Learning and the Adaptive Management of Fisheries Resources

David Tomberlin (NMFS / Santa Cruz)

and

Teresa Ish (UC Santa Cruz)

April 18, 2006

Overview

- When is learning worth the trouble?
- Adapting in a fully observable world
- Adapting in a partially observable world
- Application to habitat management

When Is Learning Worth The Trouble?

- Adaptive Management
  - different things to different people
  - new information or situation
  - stochastic dynamic optimization:
    - action* = f(situation, information)
    - trade-offs between present and future
- Learning vs doing:
  - Is it worth acting suboptimally now to be able, by virtue of better info, to do better later?
  - In general, what’s the best mix of learning and doing?

When Is Learning Worth The Trouble?

When Is Learning Worth The Trouble?

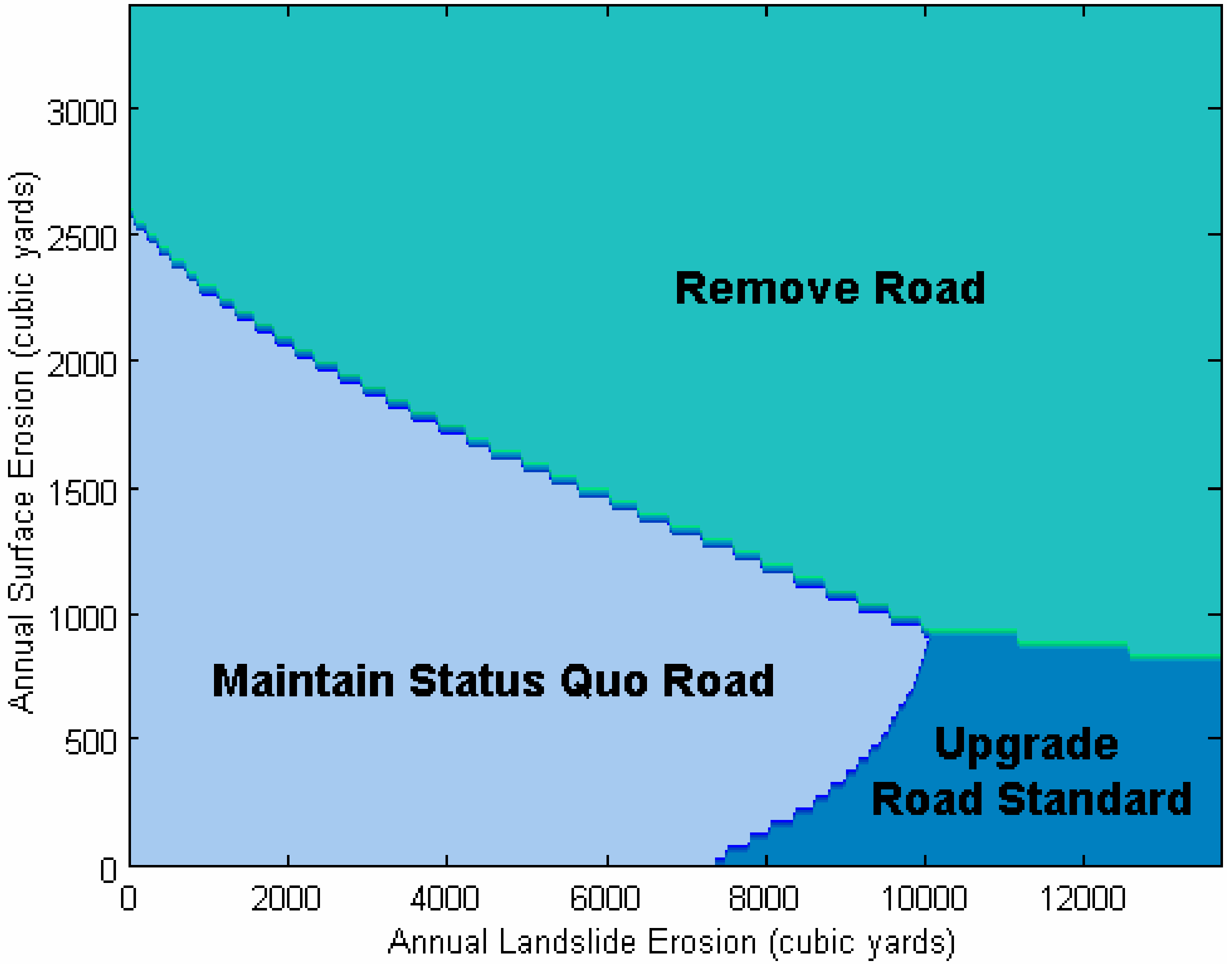
- Optimal control, Markov decision processes
  - no learning or passive learning
- Dual control
  - control engineers and a few ecologists
  - ‘active adaptive mgt’
- Partially observable Markov decision processes
  - AI and robotics
  - similar spirit to dual control, but different math

