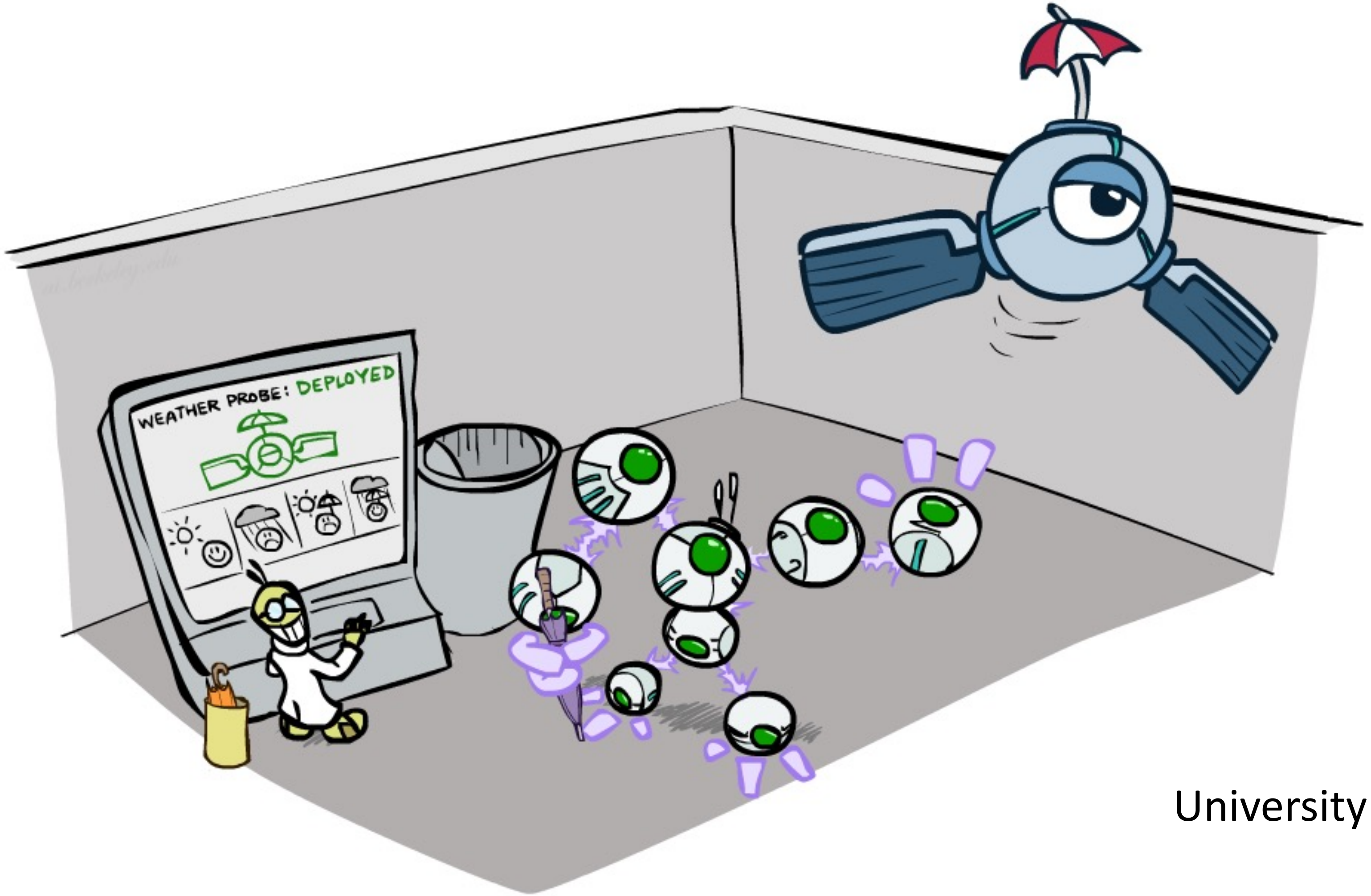
CS 188: Artificial Intelligence

Decision Networks and Value of Information

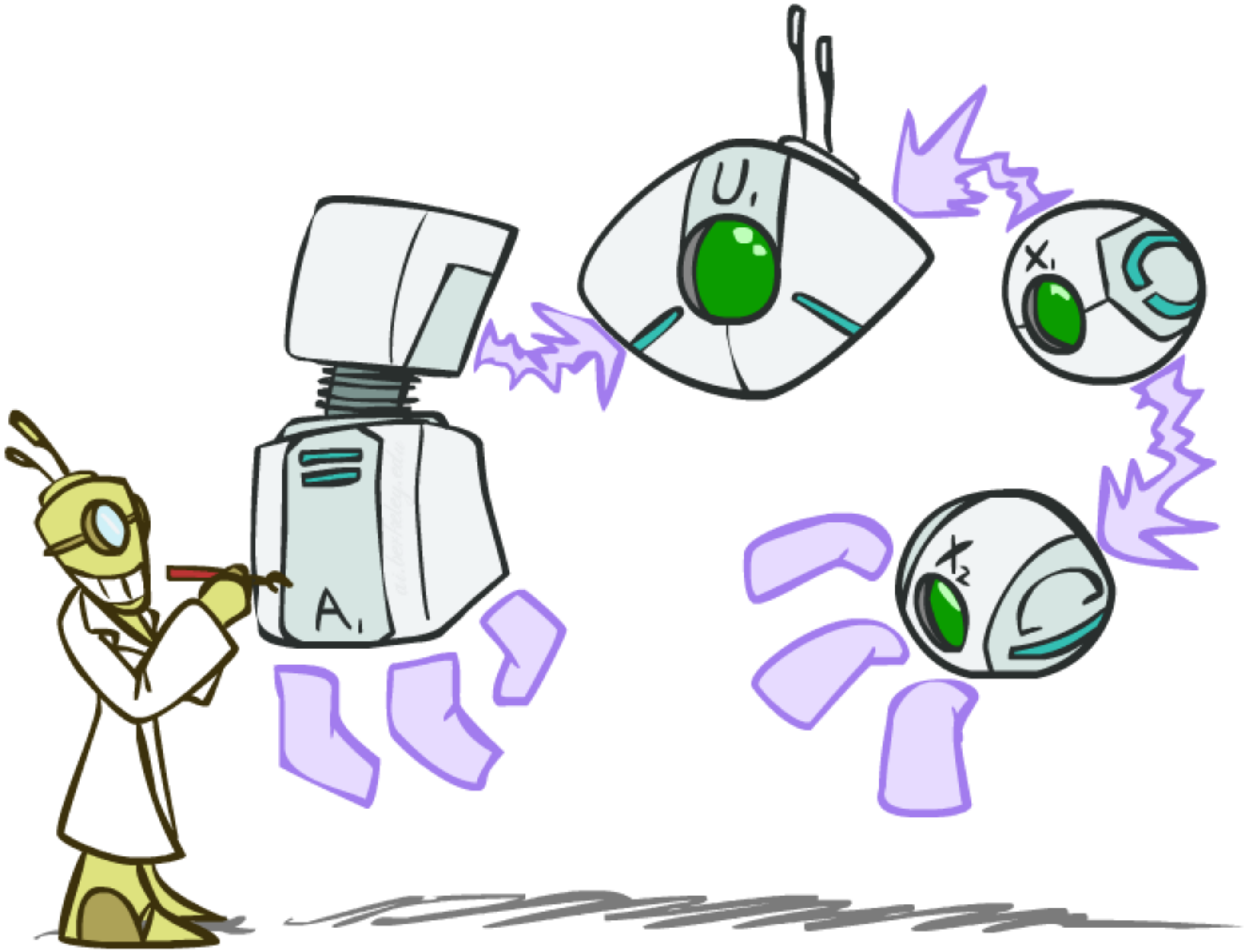


University of California, Berkeley

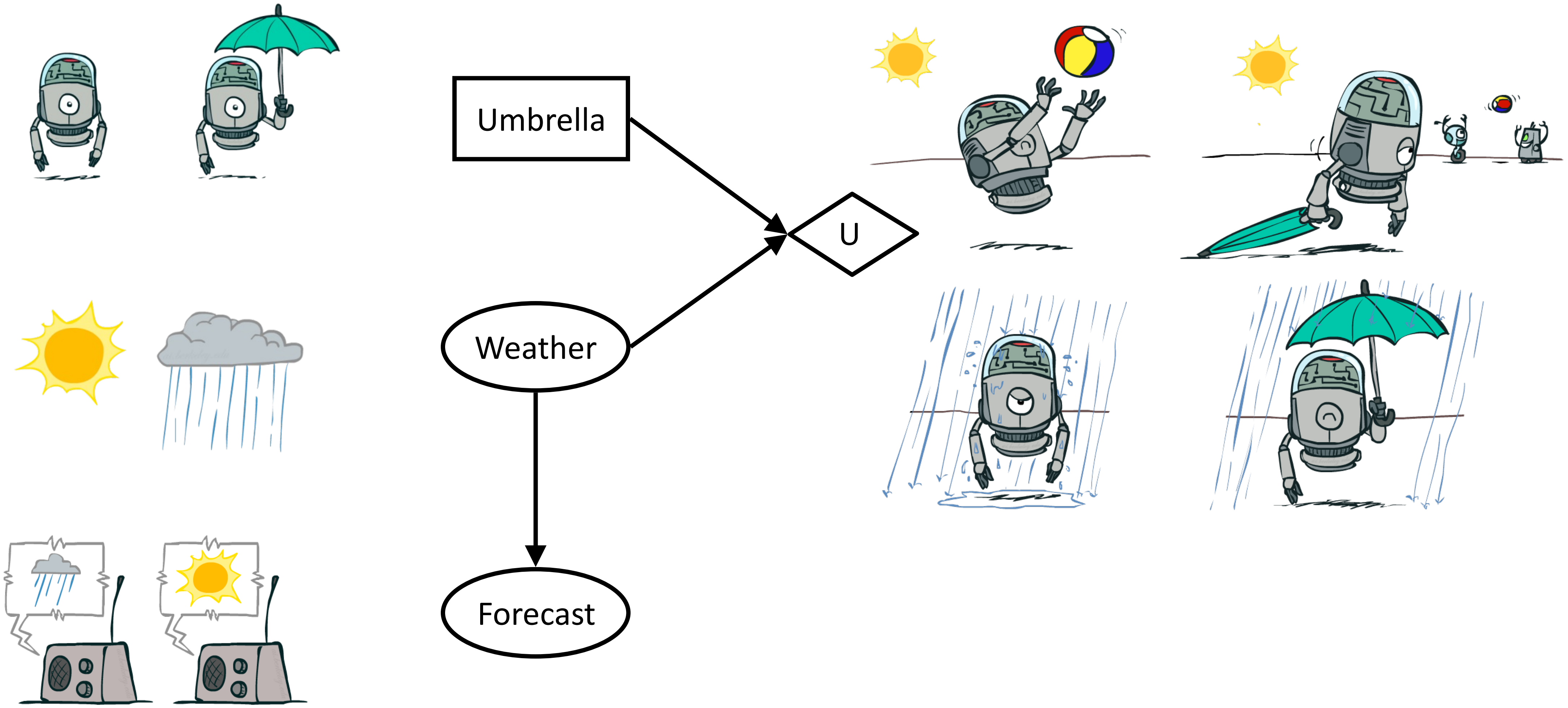
Today's Topics

- Decision Networks
- Value of Information
- (Briefly) Partially Observable MDPs

Decision Networks



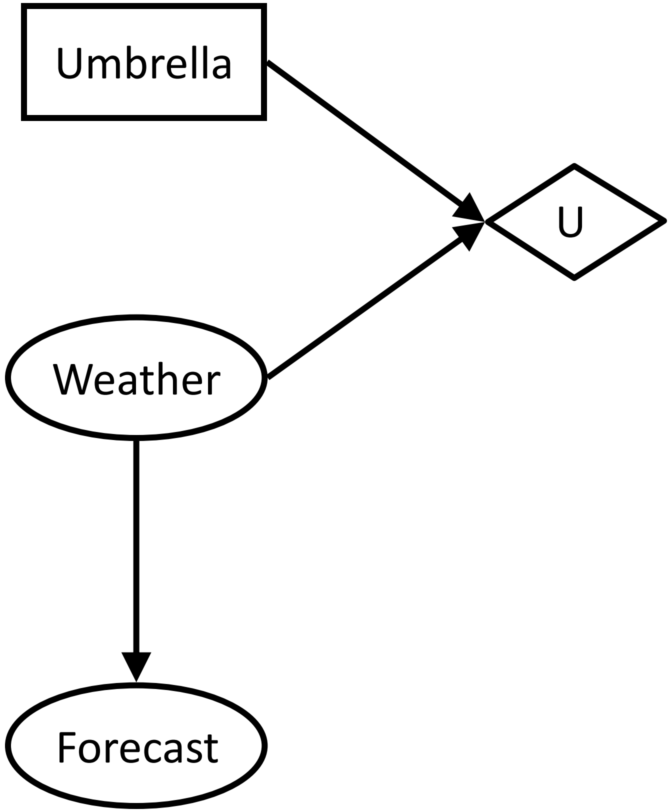
Decision Networks

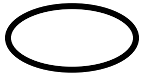
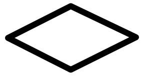


Decision Networks

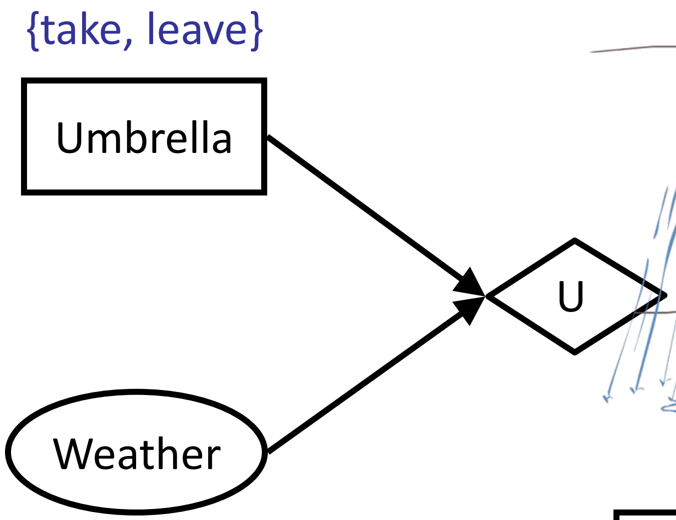
- **Maximum Expected Utility (MEU):**
  choose the action which maximizes the expected utility given the evidence

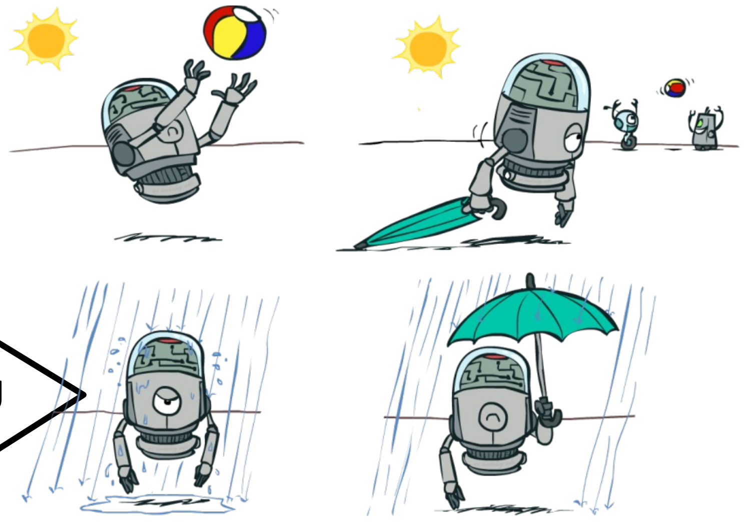
