A Behavior Analytic Paradigm for Adaptive Autonomous Agents

William R. Hutchison, Ph.D.
Seventh Generation Technologies

Behavior Analysis, the scientific study of animal and human behavior and learning described extensively on this Internet site (www.behavior.org), provides a strong conceptual framework for intelligent computer systems known as adaptive autonomous agents. Even as the weaknesses of existing approaches to computer agents are becoming more obvious, many trends in the field are moving toward the approaches in behavior analysis, including reinforcement learning, behavior-based robotics, vision (active, purposive, qualitative), empirical and grounded language, and others. The behavior analytic paradigm effectively addresses several of the key challenges of cognitive systems:

1. integration of language as the basis for cognition,
2. taking instruction, advice, and persuasion,
3. self-awareness, including reporting its own behavior, plans, and thoughts,
4. selective attention,
5. top-down contextual filtering based on understanding,
6. goal direction, and
7. dealing with surprising events

A key breakthrough, suggested by B.F. Skinner, is the notion that language is behavior in situations and is learned by the same processes as other behaviors. Cognition thus becomes an integral aspect of a consistent and relatively simple adaptive framework. The Seventh Generation Technology system, described in this paper, is an autonomous adaptive agent with cognition based on the behavior analytic paradigm.

Contents:

- Introduction
- Behavior Analysis: A Framework for Cognition in Agents
- Seventh Generation Technology: A New Paradigm for Autonomous Adaptive Agents
- The 7GT Agent as a Cognitive System
- Solving Key Challenges of Cognitive Systems
- Design Principles for Seventh Generation Technology Agents and Robots
- Core Technologies in the Seventh Generation Technology System
- Conclusion
- References and Other Publications and Presentations on 7GT
Introduction

This paper describes a paradigm for adaptive autonomous agents and robotics based on the scientific study of animal and human behavior and learning known as Behavior Analysis. This paradigm is not new as such, but its application to adaptive autonomous agents is new and revolutionary. This approach offers solutions to some of the key challenges of developing adaptive autonomous agents capable of behaving and learning in real time in complex environments.

Behavior Analysis: A Framework for Language and Cognition in Agents

Operant conditioning, a mode of learning that is characteristic of all mammals and most lower species, provides a framework for analyzing the acquisition of behavior, its control by features of the environment, and selection of behavior by consequent feedback. Thousands of researchers in organizations such as the Association for Behavior Analysis, the Society for the Quantitative Analysis of Behavior, and Division 25 of the American Psychological Association, have produced an extensive body of knowledge about behavior and learning (e.g., *The Journal of the Experimental Analysis of Behavior*, Honig & Staddon, 1977), with many important applications to human behavior (e.g., *Journal of Applied Behavior Analysis*, Cambridge Center for Behavioral Studies web site www.behavior.org).

The discipline of Behavior Analysis provides a framework not only for analyzing the acquisition of behavior, its control by features of the environment, and selection of behavior by consequent feedback; it also provides continuity of theory for the acquisition of verbal as well as nonverbal behavior. Language, both speaking and listening, is viewed as behavior in situations, learned by the same processes as other motor behaviors because of the value produced by those “verbal behaviors.” This view is completely consistent with the parsimonious framework of behavioral robotics that has produced rapid advances in nonverbal behaviors, but so far behavioral robotics theorists have not recognized the potential for extending this framework to language because it differs so radically from popular conceptions of language. This extension is the conceptual breakthrough that permits integration of language and cognition into adaptive autonomous agent systems in a consistent and relatively simple adaptive framework.