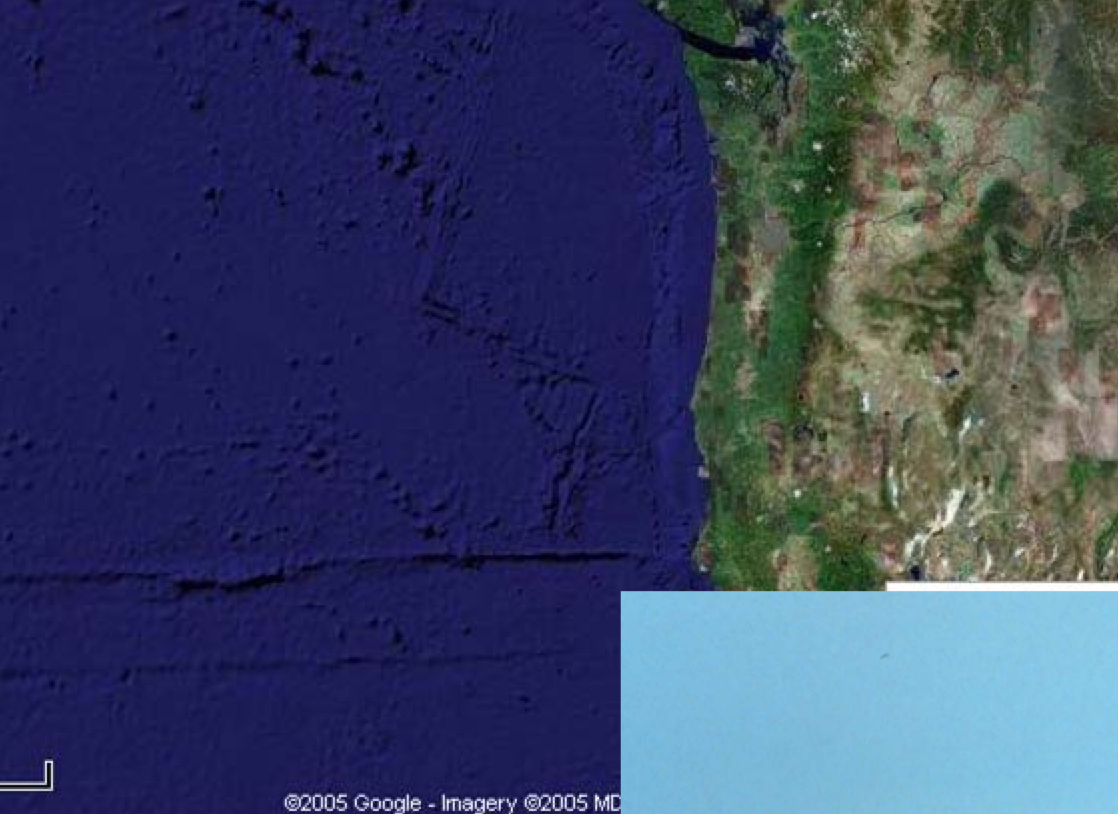
Markov Decision Processes (MDPs):

Adapting in an Observable World

- Dynamic + Stochastic: future = f(present, error)
- Solution: $\delta_t^*(x_i) = \arg\max_{a \in A} \left[q_i^a(x_i) + \beta \sum_j p_{ij}^a q_j^a(x_j) \right]$
- Process uncertainty, but not observational or model uncertainty
- Nothing about gathering info on state variables*

MDPs: adapting in an observable world




©2005 Google - Imagery ©2005 MD


Partially Observable MDPs (POMDPs): Adapting in a Noisy World

- Stochastic + dynamic + noisy
- Belief becomes a state variable
  - beliefs from priors and observations
  - uncountably infinite
- MDP: state → action

  POMDP: belief → action
- Solution: $\delta_t^*(\pi_i) = \arg\max_{a \in A} \left[\sum_i \pi_i q_i^a + \beta \sum_{i,j,\theta} \pi_i p_{ij}^a r_{j\theta}^a q_j^a \right]$

POMDP Example

North Fork Caspar Cr., plugged culvert Rd 5

When is erosion monitoring worth the expense?

