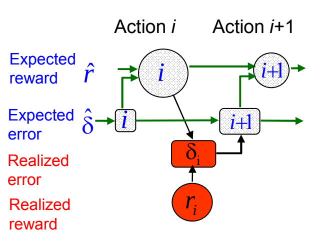
Modeling Individual and Group Behavior in Complex Environments

Dr. R. Andrew Goodwin
Environmental Laboratory

Professor James J. Anderson
Abran Steele-Feldman
University of Washington

Status:
AT-14
Continuing

Modeling Individual and Group Behavior in Complex Environments



Purpose:

- *Identify/capture robust phenomena regarding how individuals detect/respond to environmental changes.*
  - *Philosophical basis: cognitive needs of the U.S. Army will find value from novel findings in non-human studies.*

Product/Results:

- *Concise, neurologically- and physiologically-realistic algorithm replicating fundamental & pervasive psychological phenomena.*

Payoff:

- *Model can be expanded and integrated with other neurological principles / software packages to explain more complex behavior.*
- *Algorithm can be used to simulate simple behavior and multi-agent interactions.*

Modeling Individual and Group Behavior in Complex Environments

What is the Problem?

- To simulate learning and decision making in dynamic heterogeneous environments, we want algorithms that are:
  - neurologically-realistic,
  - able to generate important, observed phenomena,
  - computationally- (cognitively-) efficient,
  - able to be embedded in a hierarchal framework of self similar algorithms.

What are the barriers to solving the problem?

- Psychological models are generally descriptive or address highly structured behavioral environments.
- Neurological models are very complex and computationally intensive.
- Model of intermediate complexity needed.

What is innovative about this work?

- Algorithm frames learning as fundamentally a problem in forecasting future rewards.
- Structure naturally explains important, observed phenomena in classical and operant conditioning experiments with a simple architecture.
- Algorithm borrows from signal processing, time-series forecasting, psychological, and neurological literature.

How will you overcome these barriers?

- Use psychological literature to characterize fundamental behaviors that any model must replicate.
- Guide development and support algorithms by finding neurological equivalents to model components.
- Structure model complexity using principles from signal processing and time-series forecasting.

What are results of this research and what is its value?

- To our knowledge, no other model yet captures all three important, observed phenomena in classical and operant conditioning experiments:
  - the partial reinforcement extinction effect,
  - spontaneous recovery, and
  - latent inhibition.
- Provides a robust learning algorithm that can be embedded in multi-agent interaction models.

Collaboration across ERDC, commercial firms, and/or academia.

- Dr. R. Andrew Goodwin (EL)
- Academia Collaboration

University of Washington
- Professor James J. Anderson
- Abran Steele-Feldman

USMA
- Niki Goerger

Basis of Research

Massive, Complex Cognitive Systems

- **ACT-R (Adaptive Control of Thought – Rational)**
  - integrates theories of cognition, visual attention, and motor movement,
  - not AI programming language → not for developing expert systems,
  - learning ACT-R → visit lab where it's used for "months rather than days".
- **Soar**
  - construct integrated intelligent agents → AI researchers,
  - cognitive modeling → cognitive scientists,
  - not presently strong in low-level perception or motor control *(compared to ACT-R)*,
  - AI programming language,
  - learning Soar → visit lab where it's used for "months rather than days":
    - not good for simple rule systems.
- **Jess**
  - tool for building rule-based expert system tools → not for implementing algorithms.

Our Small Part

- Identify/capture robust phenomena regarding how individuals detect/respond to environmental changes that may be exploitable for counter-insurgency efforts.
- Develop simple algorithm with complex emergent properties.

U.S. ARMY

ERDC Engineer Research and Development Center

"One Team, One ERDC….Relevant, Ready, Responsive, Reliable"

Basis of Research

Ex: aquatic system w/fish agents | Ex: system w/generic agents

Goodwin *et al.* (2006)

- seek robustness in the algorithms that explain observed cognitive dynamics in the animal kingdom.
- seek to elevate the framework to derive simplistic but realistic analogues of the far more complex (human choice) decisions.

Basis of Research: Exploit Robust Phenomena

Cognitive needs of the U.S. Army will find value from novel findings in non-human studies

- Most biological systems adapt to different conditions & environments.
- Nervous system developed mechanisms that use prior experience to predict future events.
- Many of these mechanisms could potentially support behavioral prediction.
- Little known about which specific mechanisms are used.

