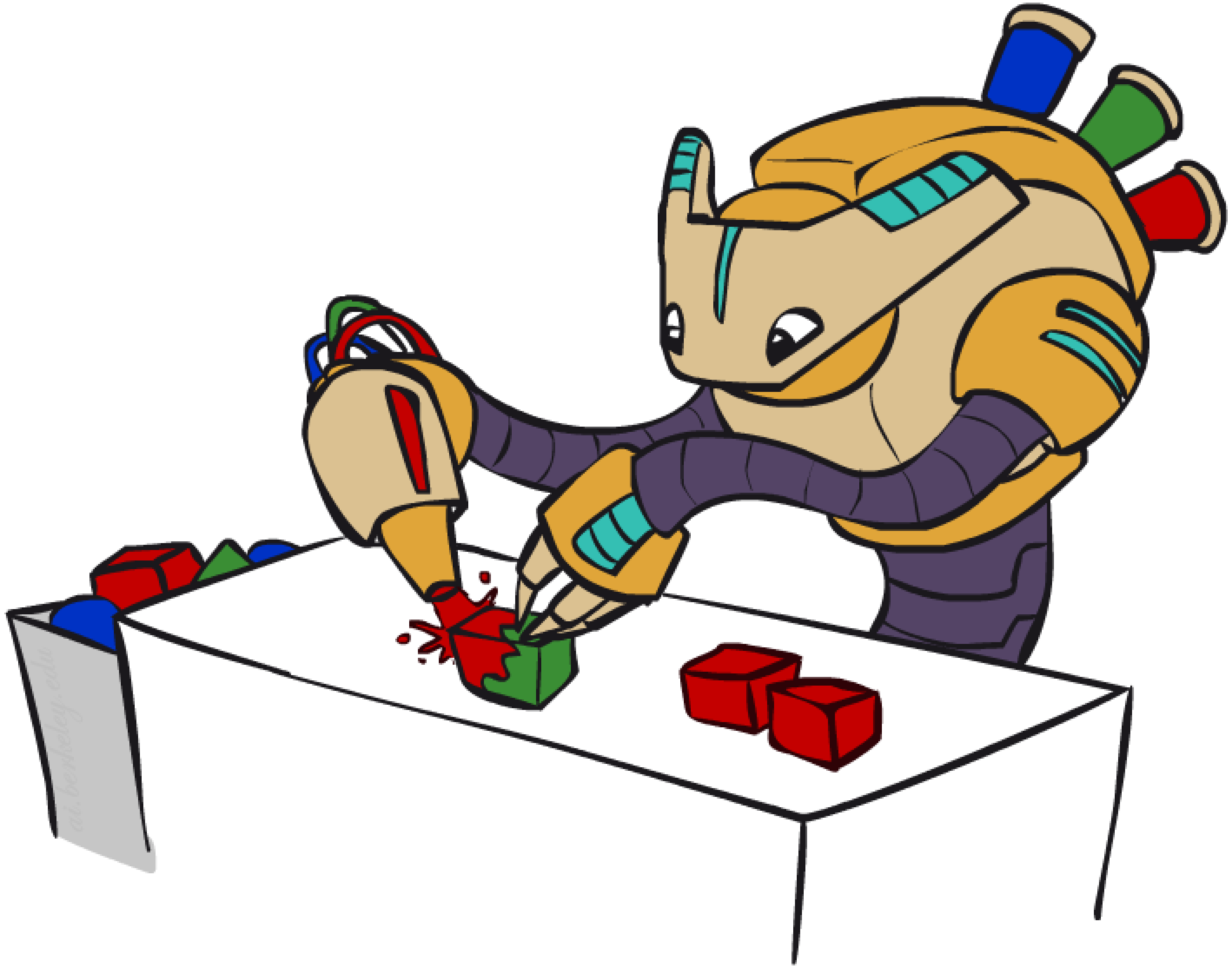
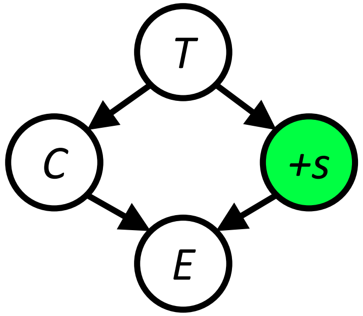
Gibbs Sampling Wrapup

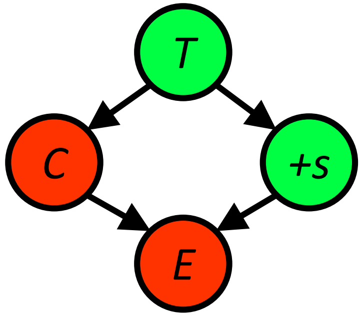


Gibbs Sampling Example: P(T, E | +s)

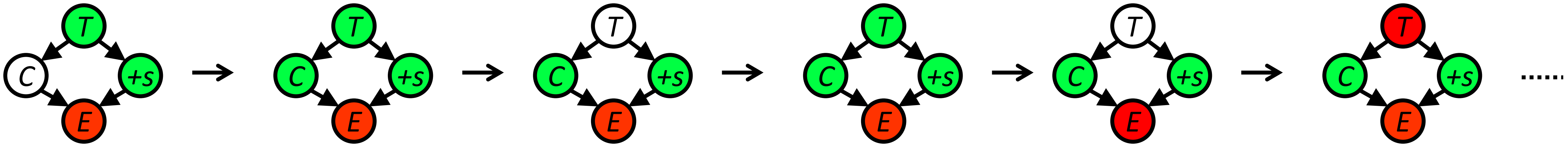
- Step 1: Fix evidence
  - S = +s



- Step 2: Initialize other variables
  - Randomly



- Steps 3: Repeat
  - Choose a non-evidence variable X
  - Resample X from P(X | all other variables)

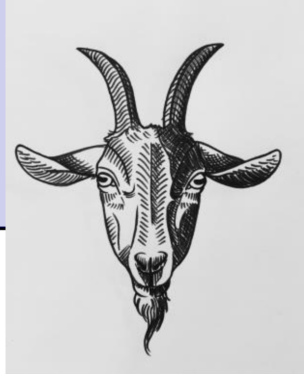


Sample from $P(C|+t,+s,-e)$

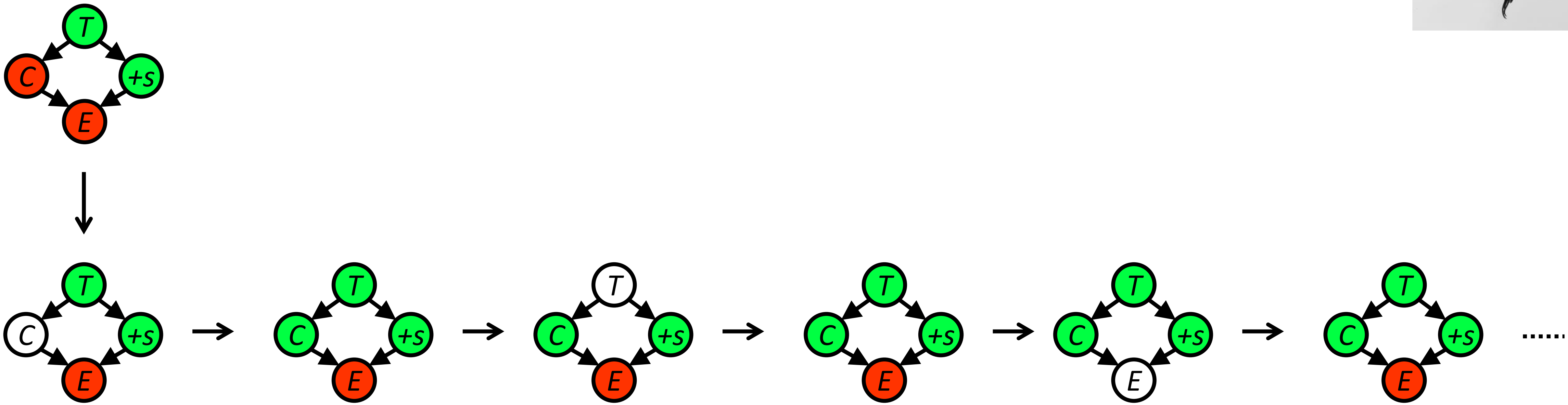
Sample from $P(T|+c,+s,-e)$

Sample from $P(T|+c,+s,-e)$

Will get to how we efficiently compute $P(C|+t,+s,-e)$ shortly.



Gibbs Sampling



- What samples are generated by the Gibbs Sampling run shown above?
- Which sample is “best”? Why? How can we get even better samples?

Gibbs Sampling

- *Procedure:* keep track of a full instantiation $x_1, x_2, \ldots, x_n$. Start with an arbitrary instantiation consistent with the evidence. Sample one variable at a time, conditioned on all the rest, but keep evidence fixed. Keep repeating this for a long time.
- *Property:* in the limit of repeating this infinitely many times the resulting sample is coming from the correct distribution
- *Rationale*: both upstream and downstream variables condition on evidence.
- In contrast: likelihood weighting only conditions on upstream evidence, and hence weights obtained in likelihood weighting can sometimes be very small. Sum of weights over all samples is indicative of how many "effective" samples were obtained, so want high weight.