Problem Setup

- When is erosion monitoring worth the trouble?
- States = {Good Road, Bad Road}
- Decisions = {Do Nothing, Monitor, Treat}
- Observations = {Good Road, Bad Road}

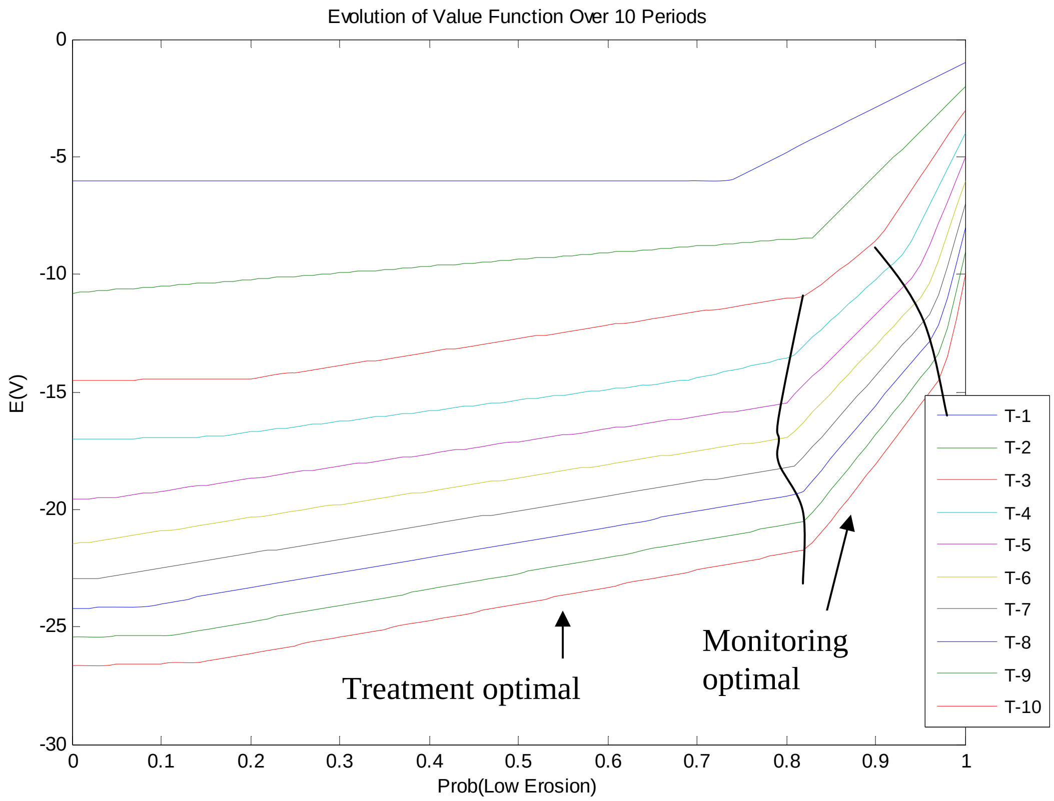
Problem Setup

$$P_{ij}^{1}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\quad P_{ij}^{2}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\quad P_{ij}^{3}=\begin{bmatrix}0.95 & 0.05\\ 0.80 & 0.20\end{bmatrix}$$

