ABSTRACT

WILLIAMS, LLOYD C. Dynamic Ontology Driven Learning and Control of Robotic Tool Using Behavior. (Under the direction of Dr. Robert St Amant.)

One of the most interesting and rich fields of recent artificial intelligence (AI) research has come from examining embodied agents, the creation of which, poses interesting challenges and opportunities. Many traditional AI approaches which have previously proven successful quickly fail in the face of the unique challenges facing embodied agents. There is extensive multidisciplinary research into solving these problems, employing ideas and theory from not just computer science, but cognitive science, psychology, philosophy, neuroscience, as well as a range of other fields. Although the nature of embodied intelligence has risen to prominence in AI research relatively recently, animal behaviorists have been examining it for decades. One of the most explored areas of research into the nature of natural embodied intelligent agents (animals) involves their use of tools. We believe that the creation of artificial tool using behaviors yields insights into the nature of intelligence.

The proposed research will survey animal tool using behaviors and argue that some form of imitation may serve as an integral part of most animal tool using behavior. This claim, for the significance of imitation in tool use, will be supported with results from ethology, psychology and neuroscience. We will present a system based on multidisciplinary research that employs action ontologies to enable robotic imitation. We will demonstrate with this research that if mechanisms for imitative behaviors are implemented on a robotic platform, these imitative mechanisms may then be employed to enable tool using behaviors. While the achievement of tool using behaviors through this type of imitative mechanism is a novel and significant technical achievement in and of itself, it's success also provides insight into how tool using behaviors may have first arisen in animals.

Dynamic Ontology Driven Control and Learning of Robotic Tool Using Behavior

by
Lloyd C. Williams

A dissertation to be submitted to the Graduate Faculty
of North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Computer Science

Raleigh, North Carolina

2009

APPROVED BY:

_____	_____
Dr. Ronald Endicott	Dr. Thomas Honeycutt
_____	_____
Dr. Robert St. Amant Chair of Advisory Committee	Dr. R. Michael Young

DEDICATION

For my parents

BIOGRAPHY

Lloyd Williams was born in Durham North Carolina. He received his Bachelors Degree from Vanderbilt University. He then served for two years in Peace Corps in the Dominican Republic. Subsequent to the Peace Corps, he worked in Information Technology before returning to North Carolina State University to successfully pursue a Masters and then a Doctorate in Computer Science.

ACKNOWLEDGEMENTS

I would like to thank my wife Laura, who helped me in innumerable ways in the completion of this thesis.

Also a great deal of thanks is due to the support to fellow members of the Knowledge Discovery Lab (KDL). Thomas Horton deserves special recognition for tireless proofreading, and support. Also, Brent Dennis, Sarat Kocherlakota, Amit Sawant, and Laura Tateosian helped and made the journey far more fun than it would have been without them. There are many other members of the KDL to whom I will always be grateful for friendship and support. It is impossible to think of KDL without mentioning Dr. Christopher Healey. Dr Healey could always be counted on as a friend, a mentor, and to provide a laugh when my frustration got the better of me.

And of course, I would like to thank my advisor, Dr. Robert St Amant, for his support and insight.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

1.1 Introduction

In the early days of artificial intelligence (AI), expectations were raised very high by the development of systems capable of impressive performance, at human levels or better, in areas such as search and game playing. During these early days there seemed to be no limit to what AI would achieve, but some problems remained remarkably persistent in the face of the early promise. One of the better known examples of this failure was robotic control software. Planners could generate good plans for robots, but in execution, early robots would fall down stairs or continuously bump into a corner more often than they would correctly execute their assigned plans. In the early 1990s there was a revolution in robotics, when researchers came to recognize the importance of the embodied nature of robots in their problem solving tasks (Brooks 1999, Clancey 1997).Recognizing the importance of embodiment has resulted in significant research not just in robotics, but

also psychology (Greeno 1994, Young, McNeese 1995), neuroscience (Clancey 1993), cognitive science (Kirsh 1995), ethology (Povinelli 2000) and philosophy (Clark 1998).

Embodiment is a somewhat fluid concept. Essentially the idea is that embodied agents are intelligent agents that interact with an environment through some body. Research on embodiment typically places a higher value on active perception (Ballard et al. 1995), being able to handle dynamic change in the environment (Brooks 1991), and biologically-inspired designs (Arkin 1998), especially in robotics; a lesser value is placed on artificial task environments, abstract problem representations, and arbitrarily long computations.

The focus of this dissertation is on a specialized category of embodied problem solving: the ability to use tools. Tool use is an interesting domain for AI research, because of the breadth of examples we find in the animal kingdom. Varying levels of tool-using competence can be seen in animals as limited (in terms of their cognitive capabilities) as ants and ant lions, and of course as complex as primates. In contrast to other aspects of intelligent behavior that are largely if not exclusively seen in human beings, such as language ability, tool-using ability ranges along a spectrum in the animal kingdom. We present and evaluate tool use as an embodied, general problem-solving task that can provide significant insight into embodied intelligence. A survey of animal tool use will demonstrate that while not all tool using behaviors require intelligence, some apparently do. We will examine these problems in detail and use examinations of learning, behavior

and representation from several fields to provide support for specific claims about how intelligent tool-using abilities can plausibly arise.

The main claim, in simple form, is the following:

Imitative mechanisms give rise to tool-using capabilities.

That is, if an agent is built with mechanisms that give it imitative capabilities, the agent also becomes capable of carrying out tool-using behaviors based on those same mechanisms. The bulk of this dissertation is devoted to collecting and producing evidence for this claim.

Evaluation of embodied cognitive systems, especially robotic systems (Brooks 1991) but also simulated physical systems, is notoriously difficult. That is, while it is straightforward to design experiments to test whether a system meets some specific performance criteria, it is much more difficult to evaluate the proposition that a given architecture or set of algorithms represents an appropriate model for some category of real-world activities. It will thus be worth explaining the research agenda of this project at the outset.

We will begin with a survey of tool use as it is exhibited in the non-human animal world. We believe that a focus on tool-using tasks can give insight into the levels of abstraction

at which it is appropriate to represent behavior, the structure of such representations, and the content of the representations. We will see that despite a startling number of exemplars of animal tool use, not many are behaviors we would describe as intelligent. The "intelligent" behaviors of some insects and some other tool users quickly break down if the tool-using task is slightly altered. It will become clear that different perspectives on tool use are appropriate, from the view that tool use can arise from blind adherence to a genetically hard-coded set of instructions, or from learning and reasoning. This recognition, that successful tool users need have little or no understanding of their own tool using behaviors, leads us to recognize that simply successfully using tools cannot be a litmus test for an intelligent agent. Although a technical challenge, hard-coding a robotic agent to use tools could be done and the agent could succeed in the task without any understanding or intelligence concerning its own behavior.

Our interest, however, is not in building agents that can use tools, except to the extent that this gives insight into the nature of the activity of tool use. We find one useful and suggestive thread of research in animal tool use: some animals appear to learn to use tools by imitation. There are often multiple groups of chimps in the same forest. It can happen that most members of one group use a certain tool while not one single chimp in the other groups employ that tool. Several researchers studying such discrepancies go so far as to cite them as evidence that imitation is the basis for learning and transfer of high-level tool-using behaviors.

One reason this hypothesis is interesting is that imitation is one of the primary methods by which young human children learn. Human imitation has been extensively studied, as reflected in a large body of research in developmental psychology. This has sparked some interest among robotics researchers as well; we find a wide range of projects in which robots and other computational systems are designed with imitative capabilities in mind. Imitation in general has received considerable attention as a possible mechanism for robots to acquire new behaviors, as an alternative to programming or to other forms of machine learning (Schaal 1999).

The research described in this dissertation pulls together these two areas, imitation (in humans and in robots) and tool use (in animals and in robots), in an argument that they can be viewed as being essentially related. The empirical aspects of the argument are reflected in a simulation of two different agents, one representing a four-fingered robot hand, the other a wheeled robot vehicle. (Relying on a simulation rather than a physical robot limits the generalizations possible from this research, but we believe that the results are nevertheless compelling.) We use the environment and its agents to demonstrate the following:

- Given a minimal spatial ontology (largely agent-centered but in a global Cartesian coordinate system) and a relatively simple clustering algorithm, the agents are capable of learning predictive linear models of the relationships between their motor signals (e.g., *flex front finger* or *forward-rotate right wheel*) and effects in the environment, based on data observed during a babbling process.

- If the ontologies of the two agents are aligned (i.e., the *north* of one agent is tied to the *north* of the other, and all other mappings are similarly supplied) then each agent becomes capable of a simple form of imitation, even though the relationships to motor effects differ between the agents because of their different embodiments.
- One agent can imitate another to achieve goals it has not experienced itself. It is possible for one agent to apply a problem solving strategy that was learned by another agent. The critical capability here is that a predictive model is imposed on a situation that is not identical to that in which it was originally learned. By representing a solution in a shared ontology it may be possible to use the strategy in other contexts or with different agents.
- Following the same process, the hand agent can predict the effects of its use of a tool, using its ontology as a bridge. Some of the same strategies that the hand has learned to manipulate objects, may be applied to a tool held in the hand, if the tools motion has been properly grounded within the ontology.

The remainder of this dissertation is structured as follows. Chapter 2 will discuss tool use as a problem domain, presenting a definition of tool use and then exploring a wide range of animal behaviors that fit the definition. These many examples of tool use will be analyzed with some being identified as more intelligent than others and imitation will be proposed as central to most of the more intelligent instances of tool use. Chapter 3 will present the concept of ontology and describe how an ontology can be made meaningful to

an agent and how, once meaningful, ontologies may be employed to create representations of effective strategies. Chapter 4 will examine imitation from a human perspective with four progressive stages of human imitation being presented. Chapter 5 will describe implementing the first three stages of human imitation on a simulated robotic platform using a core ontology as the foundation. This robot imitation will eventually lead to examples of the robot using tools. Chapter 6 will conclude the dissertation with a summery of what has been accomplished, a discussion of what is significant about the results, and present directions for future work.

Chapter 2

Tool Use as a Task Domain

2.1 Definition of Tool Use

The most widely accepted definition of tool use comes from Beck (Beck 1980):

> Thus tool use is the external employment of an unattached environmental object to alter more efficiently the form, position or condition of another object, another organism, or the user itself when the user holds or carries the tool during or just prior to use and is responsible for the proper and effective orientation of the tool.

Beyond the question of whether or not a behavior is tool use, lies the question of whether or not an instance of tool use is indicative of intelligence. To better understand tool use

and to create a foundation for considering the intelligence inherent in instances of tool use, a survey of tool use throughout the natural world will be helpful.

2.2 Survey of Tool Use in Nature

Review of literature has shown that tool use is widespread within the animal kingdom. Perhaps the best know example is Jane Goodall's famous work with chimpanzees' use of reeds to remove termites (Goodall 1963), yet far from the only example of tool use, this is one of dozens. Beck (Beck 1980) in his book Animal Tool Behavior, along with his definition of tool use, provided a survey of animal tool use which is summarized and extended here.

2.2.1 Invertebrates

Ant Lion: Neuropteran flies undergo a larval stage that may last for as long as two years. During this larval stage, they are known as ant lions. The ant lion has been observed constructing a funnel shaped depression in sand. The ant lion then waits in the bottom of the depression to catch and consume prey that fall in. The ant lion has also been observed to throw sand at prey which escapes its initial grasp. It is believed this serves the purpose of causing the escaping prey to fall back into the depression.

Myrmicine Ant: These ants have been observed using small bits of leaf, wood, mud and sand to transport soft foods, such as jelly, honey, fruit pulp, and body fluids of prey. The food is adhered to the tool and then the tool is used to transport the food back to the

colony. The use of this type of tool allows the ants to carry ten times the amount of food they could carry in their crop alone.

Sphecine Wasp: Two Species of wasp, *Ammophila urnaria* and *Ammophila aberti*, have been observed using a small pebble as a tool (Brockmann 1985). A female wasp will lay her eggs within a hole. The top of this hole is then covered with small pebbles to help protect it. The pebbles are then pounded down to make the hole less conspicuous. The wasps have been observed to hold pebbles within their mandibles and then to use the pebble as a tool to hammer down dirt and other pebbles.

Marine Crab Melia Tessellata: This crab has been observed to remove small anemones from the substrate and then brandish them one in each cheliped (Duerden 1905). The stinging properties of the anemones allows them to be used as both weapons and tools for the collection of food.

2.2.2 Fish

Toxotid and Anbantid Fish: These South American fish have been observed in the wild and captivity projecting spouts of water into to the air to disable prey and make it fall into the water for capture.

2.2.3 Birds

Bower bird: In Australia and New Guinea, the males of this species will collect all manner of flotsam and assorted colorful debris to decorate their elaborate nests, called bowers. One remarkable instance of this elaborate decoration is some birds which have been observed to use a piece of bark as a paintbrush in nest decoration. They will wet the "brush" with saliva and charcoal and proceed to use it to paint the nest.

American Robin: This bird has been observed using a twig to rummage through leaf litter looking for ants (Potter 1970).

Black-breasted Buzzard: Circumstantial evidence has been provided that these birds will use stones to break open emu eggs for consumption.

Egyptian Vulture: Demonstrate similar use of rocks to break open ostrich eggs. Interestingly only certain populations of the vultures demonstrate this behavior. It was observed that in the same environment there would be groups of vultures where many demonstrated the behavior, at the same time other populations were observed where no members demonstrated the behavior (van Lawick-Goodall 1970).

Raven: These birds have been observed in the wild using rocks to defend their nests (Janes 1976). The birds would fly into the air and drop rocks onto a human who was

approaching their nest. One scientist reported being targeted by eight golf ball sized rocks in one encounter.

Crow: New Caledonian crows have been observed in the wild creating two different types of hooked tools which aid them in capturing prey. This tool manufacture by birds has been observed by Hunt to have a high degree of standardization, distinctly discrete tool types with a deliberate and distinct imposition of form, and the use of hooks (Hunt 1996).

Geospizine Finche: These finches from the Galapagos Islands have been observed using a small twig to aid in the collection of insect prey(Gifford 1919). The finches will obtain a small and narrow tool such as a twig or cactus spine. They then take this tool in their beaks and use it to probe for insects in narrow crevices. They have also been observed to use the tool to impale and remove insects from deep holes.

2.2.4 Nonprimate Mammals

Pocket gopher: The female of the species uses stones and hard chucks of food to dig in the soil during burrow excavation. The tool is held in the forelimbs and aids in the loosening of the soil.

California Ground squirrel: Will throw sand into the face of predators such as gopher snakes and rattlesnakes. The throwing is accomplished by a shoving motion of the forearms.

California Sea Otter: These otters have been seen in the wild breaking shells with rocks (Riedman et al. 1989). It is interesting to not that although Alaskan sea otters are genetically identical to California Sea Otters, they do not demonstrate this behavior.

Elephant: Elephants are known for their intelligence and have been observed using a variety of tools. Elephants have been observed using a stick to scratch themselves (Douglas-Hamilton, Douglas-Hamilton 1975). They have also been observed constructing a tool from leafy branches to shoo flies (Hart, Hart 2001).

Dolphin: Dolphins have been observed in the wild breaking sea sponges off the ocean floor and then using then to forage for food (Krutzen et al. 2005). This is not believed to instinctual behavior but behavior which is taught from mother to offspring as evidenced by mitochondrial DNA analyses.

2.2.5 Primates

Capuchin Monkey: Capuchin monkeys have been observed in the wild dropping sticks on intruders and using sticks as a weapon in aggressive encounters with conspecifics

(Chevalier-Skolnikoff 1990), using environmental objects to open oysters (Fernandes 1991), and using a club to attack a snake (Boinski 1988).

Gibbon: Gibbons have been observed dropping sticks on intruders (Tomasello 1997).

Chimpanzee: In many ways the chimp is the holy grail of animal tool use. It was Jane Goodall's observation of chimps using narrow stalks to extract termites that exploded the myth that only humans use tools (Goodall 1963). This is surprising because as early as the 1600s there are written accounts by Portuguese explorers of chimps using stones to crack open nuts(Inoue-Nakamura 1997). Regardless of when knowledge of chimpanzee tool use became mainstream, it is not surprising that they come the closest to human-level tool use. Studies have shown chimpanzee and human DNA are 95-98% the same (Britten 2002). The chimpanzee is capable of remarkable feats of tool use. In captivity chimps have been taught to use tools ranging from hammers to complicated machinery (Hayes, Hayes 1951, Savage-Rumbaugh, S McDonald, K Sevcik, R A Hopkins, W D Rubert,E. 1986). Chimpanzees have even been taught in captivity to manufacture stone tools by knapping (Toth 1993).

2.3 Analysis of Animal Tool Use

Some of the behaviors described above appear quite intelligent while others not.

In many of the examples, if the task is altered or an action executes improperly (an unfortunately common occurrence in robots) the solution breaks down quickly. These types of failures are often seen in some insect tool using behaviors such as the wasp. It uses a stone to tamp its burrow. It uses its vision and antennae to direct the application of the pebble onto the burrow, but it is following an instinctual sequence of instructions. The wasp lacks the comprehension that it using the tool towards the goal of protecting it's young. Insect behavior often appears intelligent while not bearing out that appearance on further examination. The lack of comprehension of a task being performed by insects is often demonstrated by the way the their ability to perform a task breaks down with slight intervention; they will demonstrate strange behaviors such as continuing to "mime" the tool using task even if the tool is removed. This "apparent stupidity" of insects (Lubbock 1889) results from following a rule set in the absence of higher level comprehension.

Most of the lower levels of tool use, such as that seen in insects, appear to be driven by instincts that are hard coded by evolution. The higher-level tool-using behaviors do not seem hard-coded. They appear to be learned through some mechanism. There is controversy as to the exact mechanism employed in learning behaviors. Some researchers have said it is largely trial and error learning that accounts for tool use while others have said that they are learned by imitation (Povinelli 2000).

There do appear to be some differences in populations of tool users that social learning must explain. Only certain populations of the Egyptian vultures use stone to open eggs. Genetically similar vulture populations in the same environment do not demonstrate the egg opening behavior. Some populations of sea otters use stones to open shells while others do not. It is the chimpanzee where these discrepancies are most surprising. In the same forest there will be groups of chimps that all demonstrate a particular tool using behavior occupying the same niche groups where not a single individual demonstrates the behavior.

Boesch and Tomasello (1998) posited that there must be some primary social learning occurring if:

(1) Two groups of the same species differ in a behavior (with a countable number in each group conforming).

(2) There are no obvious differences in the environments of the two groups, making an explanation in terms of individual learning (environmental shaping) unlikely.

(3) There are no genetic differences between individuals that acquire the behavior and those that do not.

	Chimpanzee Population					
Pattern	Bossou	Taï	Kibale	Assirik	Gombe	Mahale
Foraging behavior with tool use						
Ant dip	+	+	−	+	+	−
Honey dip	−	+	−	−	+	−
Bee probe	−	+	−	−	−	−
Leaf sponge	+	+	+	−	+	−
Marrow pick	(−)	+	−	(−)	−	−
Nut crack	+	+	(−)	(−)	(−)	(−)
Termite dig	−	−	−	(−)	−	−
Pestle pound	+	−	−	−	−	−
Hook stick	+	−	−	−	−	−
Gum gouge	+	(−)	−	−	−	−
Termite fish	−	(−)	−	+	+	+
Algae fish	+	(−)	−	−	−	−

Figure 2.1 Tool using behavior observed clustering in populations (from (Boesch, Tomasello 1998))

As the figure above shows, there are in fact strong differences in tool using behaviors seen in populations of chimpanzees that meet all of Boesch and Tomasello's criteria. It seems as though there must be some learning mechanism to account for these stark divisions between genetically similar chimps in similar environments. While the exact mechanism of the learning remains controversial, many have proposed that imitation is a likely candidate to explain how genetically similar populations in similar environments would exhibit different tool using behaviors (Whiten 2000).

One of the main goals of this research is to demonstrate that if agents are given imitative mechanisms, it is possible for tool using behaviors to result from those mechanisms.

Most acts of tool use have a link to imitation. The imitation may be hard coded and passed though evolution. While not imitation in a traditional sense of the word, it is not too far a stretch to recognize the methods by which evolution repeats successful strategies as a form of imitation. In other contexts, this imitative nature is far more obvious; we will designate two forms of imitation, internal and external. In external imitation, an agent may observe another agent successfully employing a tool and imitate the behavior.

Internal imitation may be as simple as self imitation. An agent may discover a successful tool using strategy and attempt to imitate it by repeating the successful behavior. An agent may try to improve performance of a task with a tool by imitating a task it normally performs unaided but adding a tool to the act. Although it is an acknowledged stretch, it could even be said that some novel tool using behaviors involve imitating a potential solution that has been generated internally through reason. Even if a solution is discovered through trial and error, if the agent lacks an ability to self imitate, the solution must be rediscovered again and again.

There are many mechanisms by which imitation may occur. Imitation (especially human) will be discussed at great length in later sections. In theory, the simplest form of imitation involves recording every action that is done and repeating it. In practice, remembering every action perfectly quick becomes impossible, even for simple tasks. Often it is not every detail of an action which is remembered, rather it is an abstraction of the action. An abstraction of an action does not record every detail, but rather the overall

gist of an approach, with enough detail so that it may be repeated. Creating these abstractions has a long history in philosophy and is currently the topic of extensive research in computer science. The discipline has come to be known as ontology and will be examined in detail in the next chapter.

Chapter 3

An Ontological Basis for Action

3.1 The Role of Ontology

The word ontology derives from the Greek root *onto*, which means "being." Ontology is the study of being, an examination of the nature of existence. Ontology has a long and rich history in philosophy. The Greek philosophers deeply examined the nature of existence. This examination of existence is most notably attributed to Aristotle who in his *Metaphysics* and *Categories* attempted to identify an underlying foundation of categories of being that would explain reality. It was felt that choosing the correct categories would "carve life at the joints" and give a deep insight and understanding of existence.

This concept of carving life at the joints seems intuitively correct, and anyone who has carved a chicken can understand the metaphor. If one takes a cleaver and haphazardly divides a chicken there will be a disorderly pile of meat and bone fragments. If however, a knife is applied to the joints there will be a far more orderly pile with legs and thighs easily identifiable. While this metaphor is straightforward when thinking of the legs and thighs of a chicken, it is problematic when other parts are considered. The back and neck are often left whole in carving but they are filled with joints, and even though cutting the leg and thigh at their actual joints has an appeal, it would also be natural to take the legs and thighs and divide them into piles of meat, skin, fat, and bone.

The endeavor of dividing the chicken is one of categorization, classifying the substances that form the chicken and identifying natural and functional sets. Even in this straightforward example, the results are not straightforward; different divisions will more naturally fit reasoning for different tasks involving the chicken. This problem and others like it have come to be known as lying within the discipline of Knowledge Representation. Skin, fat, bone, thigh, leg etc. are all words that represent certain aspects of a chicken.