Efficient Resampling of One Variable

- Sample from P(C | +t, +s, -e)

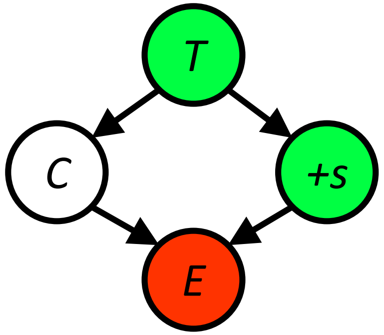
$$P(C|+t,+s,-e) = \frac{P(C,+t,+s,-e)}{P(+t,+s,-e)}$$

$$= \frac{P(C,+t,+s,-e)}{\sum_c P(c,+t,+s,-e)}$$

$$= \frac{P(+t)P(C|+t)P(+s|+t)P(-e|C,+s)}{\sum_c P(+t)P(c|+t)P(+s|+t)P(-e|c,+s)}$$

$$= \frac{P(+t)P(C|+t)P(+s|+t)P(-e|C,+s)}{P(+t)P(+s|+t)\sum_c P(c|+t)P(-e|c,+s)}$$

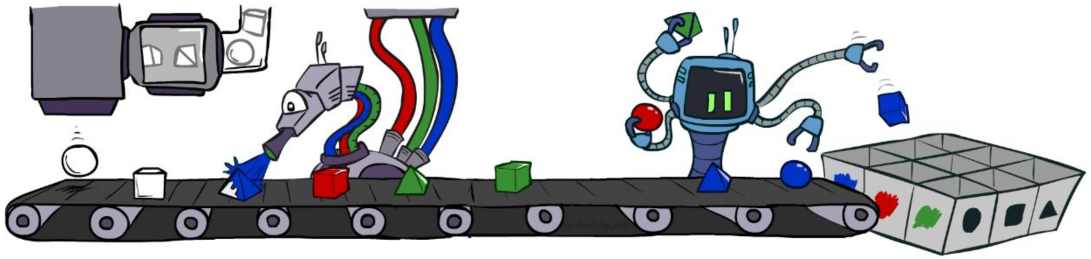
$$= \frac{P(\mathrm{C}|+t)P(-e|C,+s)}{\sum_c P(c|+t)P(-e|c,+s)}$$



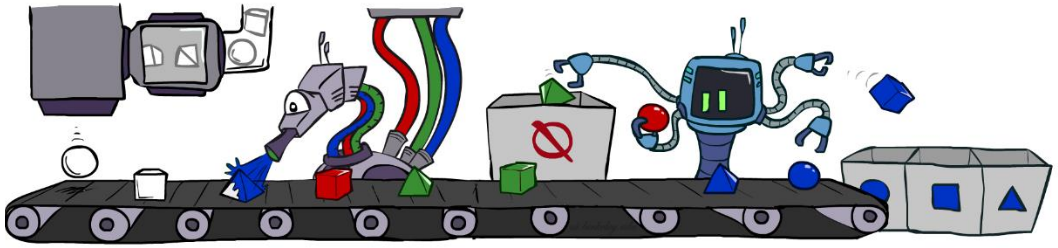
- Many things cancel out – only CPTs with C remain!
- More generally: only CPTs that have resampled variable need to be considered, and joined together

Bayes' Net Sampling Summary

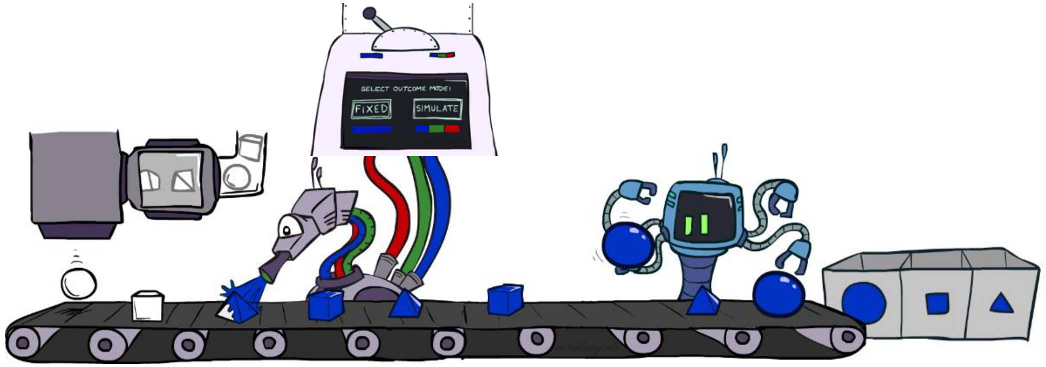
- Prior Sampling P



- Rejection Sampling $P(Q \mid e)$



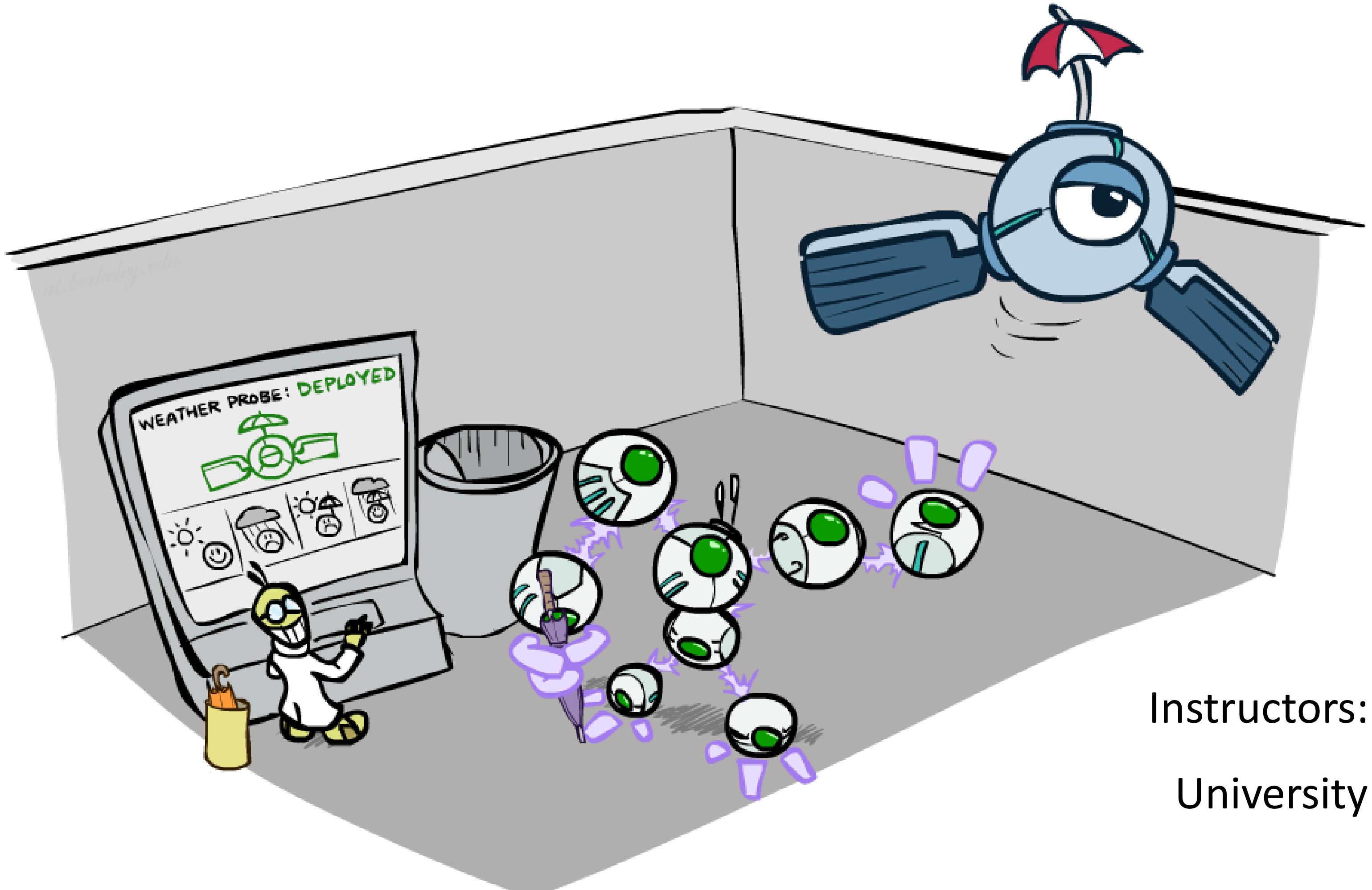
- Likelihood Weighting $P(Q \mid e)$



- Gibbs Sampling $P(Q \mid e)$

CS 188: Artificial Intelligence

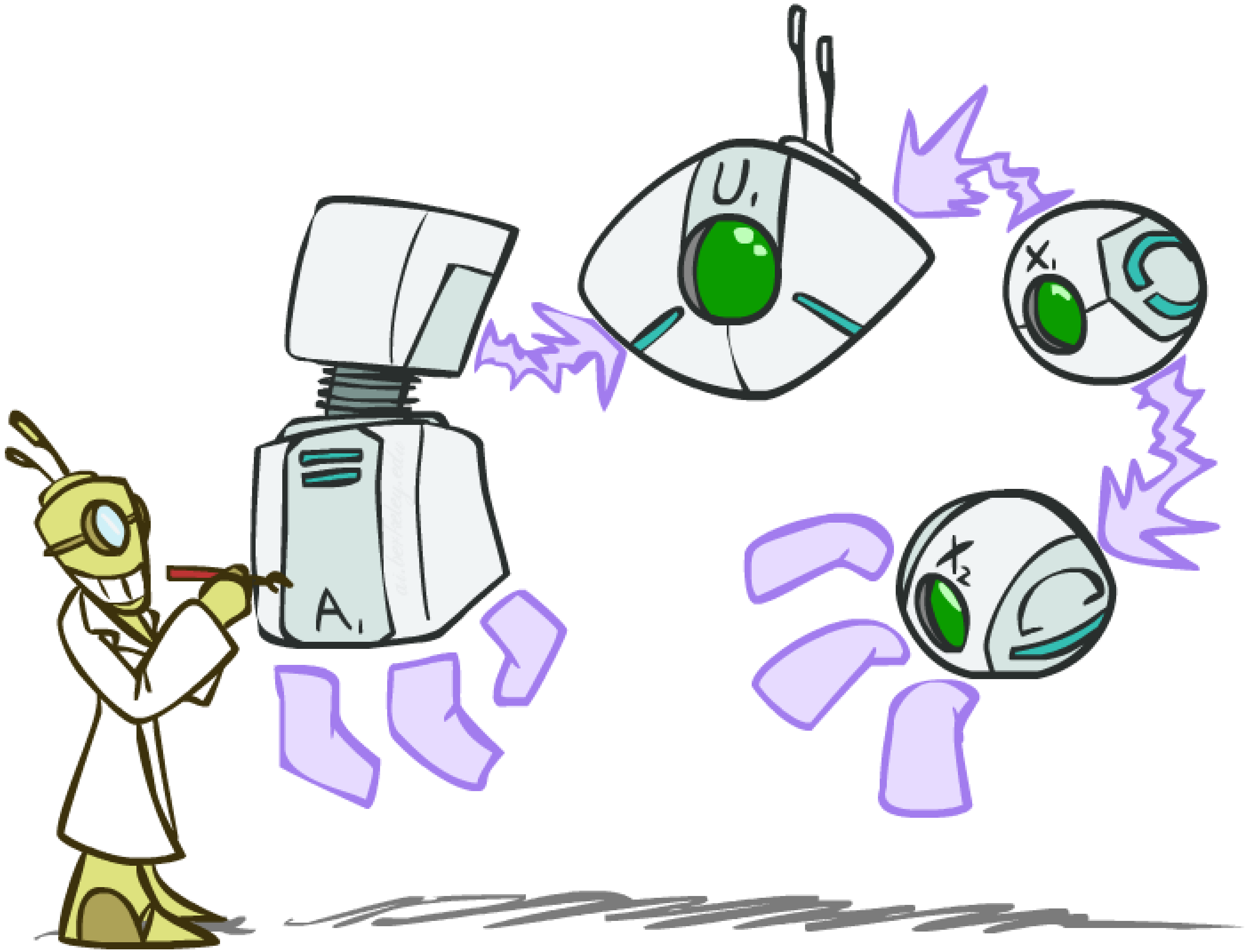
Decision Networks and Value of Perfect Information



