Learning and Dynamics

Neil Berthier
Department of Psychology
University of Massachusetts

Andrew Barto and **Matthew Schlesinger**
Department of Computer Science
University of Massachusetts

Almost all theories of human development acknowledge some role for experience-dependent changes in behavior, representation, and thinking with age. Because of the legacy of antipathy toward the behaviorists, these experience-dependent changes have not always been described as learning. It is clear, however, that learning does play a central role in development and that learning in development is not limited to associative learning. In our studies of infant learning we have a special interest in how the mechanisms of learning interact with the demands of development.

We must first recognize that the learning mechanisms used during development are probably varied and dependent on the situational context. Some situations involve perceptual learning, others involve learning to solve problems, and still others primarily involve associative learning. As we expand our investigations of learning, we are likely to discover different ways children learn during development. A situation that we have studied is motor or action learning. In this case, the child is situated in a particular context and encouraged to make some movement or take some action. We have been particularly interested in how a child learns to control and exploit the dynamics of their body and the world to achieve their goals. A key difference between our view and the view of other investigators is that we hypothesize that learning is driven by the attainment of goals and that behavior is not simply emergent or self-organizing.

Action learning has typically been characterized as interactive or exploratory learning. These terms are often used descriptively and offer limited insight into the processes of learning. In the interactive or exploratory learning task, the child is placed in a situation where they are expected to solve the task or generate some behavior on the basis of interaction with the environment. Freedland and Bertenthal (Freedland & Bertenthal 1994) and Siegler (Siegler 1994) have suggested that these tasks involve a variation and selection approach, where the child generates various actions and then repeats actions that led to solutions. One important aspect of these actions is that they cannot be the result of random variation-there are far too many possible actions that could be considered. Somehow, the child is able to generate a small set of appropriate actions upon which selection can act (Siegler 1994).

A second way of theorizing about interactive learning has been presented by Dynamic Systems Theory (e.g., (Thelen & Smith 1994)). This view characterizes the child and the environment as a dynamical system that has several modes of operation. A major advantage of this view is its emphasis on the dynamics of the composite system of the actor and environment. These dynamics should not be seen as something to be overcome, but as something to be exploited. In learning to act, humans typically find solutions that exploit the basic dynamics of their situation and find low-energy solutions (Turvey & Fitzpatrick 1993).

Thelen and Smith (Thelen & Smith 1994) view development as a series of transitions between stable attractors, each representing modes or styles of action. Solutions are said to arise by self-organization, in much the same way that atmospheric events self-organize to produce global weather. The driving force in the generation of solutions is the underlying system's dynamics.

However, we believe that as important and meritorious as this view is, it misses an important aspect of infant development. If nothing else, work in developmental psychology over the last decades emphasizes the active nature of infant behavior. Infants are active problem-solvers engaged in a wide variety of goal-directed behavior ((Bruner 1973), (Gibson 1988), (Piaget 1952), (Piaget 1954),(Willats 1990)). They explore the dynamical properties of their bodies and their environments and discover how they can bring about and maintain interesting events, situations, or sensations (e.g., (Bower 1989), (Rovee-Collier 1987), (Watson 1972)). Infants from the earliest ages show behavior that is goal-directed and is organized to secure certain ends. This is not to say that all behavior that is goal-directed involves the *explicit* representation of a problem with a specified goal and a period of planning to obtain that goal. As Piaget's discussion of the primary circular reaction makes clear, behavior that leads to interesting results will be repeated without any explicit knowledge on the part of the infant of the connection between the behavior and the interesting result.

We believe that Dynamical Systems Theory could be substantially improved if the role of the infant as an active learner is more clearly emphasized. One way to accomplish this end involves the traditional concept of *control.* However, the views of control that seem most common are not appropriate and are often misleading.

Control

In the most familiar type of control, a controller wants system's output to stay close to a given set-point, or to closely track a given reference trajectory, in the face of disturbances. These are called regulation and tracking problems respectively. Most of control theory and practice focuses on these problems, where it is assumed that the set-point or reference trajectory is provided by some unspecified seat of intelligence. This assumption is one of the reasons that this type of control has limited relevance to development (Sporns & Edelman 1993). How could a developing infant obtain all the set-points and reference trajectories that it needs?

