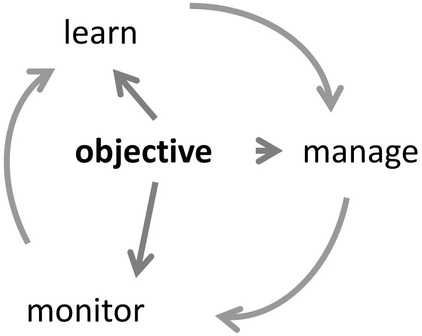
Recap Day 2 and intro to Day 3

- Day 2:
  - How to optimise sequential decision-making:
    - Markov Decision processes
    - MDPSolve
  - How to learn:
    - Sufficient statistics;
    - Model uncertainty: finite, belief.
    - Parameter uncertainty: infinite/discrete

Bayes theorem



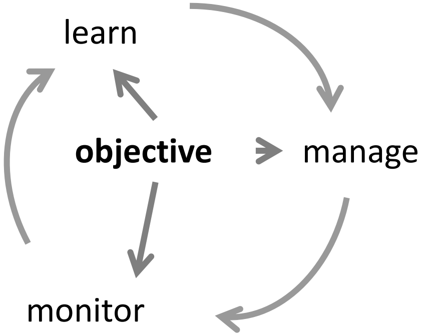
R. Bellman

Stochastic dynamic programming

Recap Day 2 and intro to Day 3

- Day 3:
  - Adaptive management model uncertainty (finite values)
  - Breakout session
  - Adaptive management parameter uncertainty (infinite values)

Bayes theorem



R. Bellman

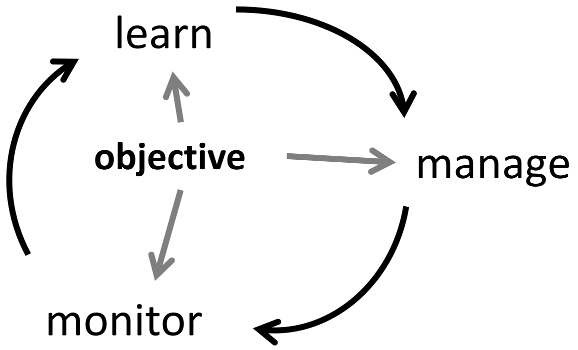
Stochastic dynamic programming

Session 5 – Discrete model adaptive management

- Introduction to model AM
- Key elements
- AM for model uncertainty as a Markov decision process
- Gouldian Finch problem
- Discussion

- MDPsolve (Paul)

In many domains, we do not have access to the dynamics but still need to make decisions



Adaptive management provides a solution. Adaptive management is "learning by doing". Decisions are selected to achieve a **management objective** while simultaneously gaining information to **improve future management outcomes** (Walters and Hilborn 1976).

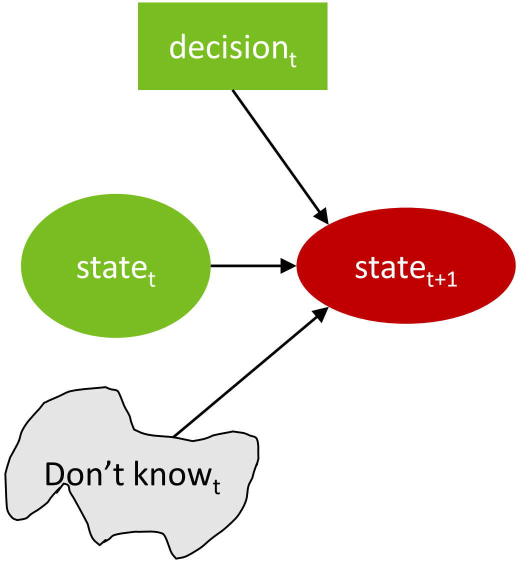
Adaptive management deals with two types of uncertainty

1) **Parameter uncertainty**:
e.g. survival, growth, probability of success

~ hidden variable takes infinite # values

1) **Model uncertainty:**
e.g. competing scenarios, Sea Level Rise, expert opinions

