Dynamic Decision-making

Adaptive Management: Structured Decision Making for Recurrent Decisions

Dynamic Decision-making
Chapter 8
Developed by: Clinton T. Moore

Session Objective: By the end of this session, participants will be able to:

- Recognize the implications of making decisions in a dynamic context
- Understand the nature of solutions to dynamic problems
- Articulate important considerations for developing objectives for dynamic decisions
- Understand the two important ways that structural uncertainty is handled in dynamic solutions

Outline

- Context
- Decisions through time
- Dynamic programming
- Structural uncertainty
  - Passive and active adaptive management
- Summary points

Context

- Here, we focus on dynamic decision processes

We also focus on making decisions under uncertainty
Why are these contexts important?

- Decisions made today have impacts on future states, future decisions, and future returns
  - Opportunities created, opportunities lost

- Uncertainty reduces management performance over the long term

- However, recurrent decisions present an opportunity to reduce uncertainty

Dynamic decision making

How do we make a good decision?

The "decision tree"

- Discrete set of possible actions
- Each action leads to an outcome
  - Outcomes are probabilistic events
  - Reflects uncertainties due to the environment and partial control
- Each consequence (action × outcome combination) has a value (utility)
Dynamic Decision-making

Adaptive Management: Structured Decision Making for Recurrent Decisions

Generalizations needed

- For dynamic decision making, we will generalize the decision tree in 2 ways:
  - **Time**
    - Decisions are linked through time
    - Today’s decisions have consequences for future decision making
  - **Structural uncertainty**
    - Probabilities of outcomes are themselves uncertain
    - Use decision making to resolve structural uncertainty over time

**Generalization 1: Time**

- Adaptive management only works in a context of sequential decision making
  - **In time:**
    - Releases of animals to establish a population
    - Harvest regulations to maximize cumulative harvest
  - **In space:**
    - Thinning of forest blocks to obtain desired understory conditions
    - Hydrologic re-engineering to restore wetland communities

Dynamic decision making – some terms

- **State variables**
  - Measurable attributes of the resource that inform “where we are”
    - May be more than one, e.g. population size and habitat condition
    - *Partial observability* – hampers management performance and ability to learn

- **Return (or reward)**
Dynamic Decision-making
Adaptive Management: Structured Decision Making for Recurrent Decisions

- Value provided for a specific action taken or for arriving in a specific state
- Model
  - Mathematical description of system dynamics that links states, actions, and returns

The system moves from state to state

Implications of sequential decisions
- Decisions should account not only for the immediate return, but for all future returns according to where the system is driven and all decisions that follow
  - Myopic decision making focuses only on the immediate future
    - Future opportunities closed off or lost
    - Unsustainable management
Dynamic optimization
• Goal is to find an optimal trajectory of decisions through time that provides greatest expected accumulated return
  o Exact approaches
  o Approximate approaches

Important to note…
• Optimization and optimal management are not technical requirements for adaptive management
  o Learning under AM can proceed by any strategy to select a decision
  o But, optimization is the only recourse for selecting actions that are most efficient for pursuing the resource objective
    • i.e., may be a trade-off between efficiency (conservation delivery) and practicality/feasibility

Exact approaches
• Continuous-time approaches
  o For systems suitably represented in continuous time domain by simple models and few controls
    • Calculus of variations
    • Maximum principle
    • Continuous-time dynamic programming

• Discrete-time approaches
  o More complex systems, or those not well represented in continuous-time domain
    • Dynamic linear programming
    • Discrete-time dynamic programming (DP)

Dynamic programming (DP)
• Finds a trajectory of actions through discrete steps of time that maximizes an objective defined over the time horizon
  o Terminal value – a return that is realized only at the end of the time horizon (i.e., a salvage or liquidation value)
  o Accumulated value – returns that occur at each decision period and are summed
The time frame

- Time interval corresponds to the interval of the recurring decision
  - Often annual, but can be shorter or longer as appropriate

- Time horizon
  - Fixed & short-term
  - Indefinite, or very long

Fixed, short-term time horizon

- Appropriate where a desired end state is to be achieved within a specified time limit
  - Terminal value formulation

- Examples:
  - “Determine the optimal 10-year sequence of actions to achieve a targeted plant community composition”
  - “Determine the optimal 20-year sequence of releases to establish a breeding population with high probability of persistence”
Indefinite, or very long time horizon

- Appropriate where a recurrent reward is sought and long-term resource sustainability is at least an implied objective
  - Accumulated value formulation
- Examples:
  - “Determine optimal sequence of regulatory actions to maximize expected cumulative harvest of waterfowl over an indefinite time horizon”
  - “Determine optimal sequence of water releases to sustain targeted diversity of an aquatic community over 100 years”

Influence of the time horizon

- A thought exercise
  - You are a manager at a forest refuge where a threatened bird occurs, and you make annual forest harvest decisions intended to sustain the population through the creation of mid-successional forest habitat
  - However, you are informed that next year, the refuge will be sold, the forest cut, and the resident population translocated
  - To best support the population until that happens, what would likely be your approach to forest management this year?
  - Scenario change: Suppose instead that you know the refuge will be liquidated 30 years from now – how would that knowledge affect your decision this year?
Discounting

- Returns in the future have less value relative to the same return today
  - May be appropriate for problems involving monetary return or where future returns are uncertain
  - High discounting is incompatible with notions of sustainability
  - But low discounting may be useful in finding optimal solutions without severely undervaluing the future

![Diagram of dynamic decision-making](image)

What are we trying to do?

