Functional Mechanisms of Motor Skill Acquisition

AUTONOMOUS LEARNING LABORATORY
www-all.cs.umass.edu

Ashvin Shah[1] and Andrew G. Barto[2]
[1]Neuroscience and Behavior Program and [2]Department of Computer Science, University of Massachusetts Amherst

ABSTRACT

As a motor skill (MS) is learned, behavior progresses from execution of movements that appear to be separately generated to recruitment as a single entity. Changes in behavior and neural activity suggest that different control strategies and systems are employed as the MS develops. We propose that as an MS is learned, control is transferred from an explicit *Planner,* which selects movements by considering the goal, to a *Value-based* controller, which selects movements based on the estimated value of each choice, to an *Automatic* controller, in which sensory cues directly elicit movement and no explicit decision is made. The *Planner* requires little experience but much computation; the *Automatic* controller requires much experience but little computation; the *Value-based* controller's requirements fall between the two.

In investigating the computational mechanisms of habit formation, which are similar in many ways to MS's, Daw et al. [1] proposed a similar model in which arbitration between a *Planner* and a *Value-based* controller was mediated according to the relative estimated uncertainty of each controller. Model behavior was similar to animal behavior in goal-devaluation simulations. In contrast, we suggest that arbitration is based on relative speed, in which the controller requiring the least computation is assumed to select a movement, if trained enough to select a movement, earlier than more computationally expensive controllers.

In this work, we describe a computational model, based on biologically plausible mechanisms and architecture, in which a learning agent must execute a series of actions (analogous to movements), elicited by the controllers, to solve tasks. We identify behavioral aspects of motor skill acquisition as seen in humans and animals and test the validity of this scheme by comparing model behavior to these aspects. We also use the model to explore how strategy changes depending on how MS recruitment affects other learning.

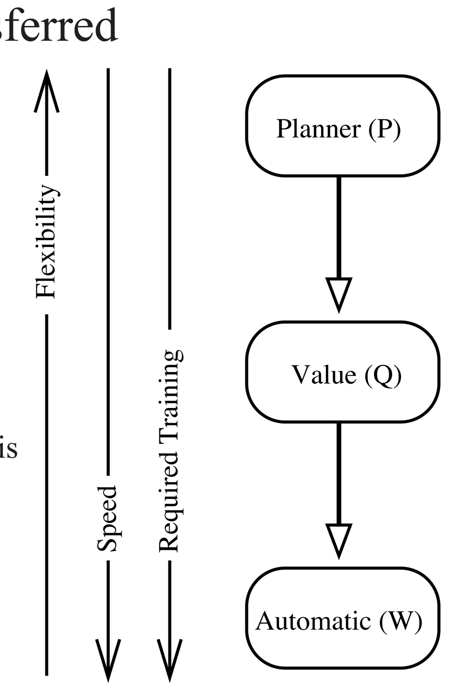
MOTOR SKILLS

Characteristics:

- behavior that arises from repeated execution of a given motor task
  - improvement in performance and speed
  - requires less attention, thought, and time
  - flexibility decreases
  - similar to *habits* [1]
- behavior common to many tasks [2]
- recruited as an entire unit, even if inappropriate [3]
- neural control transferred from cortical planning areas (e.g., frontal areas) to less cognitive areas (e.g., BG) [4]
- single neuron activity is different when movement selected in context of MS versus in isolation [5]

Acquisition scheme (cf, [1]): control is transferred

- planning areas
  - takes goal into account when planning, develops a reasonable solution
  - requires attention, thought, and time
  - planning and cognitive areas of cortex
- with repetition, simpler controllers are engaged
  - learn how "valuable" each movement taken in each context is
  - requires less resources
  - less cognitive areas of cortex and BG
- repeat same decisions and movements enough times, use simplest scheme possible: *motor skill*
  - sensory information elicits movement (similar to SR mapping)
  - requires least resources
  - thalamus to striatum