Acquisition of verbal behavior is based on theoretical analyses regarding the problems of functional language usage outlined in 1957 by B. F. Skinner in *Verbal Behavior*. The key assumption of Skinner’s “radical behaviorist” theory is that verbal behavior is not fundamentally different from nonverbal behavior. Most researchers in the area of intelligent agents have ignored this analysis, however, because linguistics theorists rejected it (Chomsky, 1959; Harris, 1993; Pinker, 1994). The issues in this debate are discussed extensively elsewhere on this Internet site (www.behavior.org), and make apparent that most criticisms by linguists are not based on actual knowledge of Skinner's analysis, but by what they have heard about it, primarily Chomsky's (1959) review. Anyone who is interested in addressing these important issues will appreciate the papers
available in David Palmer's subsection on this site. As a starting point, MacCorquodale's (1970) rebuttal of Chomsky's review documents how Chomsky grossly distorted Skinner's position in order to make it vulnerable to his own self-indulgent rhetoric. It is also important to note that Skinner's analysis appeared implausible at the time it was published because it seemed absurd to claim that learning in a network of input-output relationships (operant learning) could account for behavior of any complexity. However, the subsequent invention of modern neural networks and their successful application to a wide variety of problems has proven that Skinner was prescient in developing a remarkably detailed and disciplined framework for language learning in a neural network-based agent.

Even as the weaknesses of existing approaches to computer agents are becoming more obvious, many trends in the field are moving toward the approaches in behavior analysis, including reinforcement learning, behavior-based robotics, vision (active, purposive, qualitative), empirical and grounded language, and others. For example, there is increasing recognition of the need for “grounding” language in environmental experience (Roy & Pentland, 2002; St.John, 1992), but this is generally accomplished with ad hoc special mechanisms rather than within the consistent, parsimonious framework provided by behavior analysis.

Seventh Generation Technology: A New Paradigm for Autonomous Adaptive Agents

The Seventh Generation Technology (7GT) architecture was developed nearly 20 years ago based on general principles derived from Behavior Analysis (e.g., Hutchison, 1984, 1998). Quantitative formulations from Behavior Analysis (these are extensive, but mostly from research in the Journal of the Experimental Analysis of Behavior, the annual Advances in Behavioral Economics, and proceedings of annual meetings of the Society for the Quantitative Analysis of Behavior) provide the scientific foundation for the core elements of the 7GT design.

The 7GT system is based on an artificial neural network of the broad category known as reinforcement learning (Sutton & Barto, 1998), a well-known model for adaptive autonomous agents. Adaptive agents based on reinforcement learning learn by direct contact with environmental contingencies having three elements:

- the context or situation (input pattern),
- the behavior of the agent (actions), and
- the consequences (reinforcement or punishment).
After a behavior has been performed, the connections between the sensory inputs and that behavior are strengthened if the consequence was positive and weakened if the consequence was negative. To train a robotic agent by direct interaction, the agent is repeatedly placed in the target situations, and the agent's behaviors are observed. The consequences are delivered to the agent (if the environment does not deliver them automatically). Complexity of the learned behaviors is limited, however, due to the sheer number of interactions with the environment needed to learn through directly experienced environmental contingencies. Learning in humans is significantly more efficient than machine learning because, after infancy, much learning is mediated by language.

A major innovation of the 7GT system is its capacity to learn verbal as well as nonverbal behavior, which vastly extends the effectiveness of traditional reinforcement learning and behavior-based approaches. By seamlessly incorporating verbal behavior (language) into the set of available responses whose relations to the environment are learned through interactions with the environment experienced by the agent (Hutchison & Stephens, 1987), the 7GT system successfully implements the approach to language that has been studied for decades within the framework of Behavior Analysis. 7GT's unique implementation of language within a neural network rather than in a separate component (e.g., a hypothetical Language Acquisition Device) creates a versatile agent without any need for separate modules to implement symbolic modeling and planning.