Find these... that makes this as large as possible

**Terminal value formulation**

OR

Find these... that makes this sum of (discounted) values as large as possible

**Accumulated value formulation**
Need to account for system dynamics
- Note that the terminal reward or the time-specific rewards are dependent on the states that the system passes through
  - Must account for these transitions
- Bellman’s Principle of Optimality (1957)
  - A solution based on a recursive argument
  - Bellman suggested a way forward … by working backwards!

Walk-through of a simple DP problem
- Managing a single patch of native prairie:
  - A single state variable with 3 levels:
    - Patch is (1) mostly native composition, (2) mixed native-invasive, or (3) mostly invaded
  - 4-year decision interval
  - 2 decision alternatives at each interval:
    - Defoliate every other year for 4 years, or rest
  - Rewards
    - Certain action-outcome combinations are more favorable than others

- A simple model

- Returns and cumulative values
Recursive feature of objective function

- For each system state, find decision that maximizes

\[ V_{t0} = y_{t+1} + y_{t+2} + y_{t+3} + \ldots + y_T \]

- To solve for optimal decisions, construct the policy one decision at a time by working \textit{backwards} from the future to the present.
Simple model: Steps in optimization
1. Assign values for having arrived at each possible state at end of time frame $T$
   - Levels of satisfaction for each state
2. Move backwards 1 period – for each decision (D or R) at time $T-1$, determine return ($y$) and probability of transition ($p$) to each state at $T$
3. Calculate average value of each decision: Add current return $y$ to value associated with each state at $T$, then sum (weighted by $p$) over state outcomes
4. For each state at $T-1$, identify action yielding greatest expected accumulated return
5. Store the optimal action and its state-dependent value
   - Compute optimal values for other states
6. Return to step 2; repeat process through time frame
More iterations of this process may reveal a stationary policy, i.e., decisions sensitive only to state, not time.

DP: Summary of steps
1. Assign values for arrival at end-of-time states
2. Move back 1 time step; determine returns from each action × outcome combination
3. Calculate average value of each decision at time step
4. Identify optimal action at each state at time step
5. Store optimal actions and state-dependent value
6. Repeat (2)-(5) through time frame

DP: key points
- DP is merely a chain of decision trees
- Once a state’s optimal value is computed at any time step, the potential paths forward in time from that state are irrelevant
- Sufficient iterations may yield a stationary optimal policy, where decisions are dependent on system state but not on time
- DP provides closed-loop control
  - Today’s optimal action reflects feedback inherited from the system dynamics
Example: Invasive species control
  - Objective: Minimize discounted sum of damage, monitoring, & treatment costs
  - State: Manager's relative confidence in low, medium, or high levels of infestation (invasion state is not fully observable except through monitoring)
  - Actions: Do nothing (1), monitor only (2), treat only (3), treat + monitor (4)

Other examples
- Harvest
  - Anderson (1975) Ecology 56:1281-1297
- Reintroduction / translocation
- Habitat management / Invasive species control
- Human disturbance
Approximate approaches

- DP suffers from “Curse of Dimensionality”
  - Problem size explodes with increasing number of states, decisions, and random variables
  - Computational limits are quickly met

- Some approximate alternatives may be “good enough”
  - Simulation-optimization
  - Reinforcement learning
  - Heuristic techniques

- Again: bona fide optimization is not a technical requirement for adaptive management

Generalization 2: Structural Uncertainty

- We are often uncertain about basic dynamics of the system
  - What is the probability of transitioning to a desired community state given that burning is conducted?
  - What is the average spawning response given control of a predator?
  - What is the form of the relationship between season length and harvest rate?

- Recurrent decision making provides an opportunity to learn and adapt our management approach

Decision tree, revisited

- We earlier considered a decision problem in which carrying out the management action favored the desired outcome, compared to no action
  - \( P(\text{native} \mid \text{hydrology restoration}) = 0.7 \)
  - \( P(\text{native} \mid \text{no action}) = 0.5 \)

- But suppose that this is uncertain or in dispute; that is, a credible claim is made that restoring hydrology has no better chance than doing nothing?
Dynamic Decision-making
Adaptive Management: Structured Decision Making for Recurrent Decisions

Here, uncertainty matters

The optimal action depends on the model (hypothesis) we choose
- If we believe in H1, ‘Restore’ action is optimal (expected utility = 59)
- If we believe in H2, ‘Do nothing’ action is optimal (expected utility = 50)

