# Toward a Behavior-Grounded Representation of Tool Affordances

**Alexander Stoytchev**
Mobile Robot Laboratory, Georgia Institute of Technology
Atlanta, Georgia 30332
E-mail: saho@cc.gatech.edu

## 1. Introduction

The ability to use tools is one of the hallmarks of intelligence. Tool use is fundamental to human life and has been for at least the last two million years. We use tools to extend our reach, to amplify our physical strength, to transfer objects and liquids, and to perform many other tasks. A large number of animals have also been observed to use tools (Beck, 1980). Some birds, for example, use twigs or cactus pines to probe for larvae in crevices which they cannot reach with their beaks. Sea otters use stones to open hard-shelled mussels. Chimpanzees use stones to crack nuts open and sticks to reach food, dig holes, or attack predators. Orangutans fish for termites with twigs and grass blades. Horses and elephants use sticks to scratch their bodies. These examples suggest that the ability to use tools is an adaptation mechanism used by many organisms to overcome the limitations imposed on them by their anatomy.

Despite the widespread use of tools in the animal world, however, studies of *autonomous* robotic tool use are still rare. There are industrial robots that use tools for tasks such as welding, cutting, and painting, but these operations are carefully scripted by a human programmer. Robot hardware capabilities, however, continue to increase at a remarkable rate. Humanoid robots such as Honda's Asimo, Sony's Qrio, and NASA's Robonaut feature motor capabilities similar to those of humans. In the near future similar robots will be working side by side with humans in homes, offices, hospitals, and in outer space. It is difficult to imagine that these robots that will look like us, act like us, and live in the same physical environment like us, will be very useful if they are not capable of something so innate to human culture as the ability to use tools. Because of their humanoid "anatomy" these robots undoubtedly will have to use external objects in a variety of tasks, for instance, to improve their reach or to increase their physical strength. These problems, however, have not been well addressed by the robotics community.

Another motivation for studying robot tool behaviors is the hope that robotics can play a major role in answering some of the fundamental questions about tool-using abilities of animals and humans. After ninety years of tool-using experiments with animals there is still no comprehensive theory that attempts to explain the origins, development, and learning of tool behaviors in living organisms.

Progress along these two lines of research, however, is unlikely without initial experimental work which can be used as the foundation for a computational theory of tool use. This poster provides an overview of our dynamics simulator as well as preliminary simulation results for extension-of-reach tool tasks.

## 2. Affordances & exploratory behaviors

A simple object like a stick can be used in numerous tasks that are quite different from one another. For example, a stick can be used to strike, poke, prop, scratch, pry, dig, etc. It is still a mystery how animals and humans learn the affordances of objects and what are the cognitive structures that they use to represent them.

James Gibson (1979) defines affordances as "perceptual invariants" that are directly perceived by an organism and enable it to perform tasks. Gibson is not specific about the way in which affordances are learned but he suggests that some affordances are learned in infancy when the child experiments with objects. For example, an object affords throwing if it can be grasped and moved away from one's body with a swift action of the hand and then letting it go. The perceptual invariant in this case is the shrinking of the visual angle of the object as it is flying through the air. This very interesting "zoom" effect will draw the attention of the child (Gibson, 1979, p. 235).

The related work on animal object exploration indicates that animals use stereotyped exploratory behaviors when faced with a new object (Power, 2000, Lorenz, 1996). This set of behaviors is species specific and may be genetically predetermined. For some species these tests include almost their entire behavioral repertoire: "A young corvide bird, confronted with an object it has never seen, runs through practically all of its behavioral patterns, except social and sexual ones." (Lorenz, 1996, p. 44).

Recent human studies also suggest that the internal representation for a new tool used by the brain might be encoded in terms of specific past experiences (Mah and Mussa-Ivaldi, 2003). Furthermore, these past experiences consist of brief feedforward movement segments used in the initial exploration

of the tool. A tool task is later learned by dynamically combining these sequences.

Thus, the properties of a tool that an animal is likely to learn are directly related to the behavioral and perceptual repertoire of the animal. Furthermore, learning these properties should be relatively easy since the only requirement is to perform a (small) set of exploratory behaviors and observe their effects. The next section describes our current approach to learning tool affordances by a robot, which was inspired by the above mentioned examples.

A shortcoming of this approach is that there are tool properties that are unlikely to be discovered since the required exploratory behavior is not available to the robot. This problem has also been observed in animals, e.g., macaque monkeys have significant difficulties learning to push an object with a tool away from their bodies because this movement is never performed in their normal daily routines (Ishibashi et al., 2000). This problem should be resolved, however, if the ability to learn new exploratory behaviors is added.

## 3. Tool Representation

Let $\beta_{b_1}, \beta_{b_2}, \ldots, \beta_{b_k}$ be the binding (or grasping) behaviors available to the robot. Let $\beta_{e_1}, \beta_{e_2}, \ldots, \beta_{e_n}$ be the exploratory behaviors available to the robot. The perceptual observations of the robot form a perceptual vector: $o_1, o_2, \ldots, o_m$. A specific instance of this vector will be denoted with $\hat{o}_1^i, \hat{o}_2^i, \ldots, \hat{o}_m^i$.