The overall design of the 7GT system has sensors and actuators connected by a neural network, with a primary value function guiding the reinforcement learning process. The objective of the system is to optimize the primary value function over time through continuously learning how to behave in an environment, which may be physical or electronic. Inputs may include verbal advice or information from sources of varying reliability as well as direct or preprocessed environmental inputs. In addition to sensors for external stimuli, there are recurrent connections from responses back to sensory input, state sensors for body states including proprioceptive input and primary value deprivations (e.g., hunger), mechanisms which retain sensory inputs several time steps, and a temporal-differences type learning algorithm to enable higher-order conditioning. Most of these elements—except the sensors for primary values and change in situation value—have been described by other researchers, but 7GT researchers believe that all of them are required together to produce language and cognitive phenomena. Numerous demonstrations using the 7GT system to train language functions provide sufficiency proofs of the particular analyses used and greatly strengthen the case for the overall approach.

The 7GT Agent as a Cognitive System

An agent that learns language and can use it, not only to communicate with humans but to mediate its own "thinking," is capable of complex cognitive behavior unknown in current robots. Cognition is increasingly being defined by exclusion: whatever cannot be accomplished by "reactive behavior", where sensory inputs directly control the agent's actions. Since reactive behavior is a subset of discriminated operant behavior, cognitive capabilities seem by definition to require more than operant learning. However, this is a
simplistic assumption repudiated by Skinner's analysis of verbal behavior. An operant-based agent with verbal behavior can truly be called a cognitive system. It can:

- Learn to communicate with human users as well as with other agents,
- Learn verbally from others (very efficient; huge amount of verbal information available),
- Learn immediately from advice rather than needing many trials (and many errors); no need to construct a set of examples to train,
- Learn to discriminate reliable from unreliable sources of information by domain,
- Learn at the appropriate level of abstraction,
- Say what it knows (to explain, debate, refine),
- Learn rules as well as any kind of relations, even relations that are so remote in time or space that they would never be learned by direct experience,
- Solve novel problems by transforming into different problem or subproblems for which solutions are known,
- Solve problems by recombining responses,
- Perform verbal self-management behaviors,
- Create a verbal "model" of its environment (words, graphics, programs).

Solving Key Challenges of Cognitive Systems

The behavior analytic paradigm for autonomous agent design provides strong guidance in addressing seven key challenges of cognitive systems. These capabilities are not ad hoc mechanisms, but are inherent, emergent properties of this approach to intelligent systems.

1. **Integration of language as the basis for cognition.** Skinner's analysis gives explicit guidance for how an adaptive autonomous agent can acquire language that is functionally linked to its sensory, action, and learning systems. This integration has been implemented and tested extensively in the 7GT system.

2. **Taking instruction, advice, and persuasion.** It is not difficult to implement a system in which the agent complies with a discrete instruction from a user. However, many situations do not call for such simple instruction following, but rather a complex combination of verbal and nonverbal control. For example, a tennis coach may instruct a student to “put more spin on the ball”, but the instruction must combine in its effect with extremely complex, learned sensory-to-motor relations that control ball hitting apart from the instruction. In addition, the terms “advice and persuasion” do not imply simple compliance, but rather having an additive influence that may not result in compliance in some cases. The agent may not comply if its own experience is superior to the advice giver or if advice from that source in the past has been bad. The 7GT system is able to behave like humans in these cases because it learns to utilize instruction and advice in the same way.
3. **Self-awareness, including reporting its own behavior, plans, and thoughts.**
   The agent continually senses its own behavior as well as many internal states. It can learn to describe and report these behaviors and states in the same way it learns to describe and report external events and states.

4. **Selective attention.** Behavior analytic theory identifies two parts of the problem of selective attention.
   a. **Passive selective attention, known as “stimulus control.”** When an agent senses a scene that contains many elements competing for attention, it must have a basis for attending to the most important ones. As described below in the third item of “Design Principles”, the 7GT agent will automatically attend to the elements that have proven to be most valuable during its training, because the connections from those patterns of sensory elements have been strengthened the most. The agent may respond directly to the sensory elements or it may emit intermediate responses to orient toward them or engage in verbal behavior about them.

   b. **Active selective attention.** A process for active selective attention has been developed as an extension of behavior analytic theory, described in U.S. Patent # 6366896. The agent can be verbally instructed to “Find X” (e.g., a street sign), and this verbal instruction hypersensitizes the sensory systems to detect the target among other elements. The general process also enables a number of other valuable cognitive processes based on sameness, difference, and other relations.