~ hidden variable takes finite # values

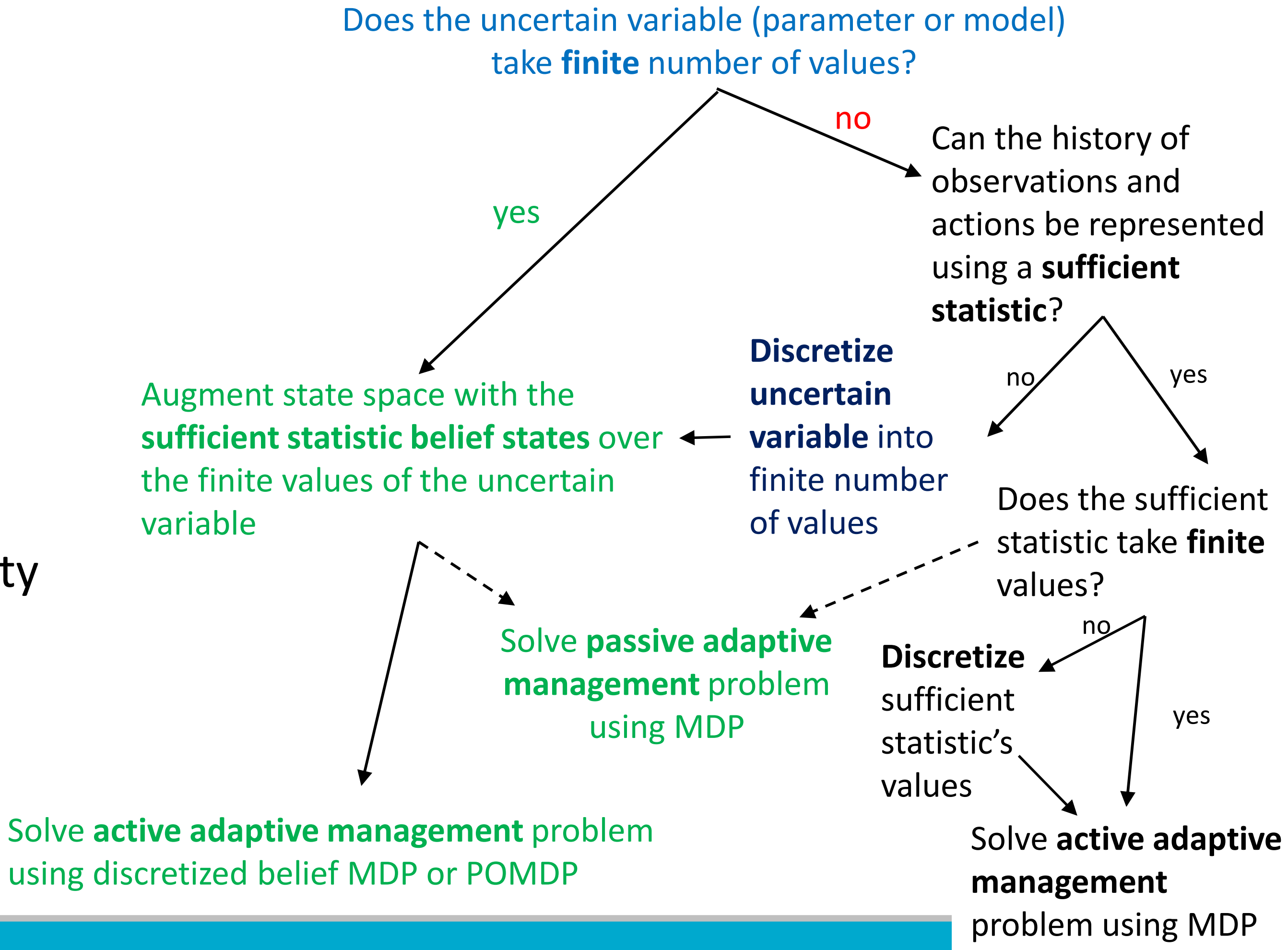


Session 5:

- Solving AM under model uncertainty

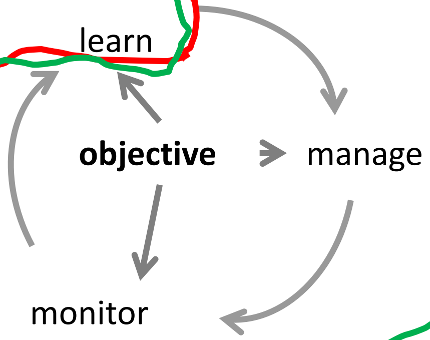
Session 6:

- Solving AM under parameter uncertainty



Passive adaptive management provides the best actions given our current knowledge … Learning occurs independently.