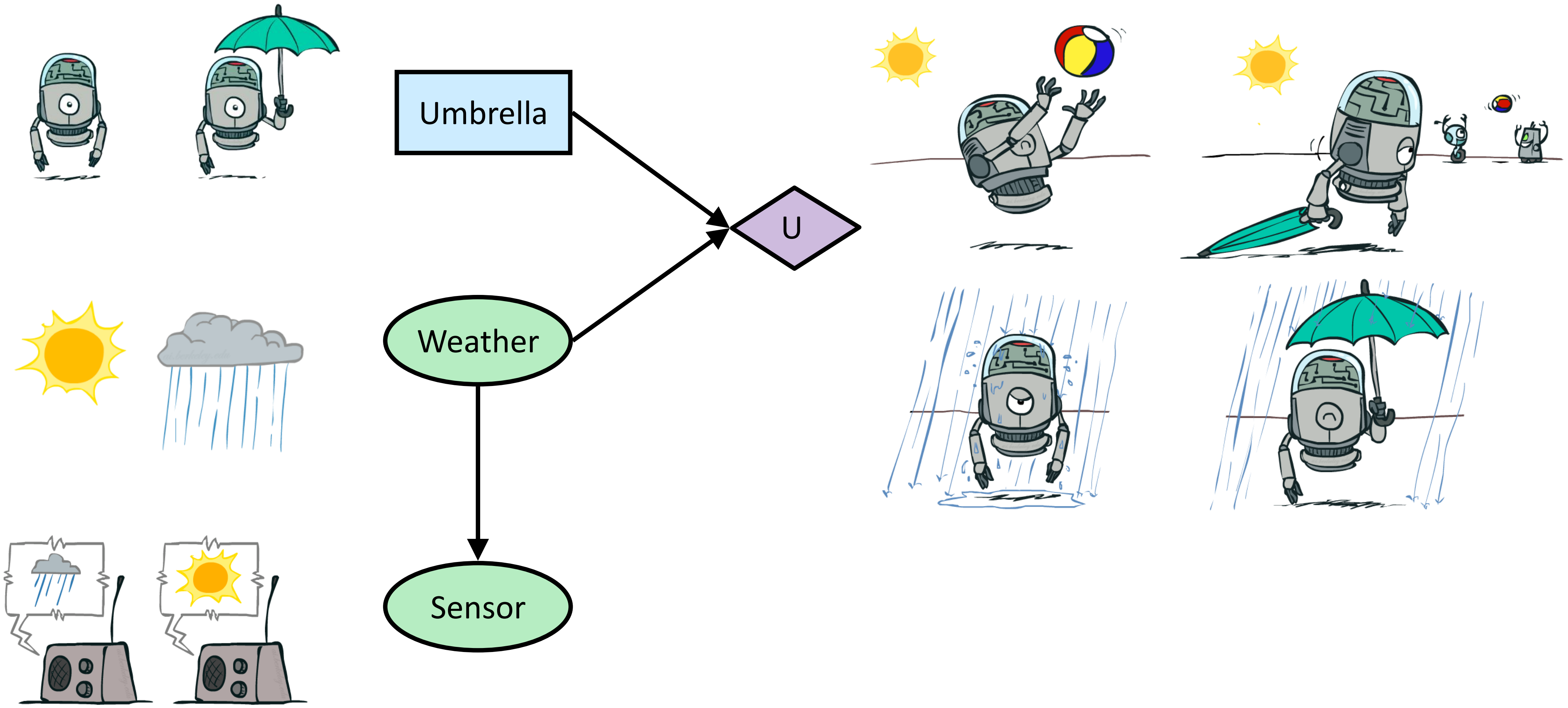
Instructors: Adam Janin & Josh Hug

University of California, Berkeley

Decision Networks



Decision Networks

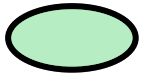


Decision Networks

- **MEU: choose the action which maximizes the expected utility given the evidence**

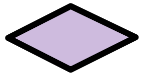
- Can directly operationalize this with decision networks
  - Bayes nets with nodes for utility and actions
  - Lets us calculate the expected utility for each action

- New node types:

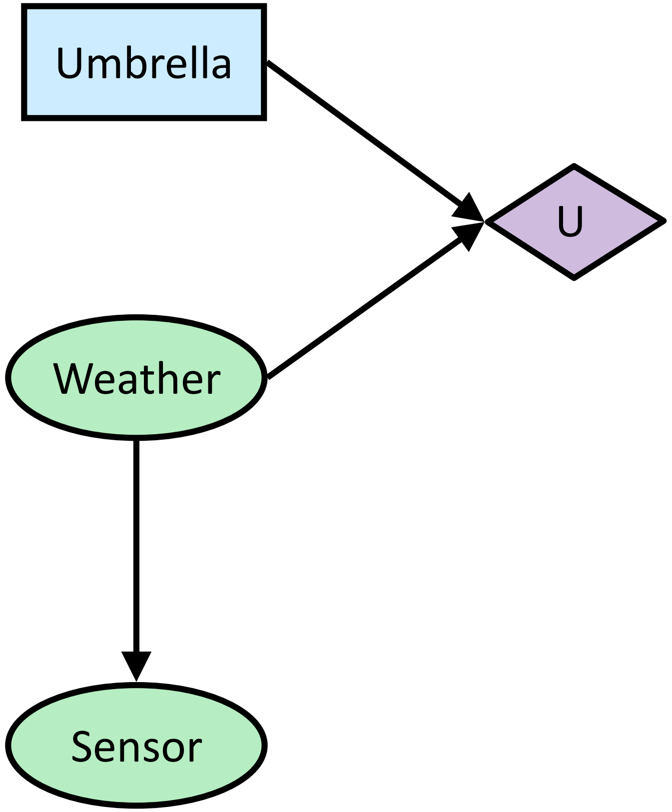


  - Chance nodes (just like BNs)

  - Actions (rectangles, cannot have parents, act similar to observed evidence)



  - Utility node (diamond, depends on action and chance nodes)



Expected Utility on Decision Networks

Umbrella = leave

$$EU(\text{leave}) = \sum_{w} P(w)U(\text{leave}, w)$$

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

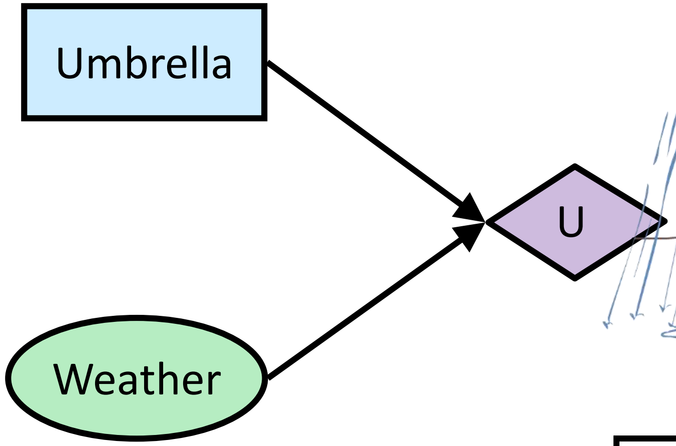
Umbrella = take

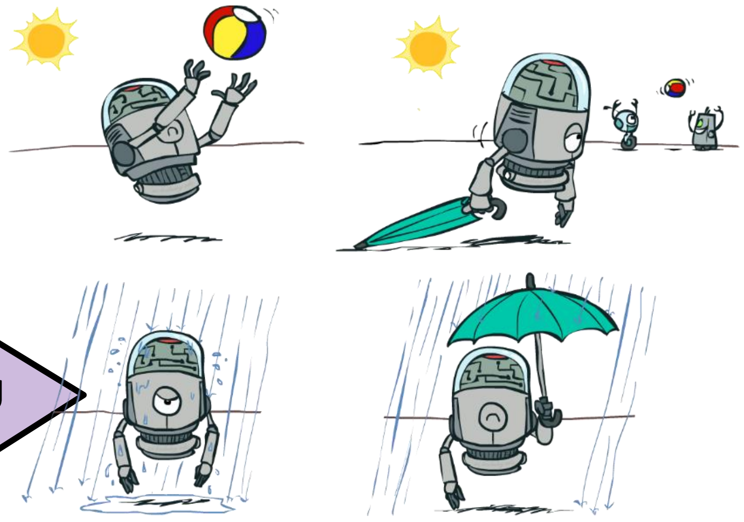
$$EU(\text{take}) = \sum_{w} P(w)U(\text{take}, w)$$