References and Acknowledgements

1. Daw, N., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature Neuroscience, 8:1704--1711.
2. Mushiake, H., Saito, N., Sato, Y., and Tanji, J. (2001). Visually based path-planning by Japanese monkeys. Cognitive Brain Research, 11:165--169.
3. Matsumoto, N., Hanakawa, T., Maki, S., Graybiel, A.M., and Kimura, M. (1999). Nigrostriatal dopamine system in learning to perform sequential motor tasks in a predictive manner. Journal of Neurophysiology, 82:978--998.
4. Doyon, J. and Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. Current Opinion in Neurobiology, 15:161--167.
5. Aldridge, J.W., and Berridge, K.C. (1998). Coding of serial order by neostriatal neurons: a "natural action" approach to movement sequence. The Journal of Neuroscience, 18:2777--2787.
6. Sutton, R., and Barto, A.G. (1998). Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA.
7. Gruber, A., Solla, S., Surmeier, D., and Houk, J. (2003). Modulation of striatal single units by expected reward: a spiny neuron model displaying dopamine-induced bistability. Journal of Neurophysiology, 90:1095--1114.

This research was made possible by NIH grant # NS 044393-01A1

HYPOTHESES

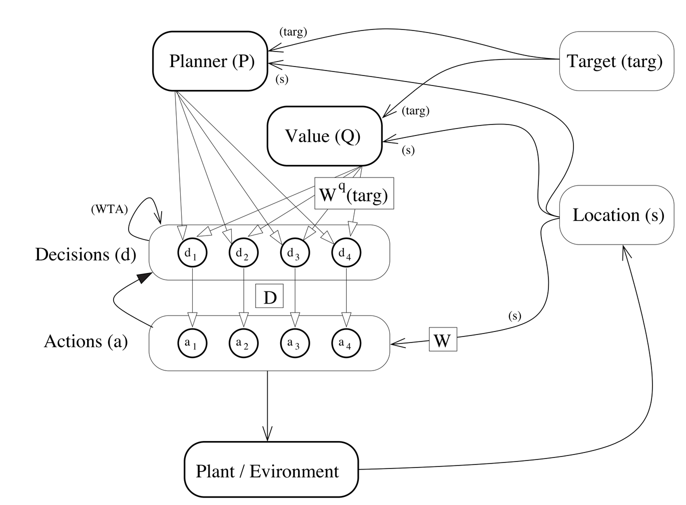
MOTOR SKILL DEVELOPMENT: the multiple controller scheme will develop a motor skill when same movements are often selected, including the common strategy of multiple tasks.

MOTOR SKILL USE: when MS's are available, they will aid in increasing performance, but may result in sub-optimal strategy.

MODEL

Decision making scheme

- movement selection is analogous to decision making
- use a discrete-state discrete-action environment
  - environment is a "grid-world," learning agent is in location (s)
  - must choose from a set of available actions (a) to navigate towards a target ($targ$)
  - each action taken incurs a context-dependent reward (r)



Connectionist model

- action taken when *action* neuron is excited past threshold
- P and Q excite *decision* neurons, which excite *action* neurons
- WTA in *decision* neuron array
- W excites *action* neurons directly

Planner (P)

- use AI algorithm to calculate best actions for a target
- excites decision neurons strongly

Value-based (Q)

- agent has an estimate, $Q_{targ}(s,a)$, of how valuable each action is for each location when moving towards a target [6]
  - thought to be mediated by dopamine in PFC and BG
- learns these values with experience (visiting locations and taking actions)
- $Q_{targ}(s_t,a_t) = Q_{targ}(s_t,a_t) + \alpha(r + \gamma Q_{targ}(s_{t+1},a_{t+1}) - Q_{targ}(s_t,a_t))$
- $Q_{targ}(s,a)$ used to train $W^q(targ)$
  - W^q excites *decision* neuron array (noise allows for exploration)
  - W^q grows from weak connections (no *decision* neuron wins WTA) to stronger connections

Automatic (W)

- $W(s,a)$, weight from s to a, is strengthened for each (s,a) experienced
- $W(s,a)$ for all actions *not* taken is weakened

Bistability

- *action* neurons may be bistable - "up" or "down" V_{rest}
  - striatal neurons may be bistable, dopamine may mediate transition [7]
- W only strong enough if *action* neurons are "up"
  - allows MS's to be turned off

