

An Architecture for General Spatial Reasoning

Samuel Wintermute

University of Michigan
2260 Hayward St.
Ann Arbor, MI 48109-2121
swinterm@umich.edu

Competence in interacting with the spatial world, the ability to move around an obstacle, or reach for a desired object, is one of the most immediate needs of any agent existing in such a world. Historically, many AI systems have focused on the higher-level parts of reasoning as a first priority. These systems represent problems in terms of abstract symbols, which can be manipulated to solve the problem. Symbolic reasoning has been successful for many applications involving high-level problem solving. However, if problems require precise reasoning over continuous spatial information, the applicability of symbolic approaches is less obvious. To steer a robot around an obstacle, signals sent to the actuators must be precisely calculated as a function of the relative position of the robot to the obstacle, among other factors. This type of reasoning is at a much lower level than what symbolic AI systems were originally designed for.

For my thesis work, I am extending a largely-symbolic AI system, the Soar cognitive architecture (Laird, 2008), to better handle spatial problems. A key aspect in the design of Soar is a commitment to generality: the goal of the architecture is to be able to solve the same breadth problems humans are able to solve. In addition, Soar is a psychologically-inspired architecture: a second goal is to solve problems in a manner similar to humans. These goals are reflected in the design of the existing architecture, and must be reflected in the design of any extension to it. Systems for spatial reasoning exist, but they are typically defined for limited domains, and in isolation from a comprehensive intelligent system. Spatial problem domains such as tasks in the blocks world, path planning, and mechanism understanding are currently addressed by very different AI techniques, but humans seem to use similar machinery for all spatial problems.

My approach to the problem derives from work in diagrammatic reasoning (e.g., Chandrasekaran, 2006) and systems exploring mental imagery (e.g., Lathrop, 2008). The system I have been developing is called Soar Visual/Spatial, or SVS. SVS augments symbolic working memory in Soar with short-term and long-term memories specialized for spatial information. Reasoning is then a process of manipulating both symbolic and lower-level perceptual data. Symbolic processing can both extract abstract qualitative information from spatial STM (high-level perception), and add information to spatial STM (imagery). A simplified architectural diagram, along with an example problem-solving trace, are shown in Figure 1.

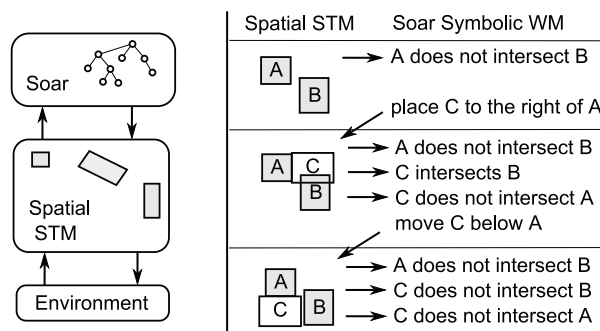


Figure 1. Soar/SVS architectural diagram (left), and a simple problem-solving trace (right)

In the example, the agent determines a non-intersecting position for a new object by inspecting symbolic information derived from high-level perception, and constructing symbolic imagery commands, which are interpreted in the spatial system.

Progress to Date

My work with SVS has had two main thrusts: defining and implementing the architecture itself, and investigating of the type of problem solving the architecture entails. These have proceeded in tandem, the theory behind the system has influenced its design, and vice versa.

One of the most difficult aspects of the design of SVS has been defining and implementing the interface between symbolic processing in Soar and the memories in SVS. This interface is crucial in defining the capabilities of the overall system. One of the first problems I looked at in this area was that of using symbolic descriptions to create new images in spatial STM (Wintermute and Laird, 2007). This work defined two classes of commands, where images can be described directly based on unambiguous geometry ("imagine the convex hull of objects A and B"), or indirectly based on higher-level properties ("imagine C to the right of A", as in the Figure). Using commands like these, along with a simple fixed system for extracting qualitative symbolic information from spatial STM (the direction and distance between objects, and whether or not they intersect), spatial problems can be solved. Problems addressed in this work include laying out a base in a real-time strategy game, and simple path planning.

Many spatial reasoning problems are not about static scenes, but about motion. In the above work, this was done by reducing motion to simple geometry. The path covered

by a convex object moving in a straight line is simply the convex hull of the object at its starting and ending locations. However, this approach does not apply generally—some movements traverse much more complicated paths. In Wintermute and Laird (2008), this issue was investigated in detail, arriving at a system where detailed motion can be simulated within spatial STM. Motion problems are solved by symbolic processing stepping through the simulation, inspecting the scene for simple qualitative properties between steps. Examples of basic path planning for a car-like vehicle and simple physical simulation of a blocks-world problem were covered in this work.