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

Optimal decision = leave

$$MEU(\emptyset) = \max_{a \in \{take, leave\}} EU(a) = 70$$

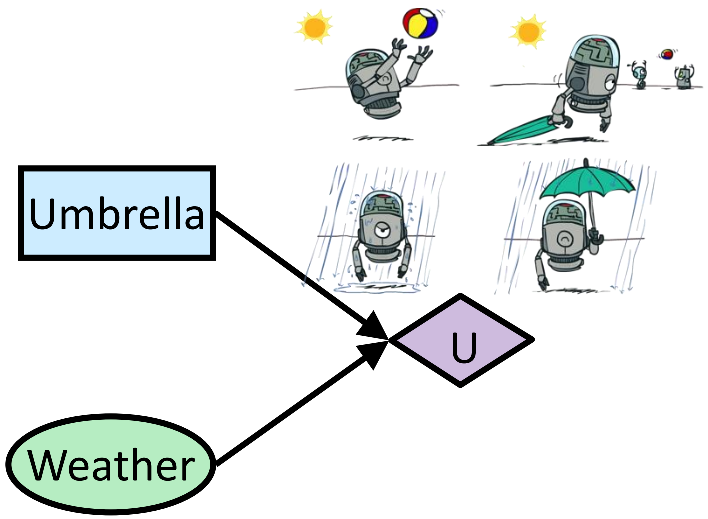


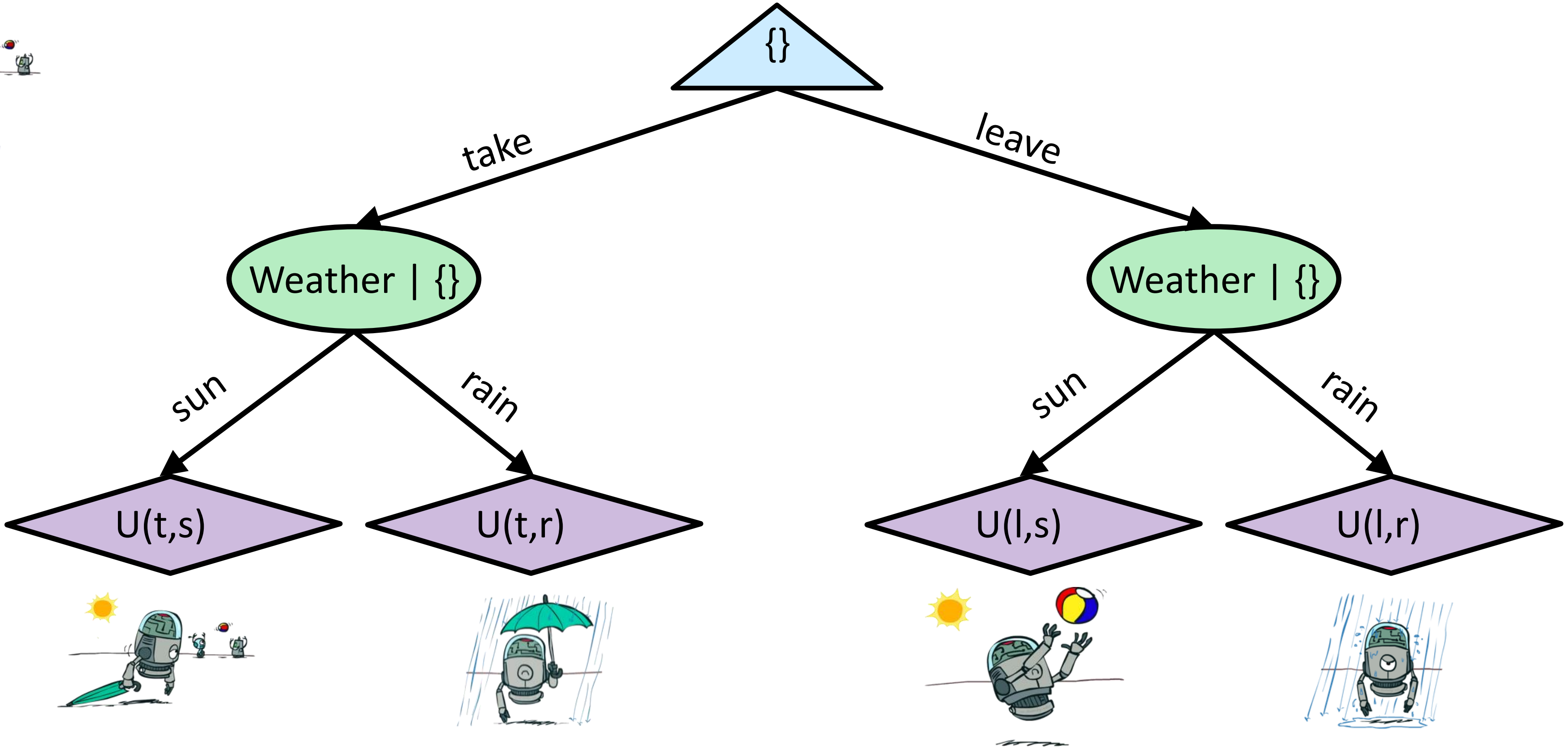


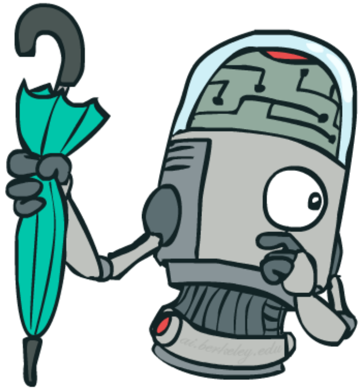
W	P(W)
sun	0.7
rain	0.3

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

Decisions as Outcome Trees



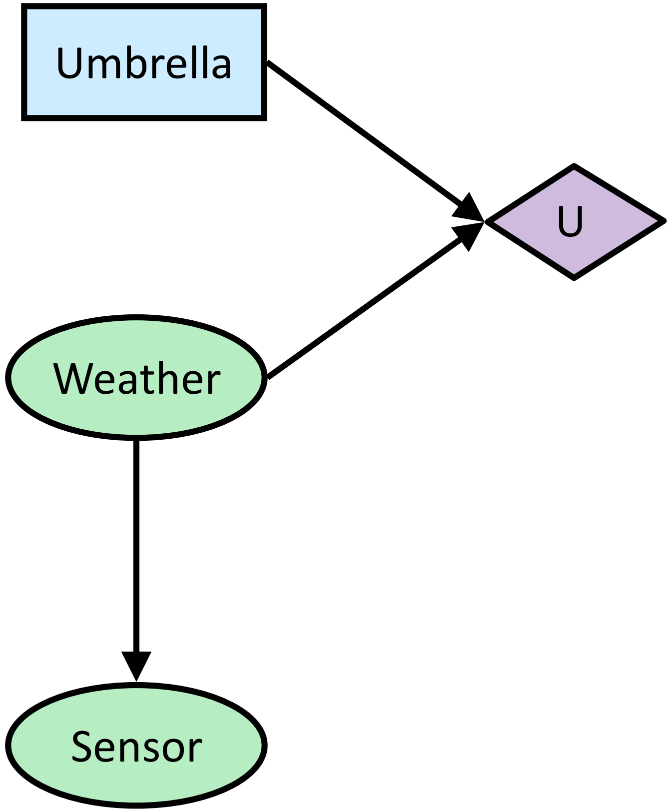




- Almost exactly like expectimax / MDPs
- What’s changed?

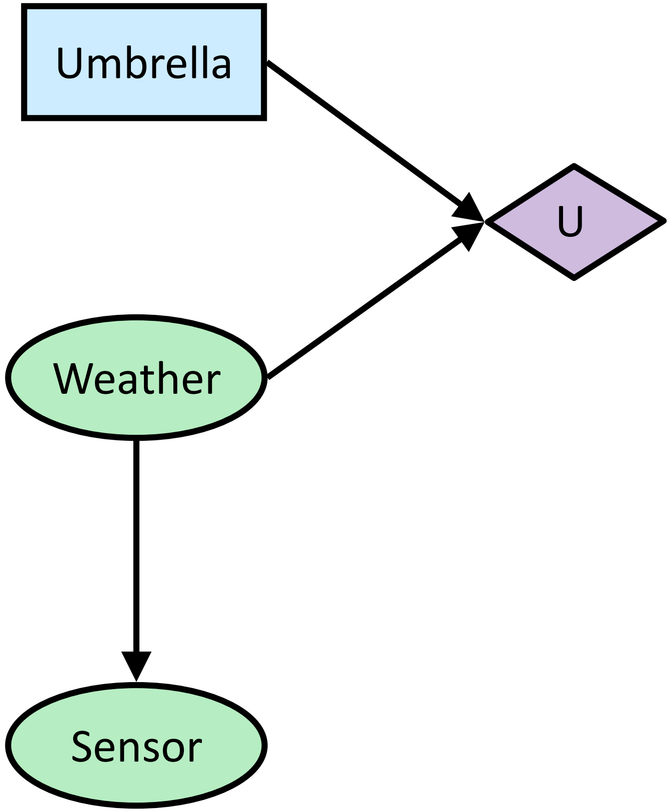
Action Selection

- Action selection
  - Instantiate all evidence
  - Calculate posterior for all parents of utility node, given the evidence (using IBE, VE, or sampling)
  - Calculate EU (expected utility) for each choice:
    - A “choice” is an assignment to every action node.
  - Choose action with the MEU (maximum expected utility).



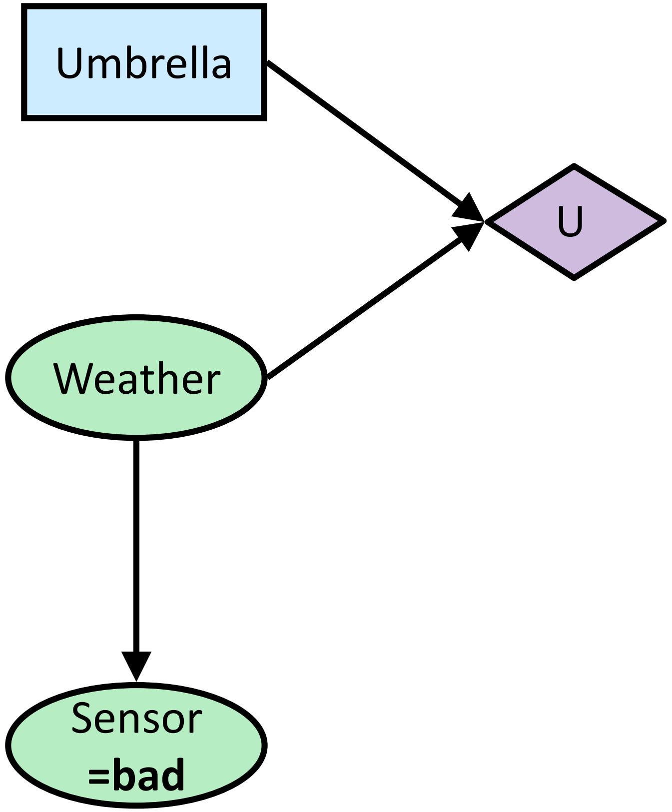
Action Selection Example

- Action selection: What action should we take if we get to see the forecast?
  - Instantiate all evidence: Sensor
  - Calculate posterior for all parents of utility node, given the evidence (using IBE, VE, or sampling)
  - Calculate EU (expected utility) for each choice: Umbrella = take and Umbrella = leave
    - A “choice” is an assignment to every action node.
  - Choose action with the MEU (maximum expected utility).