- Can directly operationalize this with decision networks
  - Bayes nets with nodes for utility and actions
  - Lets us calculate the expected utility for each action



- New node types:
  - Chance nodes (just like BNs) 
  - Actions (rectangles, cannot have parents, act as observed evidence) 
  - Utility node (diamond, depends on action and chance nodes) 

Decision Networks



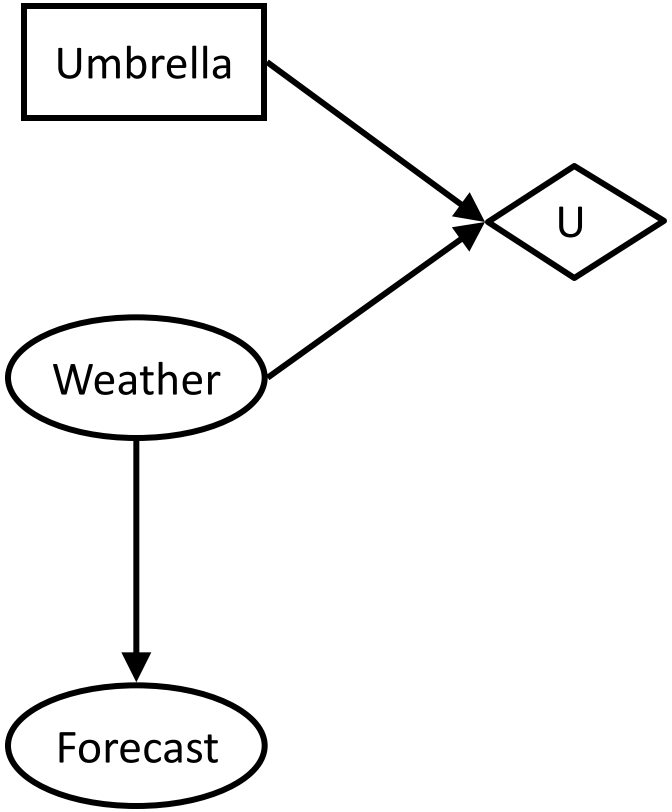


W	P(W)
sun	0.7
rain	0.3

A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

Action Selection in Decision Networks

- Instantiate all evidence
- Set action node(s) each possible way
- Calculate posterior for all parents of utility node, given the evidence
- Calculate expected utility for each action
- Choose maximizing action