Why is ontology relevant here? Knowledge representations allow us to reason and plan. They play several roles in our interaction with the environment surrounding us. In an important paper that tried to unify a wide variety of work in the field, Davis, Shrobe, and Szolovits identified five distinct roles that a knowledge representation may play (Davis,

Shrobe & Szolovits 1993). First, and most fundamentally, a knowledge representation is a surrogate for the thing itself. Once we have representations of reality, we can use them to determine consequences in the world through thought rather than action. Second, a knowledge representation is a commitment to how we will define and reason about reality. One of the most important aspects of this work will be understanding the ways in which how an agent represents reality influences its thinking regarding it. Third, a knowledge representation is a fragmentary theory of intelligent reasoning because it chooses how and when to draw inferences from partial data and to what extent those inferences are included in the knowledge set. Fourth, a knowledge representation is a medium for pragmatically efficient computation. By discretizing data and recommending some inferences as more important than others, it is possible to make many problems far more tractable. Fifth, a knowledge representation is a medium for expression. Representations of things allow reasoning, but they may also allow the sharing of that reasoning.

These roles are all important. They inform understanding of the manner in which representations allow reasoning and communicating about reality. It is fundamental that the representations an agent chooses inform what reality is to that agent. As old philosophical arguments of Descartes and Berkley and more modern brain-in-a-vat thought experiments suggest, it is not reality that is experienced, rather it is the representations chosen to employ in classifying reality that are experienced. Recognition of this supports the importance of ontology; the manner chosen to represent and reason

about reality biases the perception of it. It is not ultimate reality that is experienced, rather it is a view of reality through the lens of one particular knowledge representation. As Davis et al. describe it:

> If, as we argue, all representations are imperfect approximations to reality, each approximation attending to some things and ignoring others, then in selecting any representation, we are in the very same act unavoidably making a set of decisions about how and what to see in the world. That is, selecting a representation means making a set of ontological commitments. The commitments are, in effect, a strong pair of glasses that determine what we can see, bringing some part of the world into sharp focus at the expense of blurring other parts. These commitments and their focusing-blurring effect are not an incidental side effect of a representation choice; they are the essence (Davis, Shrobe & Szolovits 1993).

All this becomes relevant when we consider what a computational agent should represent if it is to be capable of imitation, and then, as we will demonstrate, of tool use. Essentially, the agents we build will be able to rely on a representation of objects and spatial relationships between them. Previous work has demonstrated that these spatial relations can be recognized with neural networks or other machine learning mechanisms (Schwering 2007, Regier 1996). To explain why these are important decisions, we will need to give a more detailed account of ontology research in computer science and

robotics, but first we will look at evidence from neuroscience that seems to indicate that there are mechanisms for making these kinds of abstractions inherent in the brain.

3.2 Evidence for Ontology in Neuroscience

Some of the most interesting evidence that is seen supporting the role of ontology in actions comes from recent research in neuroscience. Examination or the F5 area in the brain which occupies the rostral most part of the ventral premotor cortex showed that it was tightly linked to activities involving the hand (Matelli 1985). It came to be seen that the neurons in this area of the brain coded movement in very interesting and abstract ways. The area was only found to be active if grasping or holding was occurring and the activity was directed towards some goal. Simple hand acts not involving the goal of grasping could leave the area inactive(Gentilucci 1988, Hepp-Reymond 1994, Kurata 1986, Rizzolatti 1988, Rizzolatti 1981). This was not just a simple neuronal firing that indicated hand movement, it was an area of the brain that indicated that a goal of grasping or releasing grasp was present. These neurons were found to fire not just when monkeys would grasp objects, but when they would observe others grasping objects.

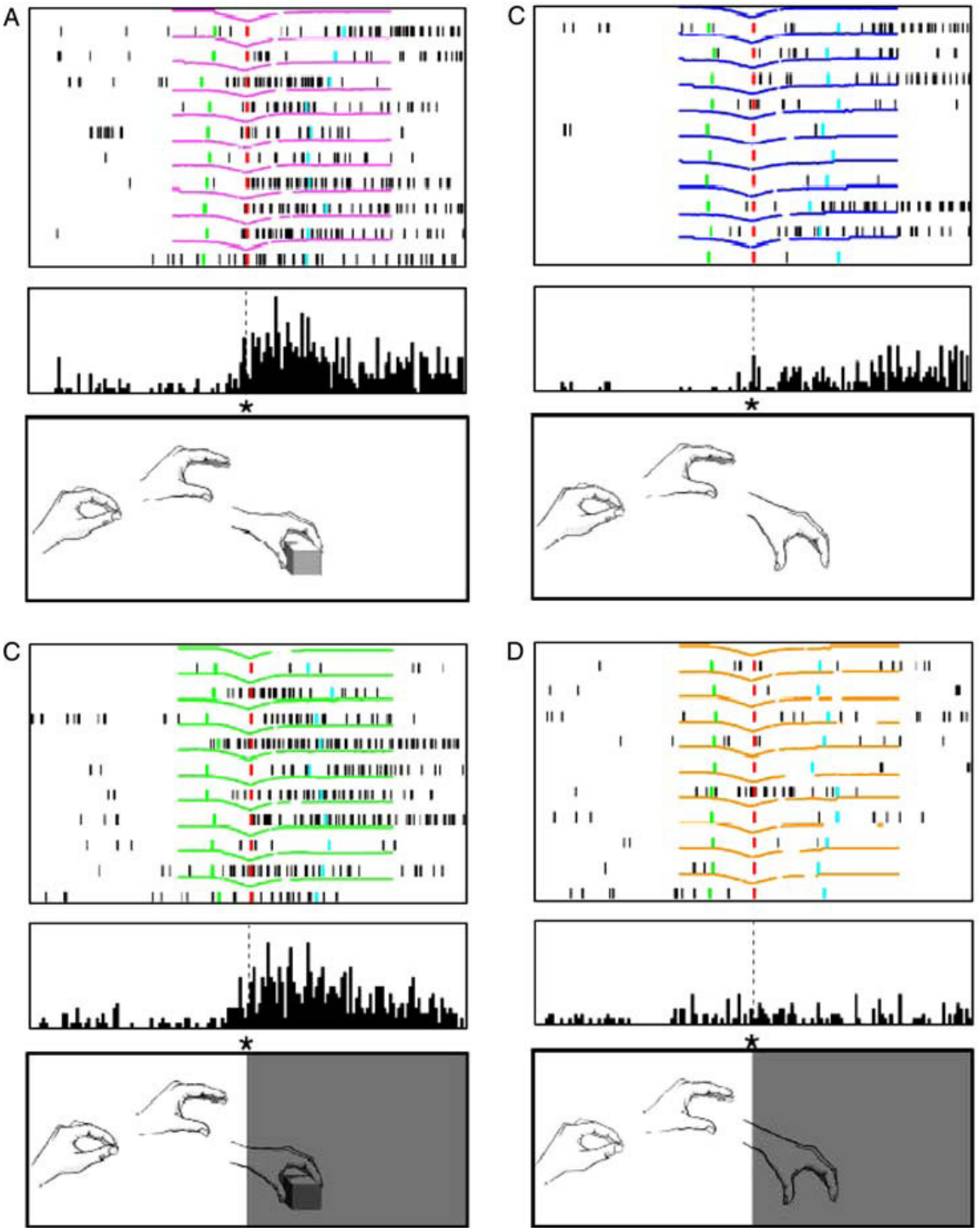


Figure 3.1 Examples of grasp task and F5 neuron firings in monkeys (image from (Metzinger, Gallese 2003))

In the figure above from Metzinger, it is possible to see the neuronal firings in a monkey observing situation **A** where the hand is observed grasping the block. When the hand is observed performing an identical action in **B**, but there is no block present, there is no firing. It is interesting to note that this relation holds, even if the view of the block is occluded. If the monkey is shown the presence or absence of the block and then a screen is placed in form of the block (represented in the image by the grey area). The same difference in neuron firing is still observed (see **C**&**D** above). It is interesting to note that these two scenarios appear identical to the monkey, but the neuron firing is different based on its knowledge of whether or not the block is present.

Further examination indicated that it was the goal of grasping being represented and not just hand movements when it was seen that different acts of grasping would excite this area regardless of the effecter employed. It was demonstrated that grasping done by not just the hand but even the mouth would excite the area (Rizzolatti 1988, Rizzolatti, Fogassi & Gallese 2000). Not just the act of grasping will excite this area, the presence of graspable objects will also do so (Jeannerod 1995, Murata 1997, Rizzolatti 1988, 2000)

The linkage of the area to certain goal based activities became clear and the neurons became known as mirror neurons. The F5 area was found to contain other neuronal areas which would fire when not just grasping occurred, but also sticking and manipulating

would occur. It was the goals of the activities that showed high and consistent correlations with neuron firings. The same neurons would associate with goals of an activity regardless of whether the hand or mouth was used to accomplish the activity. It has even been shown that the same goal neurons will fire when a tool, in the form of a stick, is used to achieve the same goal (Ferrari 2005). There appears to exist not simply neuronal control of hand movement but direct evidence for a neuronal basis for a type of goal-directed behavior model (Gallese 2003). There appears to be an innate ability for abstraction in the brain that is similar to the abstraction inherent in placing any kind of knowledge in a particular framework.

3.3 Ontology in Computer Science

Although the study of ontologies has hundreds of years of history in Philosophy, some of the most interesting work in applied ontology has come from within the field of Computer Science quite recently. Although the term ontology as it is used in Computer Science has its roots deeply in the Philosophical background, the term has evolved to take on its own separate meaning that relates to the traditional concept while being its own new and independent entity.

The first stirrings of a new Computer Science based concept of ontology were seen in the early 1980's. In these interesting instances, ontologies started to emerge as formalisms to

be employed rather than pursued as vehicles for insights into the underlying nature of being. The first example of this was in 1980 by John McCarthy in his work on Non-Monotonic Reasoning (McCarthy 1980). McCarthy described an ontology as a list of "the things that exist", this was not a complete departure from this philosophical background, but it was a shift in focus. The goal of the list of things that exist was not meant to be a description of the fundamental nature of being, but rather it was a medium, upon which logic could occur. Hayes, in 1985 wrote a book chapter entitled *Naïve Physics I: Ontology for Liquids,* where ontology was again used slightly differently than had previously been seen. In this context, ontology was again a formalism to be employed. Its members served as a representation for various states and aspects of water. These representations were then utilized to perform calculations in Predicate Calculus. John Sowa also used ontology differently in the mid 80's. Ontologies were again formalisms of underlying reality, which Sowa could use in the creation of conceptual graphs, a derivation of semantic networks (Sowa 1984).

As these examples show, there was extensive employment of ontology as a concept by AI researchers, but it was in 1986 that James Alexander and his associates at Tektronixs Laboratories clearly stated the need for knowledge engineering as a sub-discipline of Computer Science with ontology as its medium. If computation is to occur in a given domain, there must be a formal specification of that domain. Ontology in this sense is not unrelated to its philosophical background, but rather it is an extension of it.

Ontologies have emerged as ways of formalizing a particular view of some domain. The most referenced was of stating this comes from Gruber:

> An *ontology* is an explicit specification of a conceptualization. (Gruber 1993).

Ontologies serve as the foundation upon which common understandings of domains can be achieved. An ontology is an common vocabulary which is specifically designed to describe a certain domain.(Cardoso 2007, Guizzardi 2005, Sharman, Kishore & Ramesh 2007) Ontologies can provide models for the semantics that are inherent in underlying reality. They can also serve as the basis for storing informational descriptions of spacio-temporal reality (Havas, Larue & Camus 2008).

3.4 Ontology in Robotics

Bateman has described previous use of formal ontology in Cognitive Robotics as limited, but argues that a closer interaction is appropriate (Bateman, Farrar 2005). He describes three roles ontology has to play in cognitive robotics.

1) Ontology can enhance the design of robot knowledge representations by clarifying the relations among various representational levels and providing a semantically coherent account of entities used in symbolic reasoning.

2) Ontologies can be used to develop and constrain a shareable conceptualization of the environment.

3) Ontologies can provide a framework to help deal with incomplete sensor information.

Research has used ontologies as the foundation for modeling capacities, objects, verbs, colors, state, people, and space.(Havas, Larue & Camus 2008) There has been much work of modeling spatial relations with Kuipers Spatial Semantic Hierarchy (SSH) (Kuipers 2000). Ontologies have been used to create an "agent belief store" with ontology axioms acting as schema for dynamic beliefs (Clark 2007). Ontologies have served as the foundation to guide coalition formation and cooperation (Fanelli et al. 2006). Ontologies have been proposed as an appropriate foundation for guiding image recognition and object categorization (Eric Maillot, Thonnat 2008).

3.5 Grounding an Ontology

In some domains, agents can be pre-programmed with an ontology that reflects all the relevant information in their environment. For our purposes, as will become clear, it is advantageous for an agent to be able to "ground" its representation in its environment.

Dretske (Dretske 1988) examines the concept of grounding a belief with the following example. Consider sitting in your home and becoming thirsty. Eventually your thirst builds to the point where we say it causes you to get up and go to the refrigerator to get a beer. If you were asked, "Why did you go get the beer?", you would answer, "Because I

was thirsty." Dretske diagrams this relationship in the following manner:

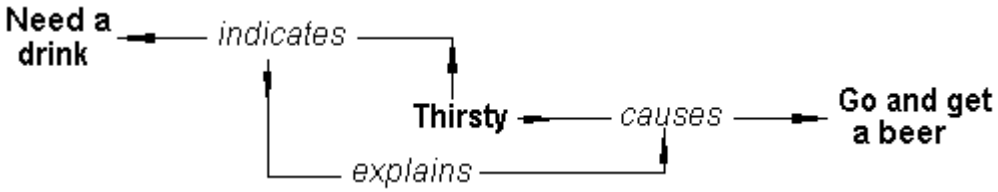


Figure 3.2 Explanation for getting a drink

I am thirsty. This indicates that I need a drink and causes me to go and get a beer. The fact that thirst indicates a need for a drink explains why it causes me to go and get a beer. This diagram clearly shows an example of a belief with semantic meaning. “Being thirsty”, has the meaning behind it of needing a drink. For contrast, consider a robot that runs on a battery. This battery eventually runs out of power. If the robot does not save his data before his battery runs out he will lose all of it. Imagine that the robot has a low battery alarm, and let us again construct a diagram of the same form as Figure 3.2:

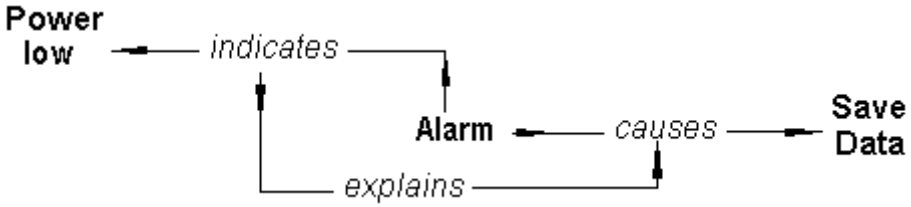


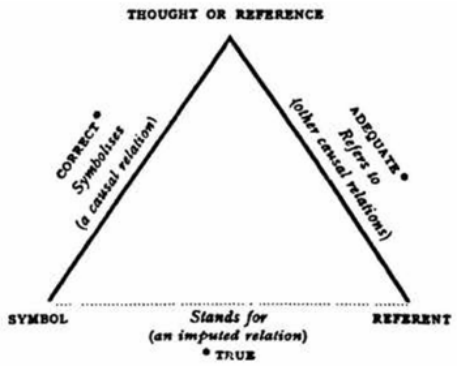
Figure 3.3 Explanation for power alarm

Humans have personal experience of the semantic value of thirst. Does this low battery alarm also have semantic value? The alarm is not merely a computer state. It has

semantic grounding in the fact that it represents the charge status of the battery. Some people might say that this alarm is just too simple a mechanism for us to describe this robot as thinking. It would be like saying that a thermostat thinks and this is certainly a valid criticism.

Learning to recognize a linkage between a idea such as "I'm almost out of power" and a sensor value such as low voltage in a battery is a problem that has received a great deal of attention from many disciplines of computer science. While it is possible to engineer such connections into a robot, making them meaningful is challenging. This problem has come to be known as the symbol grounding problem (Harnad 1990). How is it possible for an agent (artificial or otherwise) to link a symbol such as a word to an underlying meaning?

Vogt has proposed a method he calls "anchoring semiotic symbols" as a way to get around the symbol grounding problem (Vogt 2003). This method is based on a model called the "Triangle of Reference" first proposed by Ogden in *The Meaning of Meaning* (Ogden et al. 1927):



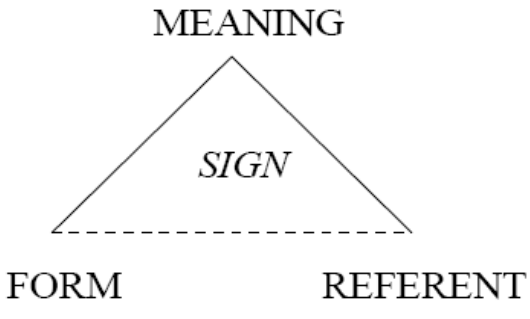


Figure 3.4 The Triangle of Meaning and Semiotic Triangle, taken from Ogden and Vogt respectively

Vogt refers to the image above as the Semiotic Triangle. The terms he uses are adopted from Pierce's definition of semiotic symbols (Peirce 1931):

Form. A form (or word) is the shape of the sign, which is not necessarily material.

Meaning. The meaning is the sense that is made of the sign.

Referent. A referent is the object that stands for the sign, which may include abstractions, actions or other signs.

The semiotic triangle is proposed as an attempt to avoid the symbol grounding problem which is still considered one of the most difficult to solve in AI and robots. Consider a simple task of recognizing a coffee cup, a task that humans perform so facilely that it is completely effortless. For an automated system, however, recognition is less straightforward: differences in perspective, orientation, lighting, and so forth may be difficult to generalize over; variations in the coffee cup itself (e.g., a cup with an oddly shaped handle or an odd pattern) may result in difficulties. While something so apparently effortless seems like it should be straightforward on a robotic platform, this ostensible ease disappears quite quickly when attempting to perform this act of simple recognition with a robot.

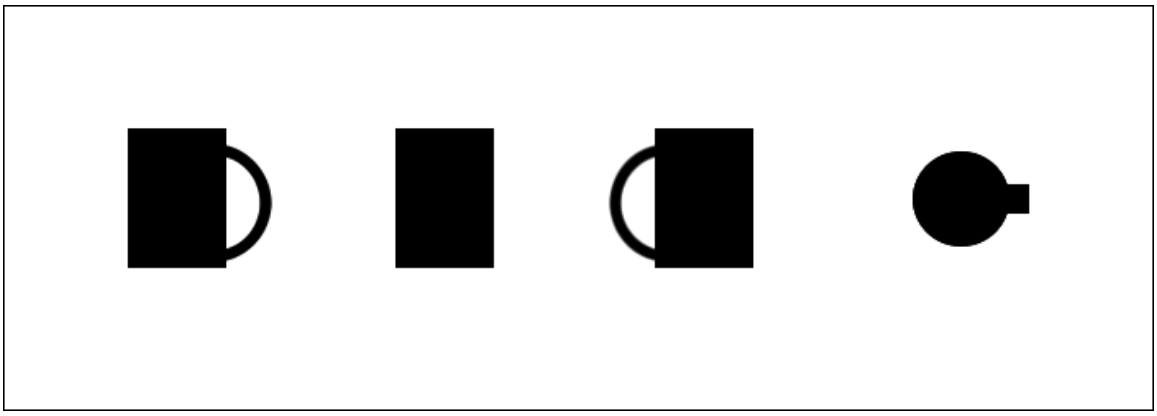


Figure 3.5 Outline of a Mug From Different Angles

Vogt's semiotic anchoring attempts to move beyond all these problems. Vogt proposes that the problems arise because the focus is on the wrong part of the problem. It is inappropriate to classify the problem of recognizing a cup as creating algorithms where every conceivable projection of a cup can be properly classified. What needs to occur is the focus needs to shift. The observer needs to recognize the objective reality that there is an actual cup sitting there and use the real object as the anchor for the concept.

Using a real object in front of oneself as an anchor for thoughts about it seems like so simple an idea that it is meaningless; it is fact quite powerful and lies at the heart of ideas such as deictic pointers, and reflexive robotics that have (in implementation) been some of the most successful strategies in effective real world control. To better understand the motivations behind this shift, again consider the two methods for solving addition problems. The brute force attempt to recognize the cup is similar to the addition machine with a lookup table. If the table contains the correct entry, it is a quick path to a solution

to the problem. What the example demonstrates is that this type of approach is doomed to fail for certain tasks. If creating a lookup table for something as simple as a coffee cup fails as the size of the table grows out of control, it is clear that we will need an alternative approach.

The approach that has often been proposed involves the use of concepts. Concepts represent the semantics of a situation. A concept is a mental category that may be employed to classify objects, events, or ideas. It is very important to note that a commitment to a particular set of concepts to represent a domain is precisely what is meant by commitment to a particular ontology as was discussed in a previous section. An ontology is, among other things, an agreement to employ a certain set of concepts when representing a particular domain.

The learning of concepts involves linking low-level perceptual information to higher level representations. Cognitive psychologist Jerome Bruner is credited with first defining the problem of concept attainment and defines it as "the search for and listing of attributes that can be used to distinguish exemplars from non exemplars of various categories."(Bruner 1956)

In a more recent paper, Lawrence Barsalou noted that agreeing on a shared definition of exactly what a concept is remains contentious, but said:

> Following psychological theories, we assume that a concept, roughly speaking, is knowledge about a particular category (e.g. birds, eating, happiness). Thus

> knowledge about birds represents the bodies, behaviors and origins of the respective entities. Knowledge plays a central role throughout the spectrum of cognitive activities. In on-line processing of the environment, knowledge guides perception, categorization and inference. In off-line processing of non-present situations, knowledge reconstructs memories, underlies the meanings of linguistic expressions, and provides the representations manipulated in thought (Barsalou et al. 2003).

For our particular application, the problem of concept learning is fortunately quite constrained by our ontology. Initially, we will only be concerned with the foundational concepts contained without our object ontology and spatial relations ontology. Once we have learned these concepts, they will serve as a basis for the creation and extension of an action ontology as is discussed in the following section.

3.6 Ontology of a Robotic Tool User

With a better understanding of ontology in hand, it is appropriate to again consider the research goal of creating a robotic tool user. Towards that goal, we propose a triumvirate of ontologies to describe the world: an Ontology of Objects, an Ontology of Spatial Relations, and an Ontology of Action.