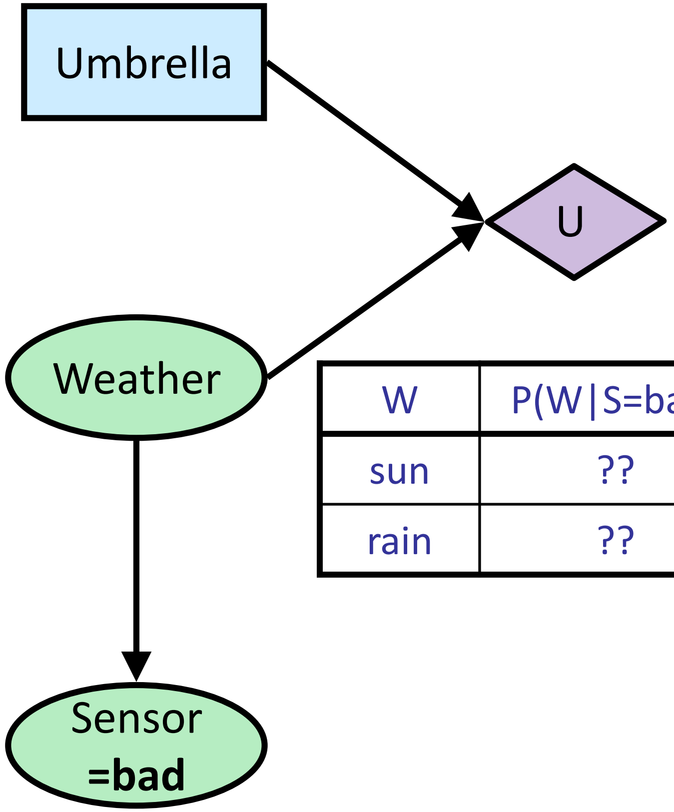
Action Selection Example, Step 1

- Action selection: What action should we take if we get to see the forecast?
  - **Instantiate all evidence: Sensor**
    - Took a measurement and got ‘bad’.



Action Selection Example, Step 2

- Action selection: What action should we take if we get to see the forecast?
  - **Calculate posterior for all parents of utility node, given the evidence (using IBE, VE, or sampling)**
    - Need: P(W | S=bad)
    - Can compute using P(W) and P(S | W)
      - Webcast viewers, try it!



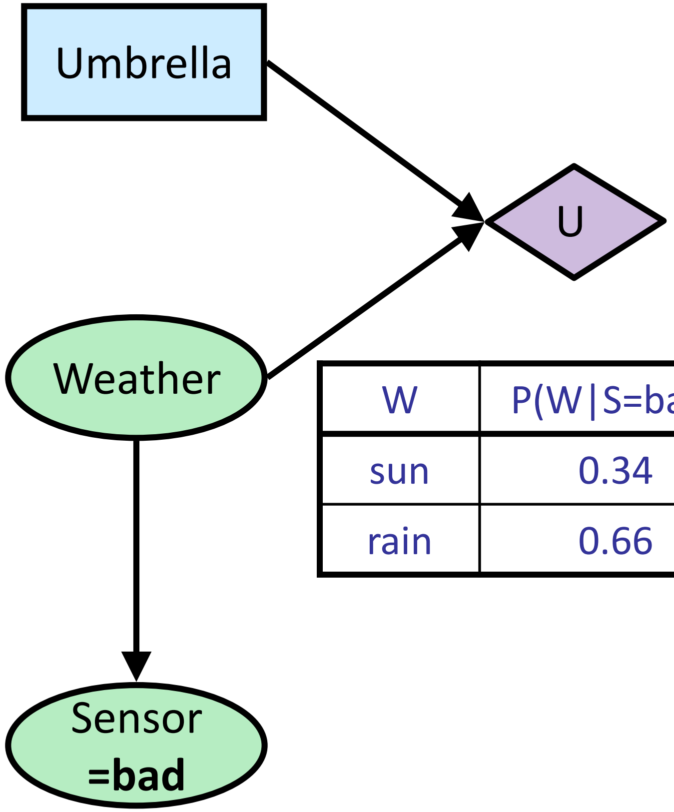
W	P(W)
sun	0.7
rain	0.3

W	F	P(S|W)
sun	good	0.8047
sun	bad	0.1953
rain	good	0.0889
rain	bad	0.9111

W	P(W|S=bad)
sun	??
rain	??

Action Selection Example, Step 2

- Action selection: What action should we take if we get to see the forecast?
  - **Calculate posterior for all parents of utility node, given the evidence (using IBE, VE, or sampling)**
    - Need: P(W | S=bad)
    - Can compute using P(W) and P(S | W)
      - Webcast viewers, try it!



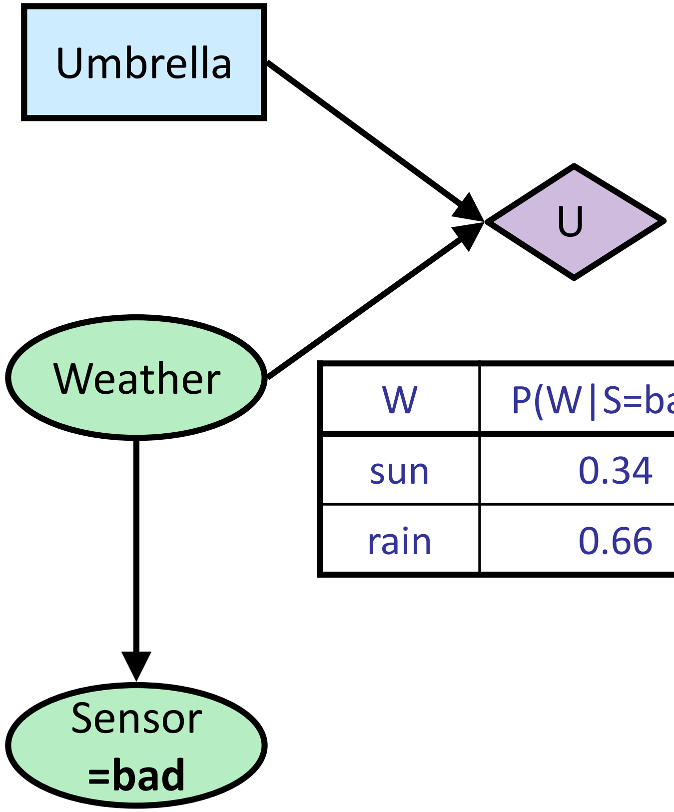
W	P(W)
sun	0.7
rain	0.3

W	S	P(S|W)
sun	good	0.8047
sun	bad	0.1953
rain	good	0.0889
rain	bad	0.9111

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

Action Selection Example, Step 3

- Action selection: What action should we take if we get to see the forecast?
  - Instantiate all evidence: Forecast
  - Calculate posterior for all parents of utility node, given the evidence (using IBE, VE, or sampling)
  - **Calculate EU (expected utility) for each choice: Umbrella = take and Umbrella = leave [next slide]**
    - **A “choice” is an assignment to every action node.**
  - Choose action with the MEU (maximum expected utility).



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

Action Selection Example, Step 3

Choice 1: Umbrella = leave

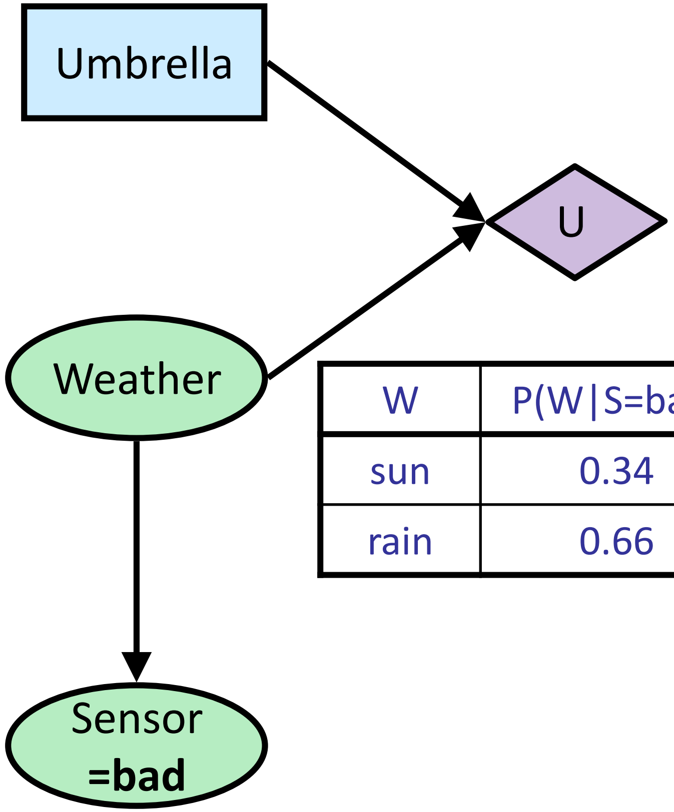
$$EU(\text{leave}|\text{bad}) = \sum_{w} P(w|\text{bad})U(\text{leave}, w)$$

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

Choice 2: Umbrella = take

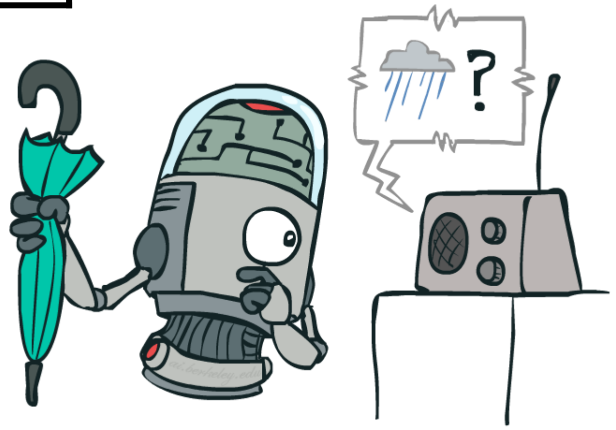
$$EU(\text{take}|\text{bad}) = \sum_{w} P(w|\text{bad})U(\text{take}, w)$$