Example: Decision Networks

Umbrella = leave

$$\mathrm{EU}(\mathrm{leave}) = \sum_w P(w)U(\mathrm{leave}, w)$$

$$= 0.7 \cdot 100 + 0.3 \cdot 0 = 70$$

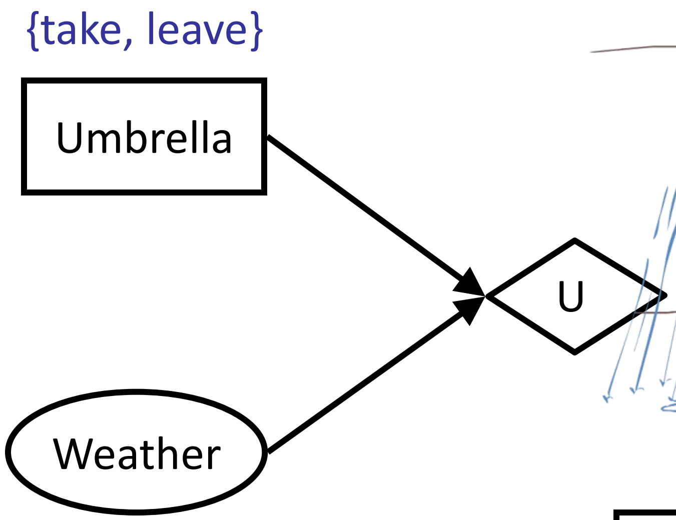
Umbrella = take

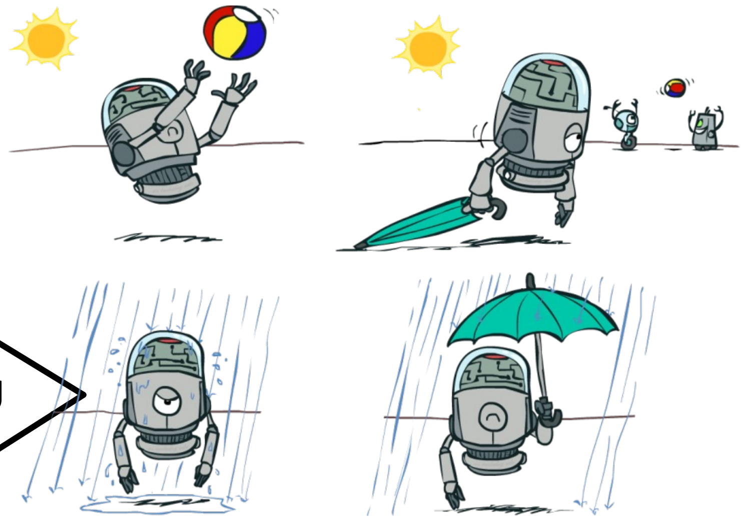
$$\mathrm{EU}(\mathrm{take}) = \sum_w P(w)U(\mathrm{take}, w)$$

$$= 0.7 \cdot 20 + 0.3 \cdot 70 = 35$$

Optimal decision = leave

$$\mathrm{MEU}(\emptyset) = \max_a \mathrm{EU}(a) = 70$$

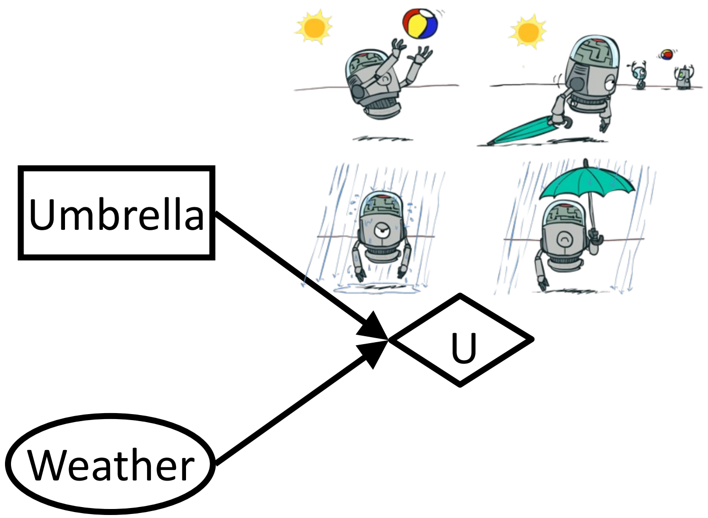


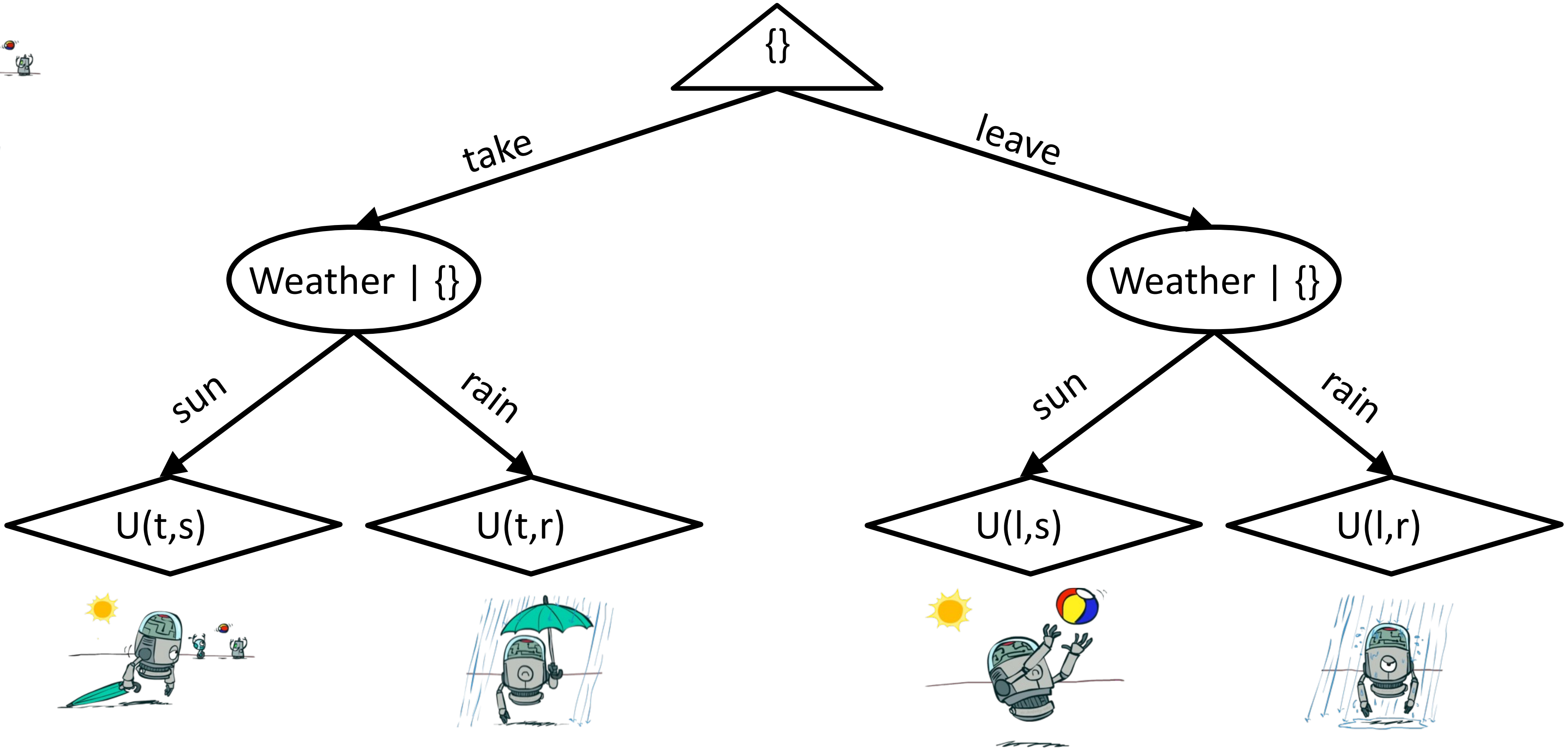


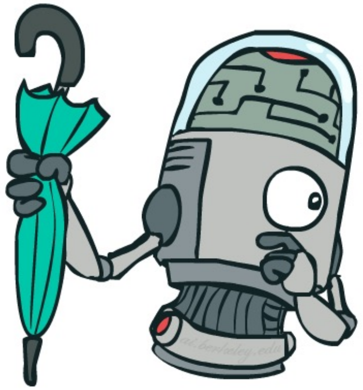
W	P(W)
sun	0.7
rain	0.3

A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

Decisions as Outcome Trees







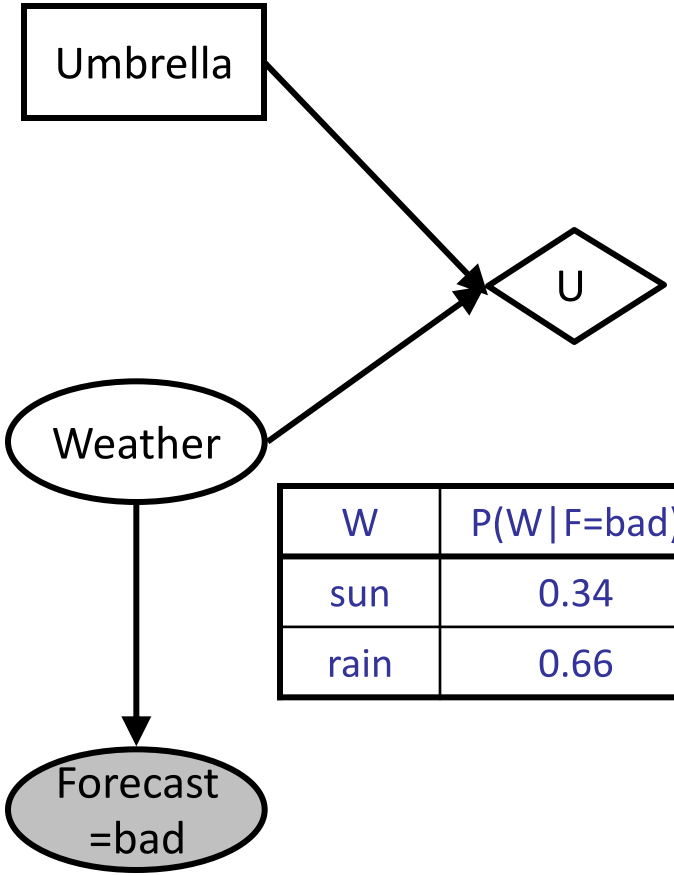
- Almost exactly like expectimax / MDPs
- What’s changed?