Bayes theorem

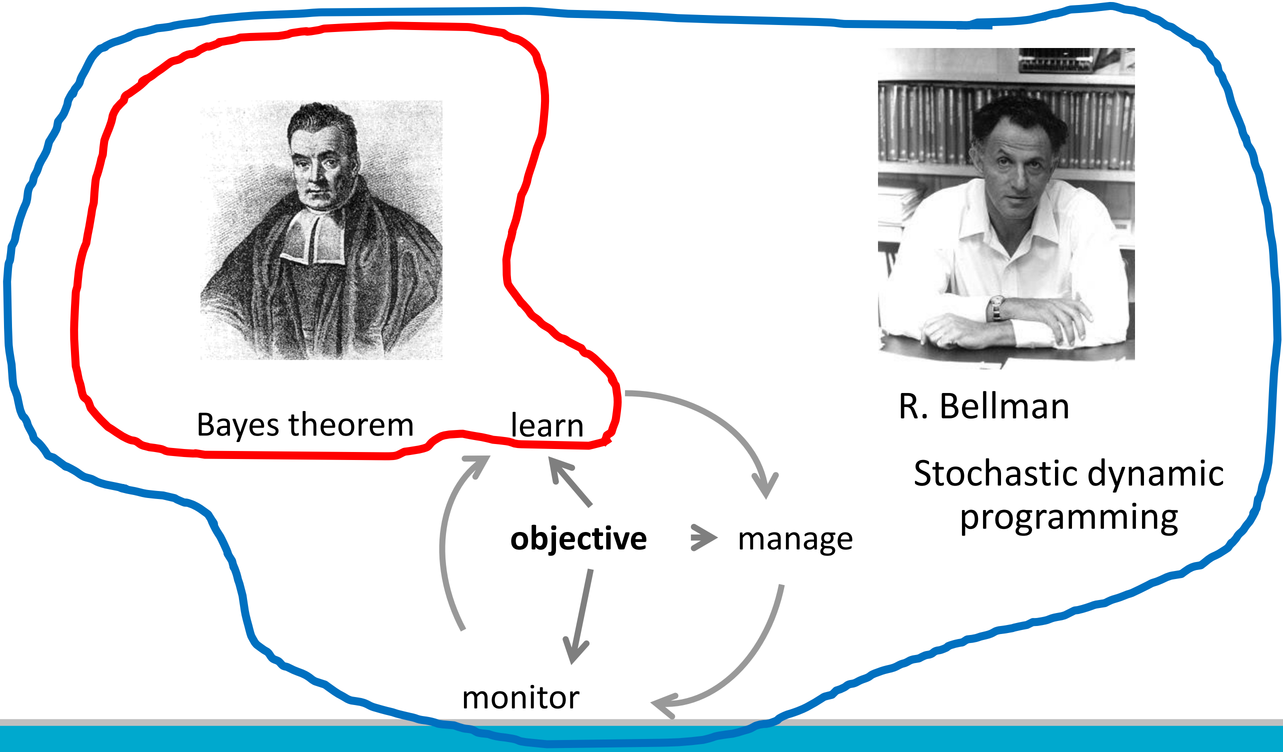


R. Bellman

Stochastic dynamic programming

heuristics
easier to solve
(certainty equivalence principles)

Active adaptive management provides the best actions given our current knowledge … AND what we will learn in the future

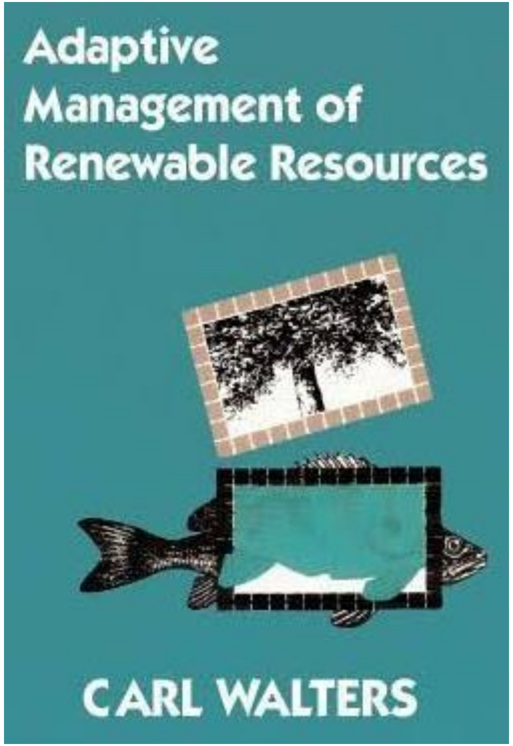


Optimal but difficult to solve

Active management requires ‘thinking ahead’ and calculating the consequences of all possible values of the unknown information before deciding the optimal action.

History

- AM tools to reduce model uncertainty were first proposed in the fisheries literature as early as 1978 (Silvert 1978), and included in (Walters 1986).
- In the mid-1990s, AM under model uncertainty was implemented by the US Fish and Wildlife Service to set harvest quotas for mallards in the USA (Johnson et al. 1997; Nichols et al. 1995)
- Plethora of other AM studies designed to reduce model uncertainty in conservation and resource management (Johnson et al. 2002; Martin et al. 2009; McDonald-Madden et al. 2010b; Moore and Conroy 2006; Smith et al. 2013; Williams 2011a) + more



