

Dynamic Symbol Grounding: Changing Referents in Engineering Analysis and Spatial Environments

Matthew Klenk and Daniel Bobrow

Palo Alto Research Center
333 Coyote Hill Rd
Palo Alto, CA, 94333
{klenk,bobrow}@parc.com

Abstract

In this paper, we propose a challenging class of problems for Cognitive Systems reasoning in spatio-temporal environments. Consider a safe area in a military engagement. The safe area refers to a region in the environment. The size, shape, and location of this region is a function of the goals and capabilities of the other agents in the environment. Furthermore, this context will change over time and thereby changing the reference of this abstraction. We call such regions context-dependent spatial regions. Through an analogy with the qualitative abstractions used in the design and analysis of engineered systems, we explore this representational problem and some possible solutions.

Context-Dependent Spatial Regions

Natural collaboration between humans and mobile robots requires both partners to communicate with spatial language. Many regions referred to by people are defined not solely by their geometry but also by their context (e.g., the functional use of the space). Consider the following regions: the front of a classroom (Figure 1), neighborhoods in a city (Figure 2), and safety in a military engagement (Figure 3).

We call such areas *context-dependent spatial regions* (Hawes *et al.* 2012). To identify the front of a classroom, one must recognize the boundaries of the classroom (e.g., the walls), the objects within the room (e.g., desks and whiteboards) as well as their functional use (e.g., students sit at the desks oriented toward a teacher). These regions are important for communication because they are tied to the goals and intentions of the actors in the environment. If someone wants to present, they stand in the front of the classroom. Neighborhoods are functional, geographic areas of cities. They provide a heuristic for people in terms of businesses, real estate prices, and crime. The Tenderloin is

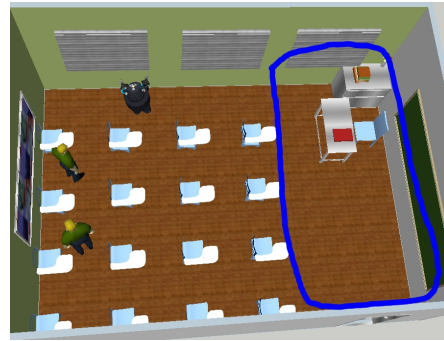


Figure 1: The front of the classroom depends on locations and orientations of the desks and other objects in the room.

only a few blocks from Union Square, but in one you will find a high crime rate and inexpensive South Asian food, and in the other you will find high-end retail. Further complicating matters, the context that defines these regions is dynamic. Consider safety in a military engagement. An

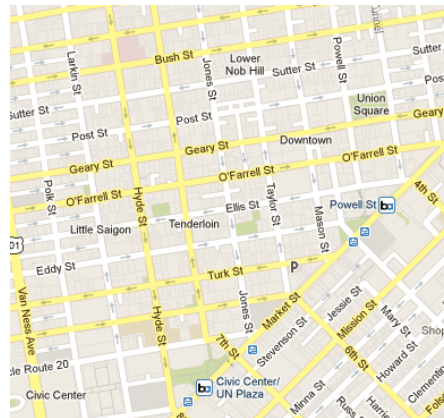


Figure 2: Neighborhoods in and around downtown San Francisco including Little Saigon, Tenderloin, and Union Square.

agent moving to safety must consider the capabilities, knowledge, and goals of the other participants. For example, in Figure 3, if the friendly tanks are advancing on the enemy tanks, the enemies may retreat and the scout is already safe. Although they operate at a different time scale, there are dynamics in the previous two examples as well. The locations of the desks in a classroom or the landmarks of a city will change over time.



Figure 3: If the Scout was ordered to safety, it would need to reason about the future actions of the Friendly and Enemy Tanks.

We hypothesize that these regions play an important role in describing continuous space and time. Therefore, including representations of these regions should improve the performance intelligent systems. Context-dependent spatial regions have been used to improve human activity recognition in a kitchen environment (Karg and Kirsch 2012). Given their functional nature, context-dependent spatial regions should be applicable to different reasoning tasks (e.g., planning, communication, intent recognition). Furthermore, learning the regions of a single environment is not very useful, because the grounding of the abstraction must change with the environment. Instead, we desire

solutions that include grounding context-dependent spatial regions in new and dynamic environments.

