Proceedings of the Twentieth National Conference on Artificial Intelligence (AAAI-05), Pittsburgh, PA. July 9-13, 2005.

Tekkotsu: A Framework for AIBO Cognitive Robotics

David S. Touretzky and Ethan J. Tira-Thompson
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213-3891
dst@cs.cmu.edu, ejt@andrew.cmu.edu

Introduction

Tekkotsu (the name means "framework", literally "iron bones" in Japanese) is an application development framework for the Sony AIBO robot dog (Tira-Thompson 2004). It provides a layer of abstraction above the Sony OPEN-R software interface and offers a variety of services, including an efficient event routing architecture, the ability to share C++ objects across processes, a hierarchical state machine formalism for constructing behaviors, and an extensive collection of wireless remote monitoring and teleoperation tools. The latter are written in Java for portability. Tekkotsu is an open source project and builds on the work of several other open source developers: it provides forward and inverse kinematics solvers based on ROBOOP (Gordeau 2005), simple object detection using CMVision (Bruce, Balch, & Veloso 2000), and two walking engines, one from CMPack-02 (Veloso *et al.* 2002) and one from the University of Pennsylvania (Cohen *et al.* 2004). Tekkotsu is currently in use at over 20 universities around the world, either in introductory robotics courses or for robosoccer. It is available at `www.Tekkotsu.org`.

Over the last two years we have been developing a new layer of Tekkotsu to support an approach to robot programming that we call "cognitive robotics". The idea is to provide a set of higher level primitives for perception and action, inspired by ideas from cognitive science, so that programmers can construct intelligent behaviors at a much more abstract level. Three components of our approach are described here: visual routines, dual-coding representations, and perceivable affordances.

Visual Routines

Ullman proposed that low-level vision might be implemented as a set of composable parallel operators he called visual routines (Ullman 1984). There is some evidence that such operations are performed in primary visual cortex (Roelfsema, Lamme, & Spekreijse 2000). Tekkotsu provides a set of visual routines that operate on 2D "sketches," starting with a color-segmented camera image (Halelamien 2004; Tira-Thompson *et al.* 2004). Operators include basic arithmetic, comparison, and boolean functions, color masking, connected components labeling, flood-fill, boundary distance, skeletonization, and neighbor sum.

Sketches are automatically organized into a derivation tree, i.e., the result of applying an operator to a sketch is a new sketch that references the original sketch as its parent. A remote viewing tool allows the programmer to "see inside the dog's head" by examining the derivation tree and any of its component sketches.

Dual-Coding Representations

Paivio's "dual coding theory" of representations posits parallel verbal and non-verbal (imagistic) systems with extensive referential connections between them (Paivio 1986). In Tekkotsu, sketches provide the imagistic representation, and "shapes" provide a complementary symbolic representation. Shapes such as lines and ellipses can be extracted from images using visual routines operators. Lines can be described symbolically in several ways, e.g., by specifying two endpoints, or a point and a slope, which provides for line segments, rays, and lines with infinite extent. Some types of computations are more easily performed in this symbolic space, e.g., constructing a line perpendicular to another.

Built-in rendering operators automatically convert from a symbolic representation to an iconic one, yielding a sketch of the shape. Thus, computations can be carried out using a mix of iconic and symbolic operations. Shapes are also recorded in the derivation tree. Figure 1 shows a derivation tree representing a tic-tac-toe board that has been parsed into a collection of sketches and shapes. Figure 2 shows a sketch containing the pixels making up the game board, with four shapes (the extracted lines) superimposed on top of it.

Symblic representations have also been used to construct a global map from multiple camera views (Tira-Thompson *et al.* 2004). It is much easier to match shapes across views then to match pixels.

Perceivable Affordances

A high-level approach to robot programming should relieve the programmer from having to perform complex perceptual or kinematics calculations to manipulate objects in the world. A perceptual system based on Gibson's theory of "affordances" (Gibson 1979) can meet this goal by representing objects in terms of the actions that can be performed on them.

We have implemented a first version of an affordance recognizer that uses visual routines plus knowledge of the body's physical constraints to determine the feasibility of actions (Edirisinghe 2005). For example, suppose the robot is looking at a segment of a line drawn on the ground. The Trace Line action moves the robot's gaze along the line in order to locate the endpoints. Any line potentially "affords" a tracing action, but if the head pan joint is already at its turning limit, the affordance recognizer detects that the action is not feasible in the current situation. Moveable objects such as balls and tic-tac-toe game pieces afford pushing actions, but these actions are only feasible when the object is within the robot's reach. Fortunately, objects also afford locomotion actions that can move the robot to a position where other actions become feasible. The idea could be extended by introducing gestalt perception, e.g., a set of lines on the ground suggesting a bounded region might be perceived as a "container" that affords moving an object into it.