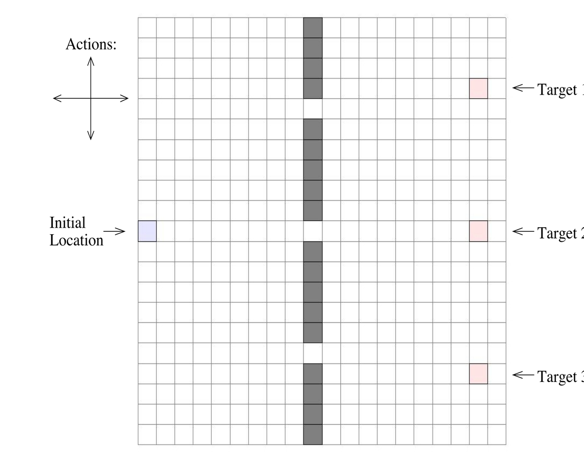
Arbitration scheme: W is faster than Q, which is faster than P

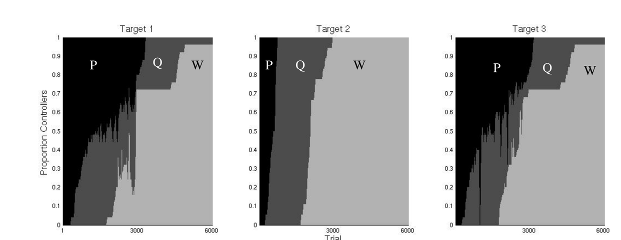
- W is engaged earliest
- if no *action* is selected, Q is engaged next
- if no *action* is selected, P is engaged

MOTOR SKILL DEVELOPMENT

Task: navigate from initial location to one of three targets

- target chosen randomly at each trial
- available actions: N,S,E,W (primitive actions)
- each step taken induces a negative r
- reaching target terminates trial and induces a positive or zero r
- objective: maximize reward during trial (reach target in minum time)





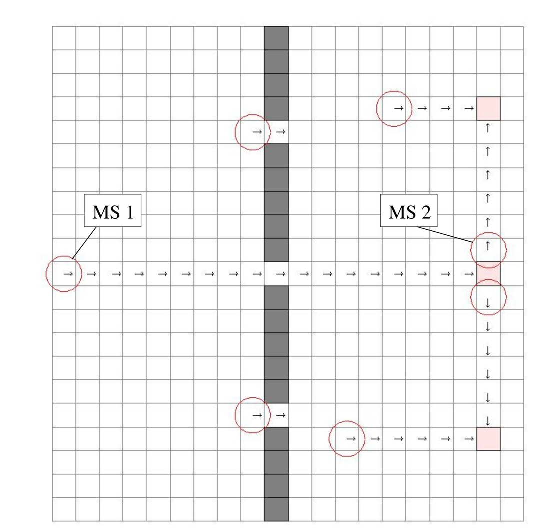
from a typical run

As task learned, control shifts from P to Q to W

- shift faster for easier target (2) - only one optimal path
- Q must still be engaged at one point for targets 1 and 3

Motor Skills (W) developed at

- common path (e.g., MS 1)
- paths where same actions are often selected (e.g., MS 2)
  - locations are rarely visited in context of other tasks



from a typical run

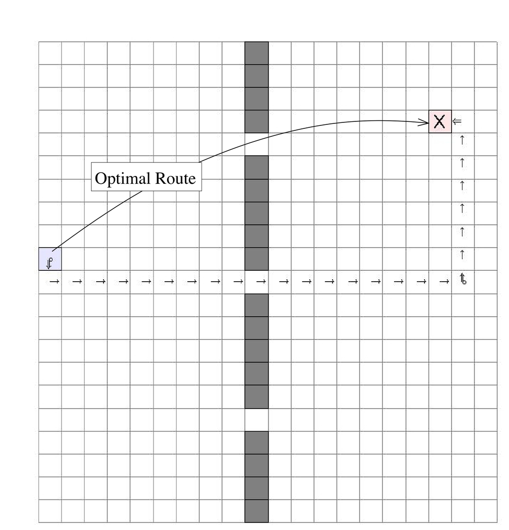
MOTOR SKILL USE

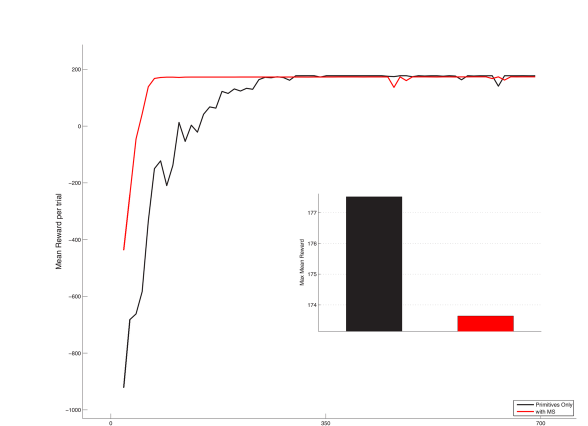
How do MS's developed in previous training affect behavior and strategy in a novel task?