Abstracting Continuous Systems

To gain insight into this problem, we consider the abstractions defined for reasoning about continuous systems. Consider a bouncing ball. It falls with increasing speed until it hits the ground. Then, it reverses direction and rises with a negative acceleration until it has zero velocity and begins to fall again. This description is in terms of abstractions in space and time. In the late 1970s and early 1980s, the field of qualitative reasoning was started to formalize this type of reasoning. The majority of the field focused on algorithms for predicting possible behaviors of engineered systems and how such predictions could support other reasoning tasks. The spatial environment introduces additional challenges of multiple dimensions. Consequently, this led to the rich field of qualitative spatial reasoning (Cohn and Renz 2008). Much of the research in this field has focused on the definition of calculi and their inferential properties. In this section, we discuss the properties of these abstractions (Figure 4) to explore the representational challenge of context-dependent spatial regions.

Abstraction of each type of system begins with the mathematical formalism. Differential equations are a common representation of engineered systems, and geometry is typically used to represent space. Differential equations and geometry are defined over real numbers, and each can be abstracted into jointly-exhaustive pair-wise disjoint sets of abstractions. Qualitative differential equations (Kuipers 1994) abstract each variable into a quantity space consisting of landmarks and intervals. The coarsest quantity space is the sign algebra consisting of three possible values: Q- if the variable is negative, Q0 if the value is 0, and Q+ if the value is positive. In spatial reasoning, researchers typically define relations between pairs and triples of entities, such as Region Connection

Engineered System	Mathematical Formalism	Spatial Environments
Differential equations	Mathematical Formalism	Geometry
Quantity Spaces	Static Physical Quantity Abstraction	Region Connection Calculi
$Q+ + Q- = \{Q-, Q0, Q+\}$	Structural Ambiguity	$DC(a,b) + EC(b,c) = \{DR,PO,PP\}(a,c)$
Qualitative Physics Trajectories	Behavioral Ambiguity	Path Trajectories
Bounding Functions	Linking Qualitative and Quantitative	Metric-Diagram/Place-Vocabulary
Requirements and scenario	Task-Dependence	Goals, actions, and objects
Landmark Introduction Corresponding Values	Dynamic Interpretation	Context-Dependent Spatial Regions

Figure 4: Alignment of properties of abstractions used in analysis of engineered systems and spatial environments.

Calculi (Cohn *et al.* 1997) that describe topological relationships. One property of both of these abstractions is that they lead to ambiguous inferences. Without additional information, we cannot know if a positive number summed with a negative number is positive, negative or zero. Similarly, if region ‘a’ is disjoint with region ‘b’, and region ‘b’ is edge connected with region ‘c’, then ‘a’ is either disjoint from, edge connected, partially overlapping, a proper part of, or a tangential proper part of region ‘c’. Abstractions also lead to behavioral ambiguities. In our bouncing ball example, there is an ambiguity concerning the relationship of end point of the upward trajectory and the initial release point. In spatial reasoning, consider the example in Figure 3. If the scout is moving toward the friendly tanks and the enemy tanks are moving toward the scout, the scout may or may not reach safety in time. These ambiguities pushed researchers to explore methods integrating the abstractions with the underlying quantitative space. For reasoning about differential equations, Kuipers (1994) introduced *bounding functions* that create a quantitative envelope around an unknown underlying differential equation. Forbus *et al.* (1991) used a metric-diagram/place-vocabulary to maintain tight integration between the underlying geometry and the abstraction space used for inference.

Representation cannot be separated from reasoning. Models of engineered systems are created to understand their behavior. In design, engineers evaluate a model of their design against requirements (e.g., accelerating 0-60 mph in under 5 seconds) in some scenario, that is, a model of how a device would be used (e.g., the driver would shift gears at particular times on a dry level road). Spatial representations are used for many tasks (e.g., navigation, planning, database retrieval). For example, when planning a route for a mobile robot, it is necessary to consider the goals and actions of other agents in the environment to avoid collisions. This task dependence has two consequences: (1) additional knowledge is added to the underlying mathematical formalism, and (2) often, the abstractions used change during the course of the inference task. We explore this second point a bit more in the next paragraph.