The first ontology is straightforward; an ontology of objects is an ontology of the objects which make up the world. This is a very simple ontology that will just be a set of object types that can exist within our robot environment. The ontology of objects would include items such as basic shapes and types of objects. This enables basic identification of items in the environment.

Once objects have been recognized it will be necessary to represent how they relate to each other. This is done with an ontology of spatial relations. Previous work has demonstrated that spatial relations can be recognized with neural networks or other machine learning (Schwering 2007, Regier 1996). In his book *Human Semantic Perception*, Regier identified seven important spatial terms that he was able to learn to recognize with a back propagating neural network. They were above, below, to the left of, to the right of, inside, outside, on, and off.

This research will employ a simple core directional ontology that is based on a Cartesian model and consists of north, south, east, west, up and down. These same directions allow for relative spatial terms to be expressed. For example, given two objects it is possible to describe on object as being either: south of another object, north of another object, or centered north/south with another object. These core directional and spatial ontologies will be presented in much greater detail when the simulation environment is described in chapter 5.

With a directional and spatial ontology to provide descriptions of our environment, it is possible to learn how an agent may manipulate that environment. This involves learning initial environment description / agent action pairs that consistently lead to certain resulting environment descriptions. Recent work by Modayil and Kuipers (Modayil, Kuipers 2007) has suggested a formalization for representing actions:

α is an action $\alpha \in Act$

$\alpha = \langle D, C, H \rangle$ Where D is the change in an object property resulting from an action and C is the context where the action occurred and H is the control signal

Given that actions are occurring in a closed world, it is possible to safely assume that any change is the result of an agent acting. The goal with this type of approach is to identify contexts and control signals that **consistently** result in certain resulting states. The resulting tuples will be the basis of an action ontology that will inform an agent how it may manipulate its world.

This type of approach to representing action was chosen in large part due to overlap with some properties seen in mirror neurons. The resulting change in an objects properties, is an inherent part of this type of representation. The importance of the results of an action in its representation fits well with theories by Gallese (2003) that point to mirror neurons as the basis for a type of goal-directed behavior model. These ontologies can be difficult

to comprehend in the abstract. In chapter 5 they will be presented again in much greater detail in the context of the environments they are being used to represent.

Chapter 4

Human and Robot Imitation

With the goal of this research being insight into tool use, the motivation for our approach lies as much in understanding human and animal cognition as it does computer science. This type of approach not only allows us to base our work on the best examples of tool users available; it will provide a path to insight and deeper understanding regarding tool using behaviors.

4.1 Imitation

As was shown empirically in the work with neuron activation in the previous section, there appears to be a strong link to what occurs when a brain observes an activity and when it carries the activity out itself. A link between observation and replication of an action seems to provide a natural pathway for imitation. Noting this linkage, led to an

examination of previous work that had been done regarding human imitation. In the late 1970's a revolutionary Science paper by Meltzoff and Moore found that newborn infants as young as two to three weeks could imitate facial gestures (Meltzoff, Moore 1977). Previously it had been thought that this capability was a landmark development and did not emerge until far later, at 8 to 12 months of age (Piaget 1962).

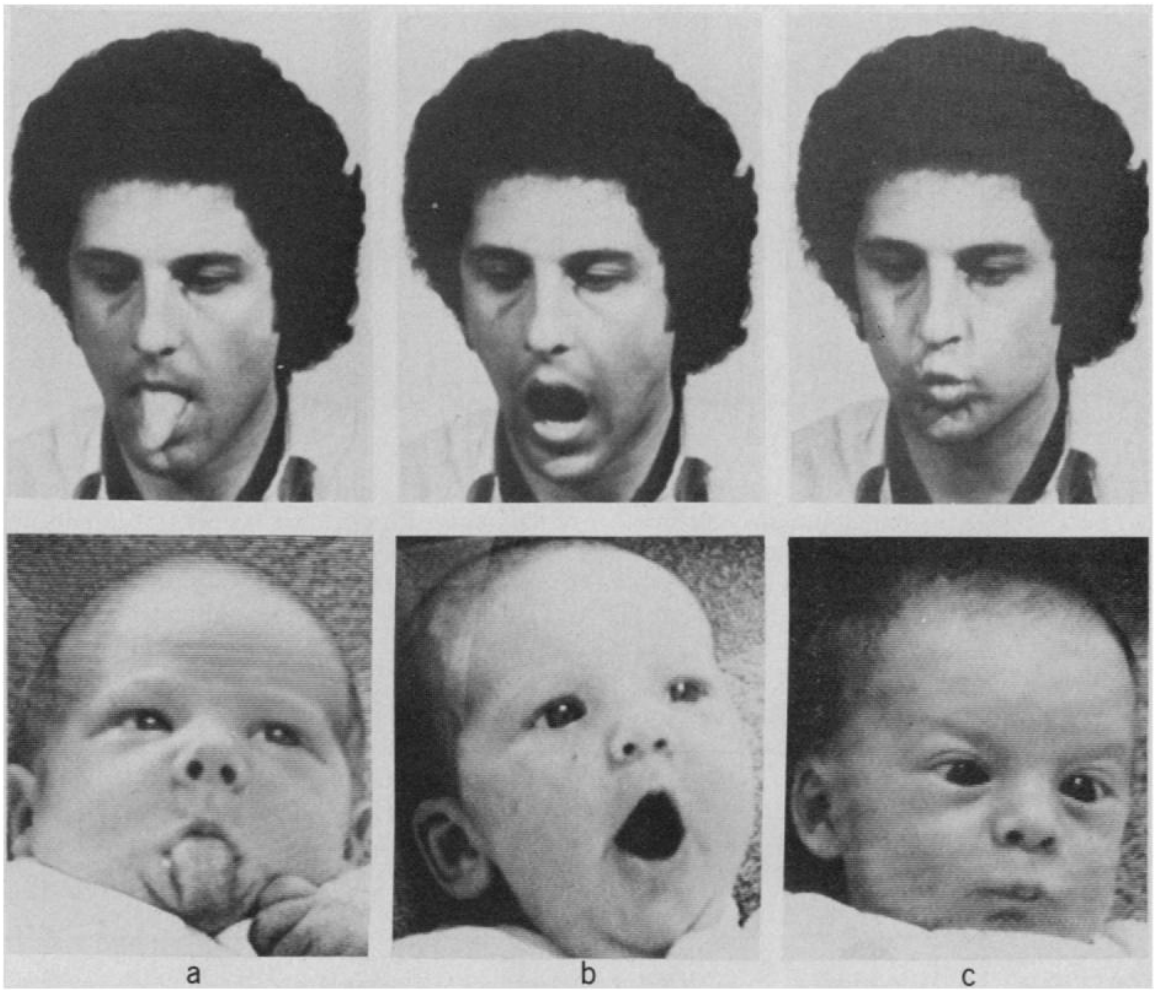


Figure 4.1 Meltzoff's video captures of newborn imitation from(Meltzoff, Moore 1977)

In subsequent years, Meltzoff continued to refine his theories of imitation, eventually settling on a four stage progression of imitative abilities (Meltzoff 1997, Meltzoff, Prinz 2002):

1) **Body Babbling**. This is the stage where an organism will learn how various motor signals influence various elementary body configurations. In this stage, random trial and error learning will link motor signals to resultant end states. This stage has been suggested by Calderon as where the construction of a body schema occurs (Acosta-Calderon, Hu 2005)and Stoytchev has used babbling behaviors to actually learn an extensible body schema for a robotic arm (Stoytchev 2005, Stoytchev 2003).

2) Imitation of Body Movements. In this stage, an organism will imitate the body movements of another organism. Meltzoff has shown extremely young infants imitating facial behaviors. The mechanisms by which they do so are not completely understood. At this most basic level, imitation cannot avoid what is called the correspondence problem which involves tracking points of interest in three dimensional space. Some have theorized that the difficulty of solving this problem might suggest that there is some direct pathway linking perception and action in this early imitative behavior. This is what has been suggested by many of the proponents of reactive robotics, but examination of experimental results of Meltzoff's earlier work, by Rao and Meltzoff (2004) was found to suggest that the observation-execution pathway in humans "**is mediated by a**

representational structure that allows infants to defer imitation and to correct their responses without any feedback from the experimenter."

Two experiments particularly supported some intermediary representation. First was the presence of delayed imitation. If a pacifier was used to restrict mouth movement, imitation would still occur after a 2.5 minute delay if the infant viewed a passive expression on the face that had made an expression before the delay (Meltzoff, Moore 1977). This delay was successfully extended to as long as 24 hours in 6-week-old infants (Meltzoff, Moore 2002) .

In other experiments agents were also found to correct imitative responses over time and converge on more accurate matches of behavior without any feedback from the experimenter (Meltzoff 1997, Meltzoff, Moore 2002).

The presence of an apparent mediating structure in human imitation strongly influenced the direction of the research into robot imitation described later in this dissertation. Recognizing the apparent presence of a mediating structure in human imitation was instrumental in the later decision to base attempts to model imitation in robots within the mediating structure of an ontology.

3) Imitation of Actions on Objects. Infants as young as fourteen months demonstrate abilities to imitate actions on objects. By observing the interactions of others with

objects, they are able to imitate that interaction. It is this form of imitation that allows the transfer of knowledge such as tool use from one individual to another.

4) Imitation Based on Inferring Intentions of Actions. This is the most sophisticated form of imitative learning. At this stage it is not the action itself being imitated but the inferred goal of the observed action. Meltzoff has also explored this type of imitation experimentally. Eighteen month old toddlers have demonstrated an ability to infer goals. After observing several unsuccessful attempts at an action the toddlers would recognize and imitate the apparent goal, not the failed attempt (Meltzoff 1995).

What is interesting to note about this four stage progression is that although it was originally developed to explain human imitation, it provides a framework to describe imitative behaviors in other organisms, and has even been used to suggest a pathway for teaching robots imitation (Rao, Shon & Meltzoff 2004). The stages have successfully served as the basis for learning robot imitation for other researchers (Acosta-Calderon, Hu 2005).

Chapter 5

Achieving Tool Using Behaviors through Action Ontology Based Imitation

In chapter two, imitation was identified as an interesting component of higher level tool using behaviors. In chapter three ontology was identified as a medium for knowledge representation and transfer. Human imitation was explored in chapter four with four individual stages identified. These disparate yet connected themes lead to both a fundamental question and a path to answering that question.

The fundamental question is whether or not imitative mechanisms are responsible for the transfer of higher level tool using behaviors. Ultimately this research cannot

conclusively answer this question, but it will offer significant insight into the discussion by demonstrating that in artificial agents, with a self learned imitative mechanism, it is possible for tool using behaviors to arise.

5.1 Simulation Environment

Due to the much lower overhead and ability to rapidly change parameters, this research was conducted in a simulation environment. The breve simulation environment was used to create simulations which would allow the exploration of basic robotic tool using behaviors. The breve environment is a small world simulation that realistically models physics with the Open Dynamics Engine (ODE). ODE models rigid body dynamics and collisions and is freely available under both the Berkeley Software Distribution (BSD) and GNU Lesser General Public License (LGPL). In addition to the physics modeling, there is an OpenGL display engine that creates a fully interactive view of all activity occurring within the simulation.

A basic breve simulation environment was created where it was possible to explore a variety of agents and scenarios. Eventually a hand based robot and a small wheeled vehicle where fully modeled within the simulation environment.

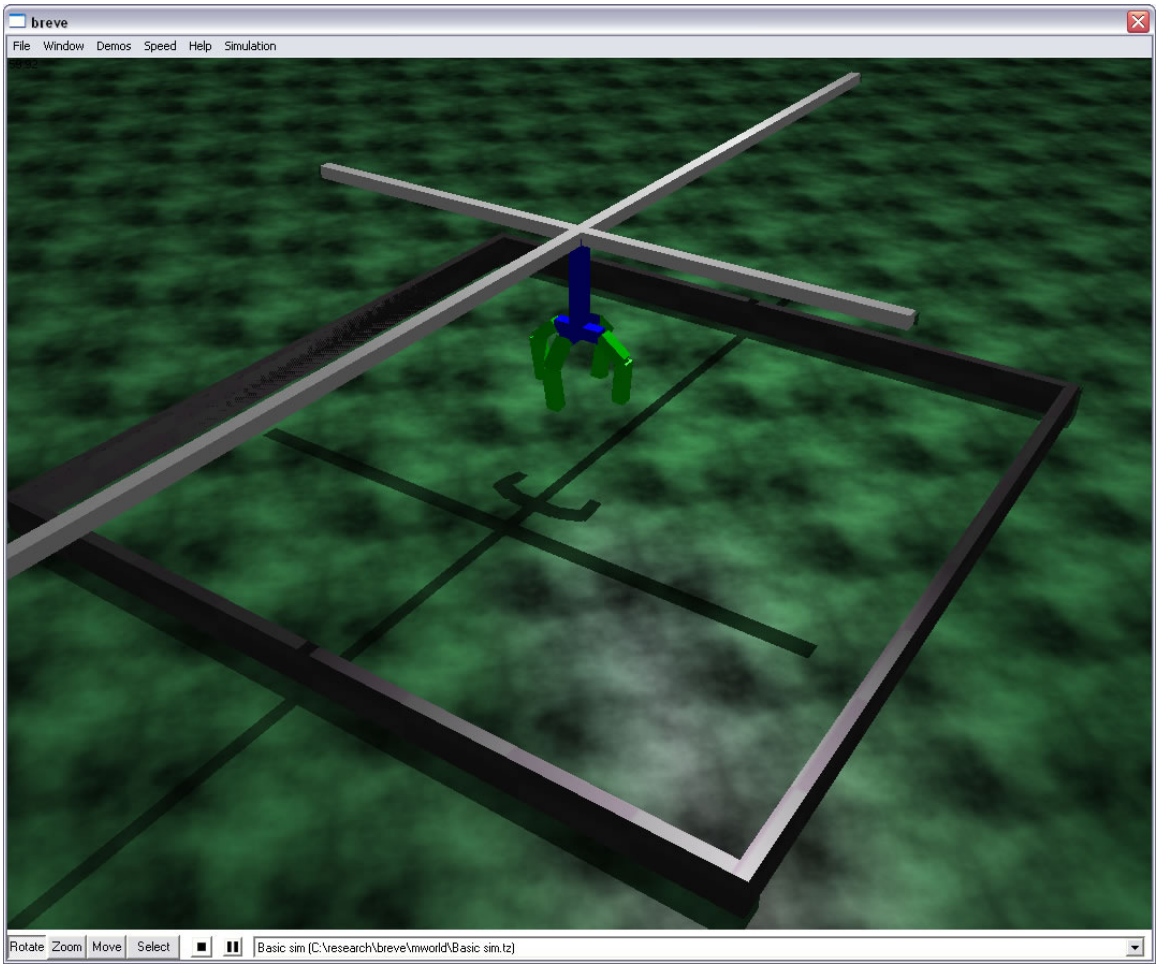


Figure 5.1 Basic simulation with robotic hand

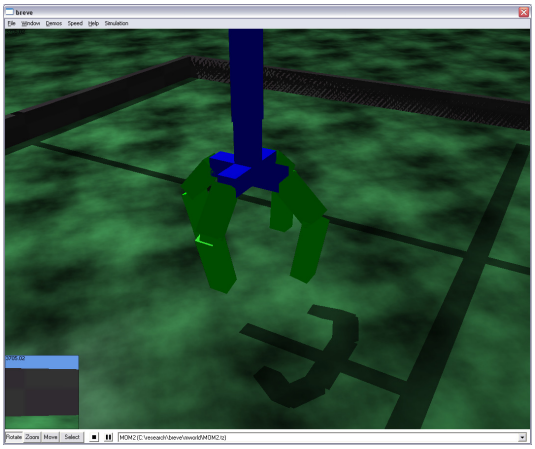


Figure 5.2 The hand

The first agent explored was a robotic hand, which was completely created within the breve simulation environment. The hand was coded within the environment using only simple rigid bodies and simple actuators with motion in two opposing directions. In the

environment, the hand is hung from two rails which act as linear actuators with one rail enabling forward and backward motion, while the other provides side to side motion. The hand itself is placed on a third linear actuator which allows it to be raised and lowered. The up and down arm actuator is connected to the palm via a wrist joint. Each of the four fingers is connected to the palm with a rotational joint which allows for each finger to rotate inwards towards the center of the hand, closing the hand in a grasp, or outwards to a limit, opening the hand. It is possible for all of these motor signals to be performed independently or simultaneously in any combination.

The environment for the simulation is a basic playpen. The simulation has a floor and the playpen area is enclosed on all four sides by barriers. It is possible to describe this environment with a basic core ontology. The direction forward from the initial camera view is North and the direction backwards is South to the right is East and to the left is West. Up is described by a line orthogonal to floor in the direction of the hand and down is described by a similar orthogonal line in the opposite direction. So to summarize briefly, we have created a simple spatial ontology consisting of North, South, East, West, Up, Down which is capable of accurately describing basic directionality in the simulation environment. In later simulations this directional sense was extended to allow relative directional classifications. i.e. describing an object as North of another object.

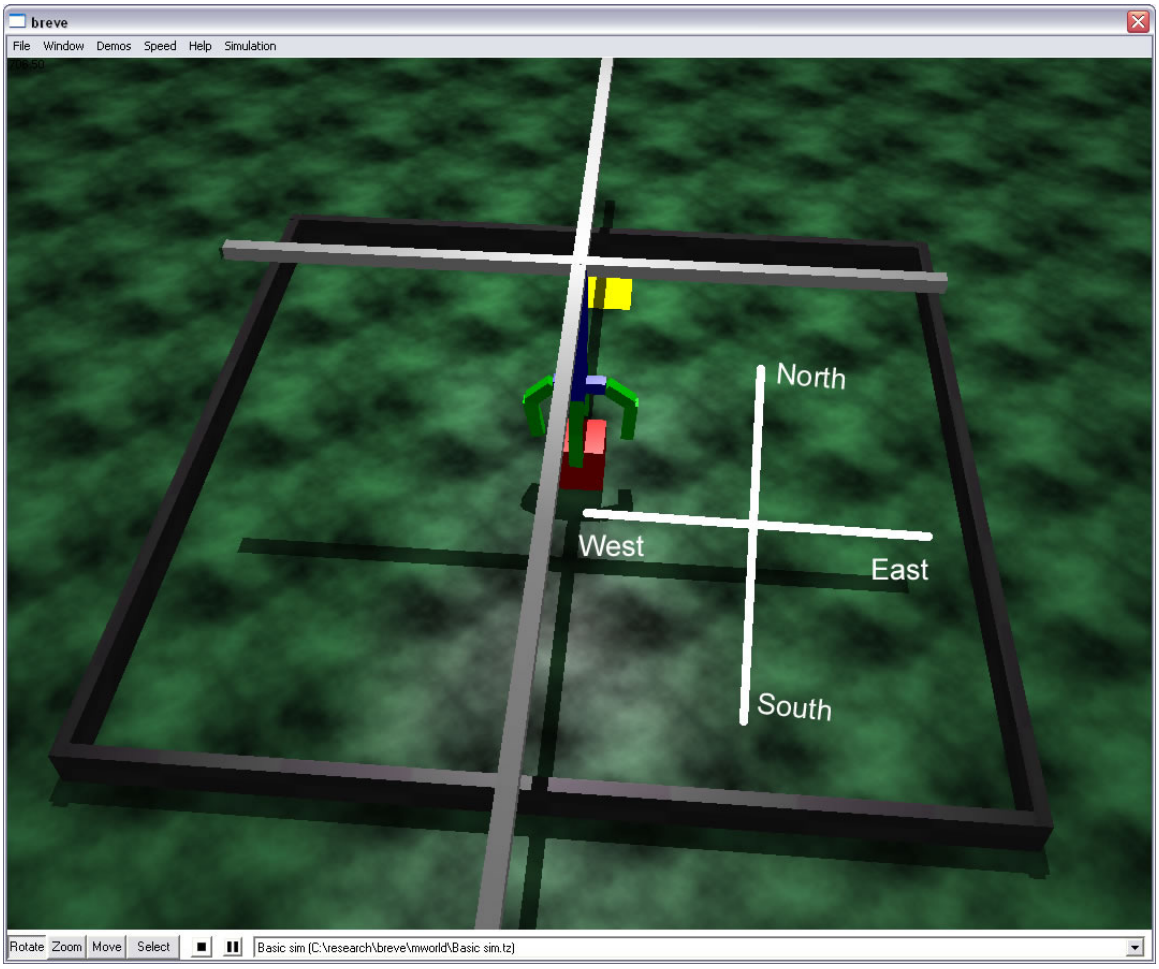


Figure 5.3 The basic simulation with the directions labeled.

The choices for the core ontology are arbitrary, but logical. They could be forward, back, right, left or even fabricated nonsense words. What is important is that what is chosen adheres to guidelines for good ontologies such as coverage, discreteness, etc. These particular terms were chosen mostly because they are similar to how humans might choose to classify the simulation and are therefore readily comprehensible to a human observer. The ontology serves as a foundation to describe the environment and actions within it. Without some framework to describe what occurs in the environment it would

very difficult to describe what was occurring within the environment at any given time. The core ontology provides an objective vocabulary to describe what is occurring within the simulation.

5.2 Semantically Grounding the Core Ontology in the Simulation Environment thorough Babbling

With a basic simulation and agent in place it is necessary for the agent to learn certain skills if it is to achieve imitation or tool use. Fortunately there is a well established road map to achieving imitative behaviors. The best understood exemplar of the development of imitative abilities is human beings, and research has shown that humans achieve competence by progressing through a series of stages. It is possible to structure the development of imitation in artificial agents around the stages of imitative development in humans that Meltzoff identified (Meltzoff, Prinz 2002). These four stages were discussed in depth in chapter 4 and consist of: body babbling, imitation of body movements, imitation of actions on objects, and imitation based on inferring intentions of actions. With extensive work, it is currently possible to successfully achieve behaviors in an artificial agent consistent with the first three stages within the simulation. Although not yet accomplished it is believed that with further work the system could achieve actions consistent with the final fourth stage of human imitation.

Using these stages of human development as the basis for robotic development has precedent. Meltzoff himself later co-authored a paper suggesting that the stages could serve as a pathway for teaching robotic imitation (Rao, Shon & Meltzoff 2004), and others have successfully executed limited imitative acts with robot based which learned to imitate based on the four stages (Acosta-Calderon, Hu 2005).