SYSTEMS NEUROSCIENCE — NATURE|Vol 448|9 August 2007

Timing is everything

Phillip Larimer and Ben W. Strowbridge

Interactions among neurons in brain circuits underlie sensory perception and information storage. Work in locusts shows how the timing of different neuronal signals is synchronized to ensure effective communication.

ANIMAL BEHAVIOUR — Vol 445|22 February 2007 — nature

Planning for breakfast

Sara J. Shettleworth

It is commonly believed that planning for the future is a skill unique to humans. Could other animals, even those as evolutionarily distant as western scrub-jays, share this skill with us?

- Until recently, biases in experimental psychology and ethology precluded interpreting animal and human behaviors with shared mechanistic constructs.
- These biases are now being questioned.
- In recent years, there has been a surge in documenting processes in animals, such as future planning, which in humans are accompanied by distinctive conscious states.

Table of Contents

March 2008

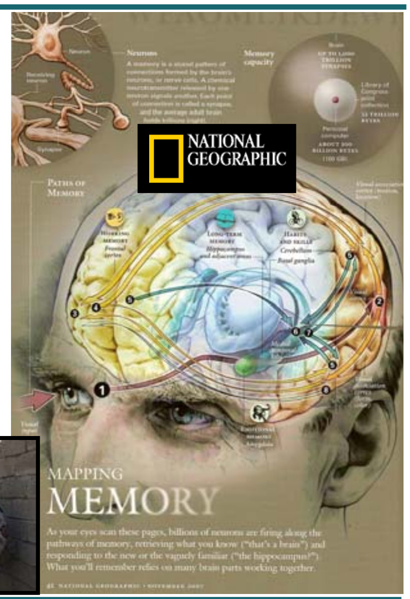
Inside Animal Minds

Brainy fellow creatures show that humans are not alone in their ability to invent, plan, or contemplate.

Basis of Research: Exploit Robust Phenomena

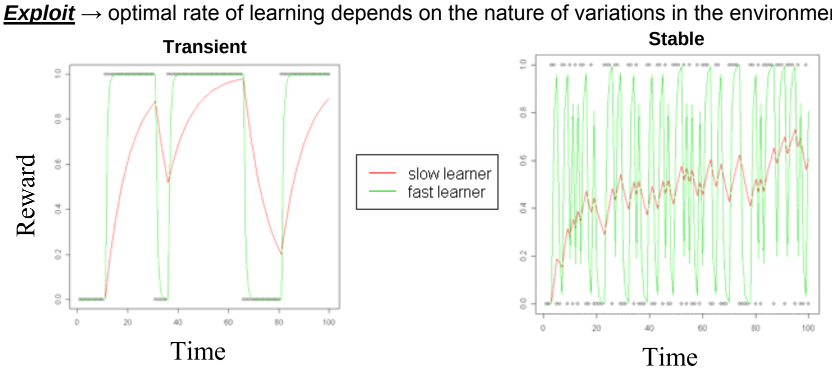
Animals have multiple memory systems:
- Basal ganglia → short-term
- Prefrontal cortex → long-term
- Single neuron → short- and long-term

- Future events not exact replicas of past events.
- A memory system storing rote records is not well-suited to simulating future events.
  - poor signal-to-noise ratio,
  - cognitively and computationally inefficient.
- More parsimonious method would draw on elements and gist of the past.
  - extract, recombine, and reassemble past events into imaginary events that never occurred in that exact form,
  - such a system will occasionally produce memory errors, but it also provides considerable efficiency and flexibility.



Our Project Objective/Hypothesis

- **Discern** how organisms differentiate between **transient** and **stable environments**.
  - **Ability to distinguish between stable and transient environments can be developed from products of two neural systems with different response rates.**
- ***Exploit*** → optimal rate of learning depends on the nature of variations in the environment.



Fast learners better predict rewards in Transient environments

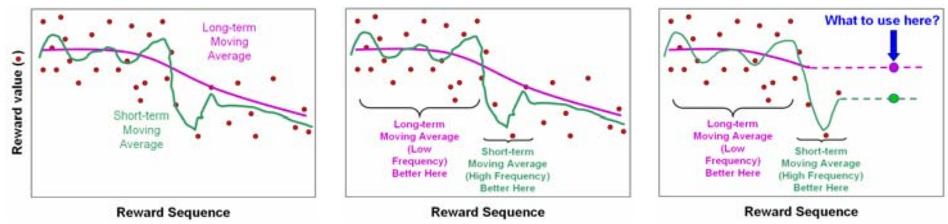
Slow learners better predict rewards in Stable environments

Product: Algorithm Process

Discern Stochastic Stationarity from ‘Real’ Change

Estimators in this example are *Exponentially-Weighted Moving Averages* (*EWMA*s):
- so pervasive in psychology → called the common model for learning (Lea and Dow, 1984).
- however, other types of estimators can be used.