5. **Top-down contextual filtering based on understanding.** This can be seen as a special case of selective attention, but it is important enough to deserve a separate description. Researchers in speech recognition have improved the accuracy of recognizing speech sounds by looking at the fit of candidate words with the meaning of surrounding words. The 7GT behavioral model automatically implements this “top-down contextual filtering” because it is actively responding to what it is hearing, i.e., it can fairly be said to understand the speech. The context provided by the previous words and the environmental situation that the agent shares with the speaker exert a strong influence through the network of associations to strengthen the correct word among competing words, even when the sounds alone are inadequate to discriminate among possible words. In the classic challenge of the “cocktail party conversation”, this context of understanding can provide the critical thread to enable picking one speaker's utterance from others. A very similar filtering process occurs in 7GT's vision, where the context provided by other objects that have been recognized by prior behaviors of the agent automatically exert a very helpful influence on recognizing an ambiguous object.

6. **Goal directed behavior.** Autonomous agents of any complexity should typically have more than one primary value/utility, because in addition to the practical goal of the agent (e.g., to collect ore samples), they must maintain their own critical resources (e.g., battery charge level, lubrication). Living organisms characteristically have primary value systems that attribute higher value to resources when their current level is low (e.g., hunger, thirst), a concept known in microeconomics as variable marginal utility. Giving higher value to activities that
restore dangerously low resources has obvious survival value, and can fairly readily be implemented in a reinforcement learning system such as 7GT. However, there is a critical additional mechanism that was missing in adaptive autonomous agents until implemented in 7GT based on behavior analysis (included in Hutchison, 2000). It is not enough that behavior that leads to recharging its battery would be more highly reinforced, because that mechanism has no influence until after the behavior has succeeded. Until that event, the fact that seeking power would be highly reinforced is unknown to the agent, and the agent could easily run completely out of power before it seeks power because it has no way to know it should be seeking power. Nearly as harmful is the fact that after succeeding, the impact of the huge reinforcement will be to cause the agent to increase its efforts to find more power even though it is no longer needed! Behavior analysis in the area known as establishing operations (e.g., Michael, 1982) asserts that deprivation has two effects, one to increase the value of the reinforcer and the other to act as a stimulus to evoke related behaviors. This theory suggested the patented innovation: give the agent a sensory input from the current levels of each primary value to its network. This connection has no a priori effect, but after a few learning experiences, deprivation of a primary value will immediately increase the relative priority of behaviors to restore it, and reducing deprivation decreases those behaviors. This extremely simple change leverages the power of the reinforcement learning process to allocate behavior among all the agent's goals in a nearly optimal way. The sensory connection also enables the agent to be aware of its deprivation, to report it, and to think about ways to remedy the deprivation, adding the potential for cognitive processes to supplement lower-level ones (e.g., to plan going to lunch rather than just getting a snack from the drawer).

7. **Dealing with surprising events.** Behavioral theory and research have actively addressed behavior in novel situations, where the subject does not have an adequate response. Much of this field can be broadly categorized as problem solving. There are two main subcases:
   a. Where the individual does not know specifically what to do, but it has learned to recognize the situation as being in a category that can be solved by intermediate responses that are known. For example, the agent may not know how to reply to “What is 23 times 117?”, but past training can enable it to break the problem into responses that are known (in this case, starting with 3 times 7). As another example, when the copier machine is not working, a learned problem solving response may be to look for a diagnostic screen or instruction manual.
   b. Where there is no response known to be effective. These situations require creativity, which can be described as performing a novel response or novel sequence of responses. One line of research has studied how organisms recombine previously learned responses into new, effective patterns to solve new problems (e.g., Epstein, 1993). A second important line of research (Neuringer, Deiss, & Olson, 2000) has demonstrated how learning not only strengthens or weakens responses in a situation but can also directly increase variability in the response system-in effect making
the individual more creative. The 7GT system has simulated a number of creative problem solving processes such as these.