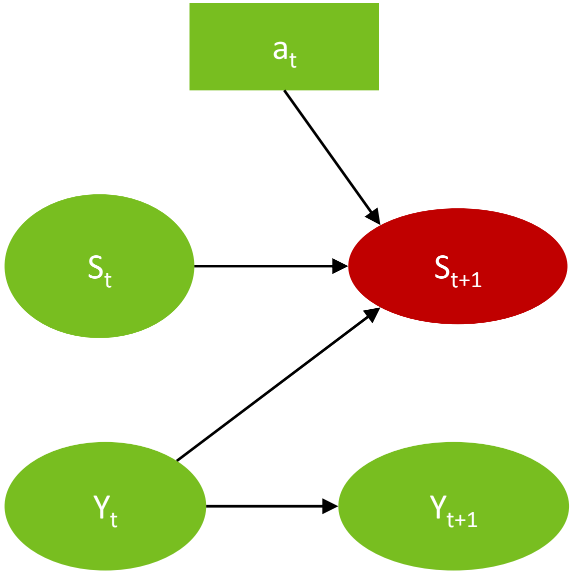
Key points to design an AM for model uncertainty

- **Plausible alternative hypotheses (models)** about system function can be articulated:

$$Y = \{model1, model2\}$$

- The models can take many forms, so long as the **transition probabilities** between states can be computed under each possible model.

$$y \in Y, P_y(s_{t+1}|s_t, a_t) = P(s_{t+1}|s_t, y, a_t)$$



Key points to design an AM for model uncertainty

- We use **belief states as sufficient** statistics. Belief states are probability distribution over states

$$b \quad = [b(model1), b(model2)]$$

- Optimal decisions depend on an observable state and belief.
- A policy is defined as: π : S, b→ A.
- Augment the state space: S x $B(Y)$
- Finding the best action:
  - Solve an MDP with a continuous belief state space (Belief MDP, today).
  - Solve a POMDP (day 4)

MDP vs belief MDP

MDP

- States S
- Actions A
- Transition matrices $P(s_{t+1}|s_t, a_t)$
- Reward function $R(s_t, a_t)$
- Discount factor

- Optimal policy $\pi^*: S \to A$
- Value function $V_\pi(s)$

belief MDP

- States S x B(Y)
- Actions A
- Transition matrices $P(s_{t+1}|s_t, y_t, a_t)$
- Reward function $R(s_t, a_t)$
- Discount factor

- Optimal policy $\pi^*: S, B \to A$
- Value function $V_\pi(s, b)$

MDP vs belief MDP

- Optimal policy $\pi^*: S \rightarrow A$
- Optimal value function $V_\pi(s)$

- Optimal policy $\pi^*: S, B \rightarrow A$
- Optimal value function $V_{\pi*}(s, b)$

$$V^*(s_t, b_t) = \max_{a \in A} \left\{ \underbrace{\sum_{y \in Y} b_t(y) r(s_t, a)}_{\text{Expected immediate reward given } s_t,\, a \text{ and } b_t} + \underbrace{\underbrace{\gamma}_{\text{Discount factor}} \underbrace{\sum_{s_{t+1}}}_{\text{future states}} \left(\underbrace{\sum_{y \in M} b_t(y)\, \mathrm{P}_y(s_{t+1} | s_t, a)}_{\text{Transition probability given } s_t,\, a \text{ and } b_t} \underbrace{V^*\left(s_{t+1}, b_{t+1}^{a, s_{t+1}}\right)}_{\text{Optimal value in } s_{t+1} \text{ and } b_{t+1}} \right)}_{\text{Expected future discounted reward}} \right\}$$

CSIRO

MDP vs belief MDP

- Optimal policy $\pi^*: S \rightarrow A$
- Optimal value function $V^*_{\pi}(s)$