Example: Decision Networks

$$P(W) \qquad P(F|W)$$

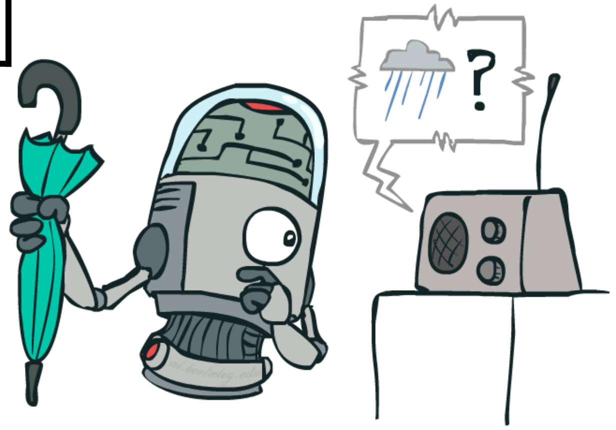
$$P(W|F) = \frac{P(W,F)}{\sum_w P(w,F)}$$

$$= \frac{P(F|W)P(W)}{\sum_w P(F|w)P(w)}$$



A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

W	P(W|F=bad)
sun	0.34
rain	0.66



Example: Decision Networks

Umbrella = leave

$$\mathrm{EU}(\mathrm{leave}|\mathrm{bad}) = \sum_w P(w|\mathrm{bad}) U(\mathrm{leave}, w)$$

$$= 0.34 \cdot 100 + 0.66 \cdot 0 = 34$$

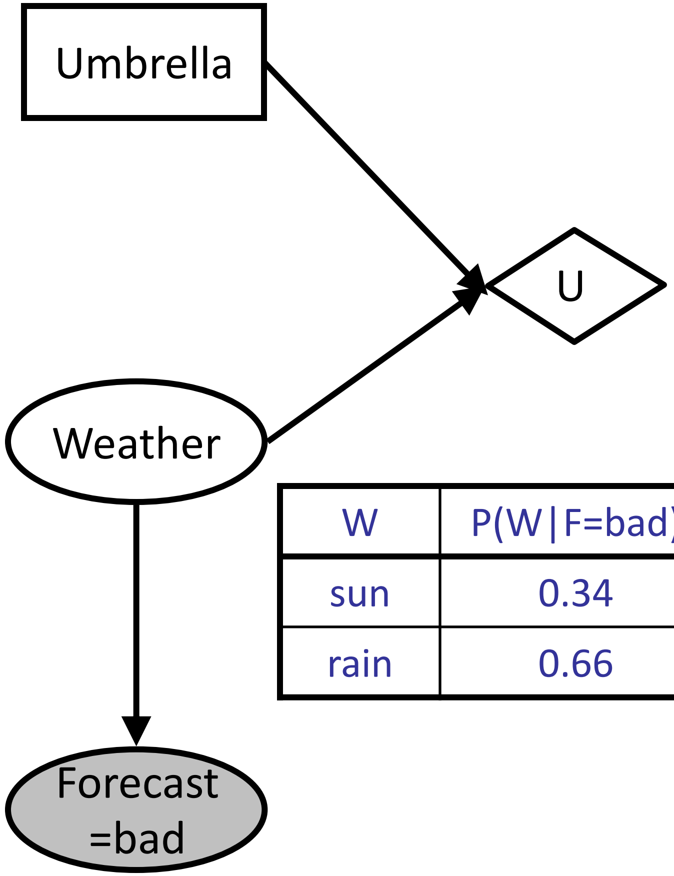
Umbrella = take

$$\mathrm{EU}(\mathrm{take}|\mathrm{bad}) = \sum_w P(w|\mathrm{bad}) U(\mathrm{take}, w)$$