Competing models
- Do we even have to choose one model over another?
  - No – Our strategy will be to compute expectations of the utilities with respect to relative confidence in the models, and choose the action with greatest expected utility
    - Let’s assume 50:50 relative confidence in the models
  - Aside: other strategies are available for one-time, non-dynamic decisions
    - e.g., minimax, info-gap theory
Incorporating model uncertainty

<table>
<thead>
<tr>
<th>Expected Utility</th>
<th>Model</th>
<th>Outcome</th>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hypothesis 1</td>
<td>Hypothesis 2</td>
<td>Native Community Established</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>0.7</td>
<td>0.5</td>
</tr>
<tr>
<td>52</td>
<td>H1: 59</td>
<td>H2: 45</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>50</td>
<td>H1: 50</td>
<td>H2: 50</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Model Belief Weight

52

Structural uncertainty in DP
- Approach #1 (passive)
  - Augment the decision tree with model belief weights, chain the trees together as before, and keep belief weights unchanged over the time steps
    - Model uncertainty is acknowledged in the optimization, but not in a way that recognizes that it can change over time
    - In application, it does change over time as decisions are made, outcomes are compared to predictions, and model weights are updated
- Strategy for approach #1:
  - Perform DP using today’s model weights throughout all time steps, pretending as though weights will never change
  - Make a decision, carry out action, and update model weights
  - Repeat (1) and (2) at next decision opportunity
- Learning is *passively* obtained as an *unplanned* byproduct of decision making
Passive adaptive management
Structural uncertainty in DP

- Approach #2 (active)
  - Alternatively, explicitly account for expected change in model weights as decisions are made
    - We track changing system knowledge (in the form of model weights) as an *information state*, alongside the physical system state
    - We use a formulation of DP that incorporates Bayes' Theorem as the model of dynamics for the information state
    - The optimization anticipates that knowledge about the system will change in response to decisions made through time and the responses they are expected to generate
  - Learning is *actively* obtained as a *planned* outcome of decision making
    - Dual control: learning is pursued to the extent that it improves long-term management

Active adaptive management

![Diagram showing dynamic decision-making](image_url)
Passive vs Active

- Both approaches provide closed-loop control of the system state, but CL control of the information state is only achieved through active AM.

- The *dual control* problem: Balancing the pursuit of management objectives against the need for information that tells us how the system works.
  - Active AM provides a balanced solution that proposes informative (but not reckless) actions when system uncertainty is high.
    - Learning (information) is pursued only to the extent that it improves management.
  - Passive AM also pursues the management objective, but under the simplifying assumption that understanding will never change.

Example: Forest harvesting for old-growth habitat

<table>
<thead>
<tr>
<th>Forest State</th>
<th>Model Weights</th>
<th>Optimal Harvest Amounts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1 (Fast)</td>
<td>F0 (Med)</td>
</tr>
<tr>
<td>Mostly young forest</td>
<td>1 0 0</td>
<td>0 1 0</td>
</tr>
<tr>
<td>Mostly old forest</td>
<td>1/3 1/3 1/3</td>
<td>0.04 0 0</td>
</tr>
<tr>
<td></td>
<td>1 0 0</td>
<td>0 1 0</td>
</tr>
<tr>
<td></td>
<td>1/3 1/3 1/3</td>
<td>0.08 0 0</td>
</tr>
</tbody>
</table>

---

Moore & Conroy (2006)
Examples

- Passive AM
  - Optimal predator control: Martin et al. (2010) Biological Conservation 143:1751-1758

- Active AM
  - Disease management: McDonald-Madden et al. (2010) Ecological Applications 20:1476-1489

Experimentation and AM

- Neither passive nor active AM defers pursuit of the management objective for the sake of learning
  - They both focus on the management objective, but they use different tactics to account for uncertainty
- In contrast, experimentation places all emphasis on learning
  - Pursuit of management returns is set aside in favor of pursuing information
- Considerations for integrating experimentation into AM
  - Maintain focus on fundamental objectives (learning is a means objective)
  - Exploit opportunities for targeted experimentation (i.e., a sample of spatial units)
  - Inferences based on model selection and parameter estimation are more useful than classical hypothesis tests
Summary points

- Decisions made in dynamic systems have consequences for future decision making
  - Today’s decision influences future states and future rewards
  - Optimal decision making should account for future system dynamics, and if possible, uncertainties about those dynamics

- Dynamic programming seeks optimal state-dependent decision policies
  - Short-term or indefinite time horizon
  - Terminal value or accumulated value
  - Uses recursion in a reverse-time perspective to account for future system dynamics
  - Solution is achieved by working through a chain of decision trees

- Structural uncertainty may matter to the decision
  - We can still make an optimal decision by computing expected decision values with respect to model confidence weights
  - Can approach this in two ways in DP:
    - Passive AM – uncertainty is recognized, but assumed to remain static through time
      → Better management occurs as an unplanned byproduct of decision making
    - Active AM – uncertainty is modeled as a dynamic state through time
      → Decision making itself can be used to elicit information that would enable better management to evolve
... and the gratuitous sports reference