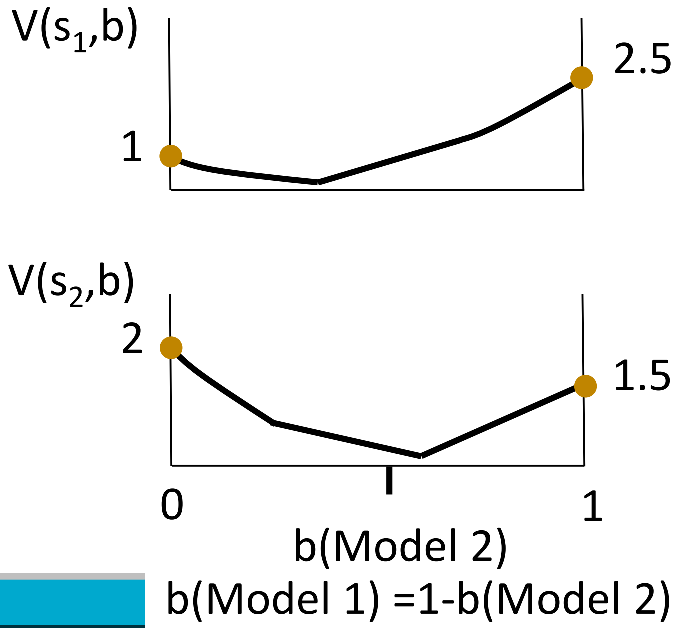
Model 1

$V(s_1) = 1$
$V(s_2) = 2$

Model 2

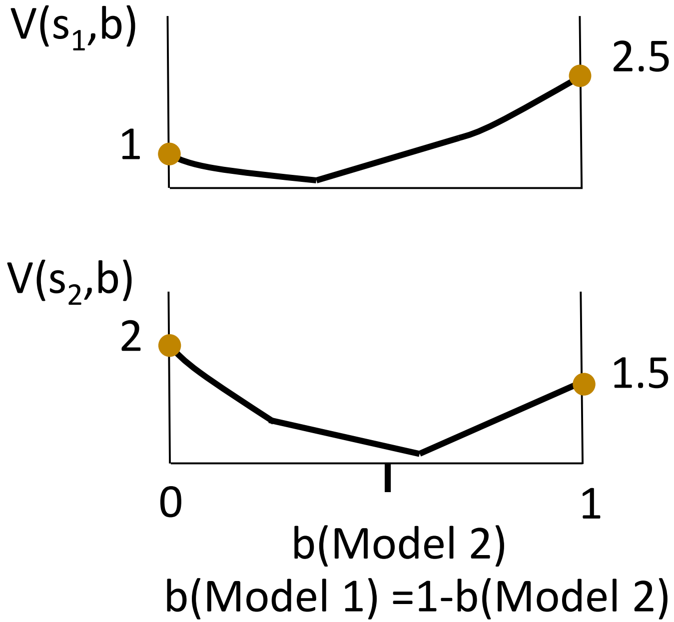
$V(s_1) = 2.5$
$V(s_2) = 1.5$

- Optimal policy $\pi^*: S, B \rightarrow A$
- Optimal value function $V_{\pi*}(s, b)$



From belief MDP to discretized belief MDP

- Continuous MDPs are computationally hard to solve and approximate solution techniques must be used to derive solutions.
- A natural way to overcome this limitation is to discretize the continuous belief state space and solve a discrete state MDP. (Paul)

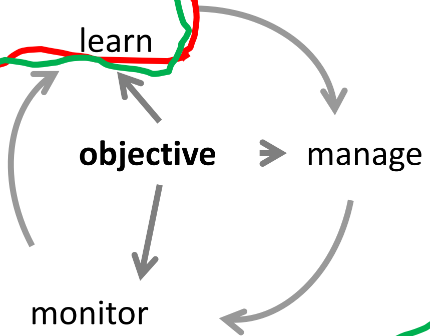


Passive adaptive management provides the best actions given our current knowledge … Learning occurs independently.

Bayes theorem

R. Bellman

Stochastic dynamic programming



heuristics

easier to solve

(certainty equivalence principles)

Passive adaptive management is solving an MDP given a fix $b_t(y)$

- Optimal policy $\pi^*: S \rightarrow A$
- Optimal value function $V_\pi(s)$

$$V^*(s_t) = \max_{a \in A} \left\{ \underbrace{\sum_{y \in Y} b_t(y) r(s_t, a)}_{\text{Expected immediate reward given } s_t,\ a \text{ and } b_t} + \underbrace{\underbrace{\gamma}_{\text{Discount factor}} \underbrace{\sum_{s_{t+1}}}_{\text{future states}} \left\{ \underbrace{\sum_{y \in M} b_t(y)\, \mathrm{P}_y(s_{t+1} | s_t, a)}_{\text{Transition probability given } s_t,\ a \text{ and } b_t} \underbrace{V^*(s_{t+1})}_{\text{Optimal value in } s_{t+1}} \right\}}_{\text{Expected future discounted reward}} \right\}$$

CSIRO

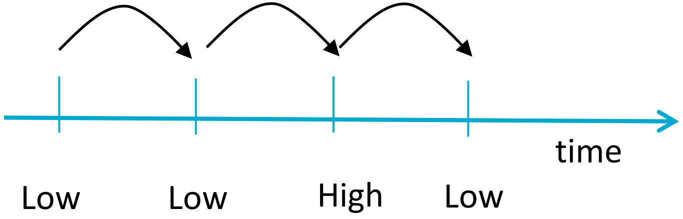
Maximise the “High” persistence state of a Gouldian finch population over time