$$= 0.34 \cdot 20 + 0.66 \cdot 70 = 53$$

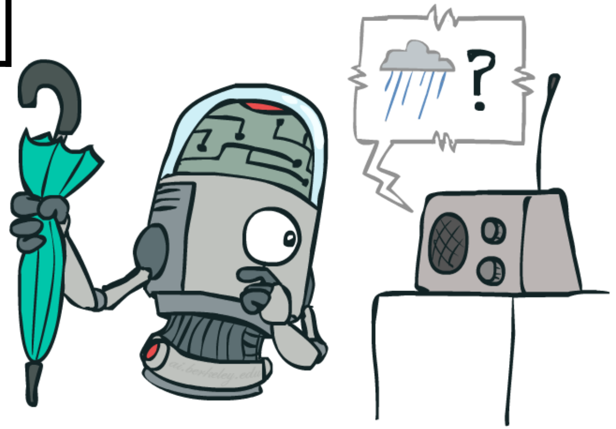
Optimal decision = take

$$\mathrm{MEU}(F = \mathrm{bad}) = \max_a \mathrm{EU}(a|\mathrm{bad}) = 53$$

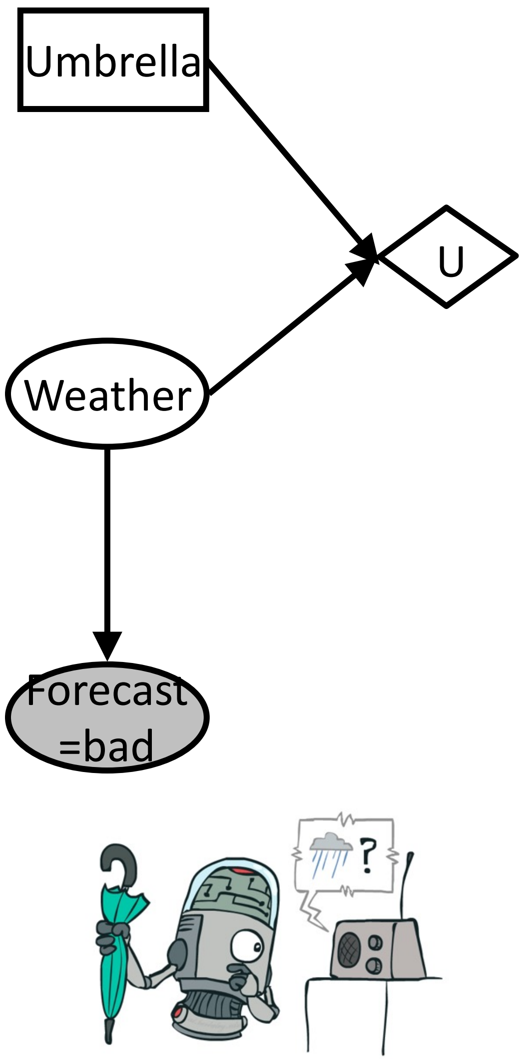


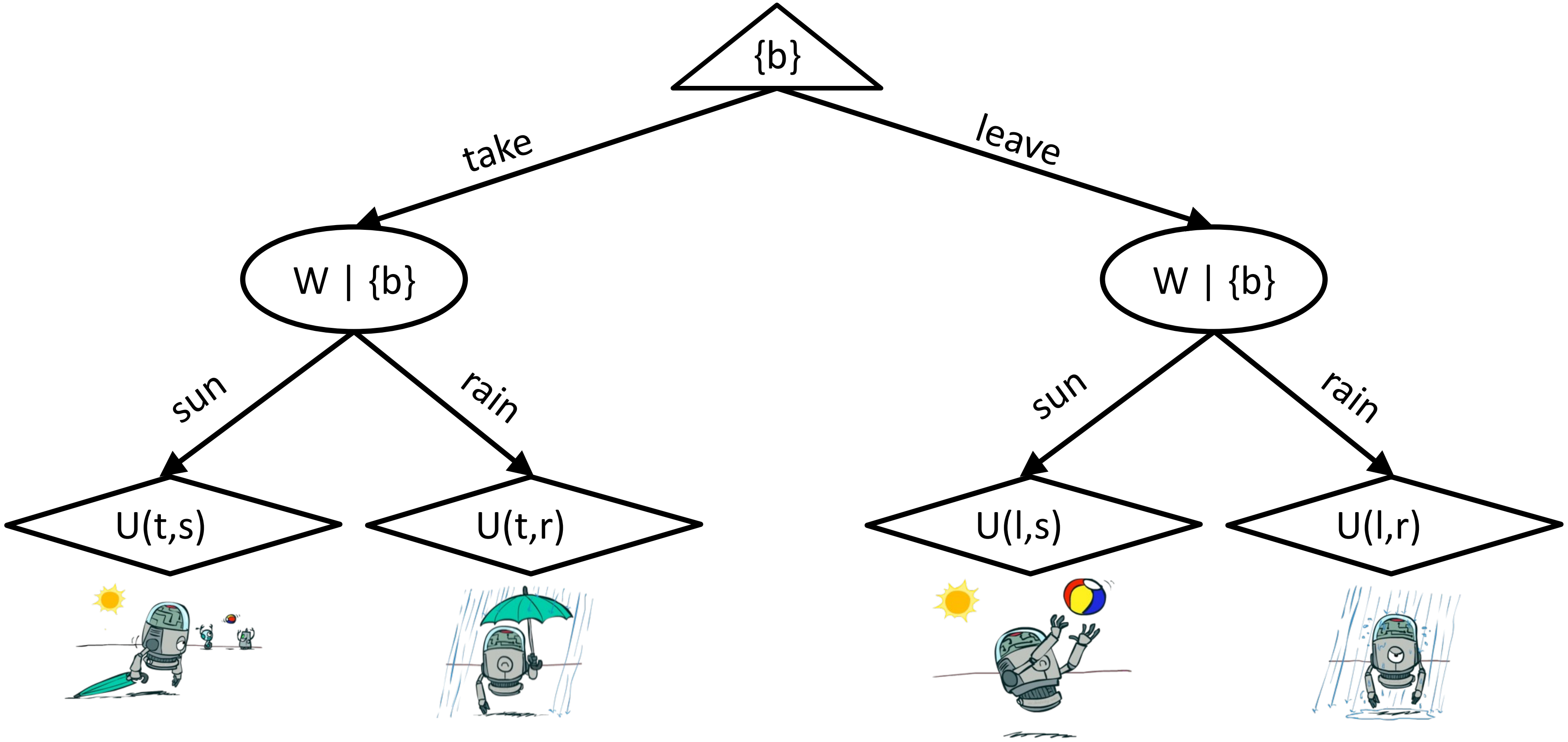
A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