In achieving ontology mediated imitation, the first task is to provide a linkage between the actions of the hand agent and the core ontology. The ontology and the agent are separate entities and if there is to be linkage between them, it must be either hard coded or learned. This is sometimes referred to as semantically grounding the ontology. The problem is that although there is a hand agent and it is only capable of performing a limited set of motor signals there is no preexisting knowledge of how motor signals and the ontology will align. This connection must be learned or hard coded, there is a very strong emphasis in this research on trying to have the agent learn all the components of the actions itself. The point is not to demonstrate that a robot agent could be programmed to perform an action, but rather to demonstrate that the agent could learn to perform that action itself. The foundation of this learning involved babbling behaviors, similar to those seen in the first stage of human imitation.

Human infants learn the relation between the signals they send their bodies and the resultant change in their bodies through babbling. Though babbling, infants learn how various motor signals will influence various body configurations. It is appropriately

described as babbling because random signals are sent and the resulting changes are observed. If there are motor signals that consistently result in certain changes then it is possible to begin to strengthen the connection between a certain signal and a certain change.

In humans, this type of babbling and learning behavior is innate. To accomplish it within simulation requires a great deal of infrastructure capable of accomplishing several tasks, first to create large amounts of data, second to analyze this data, thirdly to recognize patterns within the data, then to test these pattern, and if they test successfully, to record them as behaviors.

Within the simulation it was possible to collect huge amounts of data. This was accomplished by collecting data from a range of sources every 200 milliseconds. The data which was collected fell into several categories. For the initial babbling there were two types of data considered, motor signal and position. For the motor signals attribute, all motor signals passed to the hand were recorded. For the position attribute, virtual sensors which tracked position were placed on the hand and their position recorded. Position data was recorded for the center of the hand on the palm as well as all four of the fingertips.

With data collection in place, it was possible to babble and collect data. Babbling consisted of placing the hand in various random positions and then sending it all of its

various motor signals while recording data. With large enough amounts of data recorded, it gradually becomes possible to see certain consistent patterns within the data. With repeated firings of all motor signals it emerges in the data that there is a fixed range of motion to all of the actuators.

With each of the actuators it was possible to learn a series of values that bound their motion. Once these limits are reached the actuators no longer respond to signals directing further motion. It is important to account for these limits so that failure to respond to a motor signal at the limit of motion can be taken into account. The north south actuator was found to never move the hand farther north than 24 units and to never move it farther south than 24 units. Similarly the east west rail was found to extend the hand 20 units to the east and 20 units to west, but never any further. The up down actuator was found to go as high as 10 unit above the floor and only go as low as 1 unit above the floor.

The patterns observed in the fingers are more complex. Their absolute limits of motion are similar to that of the rails, but the fingers themselves move back and forth in a small range relative to the hand. After noting the motion of the hand from an increasingly large number of starting positions an area of coverage begins to emerge that encompasses all of the area within the limits of motion of the actuators. See the figure below that roughly shows the extent of the hands reach that emerges in testing.

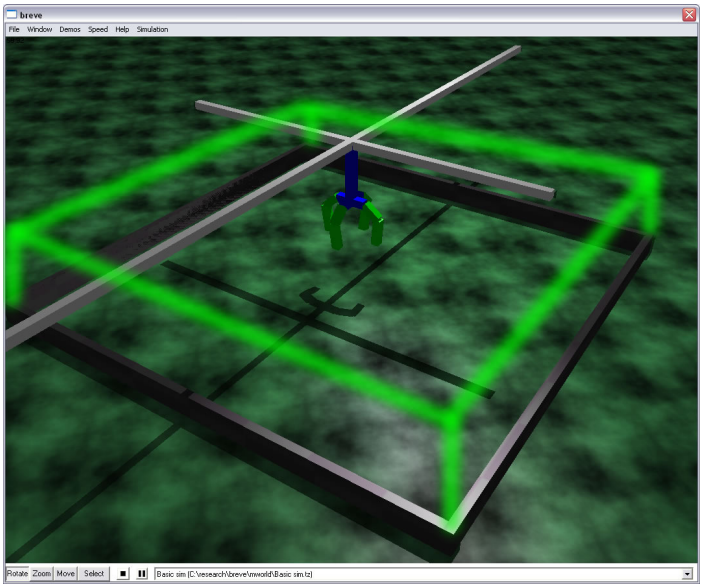


Figure 5.4 A cube, outlined in green, that extends above the floor and roughly shows the limits of the hands motion.

While exploring and learning the limits of motion is important, it is also important to learn correspondence between specific motor signals and specific motions. It was possible to see that finger signals only move the finger that is sent the signal a short distance to the finger's limit, while the rails will move the hand and all fingers a much greater distance before their much larger limit is reached. With each motor signal it was possible to create a model of motion in each direction for all position sensors. In addition to direction, these models also represent the rate of motion for each motor signal. The motion of the actuators was simple linear motion, so it was possible to represent it mathematically with equations.

The general linear equation to describe n variables is:

$$a_1x_1+a_2x_2+\ldots+a_nx_n=b$$

Each motor signal was sent to the hand and the effect of the signal on each of the hands position sensors was recorded. It was found that if the hand was not at the limit of its motion it was possible to model the motion completely using a simpler form of the equation, the general form of the linear equation for only two variables:

$$Ax+By+C=0$$

The hand was sent a single signal and this signal lead to motion in one of the axes. It was possible to collect this positional data and time data and use the data points to solve the linear equation. Once the equation had been solved it could be used to predict the position of a particular sensor, in the presence of a given signal, at a certain time.

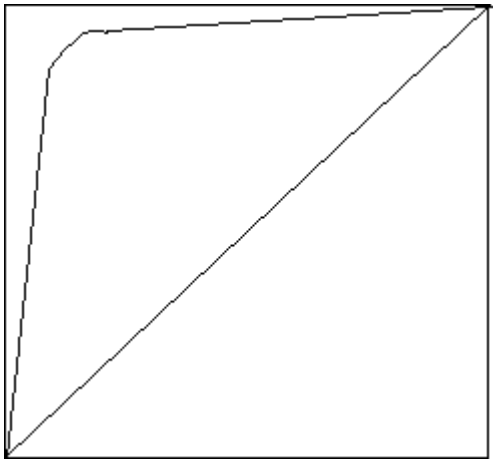


Figure 5.5 A simple ROC Curve representing learning a model for simple linear motion

Examination of the data demonstrated that each sensors motion could be accurately modeled with a linear model which was quickly learned. In practice, accurate models could be made from very few data points and performance improved only marginally if the data set was enlarged.

It was also seen that in the presence of multiple signals the effects of these models was additive with no interference between signals. Recognizing this regularity in the behavior of the hand, it was possible to create an overall software model of the hand from the data which could predict hand and finger position over time for a given set of motor signals. This software simply combined the effects on motion of each equation correlated to an active motor signal and tracked the limits of motion. In the figure below is a screen shot of the first hand model.

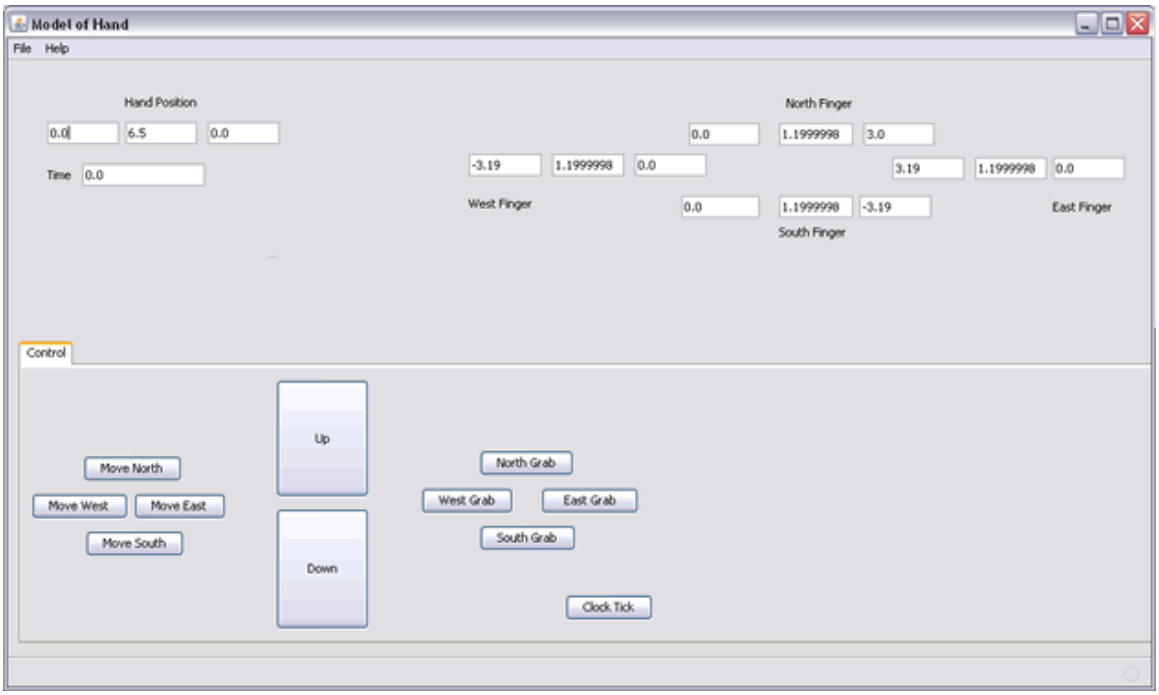


Figure 5.6 Software model of hand.

In the figure above we can see all of the position data represented by the text fields in the upper regions above as well as all of the possible motor signals represented by buttons in the lower region. This software implementation combined all the linear models to accurately predict hand behavior over time in the presence of various motor signals.

To operate the model, one or more buttons would be pressed to represent a certain set of motor signals being sent to the hand. The clock tick button would then be pressed and the time and predicted sensor position would move forward by 200 millisecond increments. This process could then be repeated with another press of the clock tick button which would again move forward time and sensor position by 200 milliseconds.

At any time motor signals could be changed or removed. In practice, it was found that the model provided accurate predictions of hand motion.

The ability of the model to accurately predict the hands movements is interesting from a technical standpoint, but it also has significance from a more semantic standpoint. Initially an instruction such as move hand north would have been meaningless to the hand agent. By exploring all the hands motor signals and their subsequent effect on position it was possible to learn from experience that to move the hand north the best method was to send the north south rail a signal to move forward. By examining the effects of motor signals on position, it was possible to learn how motor signals related to our core ontology.

Using these methods it was simple to learn which motor signals were best to move north, south, east, and west. Once this knowledge had been gained, it was possible to formalize it into behaviors. Each candidate behavior consisted of some positioning of the hand, a motor signal, and a predicted resulting motion. It was possible to test each of these by putting the hand in the appropriate position and sending it the signal, if the hand consistently moved as predicted, the behavior was considered to have been learned and was added to the hands set of behaviors. At the end of this simple babbling and learning process, the hand had learned a series of basic movement behaviors: ***Move North, Move South, Move East, Move West, Move Up, Move Down, Stop***. Although these behaviors were primarily actions based on the directional portion of the core ontology, the babbling

behavior had firmly grounded them semantically for the hand. These motions were now behaviors that were directly linked to particular motor signals.

A hand was not the only agent that was explored in the simulation environment. A model of a simple wheeled vehicle was also explored. The vehicle model created consisted of simple rectangular body with two circular wheels located towards the rear of the vehicle. The wheels may spin in the same direction causing the vehicle to move in the direction it is pointing or the wheels may rotate in opposite directions causing the vehicle rotate in place. The vehicle was designed with five simple control signals. The first control signal would rotate both wheels in the same direction, moving the vehicle forward. The other four signals would turn the two wheels in opposite directions until the vehicle was pointing in one of the four directions. The vehicle provides data in terms of a sensor that provides the position of its center and another sensor that provides the orientation of the vehicle. Given these five signals, it was possible to again engage in babbling behavior. This consisted of firing combinations of the possible motor signals at the vehicle and observing the resultant changes in state.

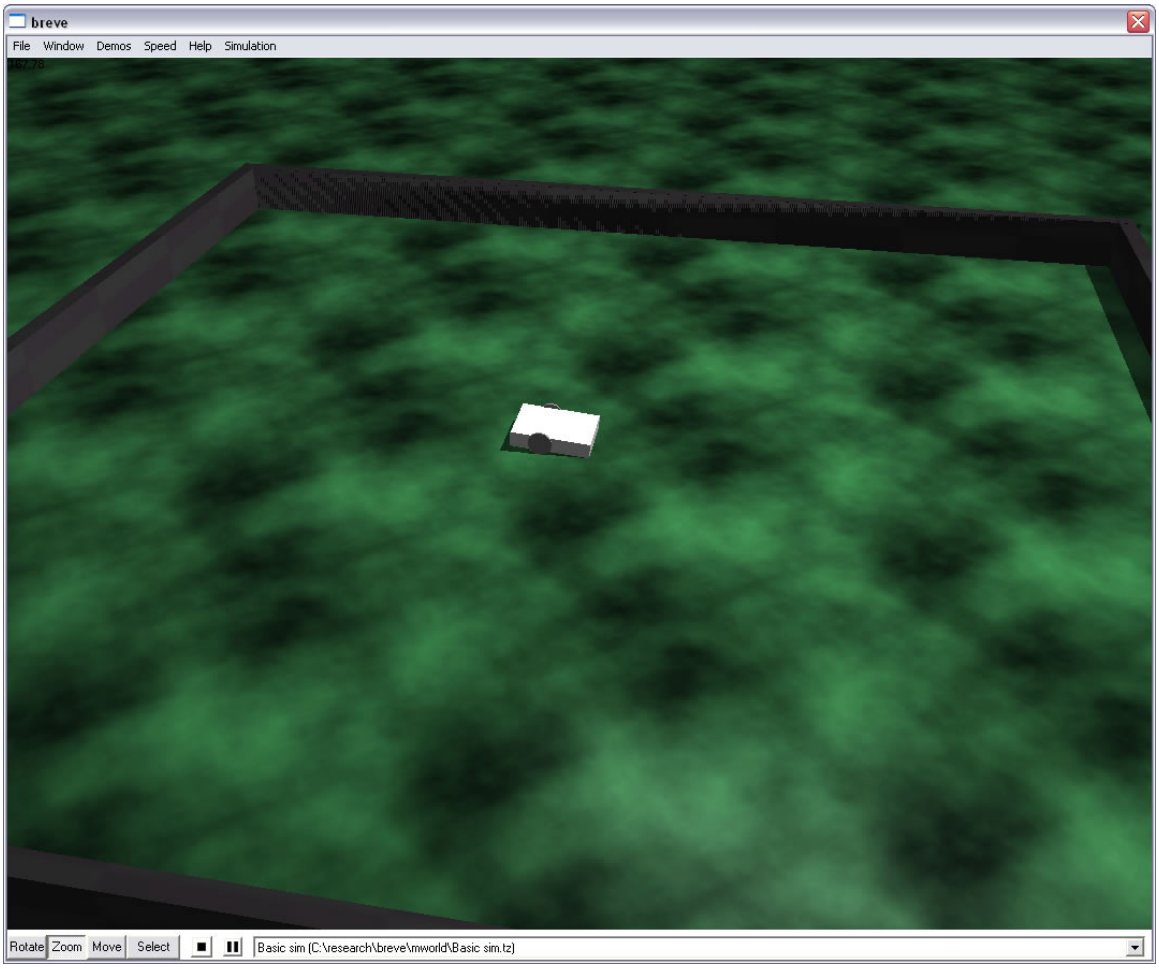


Figure 5.7 The two wheeled vehicle

In examining the data that this babbling generated it was possible to observe that there were sequences of motor signals that would lead to the vehicle moving a significant amount directly in each of the four directions. Examining the data further demonstrated that there are three successful strategies that the vehicle may employ to move directly in a particular direction. The first was for the vehicle to already be pointed in a particular direction and to then move forward. The second was for the vehicle to be pointed in any direction, then rotate to the correct direction, and then to move forward. The third

possibility involves a multiple number of rotations which end in a rotation pointing in the correct direction and then the vehicle moving forward.

It was possible to examine repeated random sequences of motor signals to the vehicle and to identify sequences that moved the vehicle directly in a particular direction. If this was solely done until three candidates where found it was fairly quick, but often several of the learned behaviors would involve spinning a few times before moving i.e. they belonged in the third category of solutions described above. If a utility function was added to that favored quicker solutions with less steps, it was found that it was possible to gradually learn strategies for moving in all directions that did not involve unnecessary turns if the data set was large enough. By increasing the size of the data set it became more likely that the random sequences of motions would contain the quickest solution for each direction.

Though the use of this process it was possible for the vehicle to learn a sequence of motor signals that would allow it to move directly in each of the four directions of our core ontology. This allowed the vehicle to be given a series of behaviors in the same manner that the hand had been. The hand was able to learn to carry out the following behaviors: ***Move North, Move South, Move East, Move West, Stop***. Through babbling behaviors both agents were able to identify motor signals that allowed themselves to change their position in a manner that correlated well with a particular description in the core

ontology. Even though the motor signals involved where quite different, both agents were able to learn sets of motor signals that allowed themselves to move in each of the four directions of our core ontology. For example, both agents were able to learn a ***Move North*** behavior that moved them northward, the motor signals each learn for the similar act varied widely. The hand learned to activate one of its rail actuators, while the vehicle learned to turn in place and then move forward. Even though their approaches to movement were quite different, both agents were able to semantically ground there own locomotion within the ontology. It is important to note that there will only be a correlation in possible behaviors between different agents in so far as their inherent abilities correlate. For example the hand has behaviors for ***Move Up*** and ***Move Down***, while there are no corresponding movements for the vehicle since it lacks any ability to move up and down.

It is interesting to note that these behaviors enable some very basic imitation between the agents where they overlap. If one of the agents moves north, it can describe its action as ***Move North***. The other agent can then roughly imitate what was done since it also has learned a behavior ***Move North***. The ontology serves as a bridge that allows the underlying semantics of one agent's actions to be roughly imitated by another. There are admitted limits to this imitation based on the inherent abilities of the agents involved. As was mentioned the hand can move up and down, while the car cannot. Therefore the car could not imitate a ***Move Up*** action performed by the hand.

5.3 Learning Actions on Objects

Once the two agents had learned basic motions the next step involve exploration of basic actions upon objects. This type of activity could be described as similar to Meltzoff's stage of "imitation of actions on objects". To explore actions upon objects, the same simulation environment was used with a cube placed within the environment along with the hand.

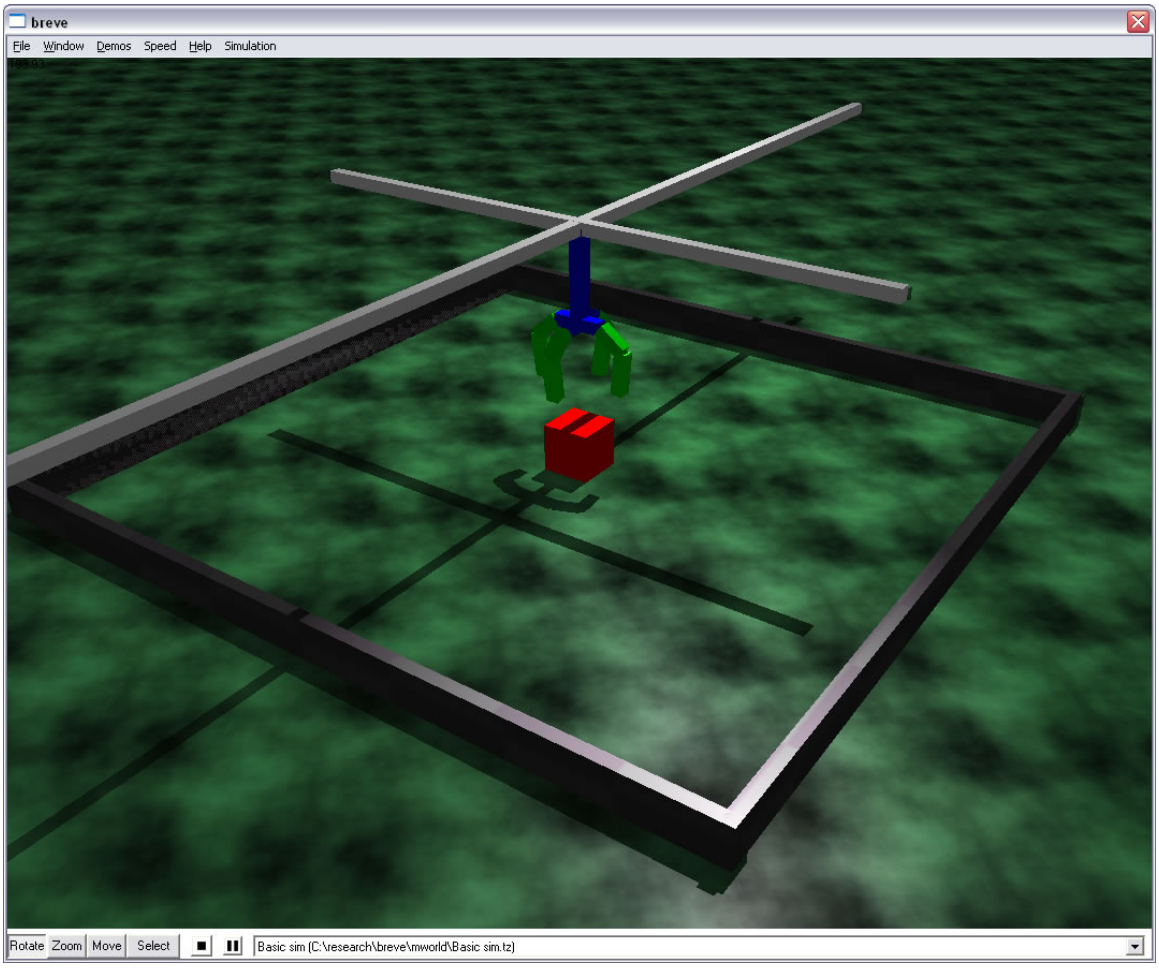


Figure 5.8 The hand in simulation now accompanied by a cube

Initially the hand will have no knowledge regarding how it might interact with the cube. In fact from the majority of positions it was seen that most of the hands motor signals have no effect on the cube. In was possible to identify locations where the hand was able to move the cube in any way through the flowing algorithm:

```
function actionExploration ( ) returns a set of interesting actions α
inputs:
static: possibleActions, The set of possible actions including combinations of actions
        maxPosition, end of range of motion
        initialObjectPosition starting position of object
        interestingActions a set of interesting actions, initially empty

while notDone
        moveAgent(randomPosition)
        for agentPosition ← 0 to maxPosition
                while possibleActionStack notEmpty then do
                        currentAction ← pop possibleActionStack
                        Execute(currentAction)
                        if currentObjectPosition notEqualTo initialObjectPosition
                                logData(agentPosition, possibleActions,
                                        initialObjectPosition)
                                interestingActions ←{agentPosition, currentAction,
                                        currentObjectPosition, initialObjectPosition}

                possibleActionStack ← possibleActions

return interestingActions
```

This algorithm performed a simple exploratory behavior where the hand would move to a random position. It would then perform all of its possible motor signals. If any of these signals resulted in the cube being moved in any way, the position was considered interesting. For each interesting position, the hand would fire every motor signal and record all resulting data to a log.

