Ready or Not, Here I Come …

Magdalena Bugajska, William Adams, Scott Thomas, J. Gregory Trafton, and Alan C. Schultz

Naval Research Laboratory, Code 5515
Washington, DC 20375
{magda, adams, thomas, trafton, schultz}@aic.nrl.navy.mil

Introduction

We have been trying to explore embodied cognition by putting computational cognitive models on physical robots. We believe that there are three primary advantages to doing so:

1. Algorithms written for traditional real-time robotic systems tend to use efficient mathematical representations such as matrices, which may not be natural for people to use. People seem to use a combination of spatial and propositional knowledge (Shepard & Metzler, 1971;Taylor & Tversky, 1992). Using common representations also eliminates complex processes of translation and decreases confusion due to loss of information.
2. Second, if a human is going to collaborate in shared space with a robot, the robot should not exhibit unnatural or "alien" behaviors. Therefore, we propose that some robot behaviors should be modeled on human performance of a task.
3. Third, we believe that in some areas of robotics, the use of computational cognitive models, rather than traditional control algorithms, is simply a better programming technique for the task.

For a domain, we have chosen hide and seek, because it forces us to work in a complex, dynamic environment, it allows us to explore embodied cognition issues (i.e., when is someone hiding?), and allows us to explore methods of combining both cognitive and robotic/AI methods into a single system. Hide and seek also encapsulates many interesting real-world tasks including urban search and rescue, homeland defense (finding nuclear, biological, or chemical weapons), and military actions such as finding hidden troops.

In our previous work (Trafton, et al., 2004), we have described a case study of a young child learning how to play hide and seek. We built an ACT-R (Anderson and Lebiere, 1998) model of the sequence of learning that the child went through, put that model on the robot, and showed that the robot is able to learn where the appropriate places to hide are, and play hide and seek. We were able to support a proposed cognitive-level hypothesis that, even though young children do not have perspective taking abilities, they are able to play hide and seek by learning object features and relations (e.g., hiding inside or under objects). Our cognitive model showed at a qualitative level the same kinds of hiding behavior that a child shows. In addition, we enabled a seeking behavior through inspection of good hiding places, which allowed the robot to play a credible game of hide and seek without perspective taking.

Clearly, there is a limit to what strategies can be learned without perspective taking abilities – a young child will not learn to simply hide behind objects, to approach the seeker without being seen, etc. For this demonstration, we built a new ACT-R model that includes spatial perspective taking, which allows for a more complex and dynamic game of hide and seek. The new model was implemented using jACT-R, the Java version of ACT-R, to allow for future extensions using the spatial component of this implementation of ACT-R, ACT-R/S (Harrison and Schunn, 2002).

Modeling Perspective Taking

Our new cognitive model uses the Player/Stage simulator to model the environment and visual perception necessary for spatial perspective taking. Each simulated agent, human or robot, is equipped with a vision system with appropriately assigned range- and field-of-view, which is capable of detecting color blobs. The robot creates a model of the real world based on the information provided by its range sensors and color camera. All perceivable objects and agents added to the simulated environment are marked with unique colors and/or markers to further simplify model of the perception process.

The cognitive model of perspective taking is written so a general question in the form "Can X see Y from Z's perspective?" can be answered. When the question is posed, the cognitive model places X at Z's location in X's model of the world, and using the vision system constrained by Z's capabilities, determines if X can see Y. This generalization allows us to keep this skill independent of the domain.

Hide and Seek with Perspective Taking

In this framework, hiding comes down to moving to different locations associated with a negative answer to the question "Can ROBOT see ROBOT from HUMAN's perspective?" i.e. places from which human cannot see the robot. As before, seeking simply requires visiting good hiding places.

To determine if a location is a good hiding place, the cognitive model consults the simulation to determine if the robot will be seen from the human's perspective if it hides at the chosen location; the robot is placed at the human's location with the human's visual capabilities, to check if it can see itself at the chosen hiding location. This process is repeated until a good hiding place is found. Hiding places can be chosen around objects as well as in empty areas between objects. The order in which locations are evaluated can be random, determined by sensing order, or other heuristics such as distance from the seeker or degree of occlusion. Extensibility of the model allows us to model a variety of mature strategies to the game as well as various military missions requiring stealthy operations.

It is easy to see that such a game of hide and seek can be quite dynamic with both hiding and seeking agents moving around. In such case, the position of the human seeker is estimated based on his last known position, which is initially known and then updated using color blob detection, since the participant is required to wear a vest of predetermined color.

Robotics

In addition to cognitive models of perspective taking and hiding and seeking, a combination of non-cognitive methods (primarily for mobility and object recognition) and cognitively inspired interactions (primarily for communicating with a person) was used. This project draws on the mobility capabilities of the previously developed WAX system (Schultz, Adams, &Yamauchi, 1999) and our human-robot interaction research (Perzanowski, et al., 2001). The CMVision package (Bruce, Balch, & Veloso, 2000) was used to provide simple color blob detection using an inexpensive digital camera mounted on the robot.

The following is a sample script of the human-robot interaction during one turn of a game of hide and seek:

1. Human initiates the game by saying, "Let's play hide and seek" or "Let's play a game" (since currently our robot only knows one game).
2. Human can elect to hide or to become a seeker, by using phrases like "I will hide" or "Go hide."
3. The seeker stands at the "home" location and waits for the other player to hide (for example, counts to ten).
4. If the robot was the one to hide, the human announces when the robot is found by saying "I can see you," "I found you," etc. If the human was the one to hide, the robot will also make similar announcements.

Conclusions

We believe that integrating a computational cognitive model with a robotic system gives us a more principled way of creating behaviors such hiding and seeking and basic skills such as perspective taking. For this demonstration, we implemented a cognitive model of hiding and seeking, which incorporates a cognitive model of perspective taking and allows for more complex and dynamic behaviors than our previous model. With our new model, we are able to further explore the integration issues involved with cognitive science and AI.

Acknowledgements

This work was supported by the Office of Naval Research under work order number N0001405WX30001. The views and conclusions contained in this document should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U. S. Navy.

References

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum.

Bruce, J., Balch, T., and Veloso, M. (2000). "Fast and inexpensive color image segmentation for interactive robots," In Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 3, pp 2061-2066, Takamatsu, Japan.

Harrison, A. M., & Schunn, C. D. (2002). "ACT-R/S: A Computational and Neurologically Inspired Model of Spatial Reasoning." In Proceedings of the 24th Annual Meeting of the Cognitive Science Society, Fairfax, VA.

Perzanowski, D. Schultz, A.C. Adams, W. Marsh, E. Bugajska, M. (2001). "Building a Multimodal Human-Robot Interface." In *IEEE Intelligent Systems*, pp. 16-20, Jan./Feb, 2001.

Schultz, A., Adams, W., and Yamauchi, B. (1999). "Integrating exploration, localization, navigation and planning with a common representation." In *Autonomous Robots* vol. 6, pp 293-308. Kluwer Academic Publishers.

Shepard, R. N., & Metzler, J. (1971). "Mental rotation of three-dimensional objects." In *Science*, 171, 701-703.

Taylor, H. A., & Tversky, B. (1992). "Spatial mental models derived from survey and route descriptions." In *Journal of Memory and Language*, 31, 261-292.

Trafton, J. G., Schultz, A. C., Cassimatis, N. L., Hiatt, L. M., Perzanowski, D., Brock, D. P., et al. (in press). "Communicating and collaborating with robotic agents." In R. Sun (Ed.), *Cognition and multi-agent interaction: From cognitive modeling to social simulation*. Cambridge University Press.