W	P(W|F=bad)
sun	0.34
rain	0.66



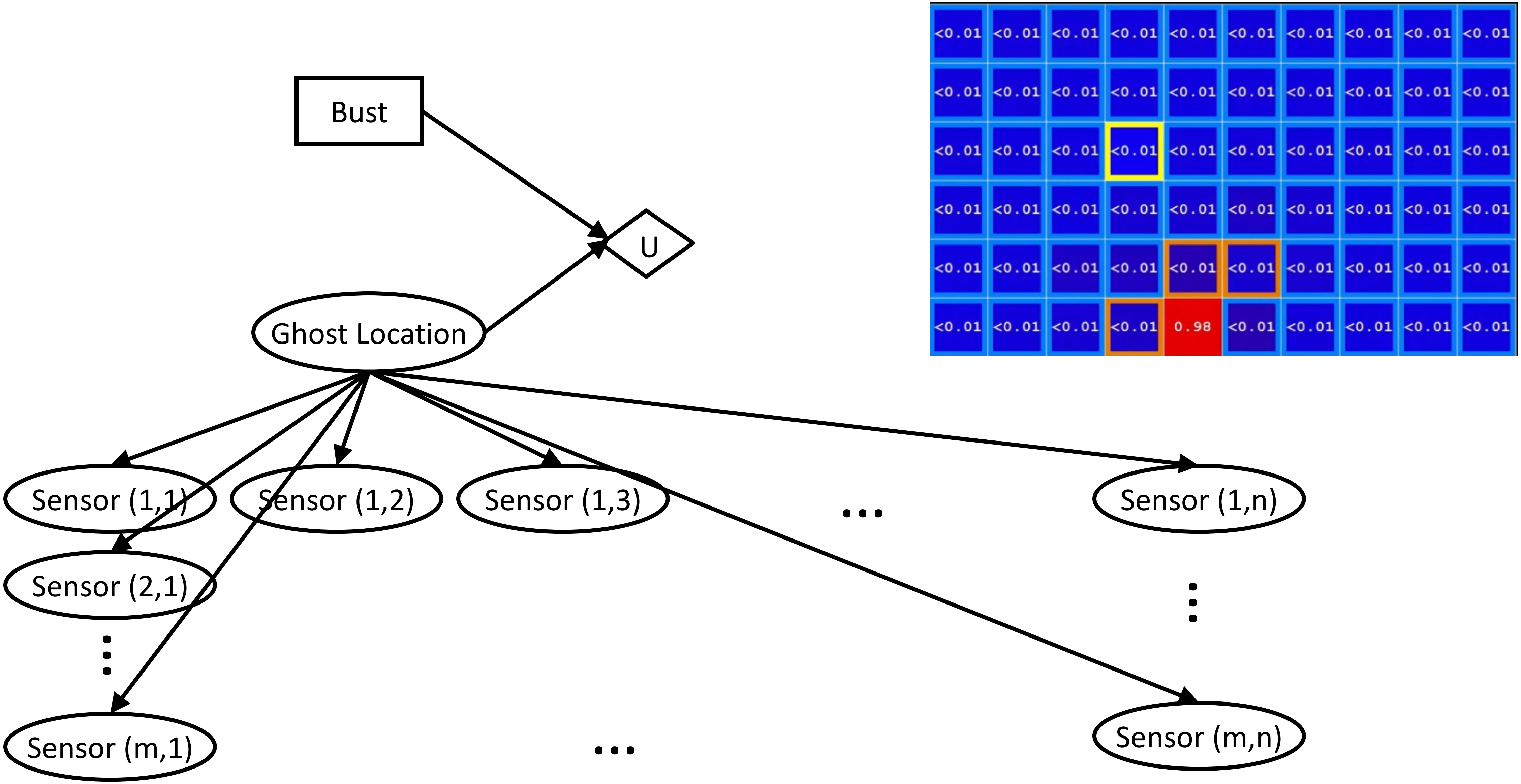
Decisions as Outcome Trees





Video of Demo Ghostbusters with Probability

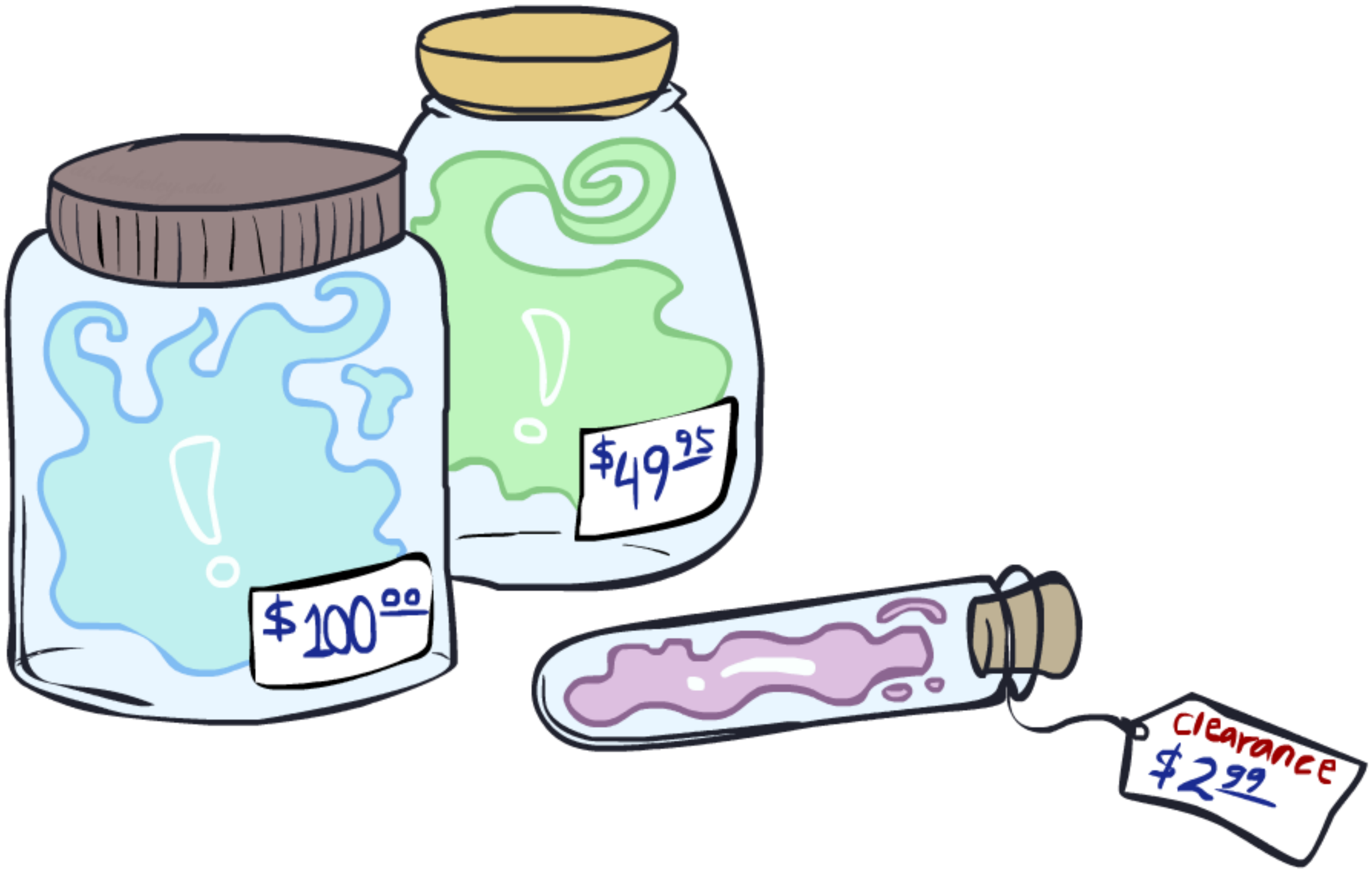
Ghostbusters Decision Network



Today’s Topics

- Decision Networks
- Value of Information
- (Briefly) Partially Observable MDPs

Value of Information



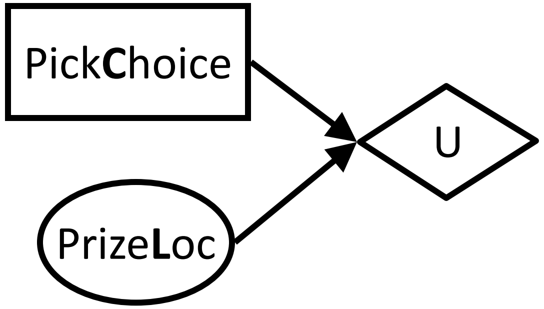
Value of Information

- Idea: compute value of acquiring evidence
  - Can be done directly from decision network

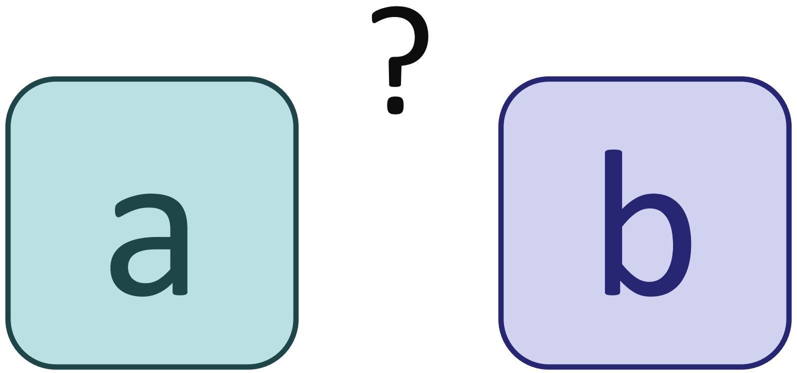
- Example: picking a box with a prize
  - Two boxes **a** and **b**, exactly one has prize, worth k
  - You can pick one box
  - Prior prize probabilities 0.5 each, & mutually exclusive
  - Picking either **a** or **b** has EU = k/2, MEU = k/2

- Question: what's the value of information of O?
  - Value of knowing which of **a** or **b** has prize
  - Value is expected gain in MEU from new info
  - Survey may say "prize in **a**" or "prize in **b**", prob 0.5 each
  - If we know **PrizeLoc**, MEU is k (either way)
  - Gain in MEU from knowing **PrizeLoc**?
  - VPI(**PrizeLoc**) = k – k/2 = k/2
  - Fair price of information: k/2

L	P
a	1/2
b	1/2



C	L	U
a	a	k
a	b	0
b	a	0
b	b	k



VPI Example: Weather

MEU with no evidence

$$\text{MEU}(\phi) = \max_a \text{EU}(a) = 70$$

MEU if forecast is bad

$$\text{MEU}(F = \text{bad}) = \max_a \text{EU}(a|\text{bad}) = 53$$

MEU if forecast is good

$$\text{MEU}(F = \text{good}) = \max_a \text{EU}(a|\text{good}) = 95$$

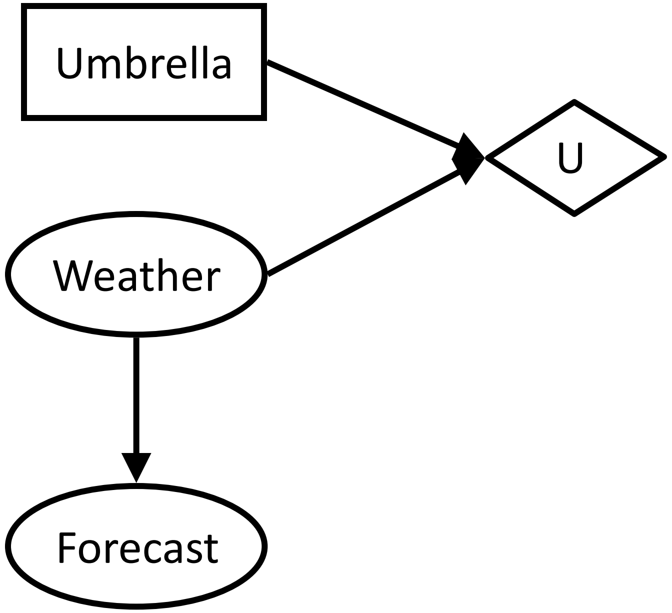
Forecast distribution

F	P(F)
good	0.59
bad	0.41

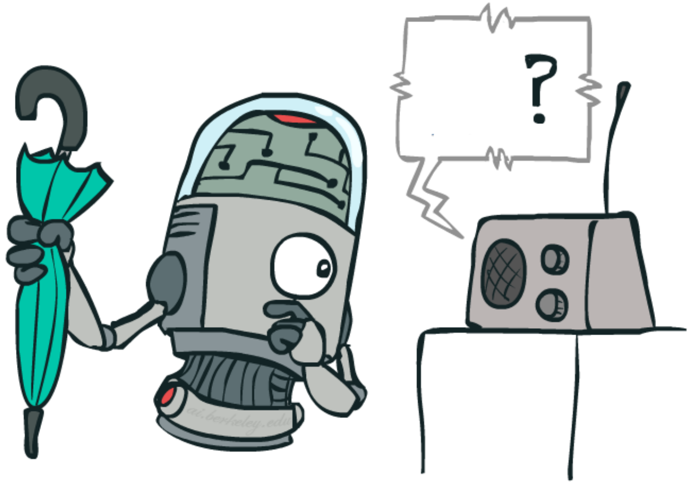
$$0.59 \cdot (95) + 0.41 \cdot (53) - 70$$

$$77.8 - 70 = 7.8$$

$$\text{VPI}(E'|e) = \left(\sum_{e'} P(e'|e)\text{MEU}(e, e')\right) - \text{MEU}(e)$$



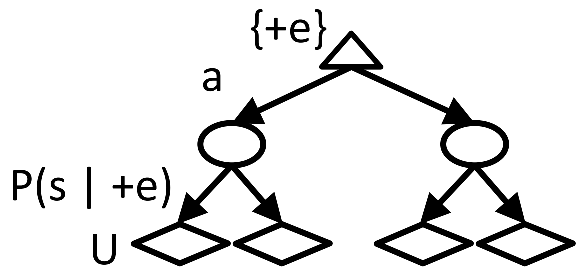
A	W	U
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70



Value of Information

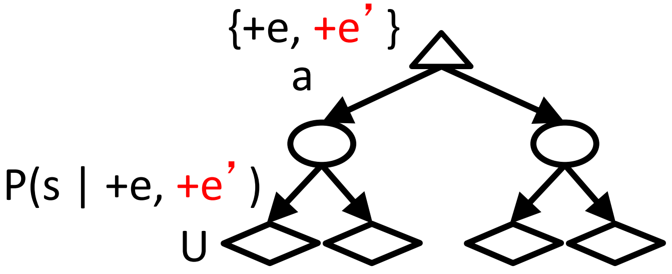
- Assume we have evidence E=e. Value if we act now:

$$\text{MEU}(e) = \max_a \sum_s P(s|e)\, U(s,a)$$



- Assume we see that E' = e'. Value if we act then:

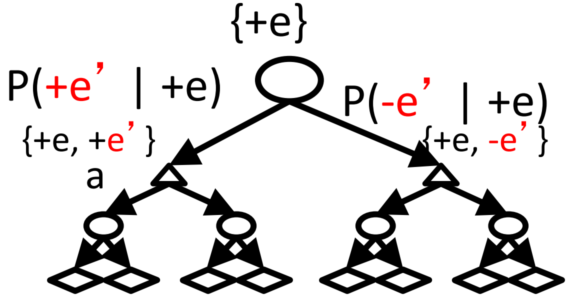
$$\text{MEU}(e, e') = \max_a \sum_s P(s|e, e')\, U(s,a)$$



- BUT E' is a random variable whose value is unknown, so we don't know what e' will be