Our approach allows the robot's behavior to be determined simply by specifying a policy for selecting from the set of affordances the robot currently perceives. When integrated with visual routines and a global map, this will provide a powerful high level approach to robot programming.

Acknowledgments

The initial development of Tekkotsu was funded by a grant from the Sony Corporation.

References

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 (IROS '00)*, volume 3, 2061 – 2066.

Cohen, D.; Ooi, Y. H.; Vernaza, P.; and Lee, D. D. 2004. The University of Pennsylvania Robocup legged soccer team. Technical report, University of Pennsylvania. Available at http://www.cis.upenn.edu/robocup/UPenn04.pdf.

Edirisinghe, D. C. 2005. Perceivable affordances and mobile robots. Senior honors thesis, Computer Science Department, Carnegie Mellon University.

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gordeau, R. 2005. Roboop – a robotics object oriented package in C++. Available at http://www.cours.polymtl.ca/roboop/.

Halelamien, N. S. 2004. Visual routines for spatial cognition on a mobile robot. Senior honors thesis, Carnegie Mellon University, Pittsburgh, PA. Available at www.Tekkotsu.org/media/thesis_neilh.pdf.

Paivio, A. 1986. *Mental Representations: A Dual-Coding Approach*. New York: Oxford University Press.

Roelfsema, P. R.; Lamme, V. A. F.; and Spekreijse, H. 2000. The implementation of visual routines. *Vision Research* 40:1385–1411.

Tira-Thompson, E. J.; Halelamien, N. S.; Wales, J. J.; and Touretzky, D. S. 2004. Cognitive primitives for mobile robots. In Schultz, A., ed., *The Intersection of Cognitive Science and Robotics: From Interfaces to Intelligence*, 110–111. AAAI Fall Symposium, report FS-04-05.

Tira-Thompson, E. J. 2004. Tekkotsu: A rapid development framework for robotics. Master's thesis, Carnegie Mellon University, Pittsburgh, PA. Available at www.Tekkotsu.org/media/thesis_ejt.pdf.

Ullman, S. 1984. Visual routines. *Cognition* 18:97–159.

Veloso, M.; Lenser, S.; Vail, D.; Roth, M.; Stroupe, A.; and Chernova, S. 2002. CMPack-02: CMU's legged robot soccer team. In Kaminka; Lima; and Rojas., eds., *RoboCup 2002: Robot Soccer World Cup VI*. Springer-Verlag.

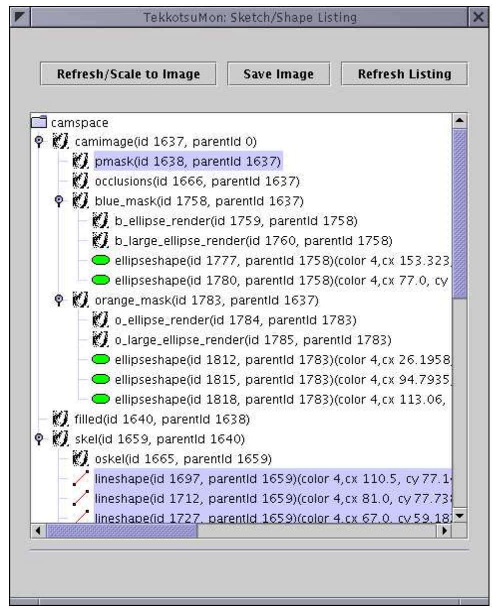

Figure 1: Derivation tree showing sketches and shapes extracted from a tic-tac-toe board image.

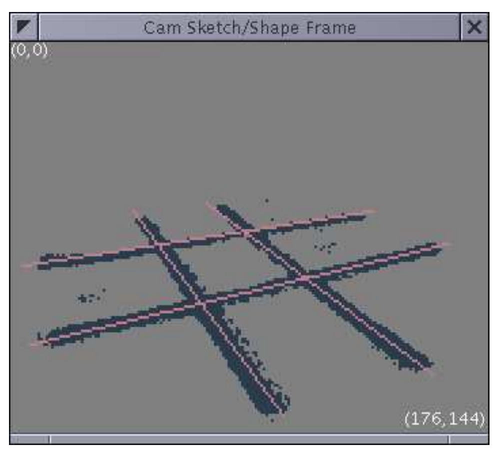

Figure 2: Display of the sketch and shapes selected from the derivation tree.