The same data from the initial babbling was recorded, the position of the hand and each fingertip, and the motor signal that was sent. Additional information was recorded as well. The position of the cube was recorded as well as the position of each sensor relative to the cube. This relative position was recorded in terms of the core ontology. For example, along the east west axis, there were five possibilities for how a sensor could be positioned relative to the cube. The possibilities were: west of the cube, on the west edge of the cube, centered east/west with the cube, on the east edge of the cube, and finally, east of the cube. This relative positioning to the cube was also recorded for the north/south and up/down axes as well.

With all of theses data objects being recorded, the hand would interact with the cube, moving to random positions and recording data when it was able to move the cube in any way. This process was carried out over hundreds of trials and gradually a data set was recorded that was representative of the ways in which the hand could move the cube if it was placed in the center of the playpen. Once the data set was sufficiently large so as to be representative, it was possible to learn from it.

To illustrate how it was possible for this learning to occur, consider the following example. In this data set the cube starts centered north south and east west for each trial. The hand is put in random positions, sent all possible motor signals, and data is recorded for positions were the cube is moved. In the graph below the Y axis represents north south movement and the X axis represents east west movement. The graph may be

thought of as a birds eye view of the cubes position on the floor. Any data point not in the center at 0,0 represents a situation were the cube has been moved by the hand.

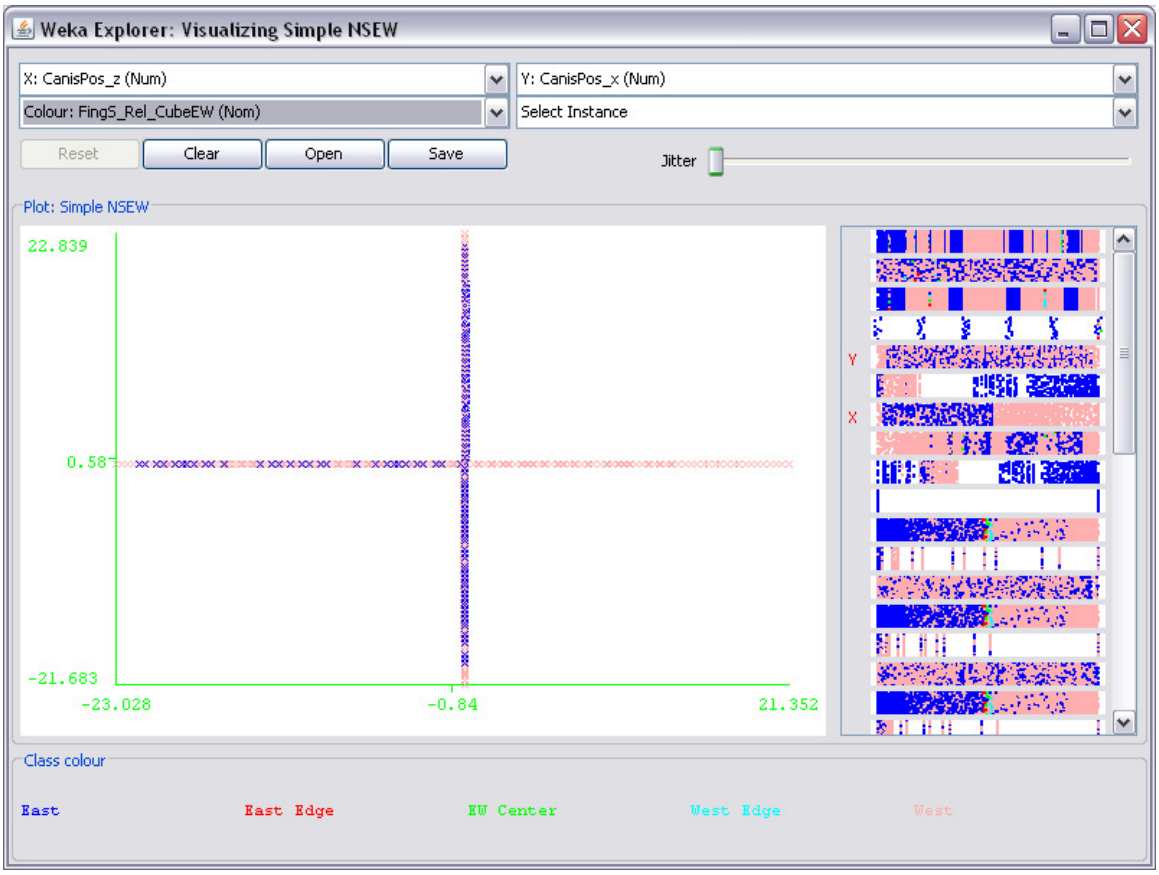


Figure 5.9 The position of the cube over time along the north/south east/west

To In this plot, the cube is either staying still or is pushed in a straight line, north, south, east or west. This plot represents all the ways that the hand may move the cube from the center of the simulation. Each plot is the result of a single motor signal from sent to the hand while it occupies some random position in the simulation. If the goal is for the hand to learn how to push the cube, this data represents example of how to do so. It is possible

to look at individual data points, but it is better to examine sets of data to see if they share characteristics.

To break the created data sets up into smaller subsets various clustering algorithms were employed. The first we employed was k-means. K-means automatically clusters a dataset into a number of clusters. The number of clusters must be specified beforehand. The algorithm operates by assuming that object attributes form a vector space and then tries to minimize intra-cluster variance calculated by the squared error function:

$$V = \sum_{i=1}^{k} \sum_{x_j \in S_i} (x_j - \mu_i)^2$$

Where there are k clusters S_i, i=1,2,…,k and μ_i is the centroid of all the pints $x_j \in S_i$. The iterative refinement heuristic known as Lloyd's algorithm may be applied until points no longer switch clusters.

In practice, this algorithm worked well with clustering the data, but there was a need to specify the number of clusters prior to the clustering occurring. Specifying the number of clusters which would appear strongly influenced the results and therefore an x-means clustering algorithm was chosen. The x-means algorithm clusters data similarly to the k-

means but does not need to be given a number of clusters. The algorithm will return a configuration and number of clusters that will optimize the Bayesian Information Criterion (BIC).

Pelleg (2000) (Pelleg, Moore 2000)presents the following method and equations to calculate the BIC for a set of data D and a family of alternative models M_j which correspond to different values of K. The models are scored by their posterior probabilities **Pr[**M_j **|** D**].** To approximate the posteriors, up to normalization, Pelleg employs the following equation from Kass and Wasserman (Kass, Wasserman 1995) also know as the Swarz criterion.

$$BIC(M_j) = \hat{l}_j(D) - \frac{p_j}{2} \cdot \log R$$

Where $\hat{l}_j(D)$ is the log-likelihood of the data according to the j-th model (taken at the maximum likelihood point) and p_j is the number of parameters in M_j. The Maximum likelihood estimate (MLE) may then be calculated as:

$$\hat{\sigma}^2 = \frac{1}{R-K} \sum_i (x_i - \mu_{(i)})^2$$

Where the point probabilities are:

$$\hat{P}(x_i) = \frac{R_{(i)}}{R} \cdot \frac{1}{\sqrt{2\pi}\hat{\sigma}^M} \exp\left(-\frac{1}{2\hat{\sigma}^2}||x_i - \mu_{(i)}||^2\right)$$

The log-likelihood of the data is

$$l(D) = \log \prod_i P(x_i) =$$
$$\sum_i \left(\log \frac{1}{\sqrt{2\pi}\sigma^M} - \frac{1}{2\sigma^2} ||x_i - \mu_{(i)}||^2 + \log \frac{R_{(i)}}{R} \right)$$

Pelleg's approach will return a number and configuration of clusters that will maximize BIC and therefore represent a good choice for clustering a given data set. When a x-means clustering algorithm is applied to the data set in the previous graph the following six clusters were identified:

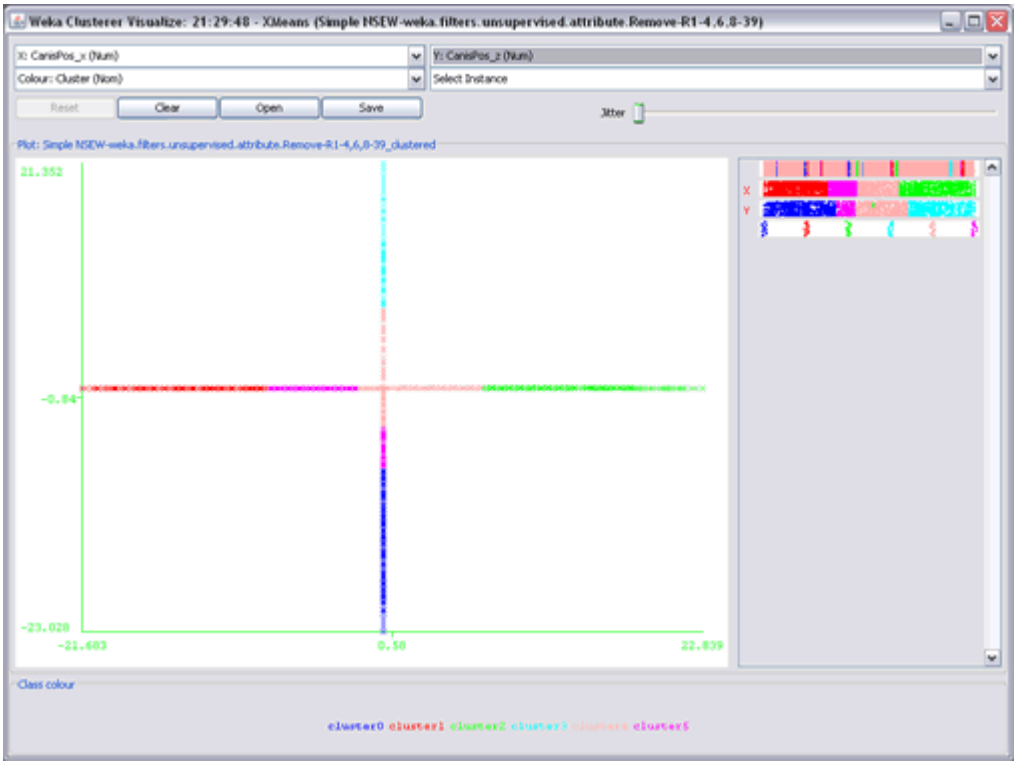


Figure 5.10 The data has now been divided into six clusters identified by color.

```
=== Run information ===

Scheme:         weka.clusterers.XMeans -I 2 -M 1000 -J 1000 -L 2 -H 7 -B 1.0 -C 0.5 -D
"weka.core.EuclideanDistance -R first-last" -S 10
Relation:     Simple NSEW-weka.filters.unsupervised.attribute.Remove-R1-4,6,8-39
Instances:    41152
Attributes:   2
```

```
              CanisPos_x
              CanisPos_z
Test mode:    evaluate on training data

=== Model and evaluation on training set ===


XMeans
======
Requested iterations            : 2
Iterations performed            : 2
Splits prepared                 : 6
Splits performed                : 4
Cutoff factor                   : 0.5
Percentage of splits accepted
by cutoff factor                : 0 %
------
Cutoff factor                   : 0.5
------

Cluster centers                 : 6 centers

Cluster 0
            -0.002904847043701792 -14.723117435732595
Cluster 1
            -13.468230047988675 -0.01093418489767108
Cluster 2
            14.173503561844885 0.007007901467505207
Cluster 3
            -0.006276841483979775 14.702864377740333
Cluster 4
            0.026043435030095267 0.07067545202494767
Cluster 5
            -3.0869459451659345 -1.5080033968253888

Distortion: 749.042159
BIC-Value : 133946.27671
Clustered Instances

0        757 (  2%)
1       1249 (  3%)
2        904 (  2%)
3        723 (  2%)
4      36637 ( 89%)
5        882 (  2%)
```

Taking the six clusters that are identified and examining them yields interesting results. The two center cluster compromises such a wide range of configurations and data types that it is impossible to see any strong patterns in the data. However if we examine the cluster to the north (i.e the light blue cluster at the top of the graph), we see a collection of data points that all share certain characteristics. Since this data was all formed with the cube starting in the center of the simulation, the pink cluster consists entirely of data

points where the hand has somehow moved that cube from the center to northern portion of the enclosure. It is possible to examine this cluster of data to gain insight into how the cube may be moved by the hand.

When the light blue cluster is examined in detail it is seen that for every data point that is present in the cluster there are certain common features. Of course there is a common position of the cube on the northern area of the floor, this is obvious given that the data was all classified based on position. What is more interesting is that there are other features that are common for the entire data set. It is interesting to note that in each instance where the hand has managed to move the cube to the north of the enclosure, there are the same spatial relations of where the hand is relative to the cube, there are some exceptions with friction or the cube flipping over creating noise in the data, but there are clear patterns. In the vast majority of cases the hand is **centered up down** with the cube. In almost all the data the hand is **centered west east** with the cube. In nearly every case the hand is to the **south** of the cube. In every single data point we found that the hand was sending a motor signal to the north/south rail to move forward. These features were common to nearly every instance where the cube was moved northward.

What is interesting to note is that this collection of data regularities is may serve as the basis for creating an Action Ontology to move the cube north. To review, action ontologies take the form :

$\alpha = \langle D, C, H \rangle$ Where D is the change in an object property resulting from an action and C is the context where the action occurred and H is the control signal.

Each of these items is known in this case. The change in object property is that the cube is moved northward. The context where the action occurred is **centered up/down**, **centered east/west**, and **south.** The control signal is to more the north/south rail forward, which in the initial babbling was learned to correlate to **move north**.

Experimentally it was possible to verify that this action ontology was valid by replicating all relative positions of the hand relative to the cube and sending a signal to the north/south rail to move forward. In nearly every instance this succeeded in pushing the cube in a northward direction. So this is a substantial extension of the hand agent's knowledge. It has learned a strategy whereby it may push the cube north. It is possible to add a new behavior **move cube north** that is described by the action ontology.

It was possible to learn similar action ontologies to move the cube in the remaining three directions based on the information in the remaining three clusters.

In the cluster to the east the hand was found to be **centered up/down**, **centered north/south,** and **west** relative to the cube and the west/east rail was receiving a motor signal to move the rail to the east.

In the cluster to the west the hand was found to be **centered up/down**, **centered north/south,** and **east** relative to the cube and the west/east rail was receiving a motor signal to move the rail to the west.

In the cluster to the south the hand was found to be **centered up/down**, **centered east/west,** and **north** relative to the cube and the north/south rail was receiving a motor signal to move the rail to the south.

Putting these in the form of our action ontology and then testing them, it was possible to learn three more behaviors: **move cube east, move cube west,** and **move cube south.**

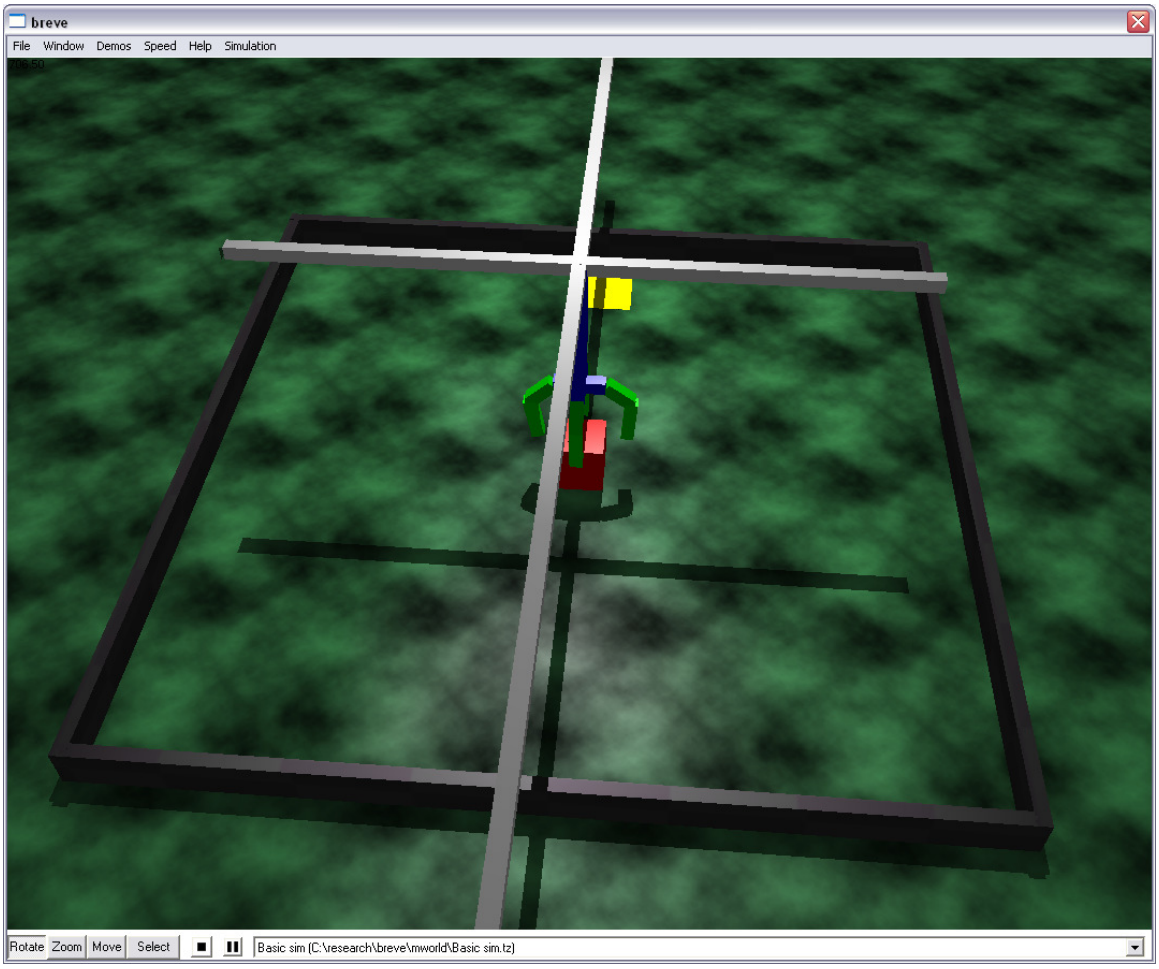


Figure 5.11 The hand presented with a simple problem.

After the hand has learned sets of behaviors, it is possible to use them to solve simple problems. In the figure above there is a simple problem to be solved. The goal is to place the cube onto the yellow goal plate to the north of the cube. Previously the hand was able to learn an action ontology that allowed it to push the cube north. If the current situation is examined the goal is to push the cube northward onto the goal plate. The hand agent knows how to push the cube northward and solving this problem is as simple as employing that ability until the cube is on top of the goal plate. The hand simply

applies the learned action ontology and moves to a position **south** of the cube, **centered east/west** with the cube, **centered up/down** with the cube and sends the north/south rail a signal to move north.

5.4 Imitating Actions on Objects

One of the prominent findings in the course of the research was that in some cases where one agent had learned a strategy to solve a problem, it was possible for another agent to use the same strategy if there was sufficient overlap between the two agent's abilities.