Design Principles for Seventh Generation Technology
Agents and Robots

1. The 7GT researchers endorse Brooks' (1990) assertion that truly intelligent agents must be embodied, but for different reasons than his: In order to fulfill the requirements of cognition, an agent must understand what it is saying about its world (including itself). This understanding can only be acquired through direct experience with the world by sensing it and acting upon it.

2. Intelligent behavior must be in tune with its environment. In certain cases, this may be achieved by programming based on knowledge identified by humans. However, it is often more efficient, and sometimes essential (e.g., visual recognition), for the robot to learn or adapt through experience. Increasingly robots will be used in new applications that will require real time learning.

3. Sensors, actions, and values are inseparably linked. Actions are guided by sensor inputs via associations learned as a function of the value produced by the actions. For example, visual recognition is not solely in the visual system, it is implicit in connections to effective actions that were learned because they produced valuable results. Value selects what visual properties are learned. Such a paradigm elegantly solves selective-attention and action-selection problems that otherwise quickly become intractable in realistically complex environments.

4. Language, both speaking and listening, is behavior in situations. Conceptualizing language in this way elegantly solves the problem of integrating cognition into adaptive autonomous agents using functional language.

5. Genetic algorithms (GAs) give us an extremely powerful design tool for adaptive autonomous agents and their training, producing designs faster, cheaper, and better than even expert humans in most cases once the key parameters have been identified. They can evolve the structural, sensory, motor, and learning parameters of individual agents. GAs have a natural synergy with artificial learning agents just like natural evolution with animal learning. To reap advantages from GA technology, we must constantly validate that their fitness measures are predictive of robot performance in the real world.
Core Technologies in the Seventh Generation Technology System

1. The core adaptive behavior system is a reinforcement learning network first implemented in 1983 based on quantitative analyses from the Behavior Analysis. The main elements are covered in U.S. patents #5802506, 6038556, and 6366896. It includes
   - higher-order conditioning (cf. temporal differences)
   - recurrent sensing of its own prior responses
   - optional hidden layers
   - lagged input nodes
   - a sophisticated primary value system based on behavioral economics
   - a maturation mechanism by which certain movements are executable only after a specified age of the agent, thereby reducing the difficulty of complex problems such as speech development in its early stages.
   - sensors including:
     - a highly developed, GA-evolved sound input system capable of providing spectral and cepstral data sufficient for speech recognition by the agent
     - active color vision with GA-evolved, software-implemented multi-level foveation with edge and motion detection, with interfaces to input from standard Virtual Reality Modeling Language (VRML), our internally-developed simple VRML, stored digital images, stored digital video, or real-time camera input from Java Media Framework (JMF)-compatible devices
     - proprioception
     - touch
     - distance (directly from IR sensors)
     - primary value deprivation levels (e.g., battery charge)
   - A patented goal-direction mechanism by which an agent optimally allocates behavior proactively based on the current status of its primary reinforcers relative to the primary reinforcement function and immediate environmental opportunities.
   - Motor movements are actuated by outputs from the neural network. The movements can be grouped into sets of mutually incompatible movements, such as a "left motor" group including "left motor move forward," "left motor move back", and "left motor no movement." At a given time, one movement from each movement group can be executed simultaneously, enabling it, for example, to move and talk at the same time.
   - The agent's speech behavior can be output as parallel motor control of speech articulators (lips, tongue, glottis, velum, diaphragm), or for convenience in many cases, as phonemes or even words. Perhaps surprisingly, the agent can learn the series of articulatory movements nearly as quickly as higher-level units, but of course with somewhat less reliability (cf. humans!). When speech articulations are output by the
agent, they currently have their effect on the world by decoding in a lookup table and feeding phonemes to a speech synthesizer. However, we have plans to interface directly to our research version of an articulatory synthesizer from Trillium Sound Research, which produces speech by calculating the physics of sounds produced by a specified series of vocal tract positions (i.e., exactly what our agent outputs). Trillium is currently preparing to release the source code in a public license, giving us the necessary access to interface directly.