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



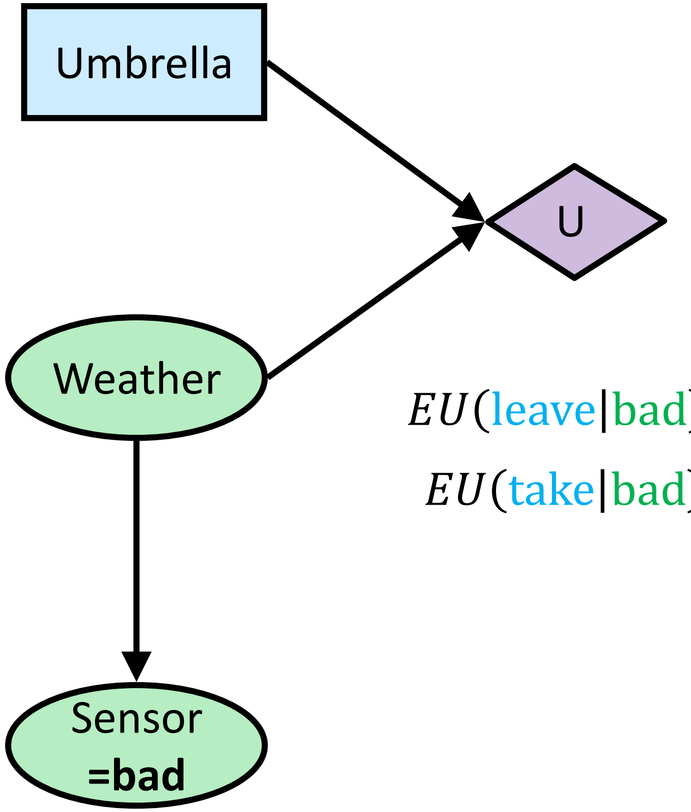
A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

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



Action Selection Example, Step 4

- Action selection: What action should we take if we get to see the forecast?
  - Instantiate all evidence: Sensor
  - Calculate posterior for all parents of utility node, given the evidence (using IBE, VE, or sampling)
  - Calculate EU (expected utility) for each choice: Umbrella = take and Umbrella = leave
    - A "choice" is an assignment to every action node.
  - **Choose action with the MEU (maximum expected utility): take**

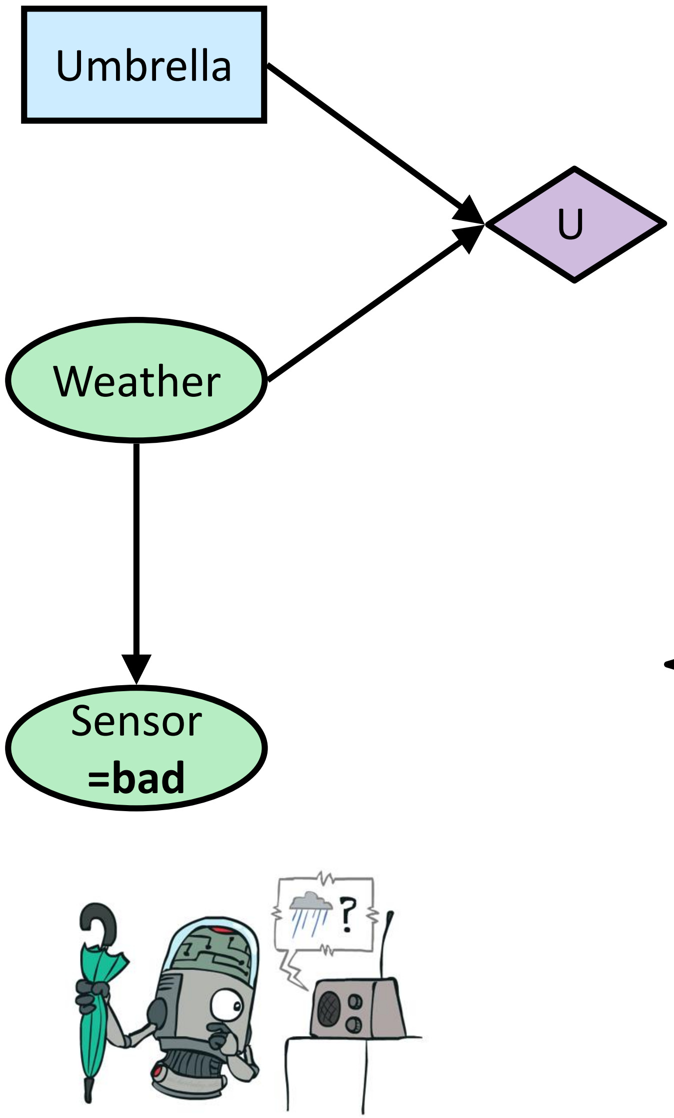


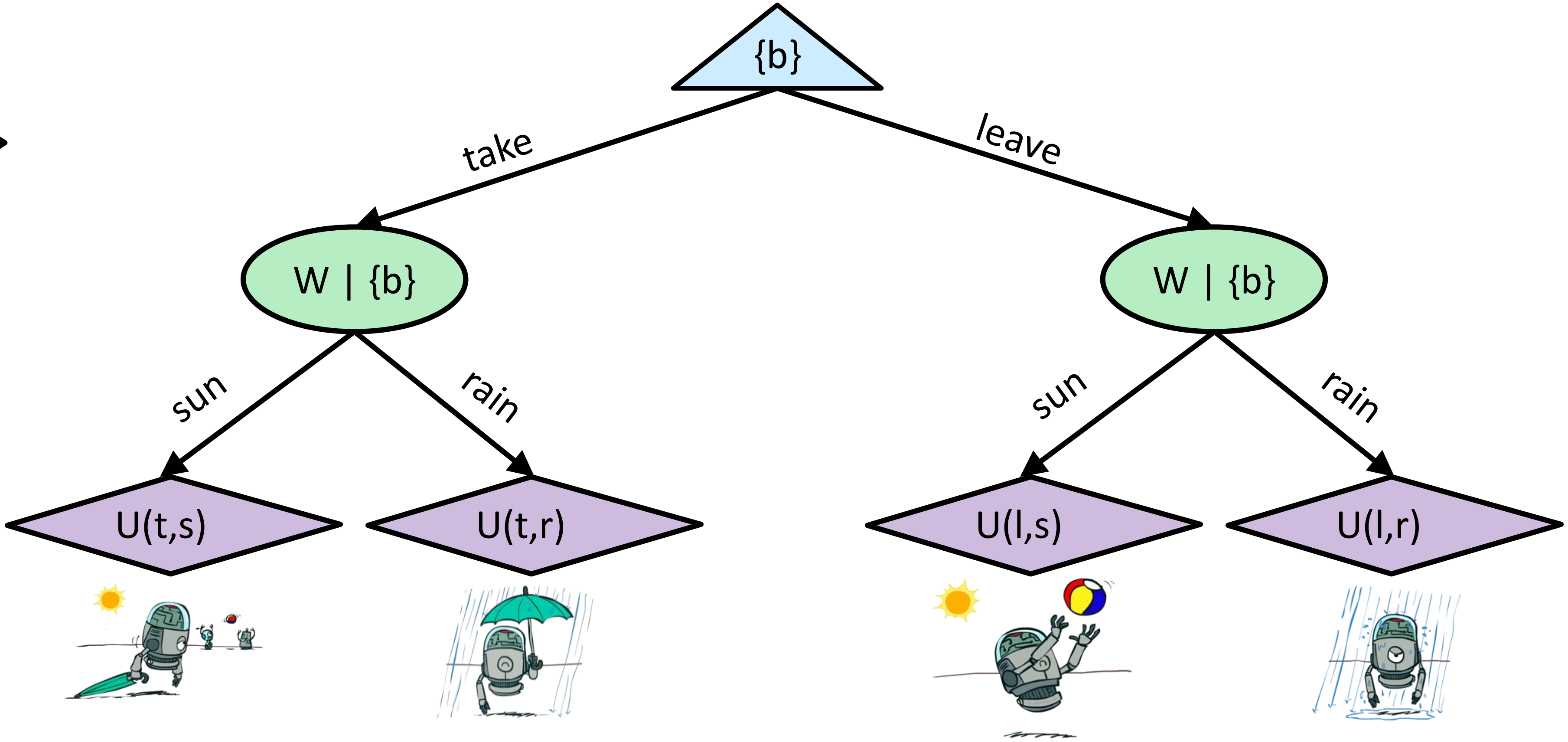
$EU(\text{leave}|\text{bad}) = 34$

$EU(\text{take}|\text{bad}) = 53$

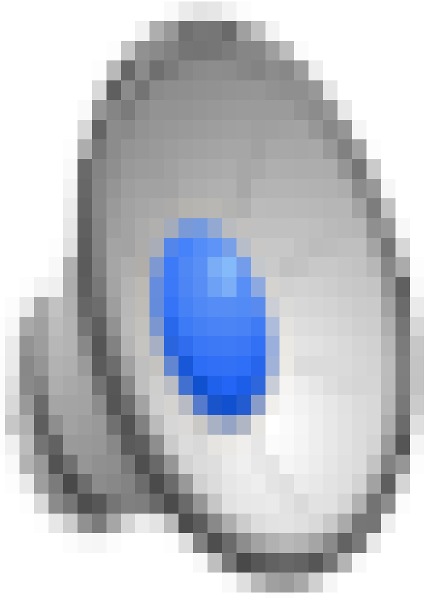
$$MEU(S = \text{bad}) = \max_{a \in \{\text{leave}, \text{take}\}} EU(a|\text{bad}) = 53$$