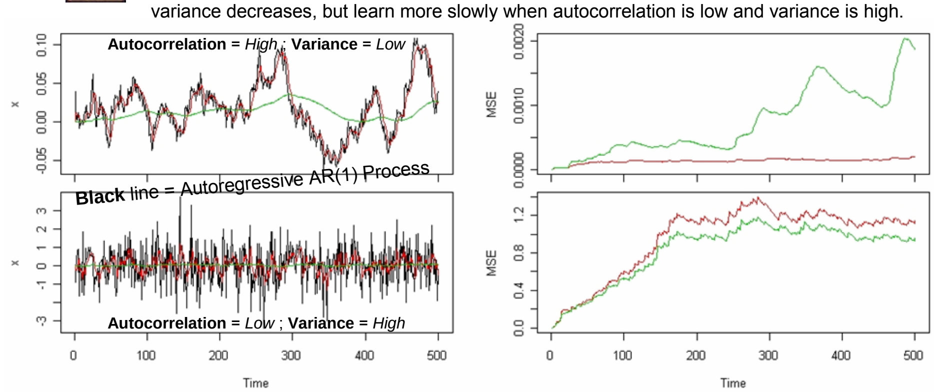
Characterize short- and long-term patterns of reward estimates and errors. → Best estimator depends on error in the estimators and a measure of the time between samples. → What if sample sequence is variable? Which stream to use to forecast reward of future sample?



“One Team, One ERDC….Relevant, Ready, Responsive, Reliable”

Product: Algorithm Process

__Principle__: Animals should learn more quickly as environmental autocorrelation increases and variance decreases, but learn more slowly when autocorrelation is low and variance is high.

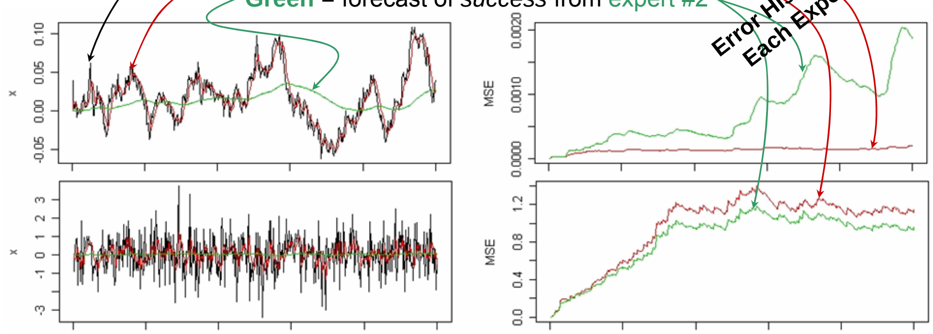


- **Left column** → AR(1) process and two different EWMA estimators:
  - **Red** is EWMA forecast with K=0.3.
  - **Green** is EWMA forecast with K=0.01.
- **Right column** → means squared error in EWMA estimators.
- **Upper-left** AR(1) process → generated with α_1=0.98 and σ^2=0.01.
- **Lower-left** AR(1) process → generated with α_1=0.01 and σ^2=1.

“One Team, One ERDC….Relevant, Ready, Responsive, Reliable”

Product: Algorithm Process

Black = underlying autoregressive AR(1) process
Red = forecast of *success* from expert #1
Green = forecast of *success* from expert #2



Experts and Filter Bank

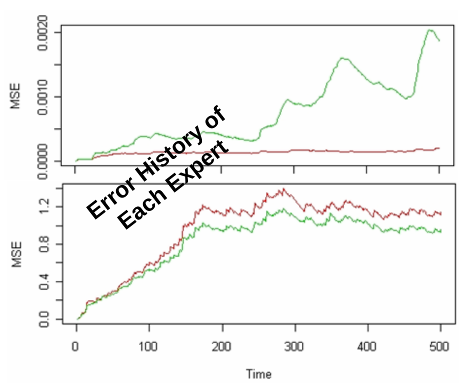
- ***Expert is a filter*** *that takes the reward sequence as input and outputs an estimate of the reward expected.*
- *Agent maintains group of N experts, each with different learning rates.*
- *EWMA is a low-pass filter (it passes low frequencies but stops high frequencies):*
  - *EWMAs that learn quickly (**large K**) pass higher frequencies than slow learning estimates (**small K**).*