Dynamic interpretation occurs when the abstractions change meaning during inference. In qualitative reasoning about engineered systems, this occurs through *landmark introduction* and *corresponding values* (Kuipers 1994). When simulating the behavior of a variable x , if its derivative goes to zero at time point t_0 , the value of that variable may be an extrema of the function (shown in Figure 5). Consequently, by adding a landmark l_{new} to the x 's quantity space and continuing simulation, our representation is better able to capture the dynamics of the function. For example, by creating a new landmark at the apex of the ball's bounce, this representation now

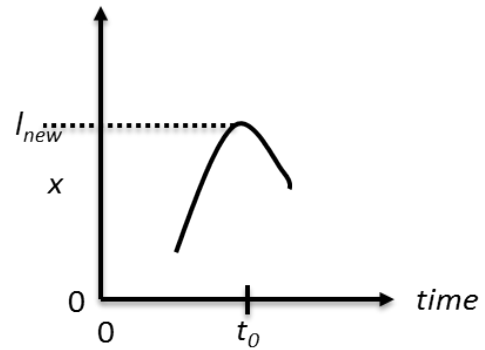


Figure 1: At t_0 , a new landmark, l_{new} , is added to x 's quantity space. Before t_0 , the value of x is the interval $(0, \infty)$. After t_0 , x is between 0 and l_{new} .

distinguishes between bouncing ball models with and without energy loss. The representation is dynamic because, before t_0 , the value of x is the interval from $(0, \infty)$, but after t_0 , the value of x is $(0, l_{new})$. Corresponding values provide another example of dynamic interpretation. Corresponding values refine a constraint during qualitative simulation. Consider the following equation, $a + b = c$. Corresponding values for this equation are triples of landmarks that satisfy the constraint, e.g., (a_i, b_i, c_i) . These can be added to a constraint during simulation, thereby increasing its discrimination for future states.

Dynamic interpretations are also being explored in spatial environments. For example, Zender *et al.* (2007) describe a socially aware robot tasked with following a human. The desired following location of the robot with respect to human changes with the context. For example, when the human approaches a door, the robot stays a bit further back to allow the human to open it. Also, when moving in a corridor, it is easier to predict the motion of the human the robot is following as well as the motion of other humans that are potentially obstacles. The desired location for the robot follower is an example of a context-dependent spatial region.

In the next section, we describe some existing approaches for dynamic interpretation of spatial abstractions, and their limitations.

Existing Approaches

Early work identifying the need to incorporate functional and geometric knowledge into a single representation comes from linguistics (Coventry 1998). Potential field models (Kelleher and Costello 2008) consider the context of distracting objects to both understand and generate locative expressions (e.g., the ball near the box). Locative expressions explicitly state the important objects in the environment (e.g., ball and box). This differs from

context-dependent spatial regions that incorporate other objects in the environment that are not included in the natural language (e.g., the desks help define the front of the classroom).

The robot follower from the previous section uses a hand authored task-specific encoding for the desired following location. One of the original motivations of abstraction was to increase the robustness of reasoning systems. Unfortunately, hand authored representations are typically brittle with respect to environment changes. Furthermore, the system lacks a general mechanism for reasoning about space. Thus, each new spatial concept will have to be encoded and refined by human engineers. Our challenge in this paper is to define an extensible spatial representation system to enable AI systems to improve their performance and extend their range of applicable tasks over time.

To alleviate brittleness, researchers train probabilistic models from human labeled examples. For example, Montello *et al.* (2003) learn a model for which points on a map are located within downtown Santa Barbara. One property of context-dependent spatial regions is that they change with the context. Unfortunately, most learned models require lots of training examples, and when the task or environment changes, a new model must be learned. One promising approach around this problem is to transfer the learned models to new environments. Karg and Kirsch (2012) use types and orientation of objects to recognize human activities in a new domestic environment.