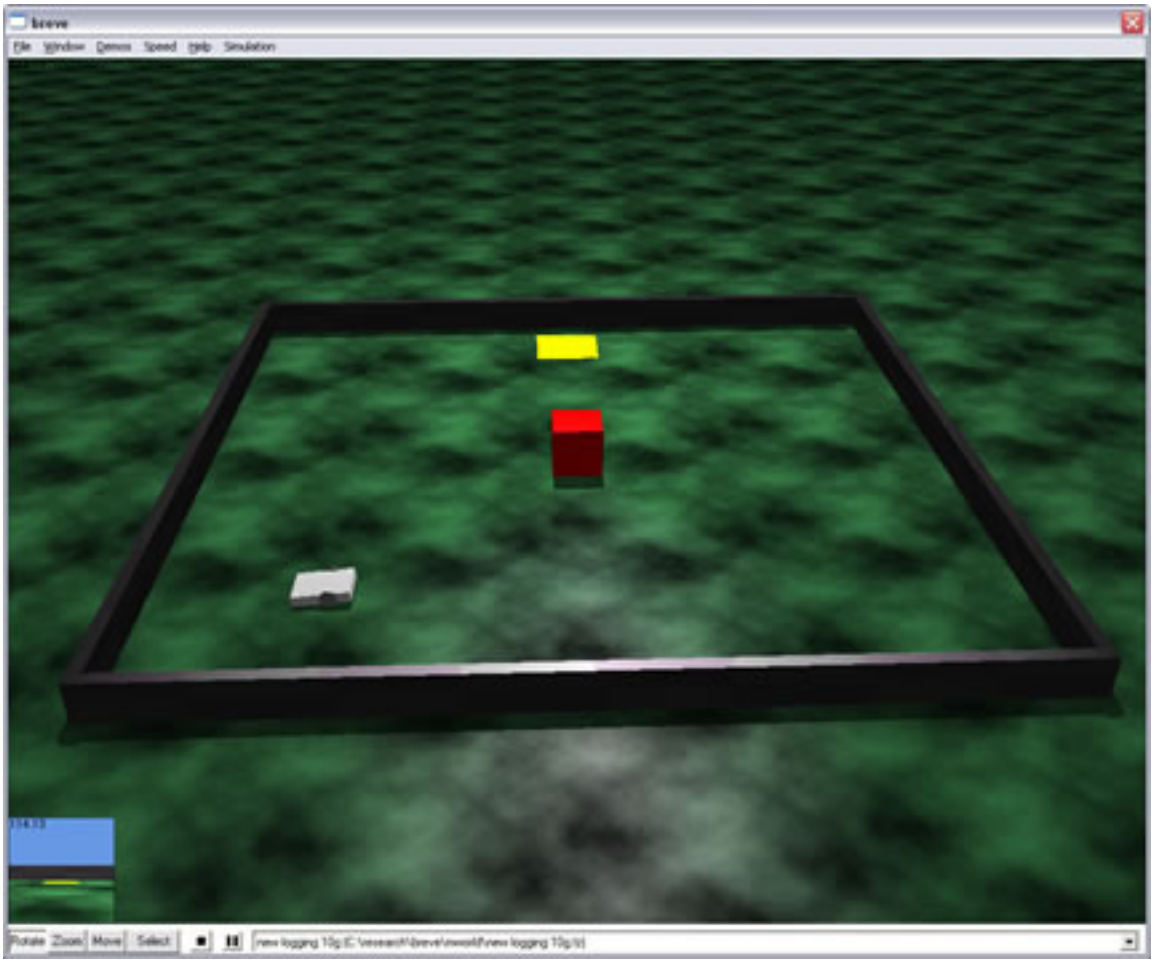
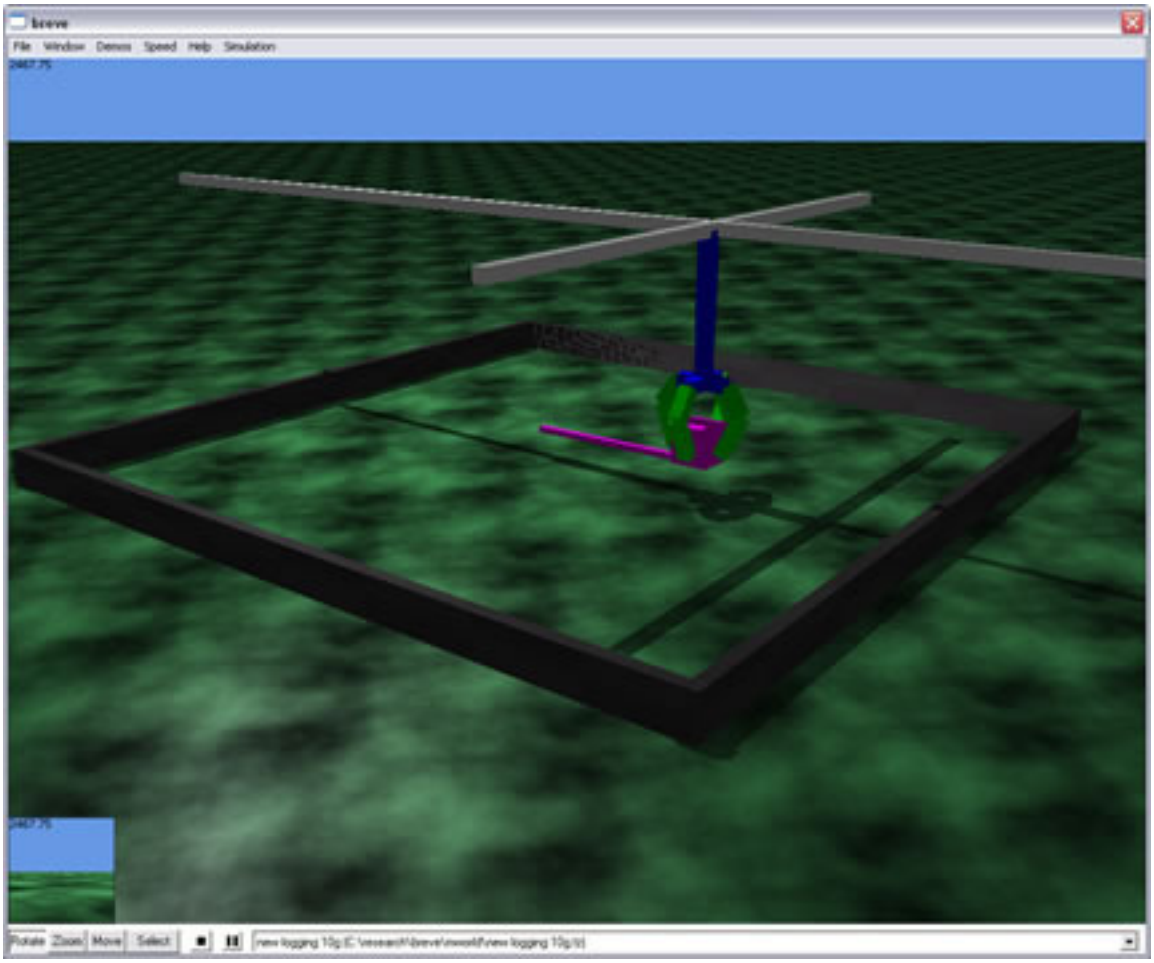


Figure 5.12 The vehicle presented with the same simple problem.

In the figure above the vehicle is placed in the same problem scenario that the hand could successfully complete. The vehicle was able to solve this problem using a problem solving strategy that had been learned entirely by the hand. Even if the vehicle was only given the very basic set of movement behaviors it had learned through babbling, it could employ the hand's solution to this problem. This was successful even if the vehicle had no benefit of having learned from babbling with the cube itself; it could still employ the hand's action ontology for pushing the cube north. The vehicle would move itself **south** of the cube. It would move itself to be **centered east/west** with the cube. It should be

noted that although the vehicle could not move up and down, it was always **centered up/down** with the cube, since they were both sitting on the floor. Once in position the vehicle could perform its known behavior and **move north** and successfully push the cube onto the goal, solving the problem scenario with a strategy learned by the hand. With both agents firmly grounded in the simple core ontology, it was able to serve as a bridge could enable the transfer of a strategy from one agent to another.

Success in implementing a strategy with one agent that had been learned by another agent set the stage for examining imitation with tools. In the original research that suggested imitation as a path to tool use, the imitation usually took the form of one agent imitating another agent employing a tool. After agent to agent imitation had been successfully implemented an interesting theory was proposed, what if rather than imitating another agent the ontology was used to have a tool itself imitate the action of an agent on an object. If a certain strategy worked for multiple agents, was it possible that the same strategy would work for a tool as well? The research with mirror neurons described in chapter three strongly suggested asking this question. Specific mirror neurons would fire when monkeys performed specific tasks. The same neurons also fired when they observed others performing the same specific task, and most interestingly, the same motor neurons would fire if the same specific task was performed with a tool.



5.13 The hand holding a poke tool.

To reuse a strategy with a tool it was necessary to ground the tool in the ontology just as the hand and the vehicle had been. It was possible to track the position of a tool tip just as the hand position had been tracked. The tool was then placed in the hand and babbling again took place. This time the fingers were not included in the babbling behavior because they were needed to hold the tool, but it was found that the same methods of modeling the motion of the hand worked for the tool tip as well. The linear modeling will not be described in detail here, but it was the same as that done with the hand as described earlier.

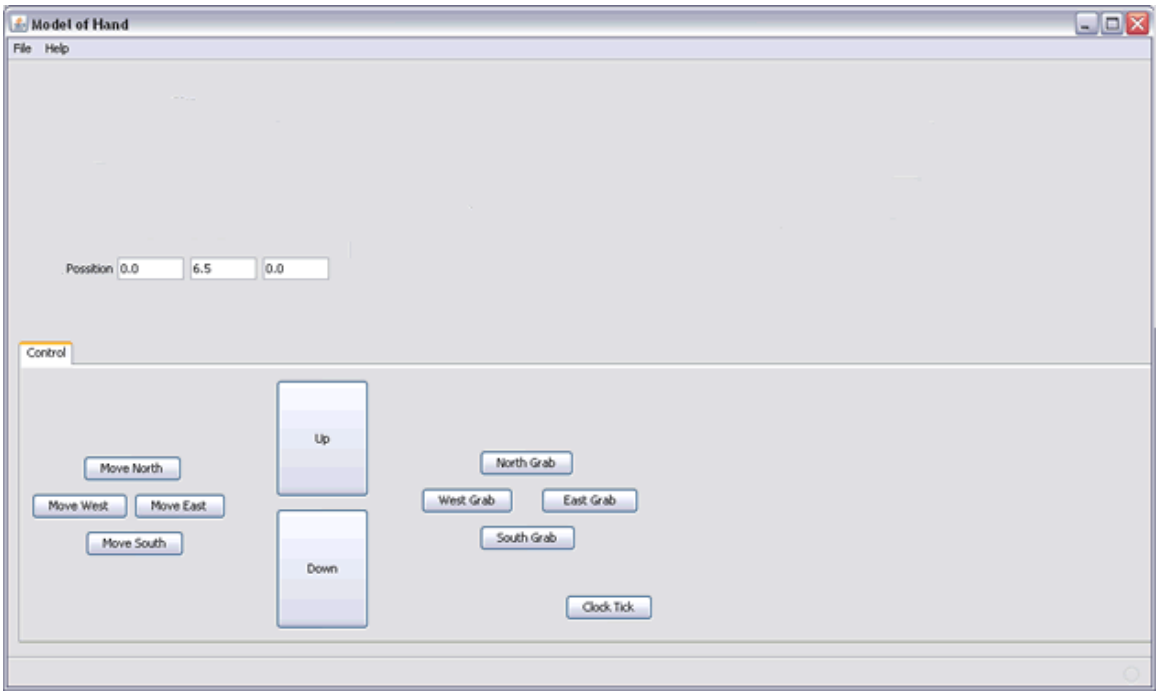


Figure 5.14 A model of the tools motion

Once equations had been learned that accurately described the tool's motion, it was possible to combine equations and predict the tool tip's motion in the same manner that the hand's had been. It was then a simple matter to learn behaviors that would move the tool tip in the four directions. At the end of this simple babbling and learning process, the hand had learned a series of basic movement behaviors that would control the tool tip: ***Move North, Move South, Move East, Move West, Move Up, Move Down, Stop***.

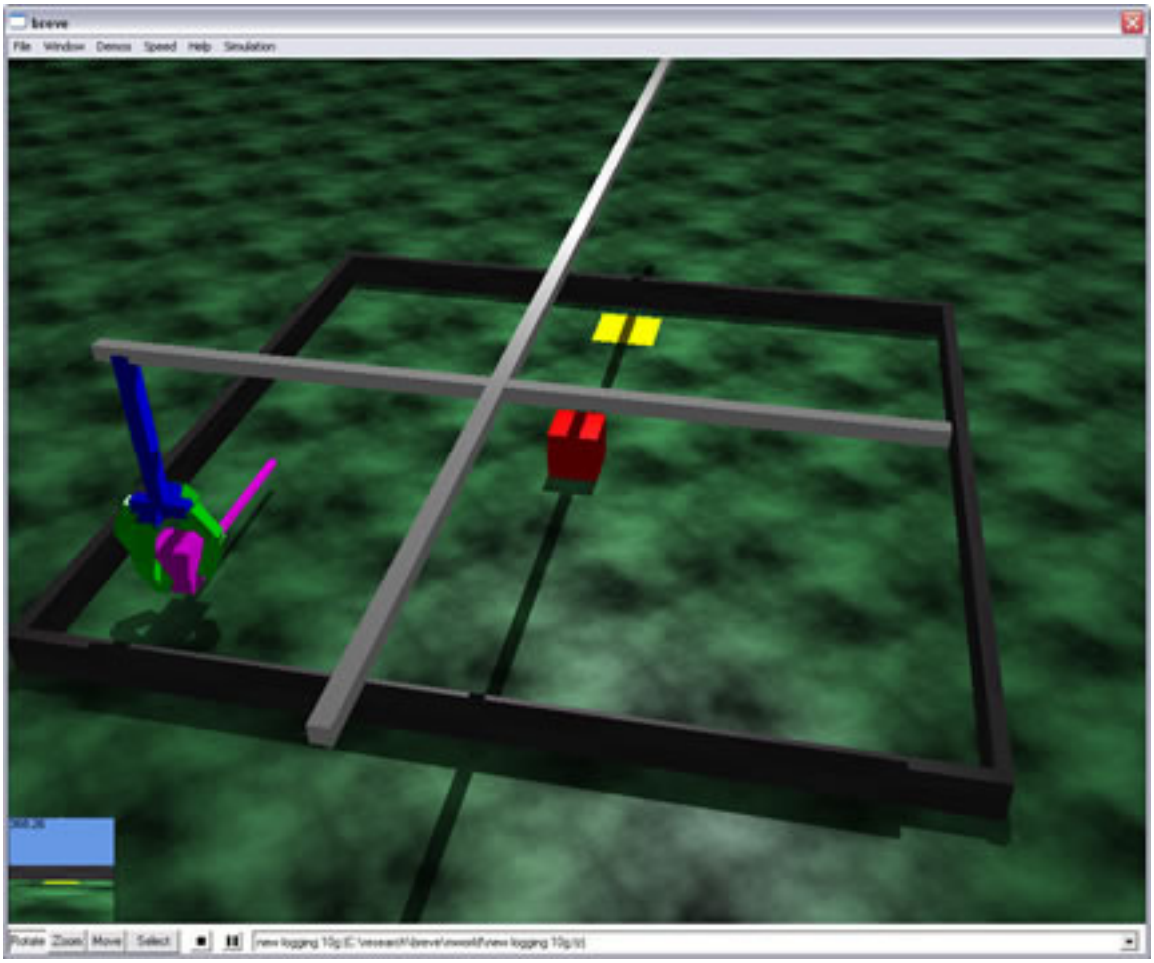


Figure 5.15 The hand holding the tool is presented the same problem

Once the hand was able to move the tool tip in a manner consistent with the core ontology, the hand was able to use the same strategy with the tool that the other agents had used. Move the tool **south** of the cube. Move the tool so it is **centered up/down** with the cube. Move the tool so that it is **centered east/west** with the cube. Move the tool north until the cube is on the goal plate. At this point it should be noted that the central claim of this dissertation has been established. Namely, that:

Imitative mechanisms give rise to tool-using capabilities.

At this point it has been demonstrated that with a simple grounded ontology, a basic set of learned behaviors can be imitated in a manner that will give rise to tool use.

5.5 Learning a Body Schema

Successfully implementing tool use by employing imitative mechanisms is a very significant accomplishment, but the previous example is a very simple scenario. The tool was effectively being used to solve a problem, but it was solving a problem that the hand could easily solve without the tool. The motivation behind many of the instances of tool use was to accomplish tasks that the animal could not accomplish unaided. It was interesting to see what happened if the task was altered slightly. Obstacles and constraints were added to the task that made it more challenging. In the next example a ledge was added that covered the cube.

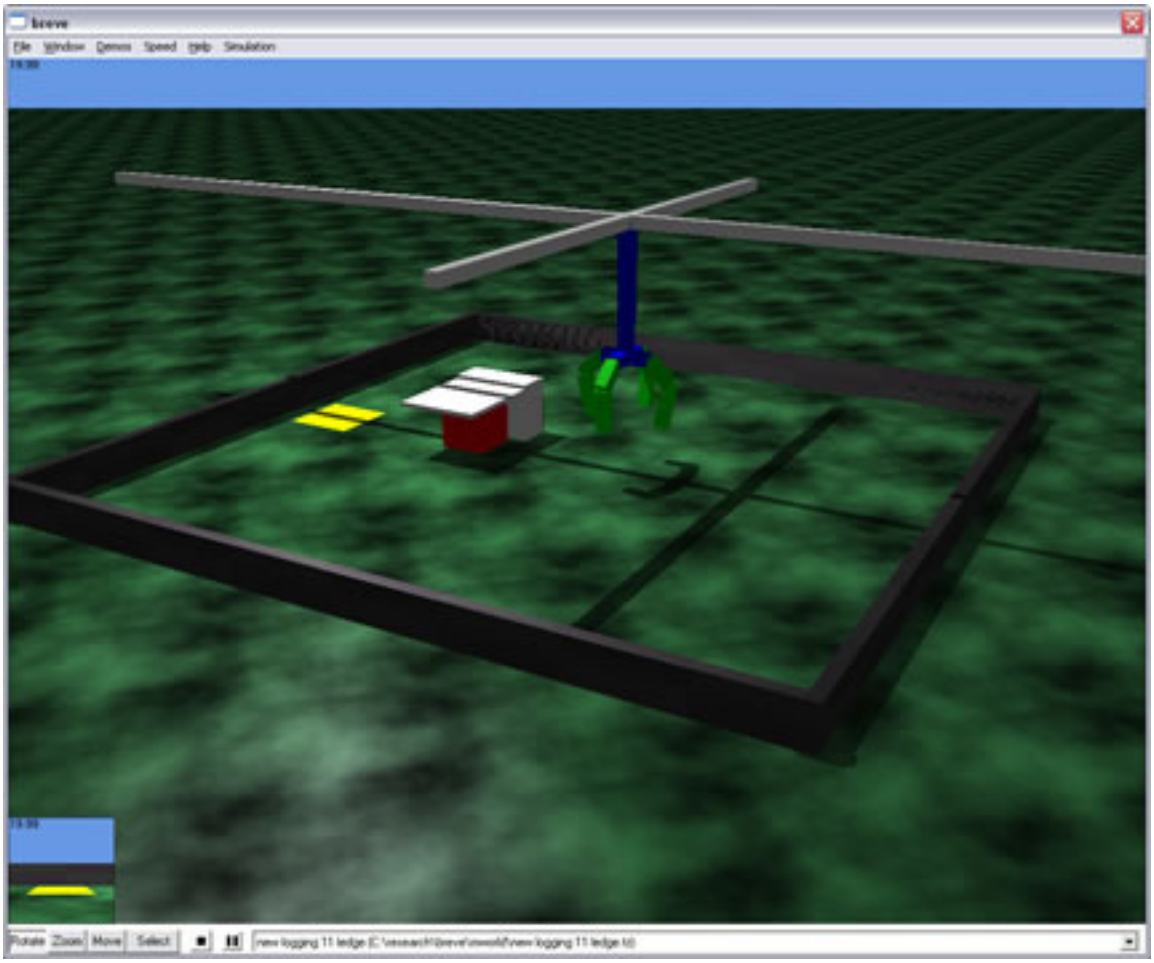


Figure 5.16 A ledge added to the cube on goal plate task

The scenario seen above is identical to the previous problem except that a constraint has been added in the form of a ledge that overhangs the cube. The goal is still to place the cube onto the goal plate. The vehicle was still able to solve the problem with the same strategy, passing under the ledge without contact. The poke tool also solved the problem without incident, sliding under the ledge without contact. The same was not true however for the hand. When the hand tried to solve the problem it would hit the ledge every time. The hand alone was unable to solve this problem.

To allow the hand to account for why it could not solve this problem it would have to be given additional capabilities. To predict when obstacles would block progress the agent needed to have some kind of innate sense of were in space its body parts where at any given time. This type of ability has been well documented in monkeys and humans and is known as a body schema (Maravita, Spence & Driver 2003). A body schema is a sense that an animal has about where in space its body parts are at any given time. Consider the following three scenarios where constraints have been added to the goal plate problem.

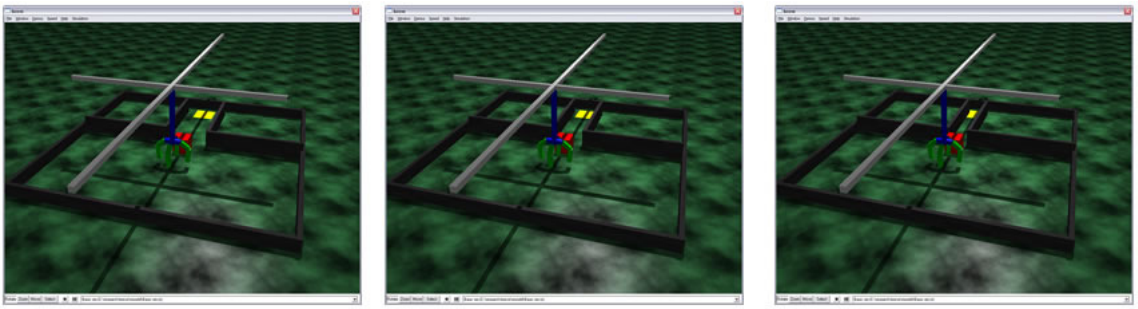


Figure 5.17 The goal plate problem with three differently sized corridors

Each of these scenarios has had a constraint added in the form of a corridor. If the hand is fully open it can fit into the widest corridor, but not the two other corridors. The smallest corridor is too small for the hand under any circumstances, but the middle may or may not be too small depending on how the hand is positioned as seen below.

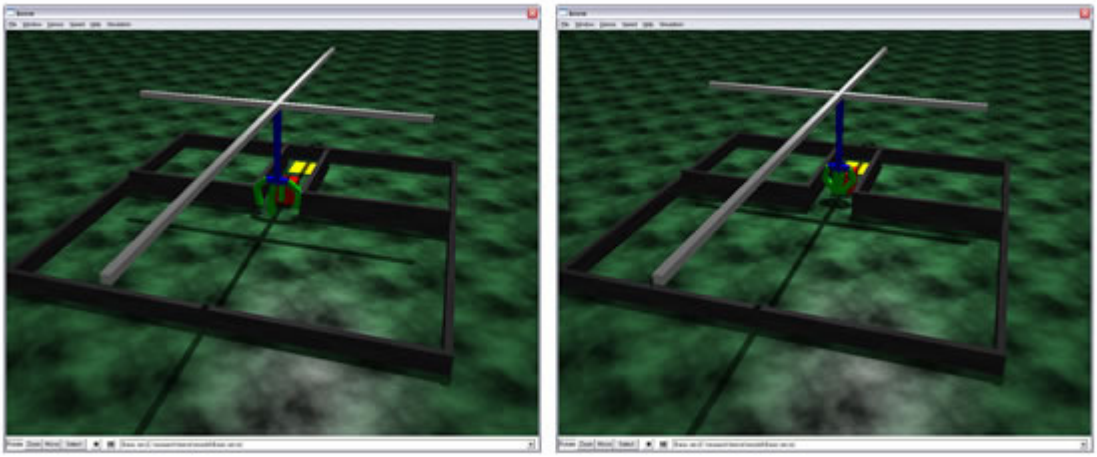


Figure 5.18 A corridor that is barely too small for the open hand to fit

It is interesting to note that in this case where the corridor is barely too small for the open hand, closing the hand will allow it to fit. A body schema will allow this to be predicted. Given that in our initial babbling we derived equations for how motor signals effected the position of limbs, it is straightforward to use this data to construct a body schema.