Decisions as Outcome Trees





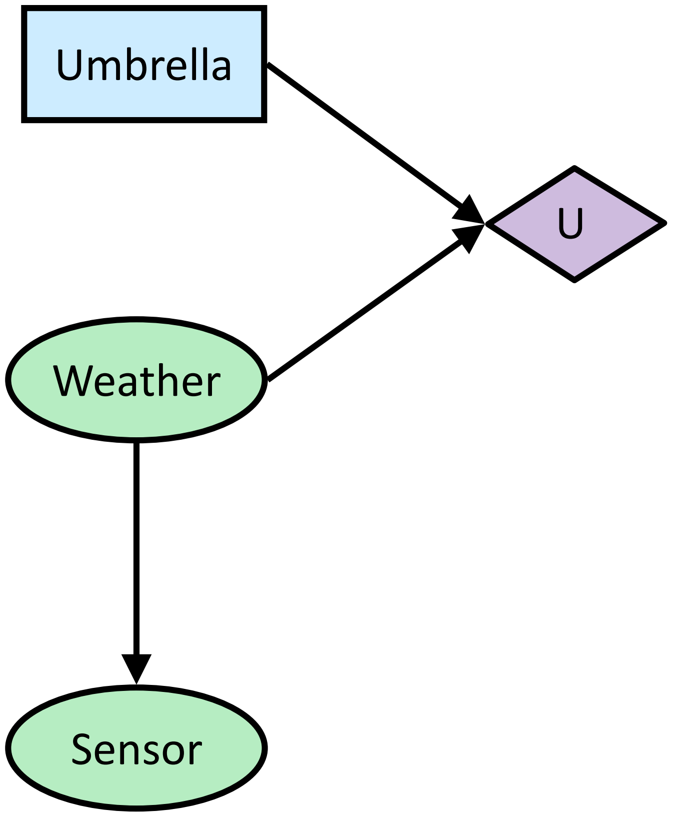
Video of Demo Ghostbusters with Probability



Demo: Ghostbusters with probability

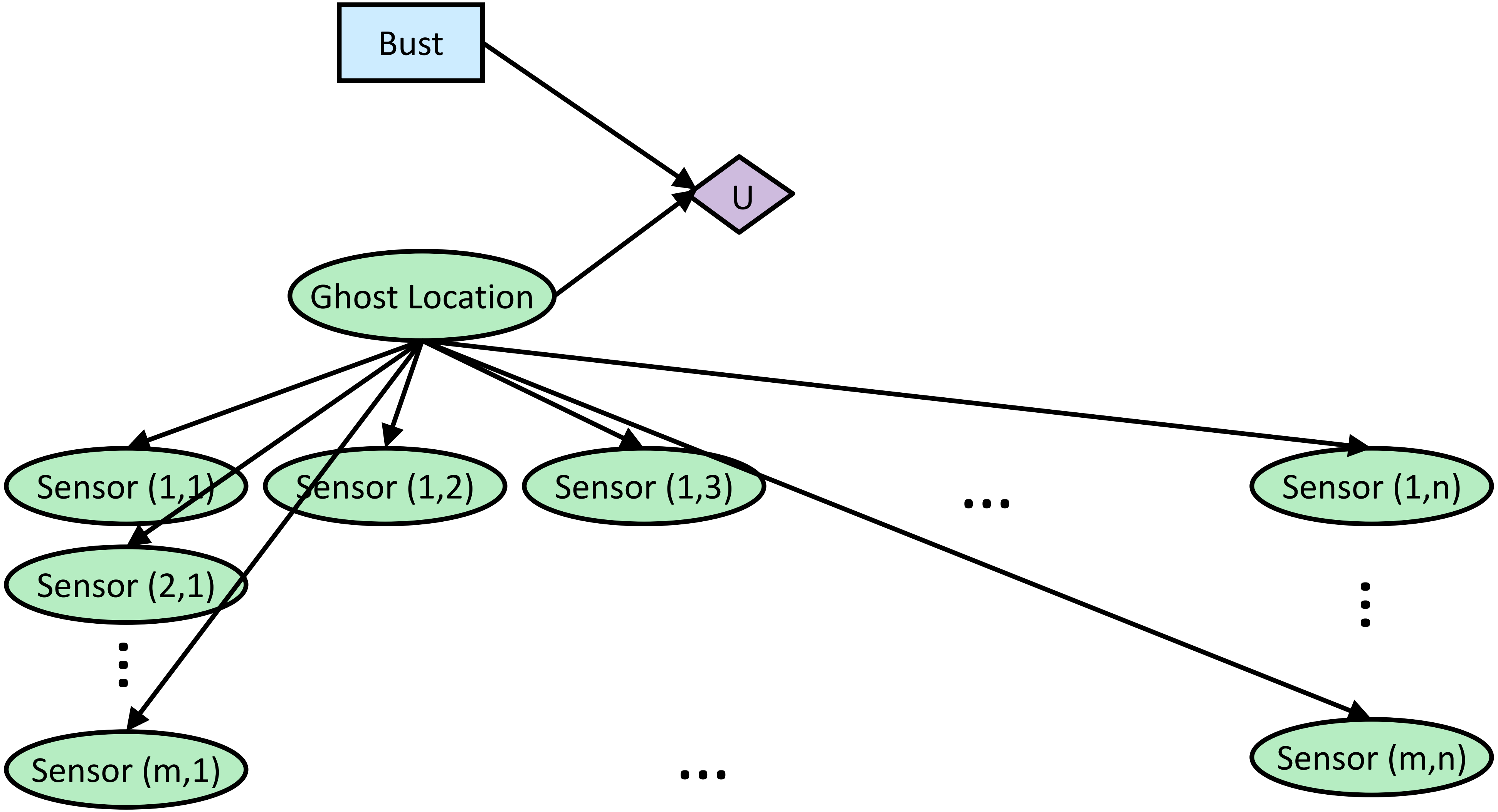
Ghostbusters Decision Network

- Draw the Decision Network for Ghostbusters

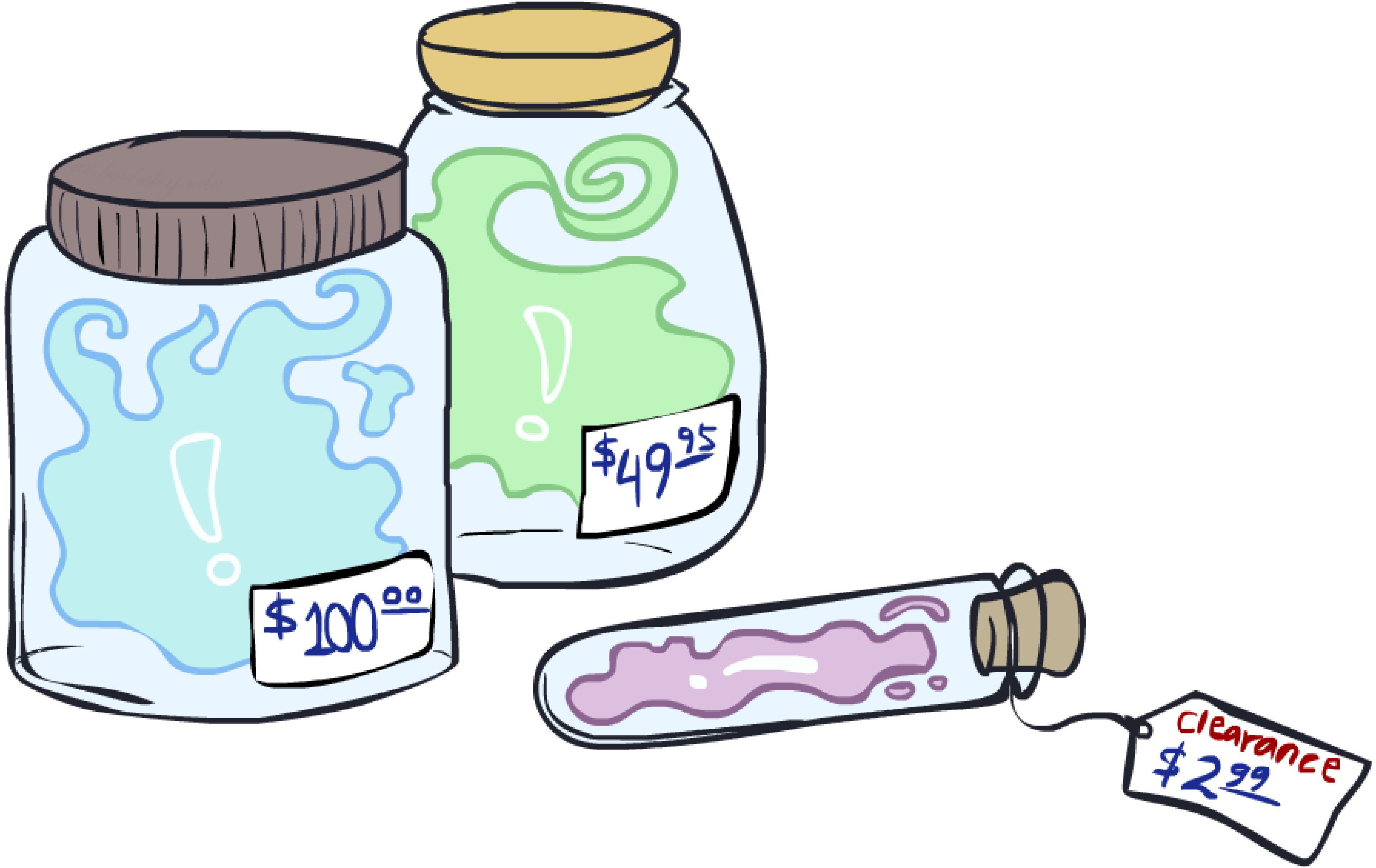


Ghostbusters Decision Network

Demo: Ghostbusters with probability



Value of Information



Value of Sensor Information

- If we don't have a sensor reading:
  - The maximum expected utility is 70, and we should leave.