Computational models of analogy provide another way of transferring instance-based definitions of regions. Lockwood *et al.* (2005) used analogy to select the appropriate spatial preposition for a new scene incorporating functional and geometric knowledge. Hawes *et al.* (2012) used *anchor points*, symbolic descriptions linking conceptual entities to perceived entities, to define context-dependent spatial regions in a classroom, and analogy to identify these regions in a new classroom. While analogy is promising because it enables learning from a single example, it is an open question to define an appropriate set of compositional spatial primitives to ground context-dependent spatial regions.

Importance of Dynamic Representations

The *poverty conjecture* states that “there is no problem-independent, purely qualitative representation of space or shape” (Forbus *et al.* 1991). We hypothesize that dynamic representations provide a way of incorporating the context of the problem into the abstract representation. We are motivated by the dynamic representations using in qualitative reasoning about engineered systems: landmark

introduction and corresponding values. These techniques expanded the expressivity of qualitative simulation.

Intelligent systems are beginning to encode dynamic representations for context-dependent spatial regions. By incorporating the goals, objects, and actions along with the geometry, current systems are able to identify fronts of rooms, follow humans appropriately, and recognize activities in domestic spaces. The challenge facing this community is to create a vocabulary that enables natural description of behavioral sequences supporting a range of inference tasks. While the meaning these abstractions is static, their referents in the spatial world change over time. For example, an agent would form the same plan to move to safety in many different environments, but the grounding of that region would differ depending on the context. Furthermore, if our reasoning systems could learn new types of regions from a few examples, that would dramatically decrease system development time while increasing the range of tasks for a given intelligent system.

References

- Cohn, A. G., Bennett, B., Gooday, J., and Gotts, N. M. (1997) “Qualitative Spatial Representation and Reasoning with the Region Connection Calculus,” *Geoinformatica*, vol. 1, no. 3, pp. 275–316.
- Cohn, A. G., & Renz, J. (2008). Qualitative spatial representation and reasoning. *Foundations of Artificial Intelligence*, 3, 551-596.
- Coventry, K. R., (1998). Spatial prepositions, functional relations and lexical specification. In P. Olivier & K. Gapp (Eds.), *The Representation and Processing of Spatial Expressions* (pp. 247-262). Lawrence Erlbaum Associates: Mahwah, NJ.
- Forbus, K., Nielsen, P., and Faltings, B. (1991). Qualitative spatial reasoning: The CLOCK project. *Artificial Intelligence*, 51(1-3), 417-471.
- Hawes, N., Klenk, M., Lockwood, K., Horn, G., and Kelleher, J. (2012). Towards a Cognitive System That Can Recognize Spatial Regions Based on Context. In *Proceedings of Twenty-Sixth AAAI Conference on Artificial Intelligence (AAAI-12)*. Toronto, ON.
- Karg, M., & Kirsch, A. (2012). Acquisition and use of transferable, spatio-temporal plan representations for human-robot interaction. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (pp. 5220-5226). IEEE.
- Kelleher, J. D., and Costello F. J. (2009). Applying computational models of spatial prepositions to visually situated dialog. *Computational Linguistics* 35, 2, 271–306.
- Kuipers, B. J. (1994). *Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge*. Cambridge, MA: MIT Press.
- Lockwood, K., Forbus, K., & Usher, J. (2005). SpaceCase: A model of spatial preposition use. *Proceedings of the 27th Annual Conference of the Cognitive Science Society*. Stressa, Italy.
- Montello, D. R., Goodchild, M. F., Gottsegen, J., & Fohl, P. (2003). Where's downtown?: Behavioral methods for determining referents of vague spatial queries. *Spatial Cognition and Computation*, 3, 185-204. Special Issue on "Spatial vagueness, uncertainty, granularity," B. Bennett, & M. Cristani (Eds.).

Zender, H., Martínez Mozos, O., Jensfelt, P., Kruijff, G.-J.M. and Burgard, W., (2008). Conceptual spatial representations for indoor mobile robots, *Robotics and Autonomous Systems*, Volume 56, Issue 6, Pages 493-502.