- Expected value if E' is revealed and then we act:

$$\text{MEU}(e, E') = \sum_{e'} P(e'|e)\text{MEU}(e, e')$$



- Value of information: how much MEU goes up by revealing E' first then acting, over acting now:

$$\text{VPI}(E'|e) = \text{MEU}(e, E') - \text{MEU}(e)$$

VPI Properties

- Nonnegative

$$\forall E', e : \text{VPI}(E'|e) \geq 0$$

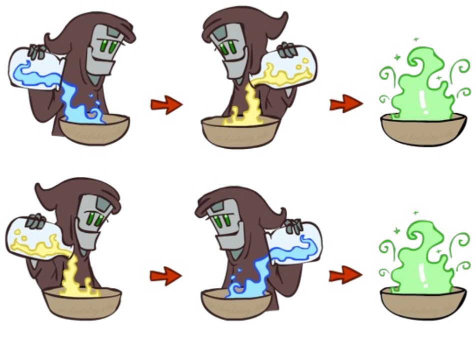
- Nonadditive

  (think of observing E_j twice)

$$\text{VPI}(E_j, E_k|e) \neq \text{VPI}(E_j|e) + \text{VPI}(E_k|e)$$

- Order-independent

$$\begin{aligned}\text{VPI}(E_j, E_k|e) &= \text{VPI}(E_j|e) + \text{VPI}(E_k|e, E_j)\\ &= \text{VPI}(E_k|e) + \text{VPI}(E_j|e, E_k)\end{aligned}$$



Value of Information

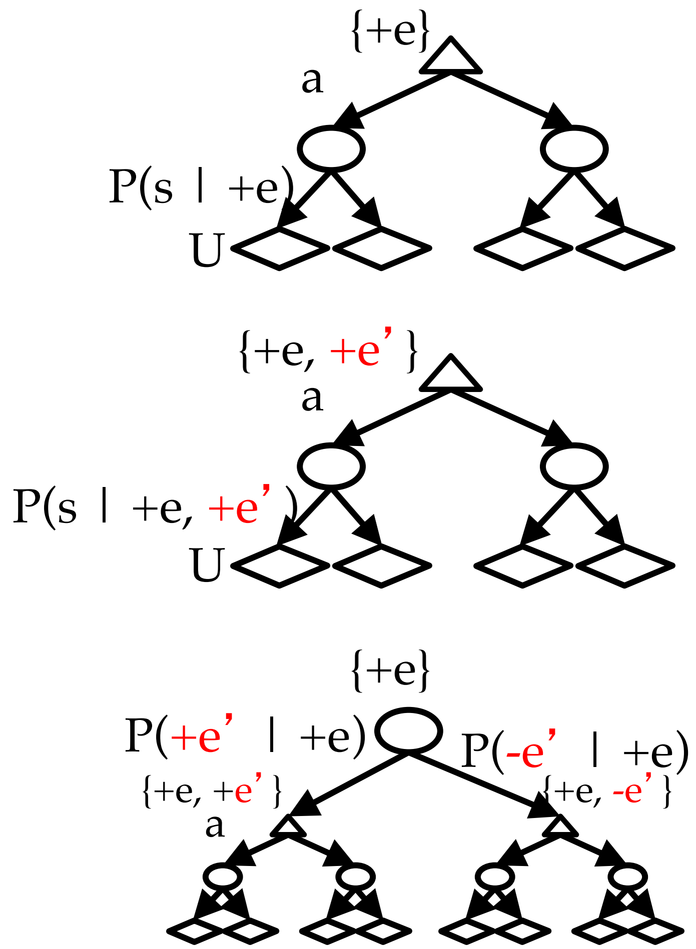
$$\text{MEU}(e, E') = \sum_{e'} P(e'|e)\text{MEU}(e, e')$$

$$= \sum_{e'} P(e'|e) \max_a \sum_s P(s|e, e')U(s, a)$$

$$\text{MEU}(e) = \max_a \sum_s P(s|e)\, U(s, a)$$

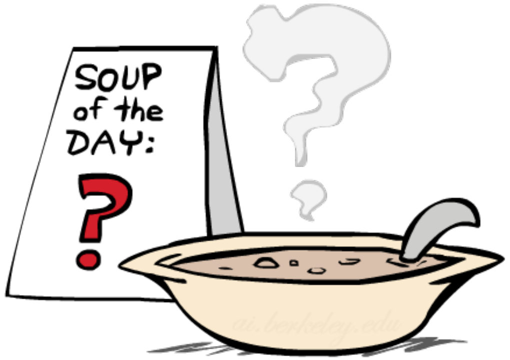
$$= \max_a \sum_{e'} \sum_s P(s, e'|e)U(s, a)$$

$$= \max_a \sum_{e'} P(e'|e) \sum_s P(s|e, e')U(s, a)$$

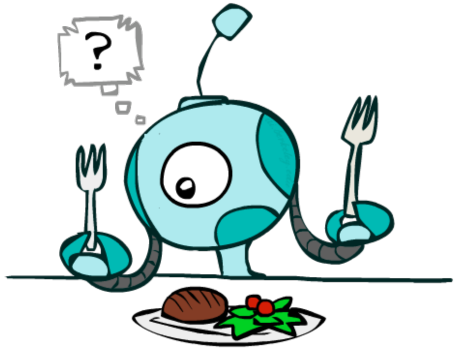


Quick VPI Questions

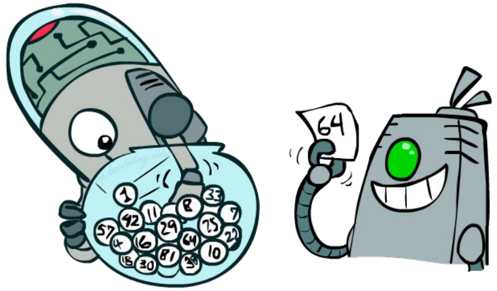
- The soup of the day is either clam chowder or split pea, but you wouldn't order either one. What's the value of knowing which it is?
  - Not valuable / slightly valuable / highly valuable?



- There are two kinds of plastic forks at a picnic. One kind is slightly sturdier. What's the value of knowing which?
  - Not valuable / slightly valuable / highly valuable?

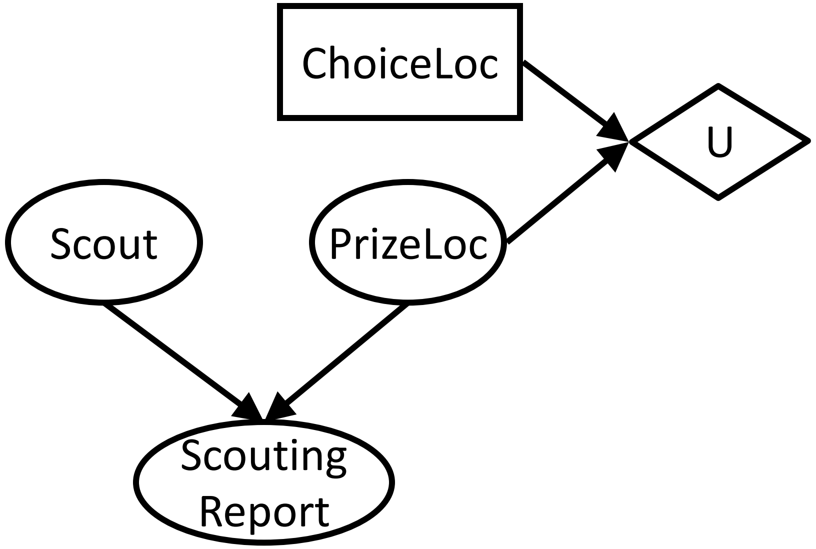


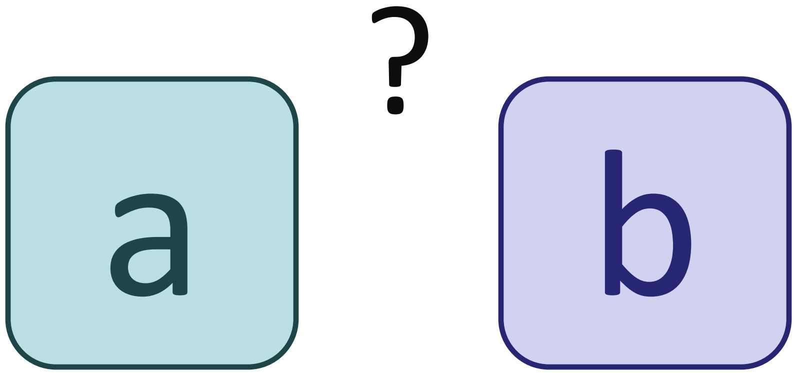
- You're playing the lottery. The prize will be \$0 or \$100. You can play any number between 1 and 100 (chance of winning is 1%). What is the value of knowing the winning number?
  - Not valuable / slightly valuable / highly valuable?



VPI Question

- VPI(PrizeLoc) ?
- VPI(ScoutingReport) ?
- VPI(Scout) ?
- VPI(Scout | ScoutingReport) ?