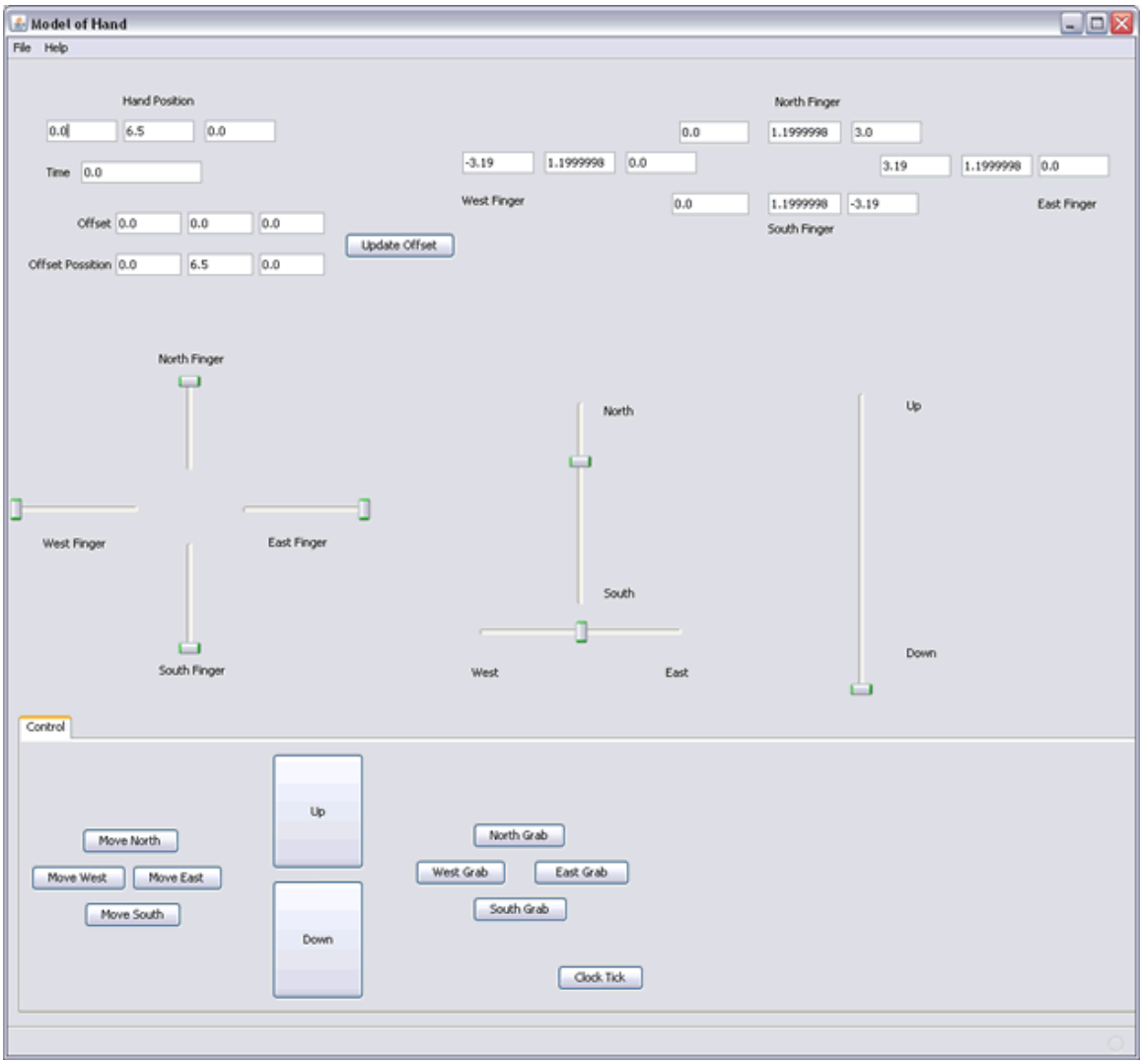


Figure 5.19 An expanded model

The same equations that predicted motion of the hands various actuators may also be employed to give the hand a good idea where in space its body parts are located. In the implementation above the positions of the fingers when viewed from above are represented by the position of the four sliders on the left. The sliders in the middle and on the right represent the long rail actuators used to move the hand. Once the hand has a sense of where its body parts are in space, it is able to experiment with various body

configurations and corridor sizes and learn that if its body is wider that the corridor, it won't fit. If it is narrower than the corridor width it will be able to fit.

Although width is sufficient to in this case to judge it was a simple matter to wrap all tracked points up within a convex hull. This gave a basic sense of were in space the agents body was at any given time. This process also readily leant itself to including a tool as well extending the body schema to incorporate the tool as seen below.

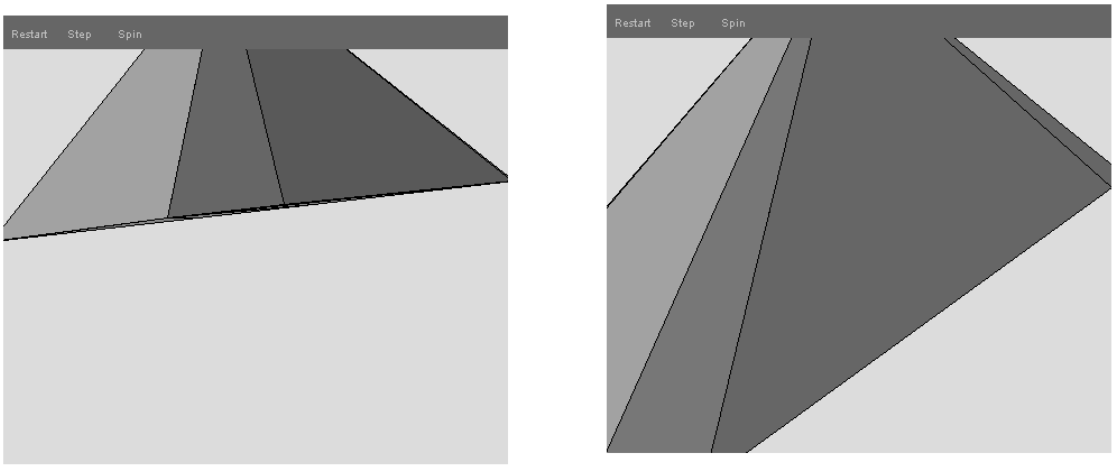


Figure 5.20 Two convex hulls

The left side of the figure above shows a basic convex hull created from the hand tracking points. The hull shown on the right has been extended to also encompass the additional volume of a tool held in the hand. It has been rotated to point slightly downward so that the additional length may be seen. As long as these hulls are made from all of the outermost vertices of objects, the hulls will completely contain all parts of the objects with the hull.

Although the convex would give a good rough implementation of fit, they were not implemented beyond a basic level, because other members of the research group are implementing a far more complex symmetry based algorithm that will match tools and openings quite precisely. In practice it was found that a simple bounding box would give rough predictions of fit that would only be improved by the using the hulls if the narrow end of an object could fit in a hole that the wider end could not. It practice bounding boxes were sufficient to demonstrate all the examples of tool use presented in the next section, if the tip of the tool was evaluated independently. Mathematically it was possible to take planar projections of tool tips and then place them in a bounding box. If this bounding box could fit in the opening, then the tip of the tool could fit in that orientation. This process was only implemented to show that it was sufficient, since the more complex symmetry based matching approach will eventually give far accurate evaluations of a tool's effectiveness.

5.6 Demonstrating Tool Use

With all of the infrastructure described in the previous sections in place, the hand could readily solve many tool using problems. The hand agent could solve the following tool using tasks using only the following capabilities:

1. An ability to babble and ground basic movements in a spatial ontology.

2. A collection of strategies for actions on objects that had been learned by the hand alone through babbling followed by clustering followed by representation as action ontologies.
3. A very basic extensible body schema that could create bounding boxes for different hand and tool configurations.

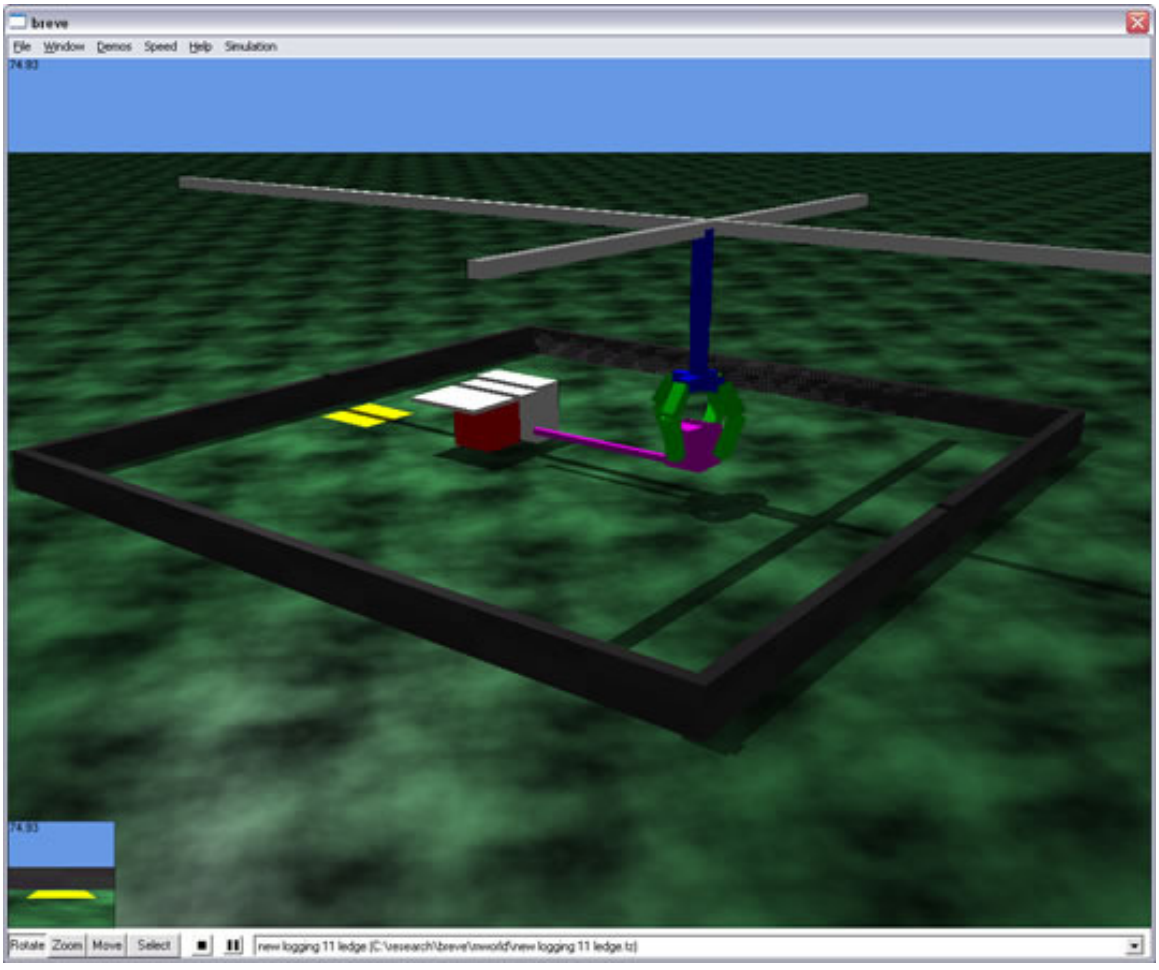


Figure 5.21 The cube is under a ledge

In this scenario the cube was placed under a small ledge. The hand alone was unable to reach the cube under this ledge. It was possible to compare the size of the ledge opening to the size of the tool end and recognize that the tool could fit under the ledge. It was then possible to push the cube with the tool onto the goal plate.

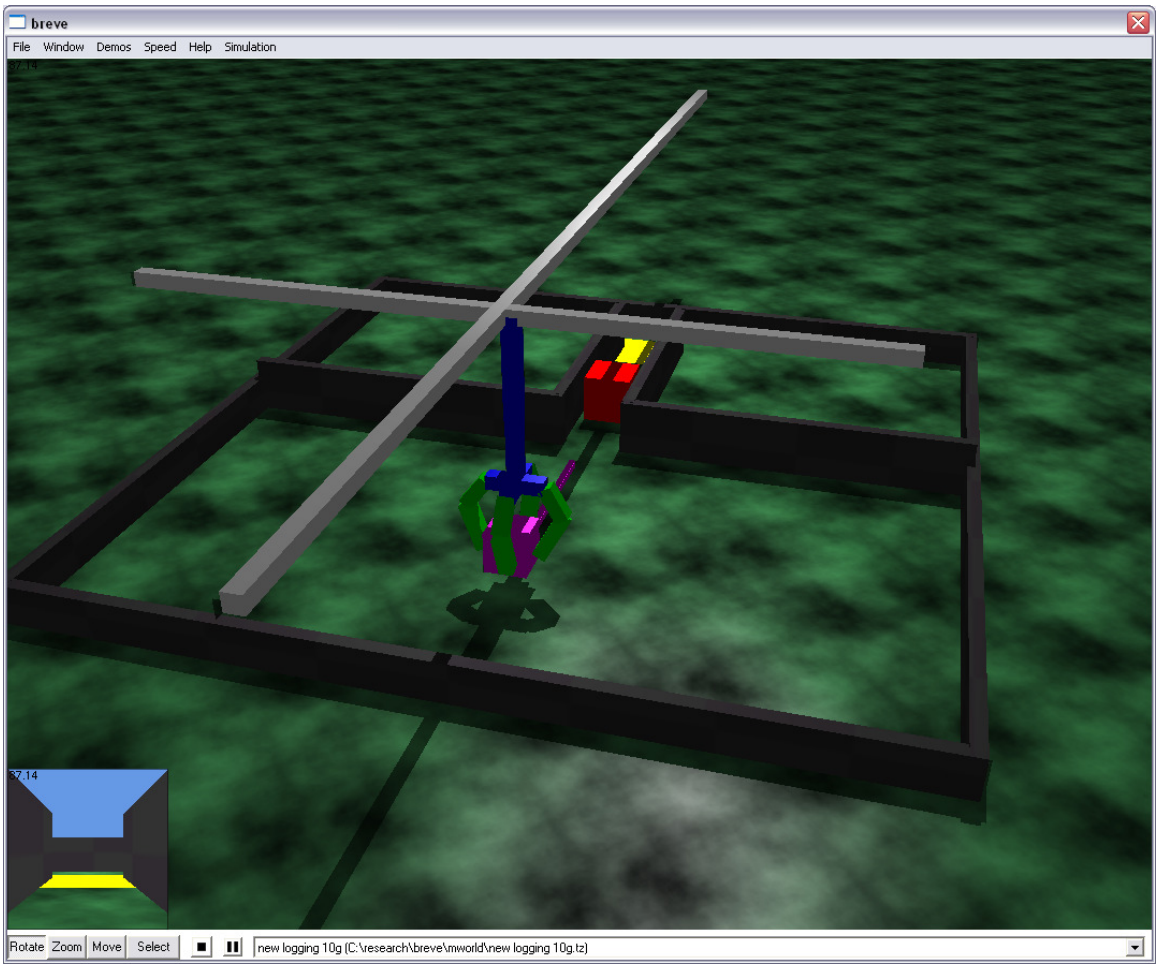


Figure 5.22 The cube in a narrow corridor

With a basic body schema, the hand could successfully predict whether it was able to fit in corridors of various sizes. In cases where the corridor was too narrow for the hand to fit, it was possible to model both a tools size and reach with bounding boxes, thereby

allowing an accurate prediction of a tool's ability to solve the problem. If a tool was properly sized, the same push cube north would solve the problem.

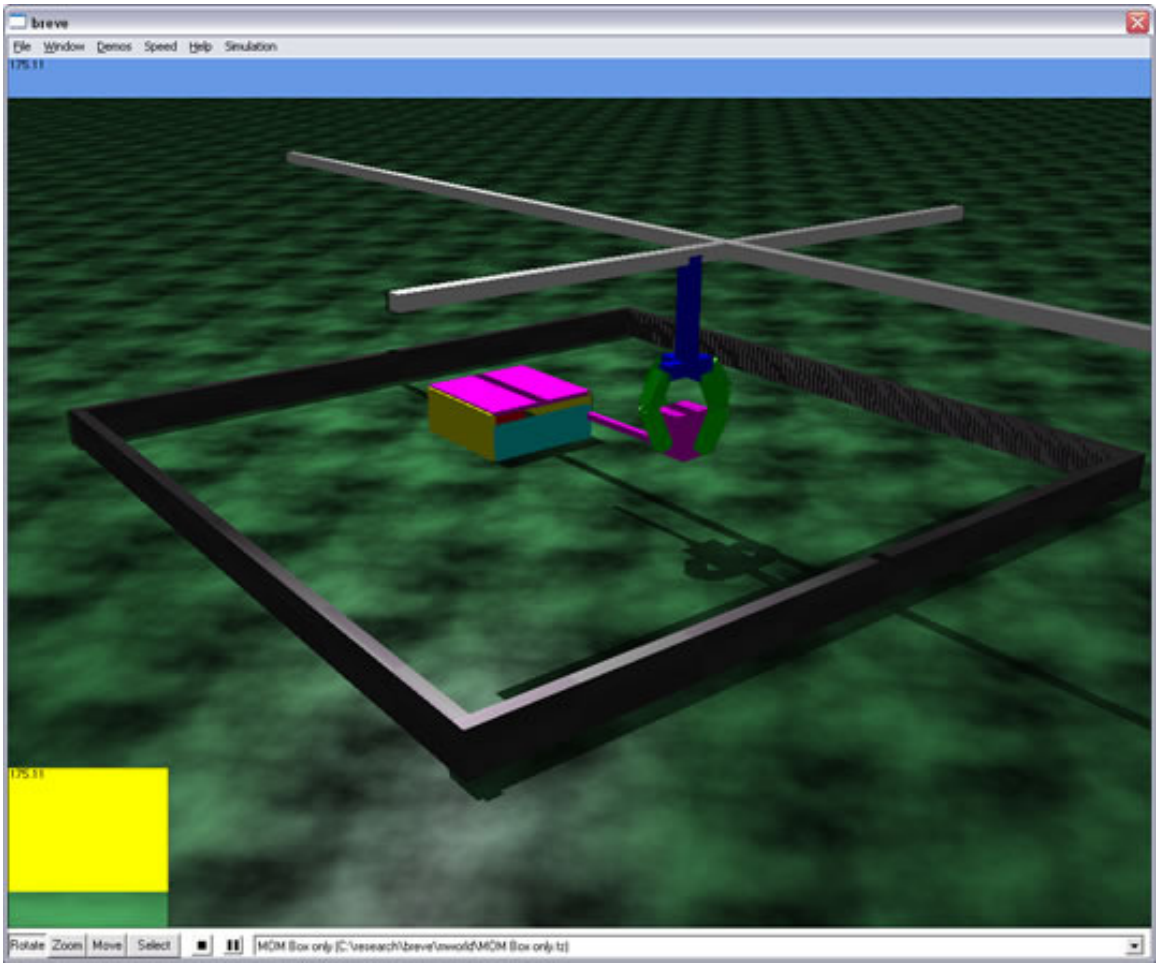


Figure 5.23 A simple box

In this problem a simple box is modeled in the simulation with a narrow opening. If the system had access to the size of the opening it could predict that the tool would fit in the opening. Once the tool was placed in the opening, several strategies were found to be effective in opening the box.

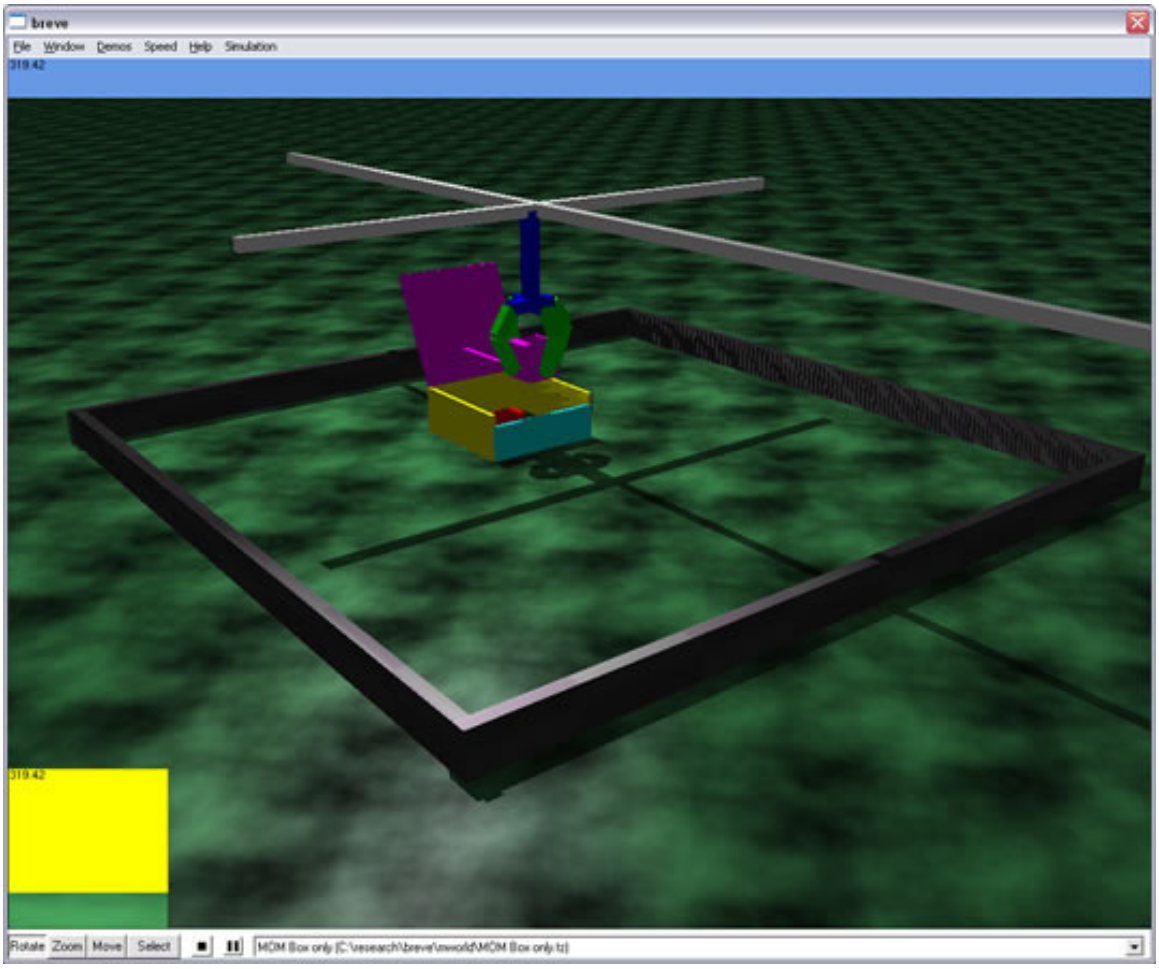


Figure 5.24 The box after opening with the tool

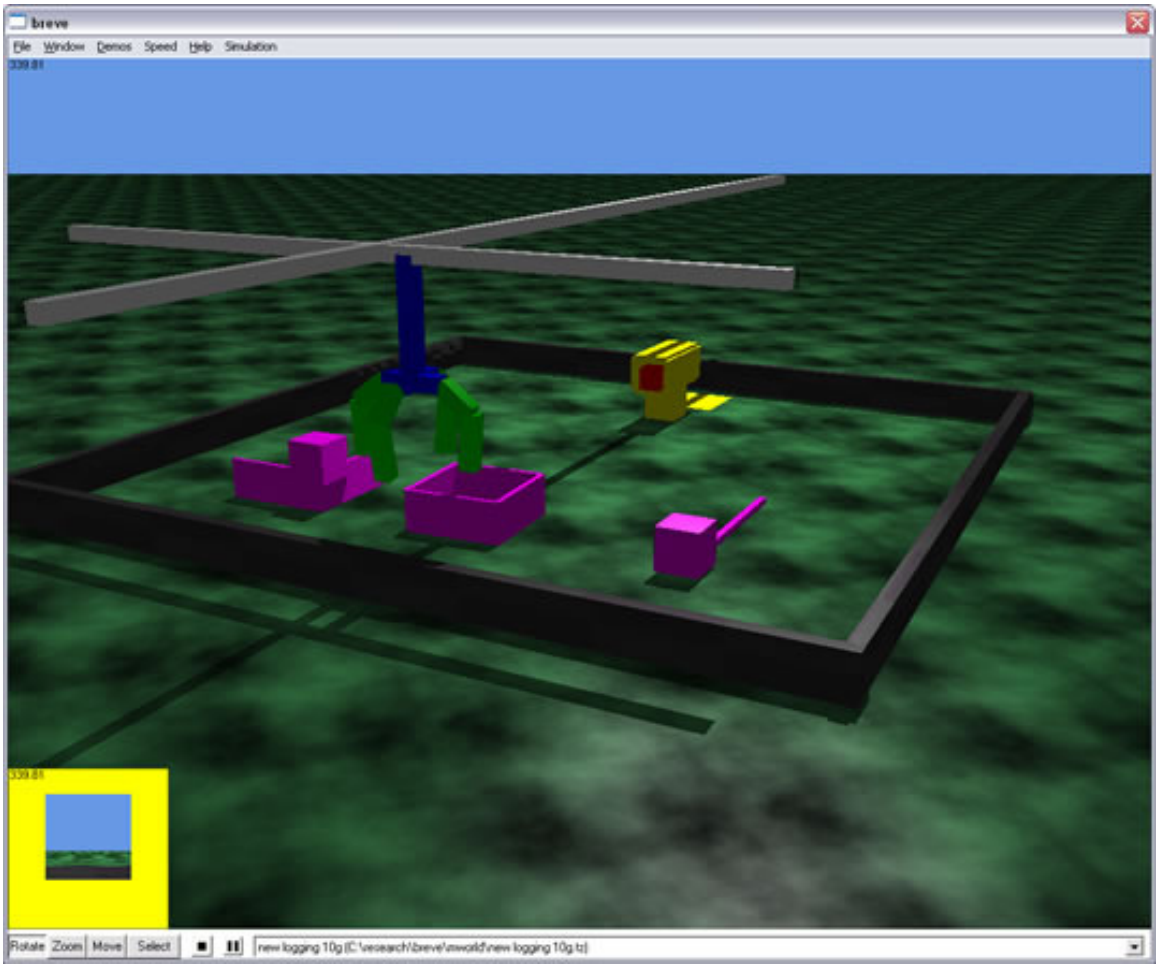


Figure 5.25 A food tube problem with three tools

This problem is similar to one seen in many animal studies where food was placed in a tube that was too small for a hand to fit. This is considered a fairly complex example of tool use and many monkeys cannot solve it. The system readily solves the problem. Comparison with the body schema indicates that the hand is too large for the opening in the tube. It is then possible to examine bounding boxes created from planar projections of each tool and identify the rightmost tool and the only tool that can fit in the tube. The learned strategy for pushing the cube may then be employed pushing the cube though the tube and solving the problem.