Product: Algorithm Process

Red = forecast of *success* from expert #1
Green = forecast of *success* from expert #2

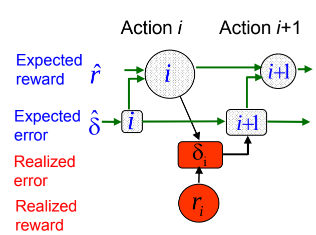
Mixture of Experts (Filter Bank)

- *Prefer as few experts as possible*



Gating Network

- *Use error histories of the experts to weigh the value of each expert's forecast.*
- *Integrate experts' forecasts into a single forecast.*

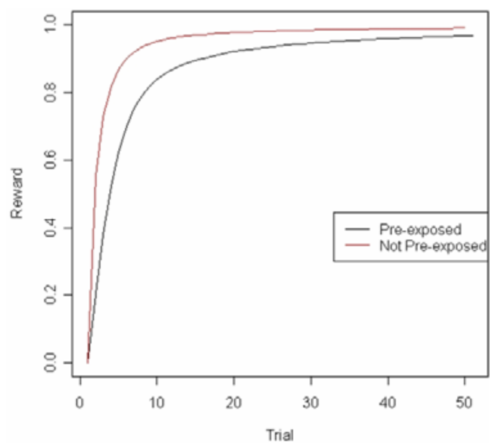


Product: Algorithm Capability

Classical and Operant Conditioning Experiments
Three Important Phenomena: 1 of 3

Latent inhibition: process by which exposure to a stimulus, of little or no consequence, prevents conditioned associations with that stimulus being formed.

- subject first exposed to repeated presentations of a stimulus that is not followed by a reward, and then after a fixed number of trials, the stimulus is paired with a reward,
- subject begins responding to the stimulus at a slower rate than does a control subject that was not first exposed to the unrewarded stimuli,
- if subject treats each stimulus as a sample, then this is a clear example of **Principle: When information is old, weight new samples more highly than when information is recent**.

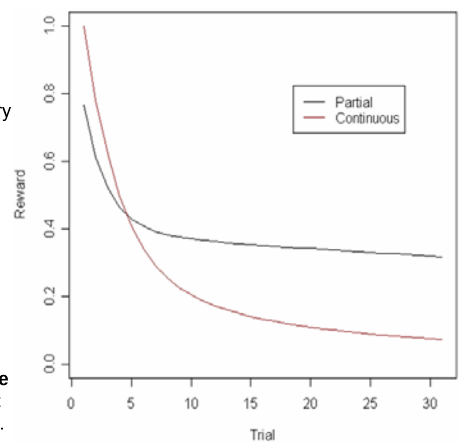


Product: Algorithm Capability

Classical and Operant Conditioning Experiments
Three Important Phenomena: 2 of 3

Partial reinforcement extinction effect (PREE):

- each trial signaled by a stimulus (e.g., a bell), then subject responds (e.g., presses lever) to receive reward,
- a continuous reward group is rewarded on every trial for taking the action,
- a partial reward group is only rewarded randomly over the trials,
- both groups receive same number of conditioning trials, followed by *extinction* where none of the trials is rewarded,
- during *extinction*, subjects eventually stop responding to the stimuli, but response rate decreases more quickly in continuous group than in partial group,
- results follow from **Principle** (repeated)**: Animals should learn more slowly when variance is high and, thus, that extinction should also proceed more slowly**.



ERDC Engineer Research and Development Center

"One Team, One ERDC….Relevant, Ready, Responsive, Reliable"

Product: Algorithm Capability

Classical and Operant Conditioning Experiments
Three Important Phenomena: 3 of 3

Spontaneous recovery: one of the most widely known phenomena in associate learning.

- subject first conditioned to respond to a stimulus by associating it with rewards, then conditioned response is extinguished by withholding rewards,
- if large period (days) passes without presentation of stimulus and then stimulus is presented again, subject will again respond, spontaneously, to the stimulus,
- if in further stimuli presentation the reward is withheld after a response, response will again extinguish but passage of more time without stimulus presentation will still lead to spontaneous recovery, although of diminishing magnitude.

  Principle: As time passes without sampling, the weight placed on the average rewards of older samples should increase.

Spontaneous recovery also possible/plausible from other types of transitions.

Question: can we exploit w/re: to TCPs? Insurgents who test TCPs frequently would *inherently* encounter less variability is sampling rewards than average public.