This car path planning example was further built upon in Wintermute (2009), and adapted to implement a modern motion planning algorithm, RRT (Rapidly-exploring Randomized Trees, LaValle, 2006). This work did not significantly change the architecture itself, as the existing system was able to implement the algorithm, but served to advance the theory behind the architecture. In particular, it was argued that allowing spatial simulation of action can simplify the constraints on the design of an overall system. The actions of the agent do not need to correspond to simple symbolic transformations to be useful, and high-level perception processes can be simpler than would otherwise be required, since the symbolic state doesn't need to be rich enough to capture implications of actions.

These arguments have been built upon and made more formal in Wintermute and Laird (2009). Here, it is argued that in a general-purpose system, the available symbolic representation of any given problem may not perfectly capture all details of that problem. That is, the representation might not be sufficient to allow the problem to be solved solely in terms of that representation. An imperfect symbolic problem representation can still be very useful, though, and if a spatial representation is also available, imagery can make up for some of the shortcomings of an imperfect symbolic representation. The basic approach is that the symbolic problem state is used to determine a potential action choice, which is then tested in an internal simulation (with its initial state grounded in perception). Internal perception of the imagined result reveals the next symbolic state the action is likely to lead to. This next state may differ from the next state the symbolic process expected (as its representation is imperfect), in which case the agent must re-evaluate its choice. I am calling this particular use of imagery *simulative imagery*.

Future Work

For the remainder of my thesis work, I plan to focus on further studying simulative imagery in the context of sequential decision making. I believe that this is a powerful technique that could be adapted to many AI architectures (it is not Soar-specific), so some work must be spent formalizing and evaluating simulative imagery outside of the context of SVS. Indeed, there are many existing AI systems that could be considered as using some form of

simulative imagery, so work must go into attempting to unify those implementations under a common framework.

That said, SVS itself, and its concerns about generality and human similarity, will remain a large part of the work. One of the key advantages of simulative imagery is that, since it relaxes constraints on symbolic problem representation, the same basic perceptual mechanisms can be used on many problems (e.g., car path planning in Wintermute (2009) and blocks world in Wintermute and Laird (2009)). In this way, it increases the generality of the agent. However, imagery in SVS allows many more capabilities than simulative imagery, such as the ability to represent non-spatial problems spatially (as is often studied in psychology), and is potentially useful for tasks other than sequential decision making (such as mechanism understanding). These areas will likely also be addressed, but probably in the form of case studies applying SVS to the problem, rather than general theories.

Since a claim about SVS is that it handles many problems, work must go into agent implementations that demonstrate this. Problem domains are more compelling (and scientifically relevant) if they are complex and not hand-designed for the purposes of evaluating the system in question. For this reason, (tentative) plans are to create an interface between SVS and an Atari video game system emulator. The graphics and controls here are simple, so perception and action are tractable, but the problems are interesting. This will provide a large library of pre-built domains in which SVS can be demonstrated and evaluated.

References

- Chandrasekaran, B., 2006. Multimodal Cognitive Architecture: Making Perception More Central to Intelligent Behavior. In *Proceedings of AAAI-06*.
- Laird, J.E., 2008. Extending the Soar Cognitive Architecture. In *Proc. First Conference on Artificial General Intelligence*.
- Lathrop, S.D., 2008. *Extending Cognitive Architectures with Spatial and Visual Imagery Mechanisms*. PhD Thesis, University of Michigan.
- LaValle, S.M., 2006. *Planning Algorithms*, Cambridge University Press.
- Wintermute, S. & Laird, J.E., 2007. Predicate Projection in a Bimodal Spatial Reasoning System. In *Proceedings of the 22nd AAAI Conference on Artificial Intelligence*.
- Wintermute, S. & Laird, J.E., 2008. Bimodal Spatial Reasoning with Continuous Motion. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*.
- Wintermute, S. & Laird, J.E., 2009. Imagery as Compensation for an Imperfect Abstract Problem Representation. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society*.
- Wintermute, S., 2009. Integrating Reasoning and Action through Simulation. In *Proceedings of the Second Conference on Artificial General Intelligence*.