Another shortcoming of regulation and tracking control is that the desired behavior does not take the natural dynamics of the controlled system into account. Reference trajectories are typically created without regard for the system's dynamics, and the controller's job is to "force" the system to follow the given reference trajectory. The objective is for the controller to make system do what the controller wants the system to do, independently of what the system might want to do on its own. This is another reason that many researchers regard regulation and tracking control as having little relevance to biological motor control, where there is ample evidence that a system's natural dynamics are often exploited rather than suppressed (e.g., (Turvey & Fitzpatrick 1993)).

These are compelling reasons to reject regulation and tracking as particularly useful formalisms, or even metaphors, for motor development. However, control is much more general than regulation and tracking. Despite its preoccupation with these more common types of control problems, control theory addresses a wide variety of other types of problems and techniques. Most relevant to development, we argue here, are optimal control problems and the collection of methods that have been developed for solving them (or for approximating their solutions).

The objective of an optimal control problem is to arrive at a means of interacting with a dynamical system that is best-optimal-according to some given way of evaluating the system's behavior over time. One kind of optimal control problem, for example, requires interacting with a system to make it go from an initial state to a goal state in a way that maximizes a given objective function (or minimizes a given cost function). For instance, in placing a satellite into orbit, the objective would be to achieve the final desired orbit while minimizing the amount of fuel used to get there. In contrast to a tracking problem, where the desired trajectory is given, here the trajectory is part of the solution: it is the result of controlling the system in the optimal way; it is not an explicit part of the problem definition. Moreover, it is essential in optimal control for the controller to take advantage of the controlled system's natural dynamics. Indeed, taking full advantage of a system's dynamics is often the very essence of the problem (as, for example, in using a planet's gravitational pull to accelerate a spacecraft most efficiently toward a new destination).

Many optimal control problems are so hard to solve that one usually has to settle for approximate solutions, something that is true in the biological context as well. We are more interested in mechanisms that successively improve behavior (according to some objective function) than we are in optimal behavior, which is unlikely ever to be achieved.

But what does the control view add that is missing from the Dynamics Systems Theory? To us, the key is that a system can "want" another system to behave in a certain way. With the framework of optimal control, the desired system's behavior is not completely specified in terms of a desired state or trajectory; it is determined by an iterative process that approaches a goal for which no explicit structural description needs to be available. In a human developmental context, this allows for the possibility that the actor, the child, has implicit objectives and goals that are organizing the child's behavior. This view of control focuses attention on the approach toward the final objective, the goal, and not on the generation and following of pre-planned trajectories. We will next show how the concept of control can improve our insight into two developmental situations.

Early control of reaching movements

The basic facts describing the development of human reaching have been know since Halverson (Halverson 1931). Human infants make arm movements at birth that are directed towards objects in the environment (von Hofsten 1982). Around 15 weeks-of-age, infants become able to reliably reach out and touch small toys presented in their workspace (e.g., (Thelen *et al.* 1993),(Berthier *et al.* 1999)). In the succeeding weeks and months, infants improve the quality of their reaching so that reaching becomes reliable and smooth (von Hofsten 1991). By 1 to 2 years-of-age infants become able to reliably grasp objects in a manner that is appropriate for manipulation (McCarty, Clifton, & Collard 1999).

Even though this normative description of the development of reaching has been available, it has only been in the last decade that progress has been made in understanding why human infant reaching follows the observed developmental trajectory. After analysis of a dense longitudinal data set, Thelen et al. (Thelen *et al.* 1993) argue that one of the primary aspects of the development of reaching is the infant's mastery of the intrinsic dynamics of the body. They note that different infants start with different general energy levels, with some infants being very active and energetic, and others showing very dampened movements. The development of both types of infants converge on a reaching pattern that results in smooth, accurate movements. While Thelen et al. (Thelen *et al.* 1993) argue that this pattern was evidence for self-organization of a dynamical system, we would argue that the developmental sequence is driven by the goal-directed nature of the infant's behavior. Infants might initially have the goal of simply obtaining the goal object. Different infants, in the context of their own dynamical proclivities, might attain this goal in different ways. Later after the infant can reliably obtain objects, infants adjust their goal so that it becomes "attain the object with the least effort." Instead of simply saying the dynamical system self-organizes as do Thelen et al. (Thelen *et al.* 1993), we argue that the infant's goal combined with the infant-environment dynamical system defines an optimal control problem. Infants approximate solutions to the particular problem using exploratory learning in much the same way that incremental methods for approximating solutions to optimal control problems improve the behavior of a system and its controller.