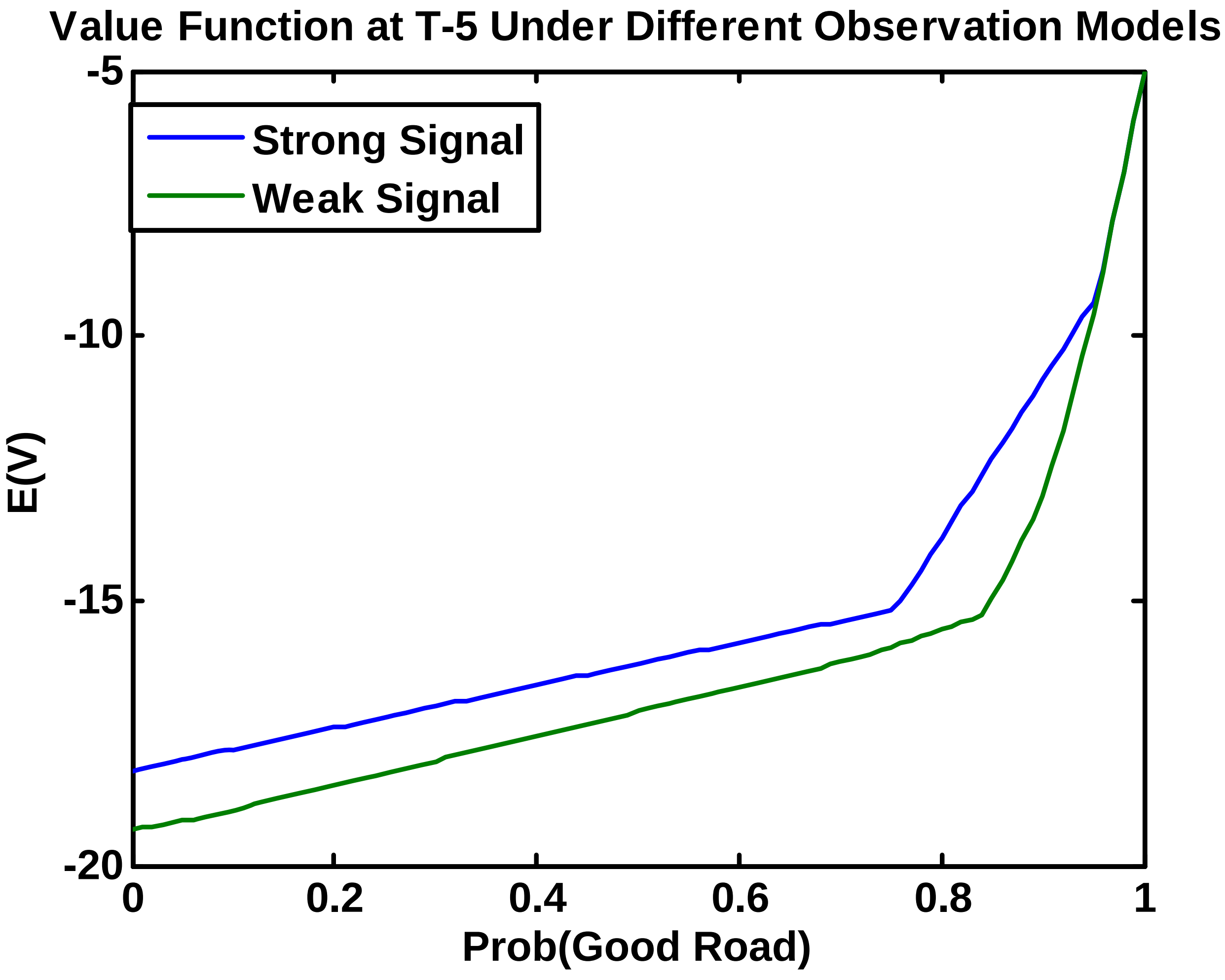
$$R_{j\theta}^{1}=\begin{bmatrix}0.6 & 0.4\\ 0.4 & 0.6\end{bmatrix}\quad R_{j\theta}^{2}=\begin{bmatrix}0.9 & 0.1\\ 0.1 & 0.9\end{bmatrix}\quad R_{j\theta}^{3}=\begin{bmatrix}0.5 & 0.5\\ 0.5 & 0.5\end{bmatrix}$$

$$W_{ij\theta}^{1}=\begin{bmatrix}-1 & -20\\ -1 & -20\end{bmatrix}\quad W_{ij\theta}^{2}=\begin{bmatrix}-3 & -22\\ -3 & -22\end{bmatrix}\quad W_{ij\theta}^{3}=\begin{bmatrix}-6 & -6\\ -6 & -6\end{bmatrix}$$

Results for a 10-period problem



Results (cont)



POMDP Assessment

- POMDP may be a good tool for fisheries mgt
  - partial observability is central in fisheries
  - monitoring funds are scarce
  - applicable to planning and behavioral models
- Drawbacks
  - computation
  - assumes (stochastic) dynamics are known
- Need to
  - increase state and decision space → heuristics
  - try state augmentation for parameter uncertainty

Odds & Ends

- POMDP and state augmentation (Fernandez-Rao)
- Reinforcement Learning (Bertsekas)
- Sequential hypothesis testing (Wald)
- [Behavioral & cognitive modeling:
  - neural basis of learning (Ishii et al.)
  - behavioral psychology (Bearden)
  - location choice (Lane)]

POMDPs: Adapting in a Noisy World

- MDP = $\{S, P, A, W\}$
  POMDP = $\{S, P, \Theta, R, A, W\}$
- MDP maps state → action
  POMDP maps beliefs → action
- Unknown state variables,
  known parameters*