With this notation in mind, the functionality of a tool when applied to a specific object can be represented with an *Affordance Table* as shown below.

| Binding Behavior | Exploratory Behavior(s) | Observation Vector |
|---|---|---|
| $\beta_{b_3}$ | $\beta_{e_1}, \beta_{e_2}, \ldots$ | $\hat{o}_1^1, \hat{o}_2^1, \ldots, \hat{o}_m^1$ |
| $\beta_{b_5}$ | $\beta_{e_6}, \beta_{e_3}, \ldots$ | $\hat{o}_1^2, \hat{o}_2^2, \ldots, \hat{o}_m^2$ |
| $\ldots$ | $\ldots$ | $\ldots$ |
| $\beta_{b_k}$ | $\beta_{e_3}, \beta_{e_7}, \ldots$ | $\hat{o}_1^i, \hat{o}_2^i, \ldots, \hat{o}_m^i$ |

In each row of the table, the first entry represents the behavior used to attach the tool to the robot. The second entry represents the sequence of exploratory behaviors applied to/with the tool. The last entry represents the observation vector.

To limit the number of observations we use the conclusions reached in animal studies that primates detect only visual *contact* between the tool and the incentive and *movement* of the incentive (Povinelli et al., 2000). Thus, the current implementation of this model supports only six types of observations: x and y position of the incentive, $\Delta x$ and $\Delta y$ of the incentive, and $t_x$ and $t_y$ the contact point between the tool and the incentive relative to the visual centroid of the tool. All variables are continuous.

The exploratory behaviors consist of different arm movement patterns: e.g., move away, move toward, sweep, shake, grasp, drop, etc. In addition to that there are two binding behaviors - grasp and regrasp for attaching the tool to the robot. The table is filled

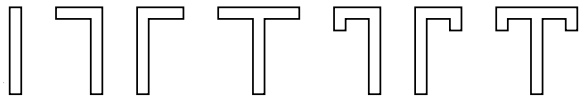

Figure 1: The seven tools used in the experiments. From left to right: straight stick, left L-rake, right L-rake, T-rake, left hook, right hook, and T-rake with hooks.

with values during learning trials in which the robot randomly chooses and performs a binding behavior followed by one or more exploratory behaviors.

Before test trials, the observations of the table are indexed pairwise into three separate kd-trees (Arya et al., 1998). These trees allow fast *approximate* nearest neighbor searches with a controllable maximum error. To solve a tool task the robot queries the table to select a behavior that is likely to move the incentive in the desired direction.

## 4. Experimental Platform

Currently we are testing this representation on an extension-of-reach task with 7 different tools (Figure 1). Extension of reach experiments have been used for the last 90 years to test the intelligence and tool-using abilities of primates (Köhler, 1931, Povinelli et al., 2000). In these experiments the animal is prevented from getting close to an incentive and must use one of the available tools to bring the incentive within its sphere of reach.

The test platform used in the experiments is a dynamics simulator for robot-tool interaction developed in house. The dynamics are calculated using the Open Dynamics Engine library (Smith, 2003). The OpenGL rendered simulation window is treated as a camera image. Color segmentation is used to detect the location and movement of the incentive, tools, and robot links. The robot is a simulated manipulator arm with two rotational joints and a gripper. A prismatic joint at the base of the robot allows it to raise and lower its hand. MPEG movies from the simulator are available at: http://www.cc.gatech.edu/~saho/tooluse/

## References


Arya, S. et al. (1998). An optimal algorithm for approximate nearest neighbor searching. *J. of the ACM*, 45:891–923.

Beck, B. B. (1980). *Animal Tool behavior: The use and manufacture of tools by animals.* Garland STMP Press, NY.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Houghton Mifflin, Boston.

Ishibashi, H., Hihara, S., and Iriki, A. (2000). Acquisition and development of monkey tool-use: behavioral and kinematic analyses. *Can. J. Physiology and Pharmacology*, 78:958–66.

Köhler, W. (1931). *The mentality of apes.* Harcourt, Brace, and Co., NY.

Lorenz, K. (1996). Innate bases of learning. In Pribram, K. H. and King, J., (Eds.), *Learning as Self-Organization.* Lawrence Erlbaum Associates, Publishers, Mahwah, NJ.

Mah, C. D. and Mussa-Ivaldi, F. A. (2003). Evidence for a specific internal representation of motion-force relationships during object manipulation. *Biol. Cybernetics*, 88(1):60–72.

Povinelli, D. et al. (2000). *Folk Physics for Apes: The Chimpanzee's theory of how the world works.* Oxford U. Press.

Power, T. G. (2000). *Play and Exploration in Children and Animals.* Lawrence Erlbaum Associates, Mahwah, NJ.

Smith, R. (2003). *The Open Dynamics Engine User Guide.* http://opende.sourceforge.net/.