- Generally:

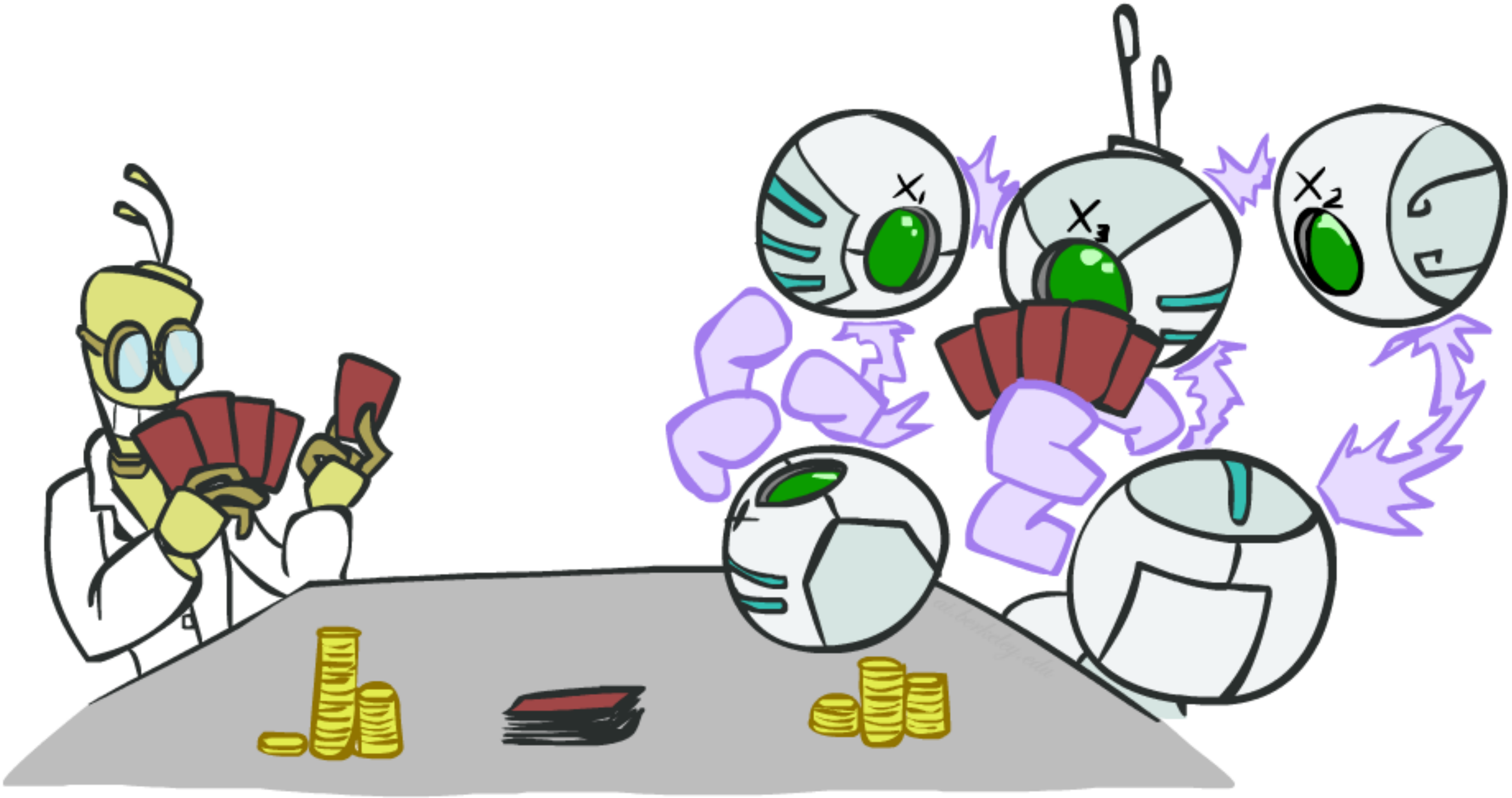
  If $\quad$ Parents(U) $\perp\!\!\!\perp$ Z | CurrentEvidence

  Then $\quad$ VPI(Z | CurrentEvidence) = 0

Today’s Topics

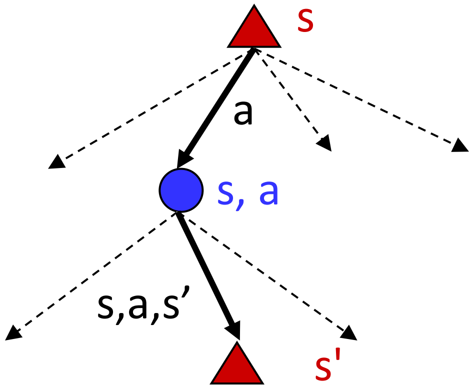
- Decision Networks
- Value of Information
- (Briefly) Partially Observable MDPs

Partially Observable MDPs (POMDPs)

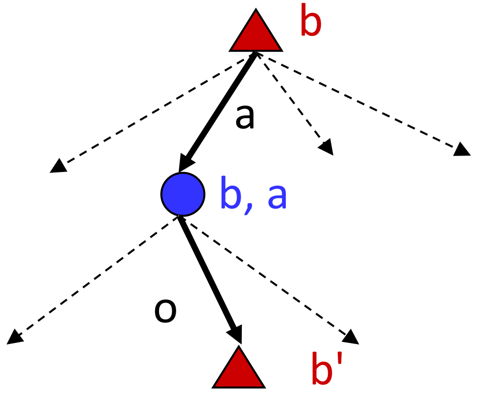


POMDPs

- MDPs have:
  - States S
  - Actions A
  - Transition function P(s’|s,a) (or T(s,a,s’))
  - Rewards R(s,a,s’)



- POMDPs add:
  - Observations O
  - Observation function P(o|s) (or O(s,o))

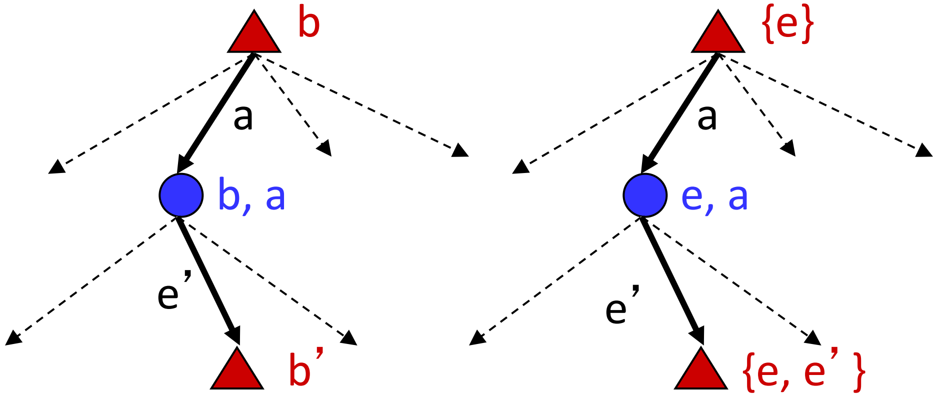


- POMDPs are MDPs over belief states b (distributions over S)

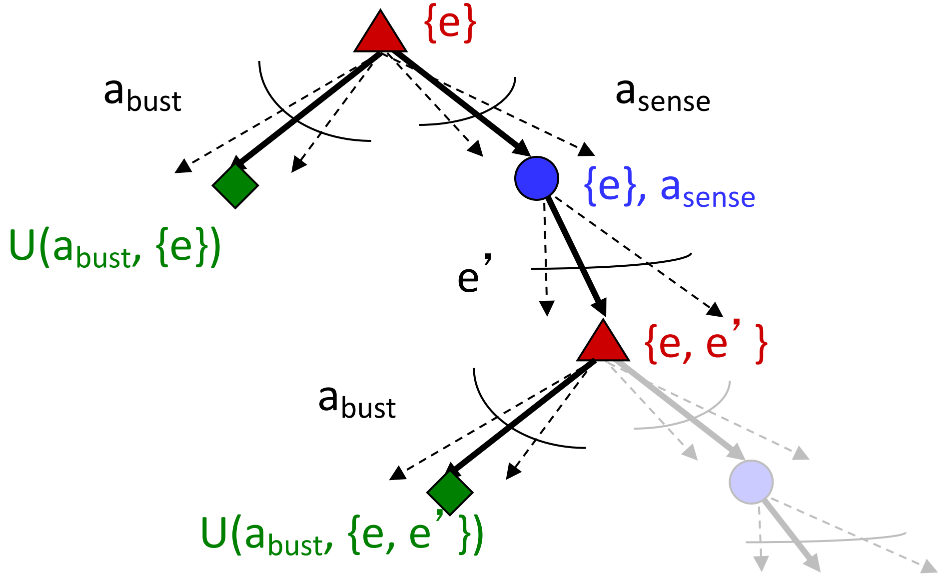
- We’ll be able to say more in a few lectures

Example: Ghostbusters

- In (static) Ghostbusters:
  - Belief state determined by evidence to date {e}
  - Tree really over evidence sets
  - Probabilistic reasoning needed to predict new evidence given past evidence



- Solving POMDPs
  - One way: use truncated expectimax to compute approximate value of actions
  - What if you only considered busting or one sense followed by a bust?
  - You get a VPI-based agent!



Video of Demo Ghostbusters with VPI

Next Time: Dynamic Models