POMDP Value Function

$$V_t(\pi) = \max_a \left[\sum_i \pi_i q_i^a + \sum_{i,j,\theta} \pi_i p_{ij}^a r_{j\theta}^a V_{t+1}[T(\pi \mid a, \theta)] \right]$$

where

π_i = probability of being in state i

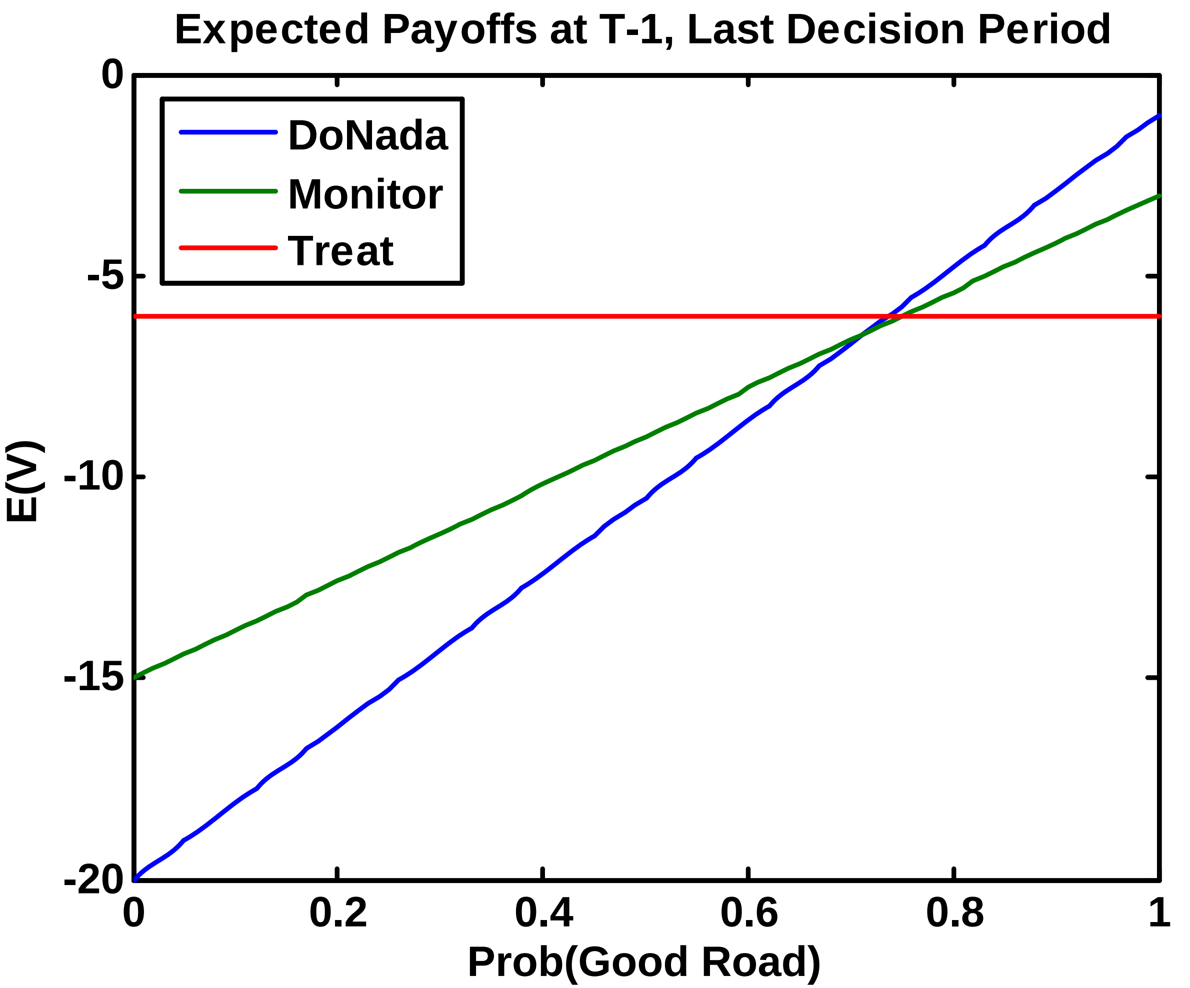
q_i^a = immediate reward for taking action a in state i

p_{ij}^a = probability of moving from state i to state j after taking action a

$r_{j\theta}^a$ = probability of observing θ after taking action a and moving to state j

T = function updating beliefs based on prior and θ

POMDP Value Function



POMDP Value Function