and machine. *Cognition* 67:1–20.

Tversky, B., and Lee, P. 1999. How space structures language. In Freksa, C.; Habel, C.; and Wender, K. F., eds., *Spatial Cognition*, volume 1404 of *LNAI*. Springer-Verlag. 157–176.

Ullmer-Ehrich, V. 1982. The Structure of Living Space Descriptions. In Jarvella, R. J., and Klein, W., eds., *Speech, Place, and Action*. John Wiley & Sons Ltd. 219–249.

Ungerleider, L. G., and Mishkin, M. 1982. Two cortical visual systems. In Ingle, D. J.; Goodale, M. A.; and Mansfield, R. J. W., eds., *Analysis of visual behavior*. Cambridge, MA: The MIT Press. 157–165.

Wilson, B.; Baddeley, A.; and Young, A. 1999. LE, A Person Who Lost Her 'Mind's Eye'. *Neurocase* 5:119–127.