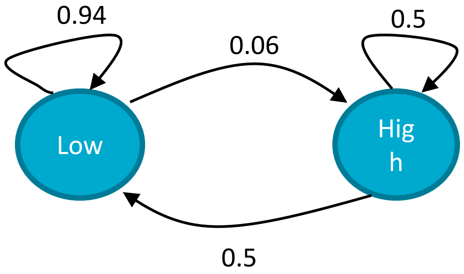
Transition matrix for each action

To

From

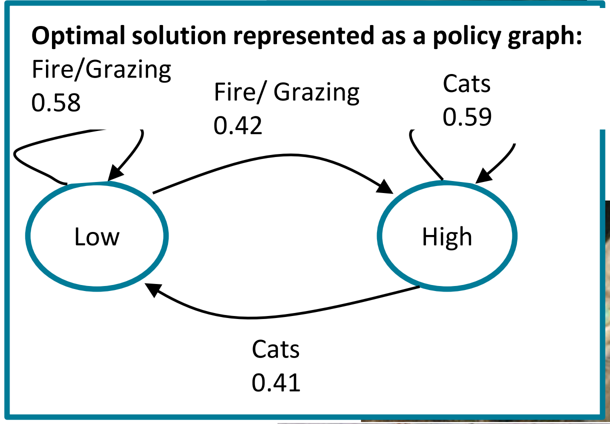
Do nothing	Low	High
Low	0.94	0.06
High	0.5	0.5





An optimal solution gives the best action to perform in each state
Same cost for each action.

do nothing



To

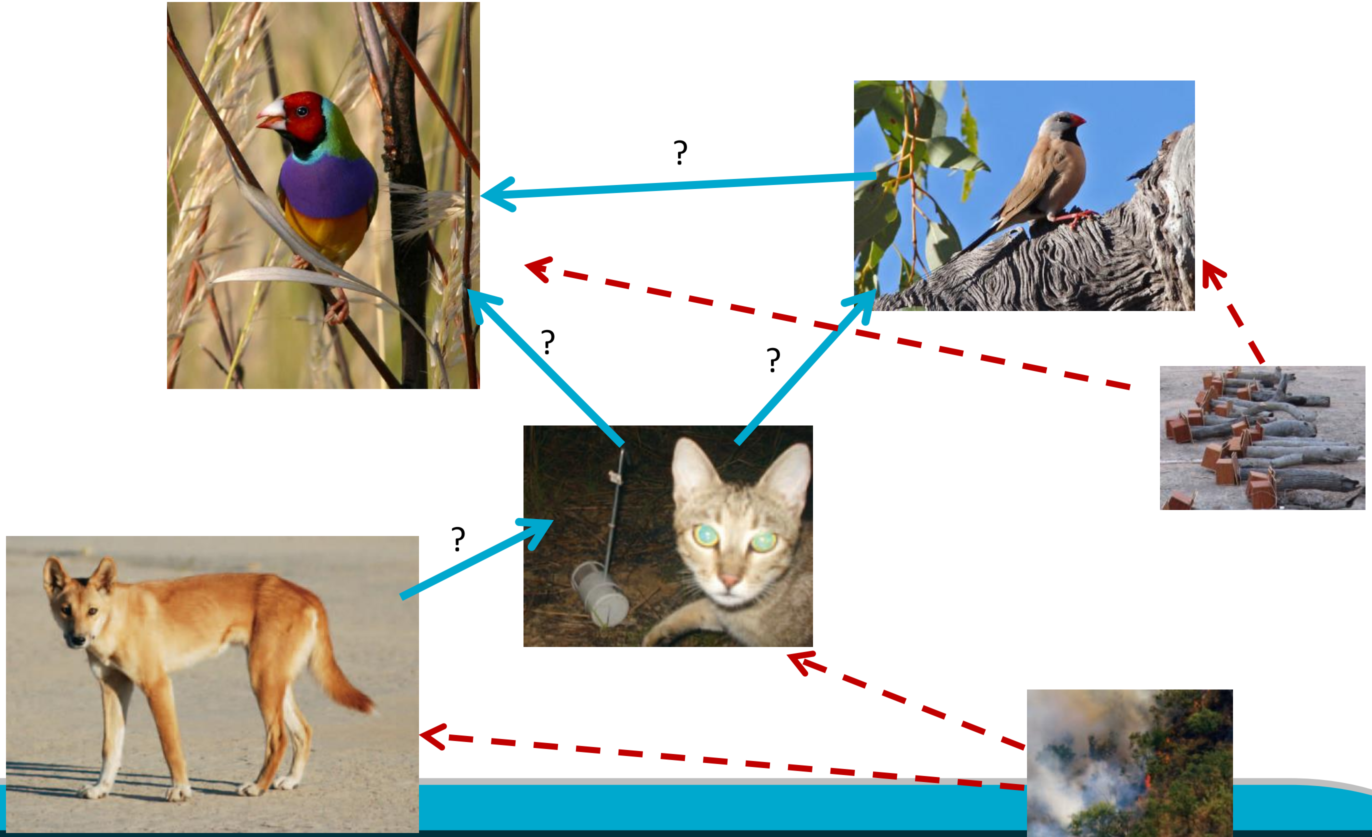
From: Do nothing	Low	High
Low	0.94	0.06
High	0.5	0.5

From: Fire Grazing	Low	High
Low	0.58	**0.42**
High	0.45	0.55

From: Cats	Low	High
Low	0.62	0.38
High	0.41	**0.59**

From: Nesting box	Low	High
Low	0.73	0.27
High	0.47	0.53

It is rare to know the efficiency of actions when managing endangered species



That’s where adaptive management becomes useful

We can define/elicit a set of plausible models

We assume that the REAL model is one of these models.