- MS's only available if transitioning into initiation locations
- MS's are treated as actions (in addition to primitive actions: N,S,E,W)
- MS's can be turned on or off by exploiting bistability of *action* neurons.
- use just Q-values to make decisions
  - no decisions made during motor skill execution

MS's help early on, but offer a suboptimal strategy:

- MS's don't take goal into account, are inflexible





from 50 runs of each condition

EFFECT OF MOTOR SKILLS ON LEARNING AND BEHAVIOR

What do we learn while executing a motor skill?

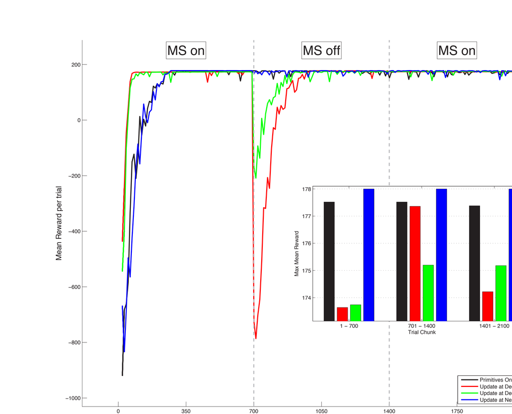
Compare four conditions:

1. primitive actions only (no MS's)
2. update $Q(s,a)$ based on next decision made
   - if a is an MS, $Q(s,a)$ updated is based on $Q(s,a)$ at end of MS
   - no other $Q(s,a)$ is updated while executing the MS
3. update $Q(s,a)$ based on next decision made if a is an MS, and based on next $Q(s,a)$ visited if a is a primitive action
   - if a is an MS, $Q(s,a)$ updated based on $Q(s,a)$ at end of MS
   - all other $Q(s,a)$'s updated based on next visited $Q(s,a)$
4. update $Q(s,a)$ just based on next $Q(s,a)$ visited
   - all $Q(s,a)$'s, including MS value, updated based on next visited $Q(s,a)$

How do these conditions affect behavior when

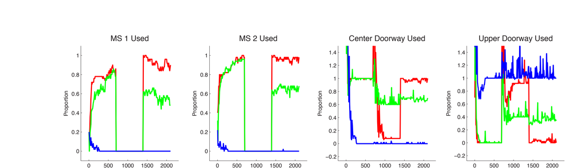
- *subject must relearn task without MS's?*
- *after relearning, subject can recruit MS's again?*

Overall effect: updating MS based on $Q(s,a)$ at end of MS profoundly affects behavior



When MS's are avoided

- performance under conditions 2 and 3 drop off, but they relearn the task
- behavior under condition 2 includes learning to use upper doorway more often than under condition 3
  - learning under condition 3 allowed agent to increase values at locations through center doorway, learning under condition 2 did not



from 50 runs of each condition

When MS's are reinstated

- increased proportion of behavior under condition 3 uses upper doorway
- behavior under condition 2 goes back to center doorway (?)

DISCUSSION

The term "motor skill" is used to describe a wide variety of learned behaviors, including a sequence of movements performed smoothly, quickly, and with less thought and attention. The underlying mechanisms of motor skill acquisition are difficult to ascertain. Is the increase in performance due to subtle changes in how we move or changes in how movement is selected? In this work, we examine the decision-making aspects of motor skills and suggest that a multiple controller scheme can account for behavioral aspects associated with motor skills. In the multiple controller scheme, computationally expensive controllers are used early in learning. However, if the same decisions are repeated, less expensive controllers are used, allowing cognitive resources to be devoted to other tasks.

The use of computational models allows us to idealize the learning agent and environment so we can more easily focus on isolated aspects of learning and behavior. In future work, we hope to clarify what aspects of motor skills are dependent on what mechanisms of motor skill acquisition.