2. The body can be a range of robot platforms. We have used a laptop computer robot base developed for us by Acroname, Inc., of Boulder, Colorado, and we are considering other platforms with infrared (IR) and 802.11-based communications to a PC where most computing is done. The platforms support motor-driven wheels, touch sensors, IR distance sensors, and the capacity for a large number of other digital and analog inputs and outputs via Acroname's Brainstem board. The Brainstem allows implementation of autonomous behavior programs and reflexes, and it interfaces with the laptop via a Java serial port interface. Camera and microphone inputs connect directly to the laptop via USB and microphone connectors. This system is designed to be simulated by a VRML system (next item) in a way transparent to the behavior system.

3. Provision for the agent to live in a VRML environment, where robot and environment properties can be made as similar as possible to a real environment of interest. This VRML provision permits us to test designs quickly, speed up computerized training programs many-fold, and to permit the use of powerful GAs tools for designing the robot, brain, and training.

4. A set of highly-developed, automated training systems using virtual reality or stored images and sounds from the robot's environment. These training systems include:
   - A system for specifying very complex contingencies to control stimulus presentations and consequences for the agent, with special emphasis on contingencies that comprise training programs
   - A higher level system to manage more than one such set of contingencies at once
   - A variation of the training system that automatically generates, adapts, and manages key elements of the training program such as order of presentation, prompting, and amount of reinforcement for target behaviors.
   - A module that automatically constructs a training program to teach an agent to imitate all the phonemes in a given set of words, then to imitate the entire words.
   - A mechanism to train and evaluate agent behavior using static digital photos captured in the robot's actual environment, where movement commands sent by the agent change the location and field of view (FOV) of the retina within the larger stored image. The training system can systematically vary the starting location and FOV of the robot to simulate realistically a wide range of situations, as validated by corresponding live tests.
Proven language training methods for phoneme and natural language word imitation, naming, command following, some generalized grammar, generalized relationship naming, and simple rule following (some of these are covered by patent #6038556).

A specialized program to train speech imitation, where the agent hears real human speech and outputs motor responses to control all its speech articulators to make the same sounds. The ability of a neural network to exhibit babbling much like children is critical for learning imitation, but a key observation from human imitation learning provides another valuable process: rather than expecting the child to successfully imitate parents, parents often encourage and imitate the child when the child vocalizes a new phoneme. This creates the conditions for imitation learning that otherwise would occur far too rarely. In addition, the learning task is simplified by our human-inspired provision for the agent's articulatory movements to mature over time rather than all being available for learning at "birth." We are also looking at concepts of resonance and automatic reinforcement to more fully replicate the human speech imitation learning process.

A language teaching system (SpeechTeach) for the robot and for humans, extensively tested with our agent and scores of developmentally delayed children and adults. This software and media are being further developed and marketed, and will provide an intensive point of interaction for comparison of the study of language development in children and our robot. The staffing, software, and training media substantially overlap between human and robot research efforts.

5. A genetic algorithm that jointly evolves agent properties (structural, sensory, motor, learning) as well as the training program that will be administered to the agent prior to fitness testing.

Conclusion

A daptive autonomous agent designs based on the paradigm of Behavior Analysis solve many fundamental challenges of machine intelligence in an elegant and parsimonious way. These basic advances are an exciting opportunity for rapid progress toward more human-like computer agents and robots as well as better understanding of human behavior.

References and Other Publications and Presentations on 7GT