Chapter 6

Conclusions

Initially, it was claimed that we would use the environment and its agents to demonstrate four things. First, we claimed that **g**iven a minimal spatial ontology and a relatively simple clustering algorithm, the two agents were capable of learning predictive linear models of the relationships between their motor signals and effects in the environment, based on data observed during a babbling process. This was accomplished successfully with models generated that accurately predicted the effects of motor signals. These models served as the bases for creating behaviors that were semantically grounded in the ontology.

Semantically grounding the agent's actions in the ontology allowed us to accomplish our second claim that if the ontologies of the two agents are aligned then each agent becomes

capable of a simple form of imitation, even though the relationships to motor effects differ between the agents because of their different embodiments. Once move north, south, etc. was meaningful to both the hand and the vehicle imitation was straightforward.

Our third claim was that one agent can imitate another to achieve goals it has not experienced itself. The vehicle demonstrated this ability when it was able to successfully solve problems using strategies learned by the hand. Our fourth and final claim was that the hand agent could predict the effects of its use of a tool, using its ontology as a bridge. This was demonstrated when the same strategies that had been imitated by the vehicle, were imitated by the tool as well.

Ultimately all these achievements were aimed at establishing our main claim that imitative mechanisms would give rise to tool using behaviors. The work described in chapter 5 does indeed demonstrate that an artificial agent, with an imitative mechanism is able to develop sufficient knowledge to allow it to achieve impressively complex examples of tool use.

Ultimately it was surprising the degree of task complexity that could result from a very simple ontology. At first it was thought that to achieve tool use a very large and complex ontology would be required. In the end, complex behaviors emerged with no need to extend the ontology past basic representations.

Evaluating any ontology is difficult. For example, in a study of methods for ontology evaluation in industrial applications, Hartman et al. (2005) write, "No industrial method currently exists for helping an ontology engineer to evaluate and select ontologies that best matches his/her needs. Even in academia, few methods have been proposed..."

Brank has surveyed techniques used to evaluate ontologies and identified the following four categories into which evaluation strategies for ontologies may be classified (Brank et al. 2005):

1) Comparison of ontology to a "golden standard"
2) Employment of ontology in a particular application and subsequent evaluation of results
3) Comparison with a corpus of source data from the domain described
4) Evaluation by human subjects who try to determine how well the ontology meets some set of predefined criteria, standards, requirements etc.

Three critical questions can be posed for an ontology designed for an autonomous embodied agent in AI research: What does the ontology enable the agent to do, in its reasoning or its behavior? Does the ontology meet criteria for being well-designed, from a knowledge representation perspective? How does the ontology compare with what is known about human cognitive representation, to the extent that this can be inferred?

These questions support three avenues for examining an ontology: consistency, efficacy and anthropomorphic similarity.

The most straightforward test of an ontology is one for soundness and consistency. The algorithms to evaluate an ontology to identify and eliminate inconsistencies are well known and may be applied throughout the creation of our ontologies. There are mature software packages such as OntoClean to evaluate the adequacy of ontologies (Guarino, Welty 2004). There are several description logic reasoners available which are capable of evaluating ontologies (Sim et al. 2006). These tools are mature and aid in the construction of our ontologies, but in terms of scientific interest, their use is too well established to deserve extensive attention. In the end the ontology employed was sufficiently simple that it could easily adhere to these types of criteria.

The second possible test of the ontology is one of efficacy for a particular application. St. Amant et al. have defined a reference set of tool using tasks for evaluation, based on an analysis of the animal tool use literature (St Amant, Horton 2008). This set of tasks can be viewed as being analogous to the NIST standard test bed for urban search and rescue (USAR).

Initially, it was thought that performance on a reference tool set would be one of the best evaluation methods for this research. However as the problem was examined at deeper levels it became clear we could create tool using agents and have only succeeded in

surface solutions to the problems. There exist many examples of quite complicated tool use by factory robots. A modern robotic arm could use almost any tool. While solving the problems necessary to make a robot use a particular tool is an interesting and worthwhile pursuit, the problems that arise are engineering problems where a programmer will hard code a solution.

What emerged as more interesting than succeeding in a reference set of tasks was demonstrating that a certain imitation based approach to learning could lead to tool use. This was interesting not because it demonstrated that a robot could use tools, but because it demonstrated that imitation methods could be employed to generate sufficient knowledge to carry out the tool use. There is still much debate about how tool using behaviors arise in some animals. By creating a system that achieves fairly high level tool use primarily through imitation, this research has demonstrated that imitative mechanisms are sufficient to account for many example of tool use.

The third possible test for an ontology is anthropomorphic similarity. There are two main arguments that support the anthropomorphic similarity of our system. Firstly, the approach to representing action was chosen in large part due to overlap with some properties seen in mirror neurons. The resulting change in an object's properties, is an inherent part of the action ontology representation. The importance of the results of an action in its representation fits well with theories by Gallese (2003) that point to mirror neurons as the basis for a type of goal-directed behavior model. Ultimately the system

as a whole demonstrated an ability to recognize overlap between the hand pushing the cube, the vehicle pushing the cube and a tool pushing the cube. This is very similar to mirror neurons firing when a particular action is carried out, observed or accomplished with a tool. A second argument for anthropomorphic similarity comes from the approach taken. By building up the system's capabilities though the same set of stages that Meltzoff has identified in humans it is possible to claim a degree of similarity to humans.

There are several paths of further research that suggest themselves for continuing this research. One concerns extending the development of the system to account for Meltzoff's fourth stage of human imitation. This stage concerns imitation based on inferring intensions of actions. Consider that the system represented actions as a tuple, consisting of the expected result of an action, the context where the action would occur, and the motor signal. To accomplish basic imitation, the context and the motor signal were mimicked. It would be interesting to explore if the learned contexts could be used in some manner to guess what result is trying to be obtained.

For example consider the example of the hand pushing the cube to the north. To do so the system learned that the hand should be **centered up/down**, **centered east/west**, **south** and then sent a certain motor signal**.** It is easy to imagine that the hand might be in this configuration and fail to successfully push the cube due to friction or the cube flipping over. Even if the hand failed in the task would it be possible for the system to infer what it was trying to do?

Another interesting path of research would involve taking the knowledge that was learned in the simulation and using it with a real robot. It would be interesting to see the extent to which knowledge learned in simulation transferred to a real world robot. If it could be demonstrated that some of the strategies learned by the system in the simulation could work with a real world robot, it could strengthen some of the claims made in this research.

References

Acosta-Calderon, C.A. & Hu, H. 2005, "Robot imitation: Body schema and body percept", *Applied Bionics and Biomechanics,* vol. 2, no. 3-4, pp. 131-148.

Arkin, 1998, *An Behavior-based Robotics,* .

Ballard, D.H., Hayhoe, M.M., Pook, P.K. & University of Rochester. National Resource Laboratory for the Study of Brain and Behaviour. 1995, *Deictic codes for the embodiment of cognition,* University of Rochester, New York.

Barsalou, L.W., Kyle Simmons, W., Barbey, A.K. & Wilson, C.D. 2003, "Grounding conceptual knowledge in modality-specific systems", *Trends in Cognitive Sciences,* vol. 7, no. 2, pp. 84-91.

Bateman, J. & Farrar, S. 2005, *Modelling Models of Robot Navigation Using Formal Spatial Ontology*.

Beck, B.B. 1980, *Animal tool behavior : the use and manufacture of tools by animals,* Garland STPM Pub., New York.

Boesch, C. & Tomasello, M. 1998, "Chimpanzee and Human Cultures", *Current anthropology,* vol. 39, no. 5, pp. 591-614.

Boinski, S. 1988, *Use of a club by a wild white-faced capuchin (Cebus capucinus) to attack a venomous snake (Bothrops asper)*.

Britten, R.J. 2002, "Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels", *Proceedings of the National Academy of Sciences,* vol. 99, no. 21, pp. 13633-13635.

Brockmann, H.J. 1985, "Tool use in digger wasps (Hymenoptera: Sphecinae)", *Psyche: A Journal of Entomology,* vol. 92, no. 2-3, pp. 309-330.

Brooks, R.A. 1991, "New approaches to robotics", *Science,* vol. 253, no. 5025, pp. 1227.

Brooks, R.A. 1999, *Cambrian intelligence: the early history of the new AI,* MIT Press, Cambridge, MA, USA.

Bruner, J.S. 1956, *A study of thinking,* Wiley, New York.

Cardoso, J. 2007, *Semantic Web services : theory, tools, and applications,* Information Science Reference, Hershey, PA.

Chevalier-Skolnikoff, 1990, *Tool use by wild cebus monkeys at Santa Rosa National Park, Costa Rica*.

Clancey, W.J. 1993, "Situated action: A neuropsychological interpretation response to Vera and Simon", *Cognitive Science: A Multidisciplinary Journal,* vol. 17, no. 1, pp. 87-116.

Clancey, W.J. 1997, *Situated Cognition: On Human Knowledge and Computer Representations,* Cambridge University Press, New York, NY, USA.

Clark, A. 1998, "Embodied, situated, and distributed cognition", *A companion to cognitive science,* , pp. 506-517.

Clark, K.L. 2007, "Ontology schema for an agent belief store", *International journal of human-computer studies,* [Online], vol. 65, no. 7, pp. 640.

Davis, R., Shrobe, H. & Szolovits, P. 1993, "What Is a Knowledge Representation?", *AI Magazine,* vol. 14, no. 1, pp. 17-33.

Douglas-Hamilton, I. & Douglas-Hamilton, O. 1975, *Among the elephants,* Viking Press, New York.

Dretske, 1988, *Explaining behavior,* .

Duerden, J. 1905, "On the habits and reactions of crabs bearing actinians in their claws.", *Proceedings of the Zoological Society of London,* vol. 2, pp. 494-511.

Eric Maillot, N. & Thonnat, M. 2008, "Ontology based complex object recognition", *Image and Vision Computing,* vol. 26, no. 1, pp. 102-113.

Fanelli, L., Farinelli, A., Iocchi, L., Nardi, D. & Settembre, G.P. 2006, "Ontology-based coalition formation in heterogeneous MRS", *PCAR '06: Proceedings of the 2006 international symposium on Practical cognitive agents and robots*ACM, New York, NY, USA, pp. 105.

Fernandes, 1991, *Tool use and predation of oysters(Crassostrea rhizophorae) by the tufted capuchin, Cebus apella apella, in brackish water mangrove swamp.*

Ferrari, 2005, "Mirror Neurons Responding to Observation of Actions Made with Tools in Monkey Ventral Premotor Cortex", *Journal of cognitive neuroscience,* [Online], vol. 17, no. 2, pp. 212.

Gallese, 2003, "Motor ontology: the representational reality of goals, actions and selves", *Philosophical Psychology,* [Online], vol. 16, no. 3, pp. 365.

Gentilucci, M. 1988, "Functional organization of inferior area 6 in the macaque monkey. I. Somatotopy and the control of proximal movements.", *Experimental brain research,* [Online], vol. 71, no. 3, pp. 475.

Gifford, E. 1919, "Field notes on the land birds of the Galapagos Islands and of Cococ Islands.", *Proceedings of California Academy of Sciences*California Academy of Sciences, , pp. 189.

Goodall, J. 1963, "Tool-using and aimed throwing in a community of free-living chimanzees.", *Nature,* vol. 201, pp. 1264-1266.

Greeno, J.G. 1994, "Gibson's affordances", *Psychological review,* vol. 101, no. 2, pp. 336-342.

Gruber, T.R. 1993, "A translation approach to portable ontology specifications", *Knowledge Acquisition,* vol. 5, no. 2, pp. 199-220.

Guarino, N. & Welty, C. 2004, "An overview of OntoClean", *Handbook on Ontologies,* , pp. 151–159.

Guizzardi, G. 2005, *Ontological Foundations for Structural Conceptual Models*.

Harnad, S. 1990, "The symbol grounding problem", *Physica D: Nonlinear Phenomena,* vol. 42, no. 1-3, pp. 335-346.

Hart, B. & Hart, L. 2001, *Cognitive behaviour in Asian elephants: use and modification of branches for fly switching*.

Havas, C., Larue, O. & Camus, M. 2008, "Knowledge organization with pattern recognition in an auto-adaptive system", *7th WSEAS International Conference of Artificial Intelligence, Knowledge Engineering and Data Bases*.

Hayes, K.J. & Hayes, C. 1951, "The Intellectual Development of a Home-Raised Chimpanzee", *Proceedings of the American Philosophical Society,* vol. 95, no. 2, pp. 105-109.

Hepp-Reymond, M.C. 1994, "Force-related neuronal activity in two regions of the primate ventral premotor cortex.", *Canadian journal of physiology and pharmacology,* [Online], vol. 72, no. 5, pp. 571.

Hunt, G.R. 1996, *Manufacture and use of hook-tools by New Caledonian crows*.

Inoue-Nakamura, N.M.,T. 1997, *Development of stone tool use by wild chimpanzees (Pan troglodytes).*

Janes, S. 1976, "The apparent use of rocks by a raven in nest defense.", *Condor,* vol. 78, no. 78, pp. 409.

Kass, R.E. & Wasserman, L. 1995, "A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion", *Journal of the American Statistical Association,* , pp. 928-934.

Kirsh, D. 1995, *The intelligent use of space*.

Krutzen, M., Mann, J., Heithaus, M.R., Connor, R.C., Bejder, L. & Sherwin, W.B. 2005, "Cultural transmission of tool use in bottlenose dolphins", *Proceedings of the National Academy of Sciences,* vol. 102, no. 25, pp. 8939-8943.

Kuipers, B. 2000, "The Spatial Semantic Hierarchy", *Artificial Intelligence,* vol. 119, no. 1-2, pp. 191-233.

Kurata, K. 1986, "Premotor cortex neurons in macaques: activity before distal and proximal forelimb movements.", *The journal of neuroscience,* [Online], vol. 6, no. 2, pp. 403.

Lubbock, 1889, *On the Senses, Instincts, and Intelligence of Animals: With Special Reference to Insects,* .

Maravita, A., Spence, C. & Driver, J. 2003, "Multisensory integration and the body schema: close to hand and within reach", *Current Biology,* vol. 13, no. 13, pp. R531-R539.

Matelli, M. 1985, "Patterns of cytochrome oxidase activity in the frontal agranular cortex of the macaque monkey.", *Behavioural Brain Research,* [Online], vol. 18, no. 2, pp. 125.

McCarthy, J.L. 1980, "Circumscription - A Form of Non-Monotonic Reasoning", *Artificial Intelligence,* vol. 13, no. 1-2, pp. 27-39.

Meltzoff, A.N. 1997, "Explaining Facial Imitation: A Theoretical Model", *Early development parenting,* [Online], vol. 6, no. 3-4, pp. 179.

Meltzoff, A.N. 1995, "Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children", *Developmental psychology,* vol. 31, no. 5,838-850.

Meltzoff, A.N. & Moore, M.K. 2002, "Imitation, memory, and the representation of persons,", *Infant Behavior and Development,* vol. 25, no. 1, pp. 39-61.

Meltzoff, A.N. & Moore, M.K. 1977, "Imitation of Facial and Manual Gestures by Human Neonates", *Science,* vol. 198, no. 4312, pp. 75-78.

Meltzoff, A.N. & Prinz, W. 2002, *The imitative mind development, evolution, and brain bases.* Cambridge, U.K. Cambridge University Press 2002.

Metzinger, T. & Gallese, V. 2003, "The emergence of a shared action ontology: Building blocks for a theory", *Consciousness and Cognition,* vol. 12, no. 4, pp. 549-571.

Modayil, J. & Kuipers, B. 2007, "Autonomous Development of a Grounded Object Ontology by a Learning Robot", *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE,* vol. 22, no. 2, pp. 1095.

Ogden, C.K., Richards, I.A., Malinowski, B. & Crookshank, F.G. 1927, *The Meaning of Meaning: A Study of the Influence of Language Upon Thought and of the Science of Symbolism,* K. Paul Trench, Trubner & Co., ltd.

Peirce, C.S. 1931, *Collected Papers of Charles Sanders Peirce,* Harvard University Press.

Pelleg, D. & Moore, A.W. 2000, "X-means: Extending K-means with efficient estimation of the number of clusters", *Proceedings of the Seventeenth International Conference on Machine Learning table of contents*Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, , pp. 727.

Piaget, J. 1962, "Play, dreams, and imitation in childhood", .

Potter, E. 1970, "Anting in wild birds, its frequency and probable purpose.", *Auk.,* vol. 87, no. 87, pp. 692-713.

Povinelli, D.J. 2000, *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works,* Oxford University Press.

Rao, R.P.N., Shon, A.P. & Meltzoff, A.N. 2004, "A Bayesian Model of Imitation in Infants and Robots", *Imitation and Social Learning in Robots, Humans, and Animals,* .

Regier, 1996, *The Human Semantic Potential: Spatial Language and Constrained Connectionism, .*

Riedman, M.L., Staedier, M.M., Estes, J.A. & Hrabrich, B. 1989, "The transmission of individually distinctive foraging strategies from mother to offspring in sea otters (Enhydra lutris).", .

Rizzolatti, G. 1988, "Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements.", *Experimental brain research,* [Online], vol. 71, no. 3, pp. 491.

Rizzolatti, G. 1981, "Afferent properties of periarcuate neurons in macaque monkeys. II. Visual responses.", *Behavioural Brain Research,* [Online], vol. 2, no. 2, pp. 147.

Rizzolatti, G., Fogassi, L. & Gallese, V. 2000, "Cortical mechanisms subserving object grasping, action understanding, and imitation" in *The Cognitive Neurosciences III*, ed. M.S. Gazzaniga, Third edn, pp. 427-440.

Savage-Rumbaugh, S McDonald, K Sevcik, R A Hopkins, W D Rubert,E. 1986, *Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (Pan paniscus).*

Schaal, S. 1999, "Is imitation learning the route to humanoid robots?", *Trends in cognitive sciences,* vol. 3, no. 6, pp. 233-242.

Schwering, A. 2007, "Evaluation of a Semantic Similarity Measure for Natural Language Spatial Relations", *LECTURE NOTES IN COMPUTER SCIENCE,* vol. 4736, pp. 116.

Sharman, R., Kishore, R. & Ramesh, R. 2007, *Ontologies : a handbook of principles, concepts and applications in information systems,* Springer Science+Business Media, New York.

Sim, J., Kim, J., Kwon, O., Lee, S.S., Kim, J., Jang, H.K. & Lee, M. 2006, "Applying Inference Engine to Context-Aware Computing Services", *Proceedings of the Second International Workshop on Personalized Context Modeling and Management for UbiComp Applications, Orange County, CA,* .

Sowa, J.F. 1984, *Conceptual structures: information processing in mind and machine,* .

Stoytchev, A. 2005, "Behavior-Grounded Representation of Tool Affordances", , pp. 3060.

Stoytchev, A. 2003, *Computational Model for an Extendable Robot Body Schema*, Georgia Institute of Technology.

Tomasello, 1997, *Primate Cognition,* .

Toth, 1993, *Pan the tool-maker: Investigations into the stone tool-making and toolusing capabilities of a bonobo (Pan paniscus).*

van Lawick-Goodall, J. 1970, "Tool-using in primates and other vertebrates", *Advances in the Study of Behavior,* vol. 3, pp. 195-249.

Vogt, P. 2003, "Anchoring of semiotic symbols", *Robotics and Autonomous Systems,* vol. 43, no. 2-3, pp. 109-120.

Whiten, A. 2000, "Primate Culture and Social Learning", *Cognitive Science: A Multidisciplinary Journal,* vol. 24, no. 3, pp. 477.

Young, M.F. & McNeese, M.D. 1995, "A situated cognition approach to problem solving", *Local applications of the ecological approach to human-machine systems,* vol. 2.