A recent investigation by Berthier et al. (Berthier *et al.* 1999) examines the reaching of a group of young infants who are just learning to reach for and contact objects. The investigation focused on an aspect of exploratory learning that is problematical. In the context of learning to control a dynamical system as complicated as a human arm, exploratory learning methods could fail because of the immense complexity of the space that must be searched for solutions and because exploratory learning might cause movements of the arm and body that might damage the infant.

Berthier et al. (Berthier *et al.* 1999) found that beginning reachers adopted a pattern of reaching with the elbow locked, presumably because of muscular cocontraction. This pattern vastly simplifies the learning problem that must be solved by the exploratory learning algorithm by reducing the number of degrees-of-freedom to be controlled, by simplifying the dynamical system to be controlled, by improving the quality of feedback from movement, and by reducing the likelihood of damage to the arm. These simplifications make it reasonable to assume that simple exploratory learning methods might be effectively employed to solve the initial problems involved in learning to reach.

Controlling a dynamical system: The Jolly Jumper

The Jolly Jumper is a device consisting of an elastic-type spring, suspended from a ceiling or door frame, with a supporting harness in which the infant sits. Gravity exerts a downward force; the Jumper pulls upward when the elastic is stretched, and the infant can provide an additional upward force when the feet are on the floor and the legs kick outward. Goldfield et al., (Goldfield, Kay, & Warren 1993) have studied the behavior of human infants as they learn to bounce in the device. We used simulations of an infant interacting with the device to investigate how optimal control methods might be used to understand infant learning. Following Goldfield et al. (Goldfield, Kay, & Warren 1993), we modeled the infant as a mass-spring system. The legs of the infant were represented by a single spring, and the infant's kicking movements were represented as changes leg-spring stiffness. To simplify the task, we allowed the infant only two levels of stiffness so that at each time step the infant decides whether to "kick" or "relax" its legs.

We had two goals in this research: First, we wished to determine if optimal control policies led to behavior of the model system that resembled infant behavior. Second, we sought to determine if on-line learning methods, such as reinforcement learning (RL(Berthier 1996)), that lead to approximate optimal control policies also lead to control policies that parallel infant behavior. An interesting question in regard to the latter is whether the trajectory of learning using the on-line learning algorithms was a realistic model of human infant learning.

Optimal Control of the Jolly Jumper. We investigated several objective functions. (1) bouncing quickly, (2) absolute acceleration, (3) absolute distance from equilibrium, and (4) total energy (i.e., kinetic plus potential). The resulting optimal control problems were simple enough to solve using a standard method known as dynamic programming. The solutions reveal that all four objective functions lead to virtually identical optimal policies. In each case, height, velocity (kinetic energy), acceleration, and total energy are maximized by kicking the legs during upward movements and relaxing them on downward movements. Second, because the optimal policy selects actions as a function of movement direction (i.e., velocity sign), it is automatically phase-dependent. This phase dependence means that the kicking actions of the legs and the oscillations of the jumper-infant system are guaranteed to be in resonance.and solved for the optimal control policy using dynamic programming.

An additional consequence of phase dependence is that the optimal policy also exhibits both stable limit cycle behavior and resistance to perturbations. Position-velocity phase plots of an jumper-infant system that maximizes its kinetic energy show that its behavior displays a stable limit cycle, returning to the same oscillation pattern reaches by a jumper that begins at equilibrium.

On-line Learning. Dynamic system theorists stress the importance of learning by exploration. The distribution of random actions over time plays a central role in dynamic systems accounts of motor skill acquisition. The idea of "self-organization" can be misinterpreted to suggest that dynamic modes of activity are uniquely determined by the physical task constraints, and require only disorganized, random activity in order to be generated. Purely random or spontaneous actions, however, may not be sufficient for discovering and then maintaining even weakly stable dynamic patterns. Rather, some type of selection mechanism must operate in parallel with the generation of exploratory actions.

Many computational methods exist that have these properties, and some of them can be used to successively approximate solutions to optimal control problems. Some of these take the form of relatively simple online learning algorithms called reinforcement learning algorithms (Sutton & Barto 1998). Some of these algorithms have been shown to eventually lead to the same optimal control solutions that are computed by dynamic programming methods, if operating under suitable conditions. We are exploring the suitability of these algorithms as models of the mechanisms infants might use for goal-directed learning.