We are uncertain about which model is the real model but we can observe the state perfectly

We asked **4 anonymous experts** to provide their state transition model.

Expert 1

Nesting box	Low	High
Low		

Cats	Low	High

Fire/	Low	High

Do nothing	Low	High
Low	0.94	0.06
High	0.5	0.5

Expert 2

Nesting box	Low

Cats	Low

Fire/ grazing	Low

Do nothing	Low	Hig
Low	0.94	0.0
High	0.6	0.4

Expert 3

Nesting box	Low	High
Low		

Cats	Low	Hi

Fire/	Low	High

Do nothing	Low	High
Low	0.84	0.16
High	0.2	0.8

Expert 4

Nesting box	Low	High
Low		

Cats	Low	High

Fire/	Low	High

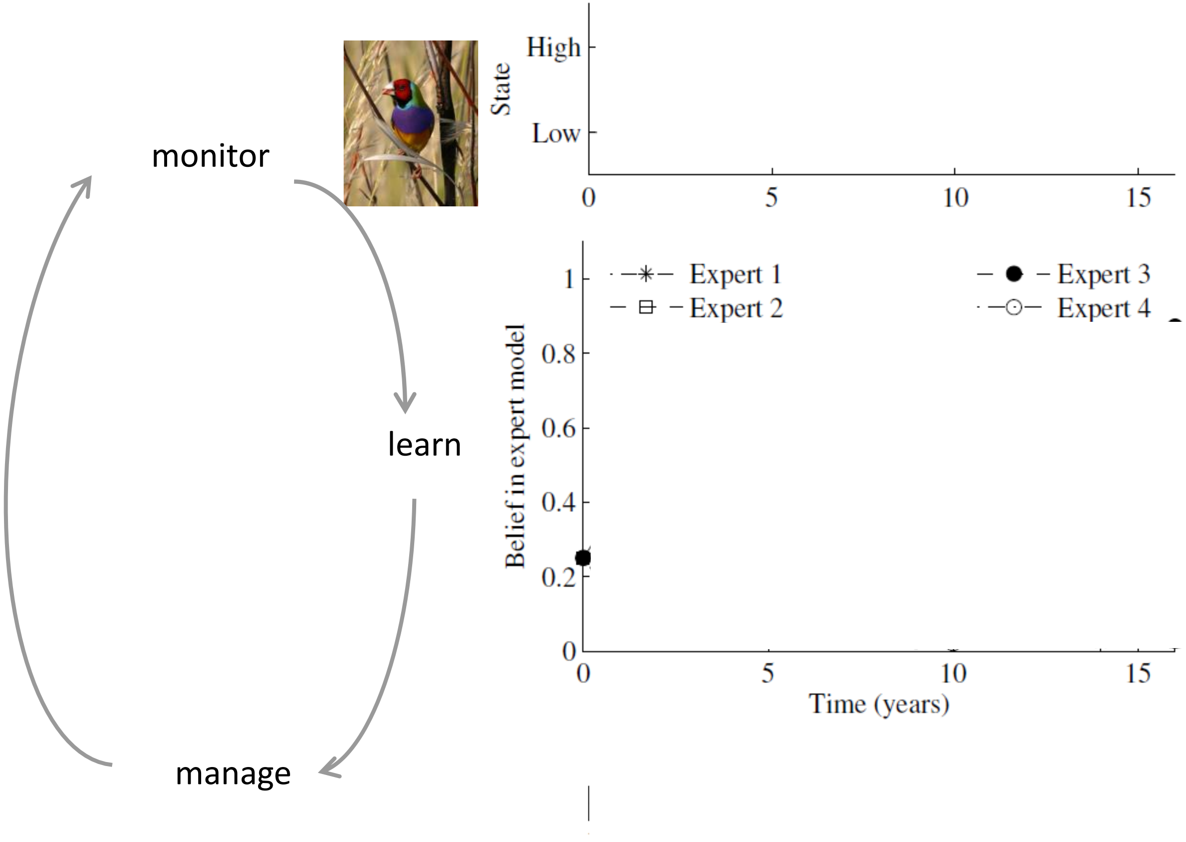
Do nothing	Low	High
Low	0.9	0.1
High	0.5	0.5

We use belief state to represent "where we think we are" at anytime.