The Jolly Jumper provides a good starting point. Our purpose is not to propose a specific model of infants as they learn to bounce, but rather to provide a basic example of the how trial-and-error learning can work when applied to a particular dynamic system. We used a specific reinforcement learning algorithm called "Sarsa" (Sutton & Barto 1998), which uses experienced sequences of states, actions, and rewards to estimate the long-term value of doing each action in each state. These estimated values then form the basis of the controller's decision-making process.

The jumper-infant system is placed in equilibrium (e.g., at a height of 25 cm, with an zero initial velocity and acceleration), and then the system is given unlimited time to explore the effects of both kicking and relaxing its legs. Rewards were defined in accord with the different objective functions described above. For example, the rewards can be set so that the goal of the learning process is to maximize the long-term kinetic energy of the jumper-infant system. We can also study learning under a variety of other types of rewards that are perhaps more plausible in terms of what an infant can sense. Although the performance of the jumper-infant system does not become optimal (as compared with the solution computed by dynamic programming), it learns to produce actions well-synchronized with the movements of the jumper, approximating the optimal strategy of kicking during upward movements and relaxing during downward movements. This type of behavior is not the result of self-organization; it arises due to an active selective mechanism operating on behavioral variety.

Studies of this kind show show that the concept of control does not need to rely on pre-defined set points or reference trajectories. It can be goal-directed in more general ways that, we think, are well represented by the optimal control framework. Moreover, despite the computational difficulty of solving most optimal control problems, there are relatively simple on-line trial-and-error learning algorithms that can cause behavior to improve toward optimal control solutions that exploit natural dynamics.

References

Berthier, N. E.; Clifton, R. K.; McCall, D.; and Robin, D. 1999. Proximodistal structure of early reaching in human infants. *Experimental Brain Research* 127:259–269.

Berthier, N. E. 1996. Learning to reach: A mathematical model. *Developmental Psychology* 32:811–823.

Bower, T. 1989. *The rational infant.* New York: Freeman.

Bruner, J. 1973. Organization of early skilled action. *Child Development* 44:1–11.

Freedland, R., and Bertenthal, B. I. 1994. Developmental changes in interlimb coordination: Transition to hands-and knees, crawling. *Psychological Science* 5:26–32.

Gibson, E. J. 1988. Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology* 39:1–41.

Goldfield, E. C.; Kay, B. A.; and Warren, W. H. 1993. Infant bouncing: The assembly and tuning of action systems. *Child Development* 64:1128–1142.

Halverson, H. M. 1931. An experimental study of prehension in infants by means of systematic cinema records. *Genetic Psycholgy Monographs* 10:107–286.

McCarty, M. M.; Clifton, R. K.; and Collard, R. R. 1999. Problem solving in infancy: The emergence of an action plan. *Developmental Psychology* 35:1091–1101.

Piaget, J. 1952. *The origins of intelligence in children.* New York: Norton.

Piaget, J. 1954. *The construction of reality in the child.* Basic Books.

Rovee-Collier, C. 1987. Learning and memory in infancy. In Osofsky, J. D., ed., *Handbook of infant development, 2nd Edition.* New York: Wiley. 98–148.

Siegler, R. S. 1994. Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science* 3:1–5.

Sporns, O., and Edelman, G. 1993. Solving berstein's problem: A proposal for the development of coordinated movement by selection. *Child Development* 64:960–981.

Sutton, R. S., and Barto, A. G. 1998. *Reinforcement learning: An introduction.* Cambridge MA: MIT Press.

Thelen, E., and Smith, L. 1994. *A dynamic systems approach to the development of cognition and action.* Cambridge MA: MIT Press.

Thelen, E.; Corbetta, D.; Kamm, K.; Spencer, J. P.; Schneider, K.; and Zernicke, R. 1993. The transition to reaching: Mapping intention to intrinsic dynamics. *Child Development* 64:1058–1098.

Turvey, M. T., and Fitzpatrick, P. 1993. Commentary: Development of perception-action systems and general principles of pattern formation. *Child Development* 64:1175–1190.

von Hofsten, C. 1982. Eye-hand coordination in the newborn. *Developmental Psychology* 18:450–461.

von Hofsten, C. 1991. Structuring of early reaching movements: A longitudinal study. *Journal of Motor Behavior* 23:280–292.

Watson, J. W. 1972. Smiling, cooing, and the game. *Merrill-Palmer Quarterly* 18:323–329.

Willats, P. 1990. Development of problem-solving strategies in infancy. In Bjorklund, D. F., ed., *Children's strategies: Contemporary views of cognitive development*. New Jersey: Erlbaum. 23–66.