	Expert 1	Expert 2	Expert 3	Expert 4
$b_t(y)$	0.25	0.25	0.25	0.25
$b_{t+1}(y)$	0.1	0.15	0.65	0.1

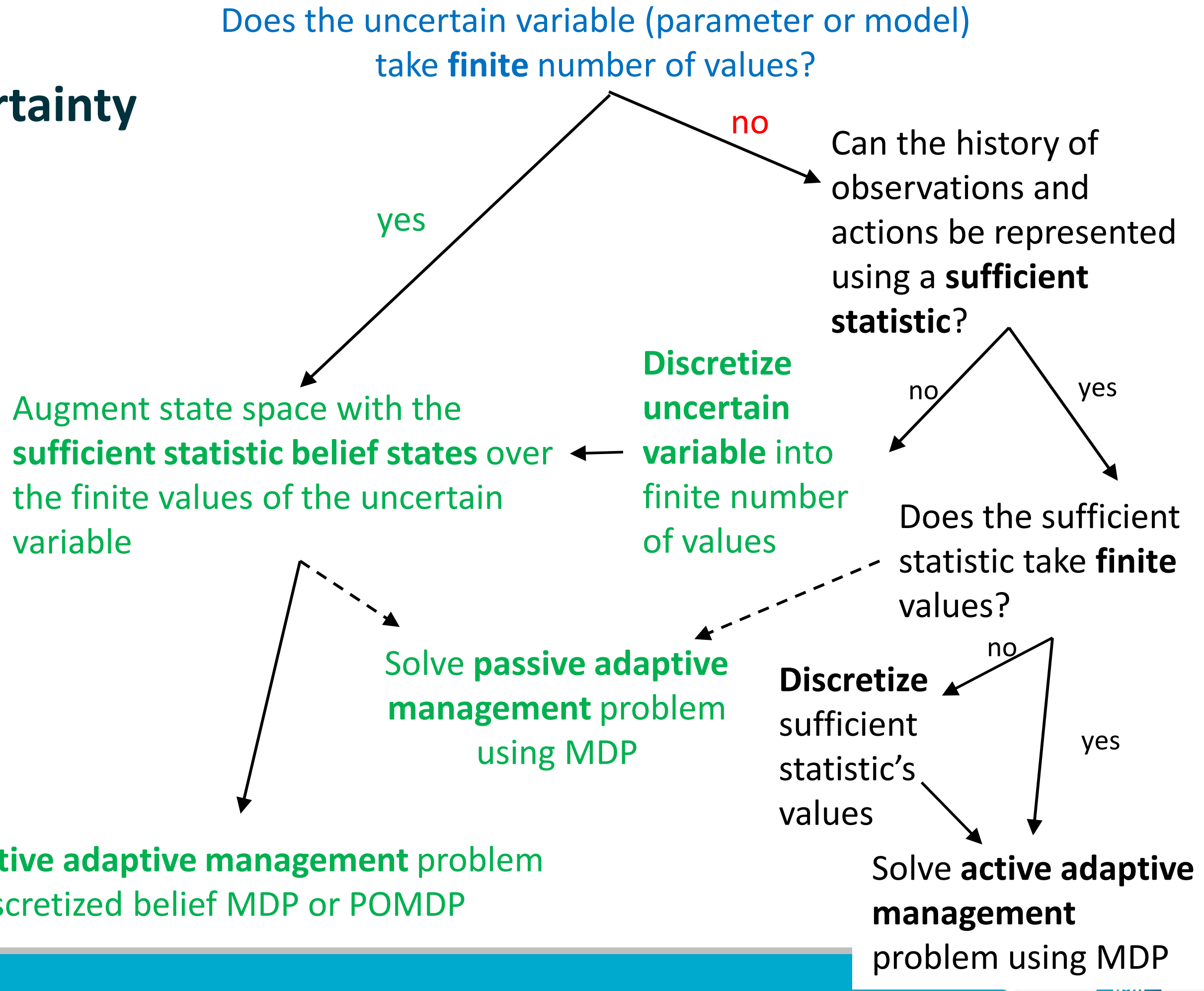
What is the best conservation action to perform for **every combination of state of the species and belief in models** to maximize a species abundance?

What are the best conservation actions to perform when expert 3 is the true model?



Discrete approach is also be relevant for parameter uncertainty

- **Convenient sufficient statistics may not exist** distinguish between a small number of values of a single parameter (McDonald-Madden et al. 2010b; Moore et al. 2011; Runge 2013)
- **Multiple parameters** are uncertain and key hypotheses need to be tested, model uncertainty is currently the only tractable approach (Williams 2009) (Moore et al. 2008).



CSIRO

Important caveats

- One of the candidate models must be close to the true model (sub-optimal).
- The computational complexity is proportional to the number of models.
- Objective is to maximise the expected sum of rewards not finding out which model is true.
- Value of information: Models must be different enough to require alternative optimal management strategies (and outcomes). Selecting the minimum set of models?
- Markov property (delayed benefits, complex life cycles).
- Interpretation, explanation and communication!