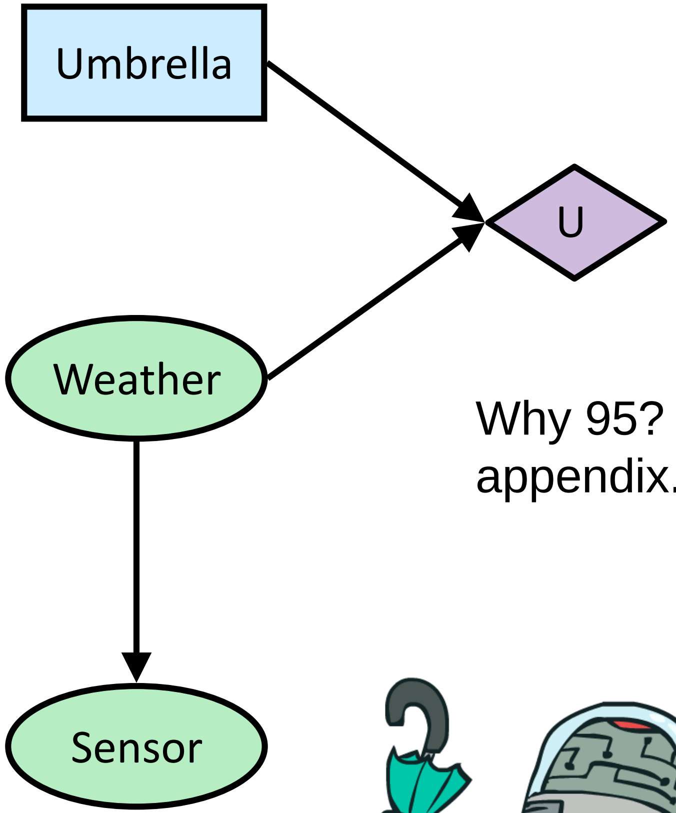
$$MEU(\emptyset) = \max_a EU(a) = 70$$

- If we have a sensor reading of bad:
  - The maximum expected utility is 53, and we should take.

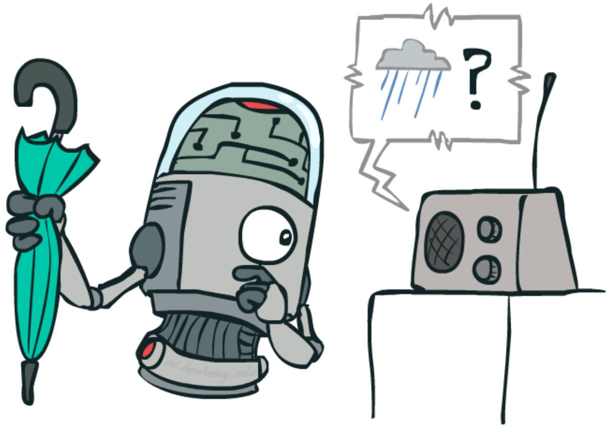
$$MEU(S = \text{bad}) = \max_a EU(a|\text{bad}) = 53$$

- If we have a sensor reading of good:
  - The maximum expected utility is 95, and we should leave.

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



Why 95? See appendix.

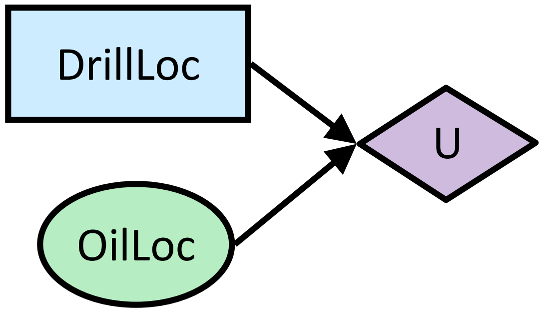


- Suppose using the sensor isn't free (as in ghostbusters).
  - How much utility are we willing to sacrifice to read the sensor?
  - In ghostbusters: Should we take another reading or bust?

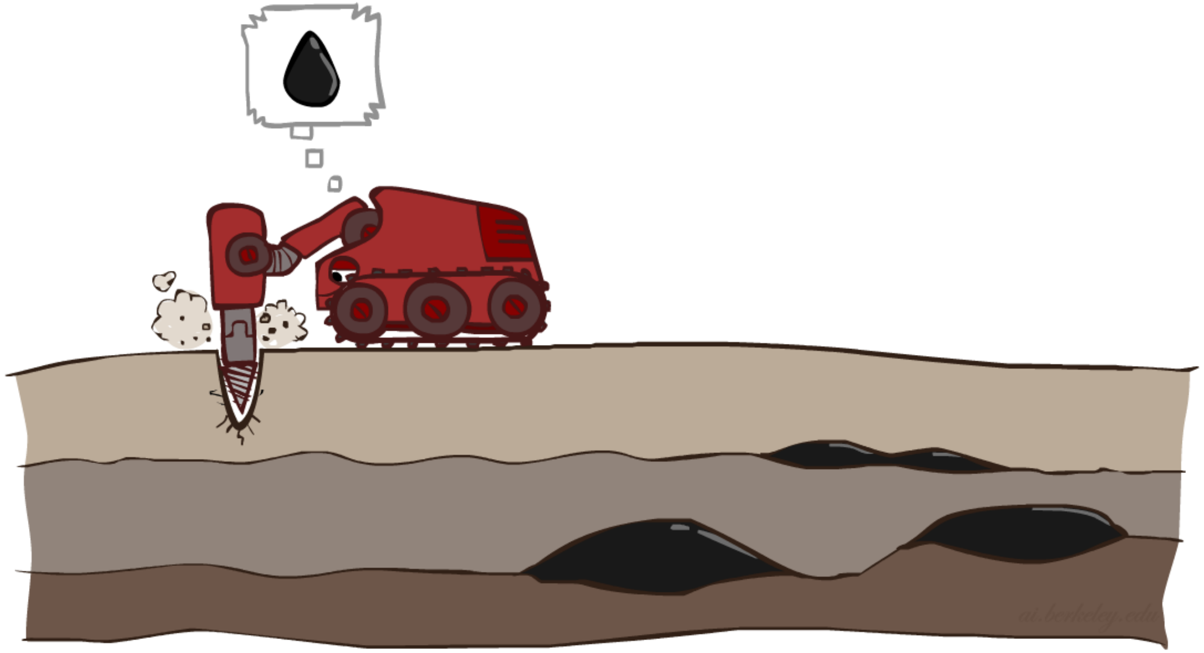
Value of Perfect Information

- Idea: compute value of acquiring evidence
  - Can be done directly from decision network
- Example: buying oil drilling rights
  - Two blocks A and B, exactly one has oil, worth k
  - You can drill in one location
  - Prior probabilities 0.5 each, & mutually exclusive
  - Drilling in 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 oil
  - Value is expected gain in MEU from new info
  - Survey may say "oil in a" or "oil in b," prob 0.5 each
  - If we know OilLoc, MEU is k (either way)
  - Gain in MEU from knowing OilLoc?
  - VPI(OilLoc) = k/2
  - Fair price of information: k/2

O	P
a	1/2
b	1/2



D	O	U
a	a	k
a	b	0
b	a	0
b	b	k



Value of Weather Information

- Goal: Find the value of information about the weather.

  MEU if we don't know anything:

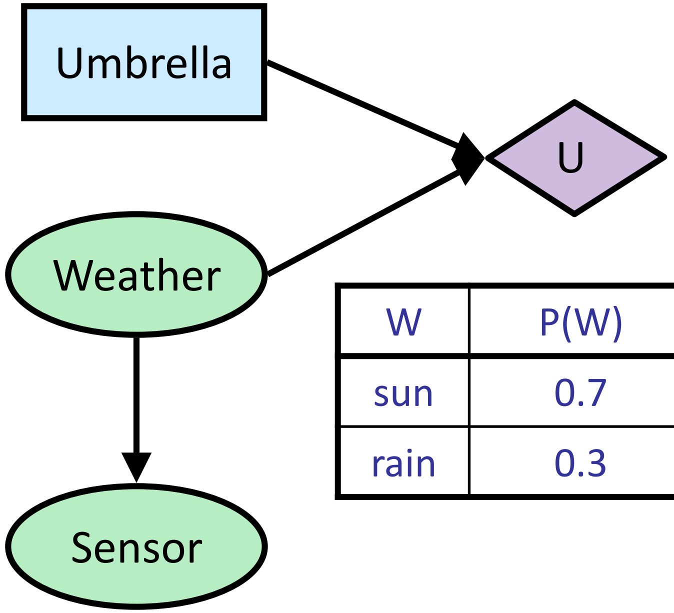
  $$MEU(\emptyset) = \max_a EU(a) = 70$$

  MEU if weather is sunny:

  $$MEU(W = \text{sun}) = \max_a EU(a|W = \text{sun}) = 100$$

  MEU if weather is rainy:

  $$MEU(W = \text{rain}) = \max_a EU(a|W = \text{rain}) = 70$$



W	P(W)
sun	0.7
rain	0.3

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

- What is VPI(weather)?
  - Recall, the VPI is the gain resulting from having the information.

Value of Weather Information

- Goal: Find the value of information about the weather.

MEU if we don't know anything:

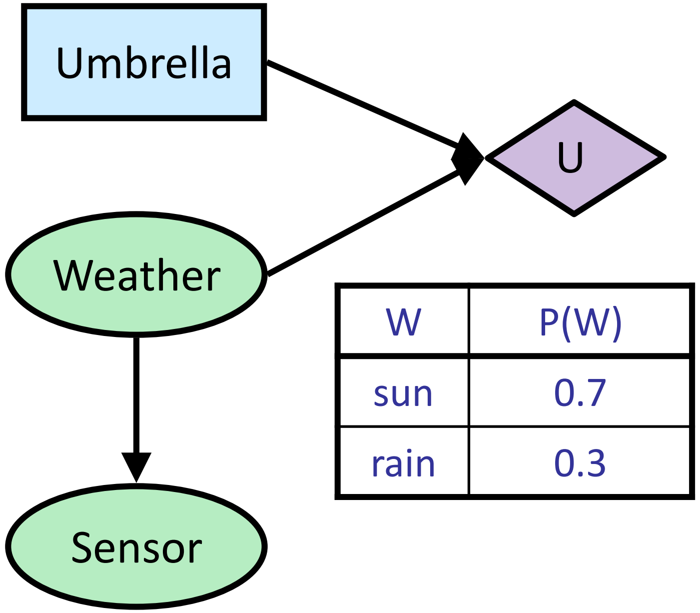
$$MEU(\emptyset) = \max_a EU(a) = 70$$

MEU if weather is sunny:

$$MEU(W = \text{sun}) = \max_a EU(a|W = \text{sun}) = 100$$

MEU if weather is rainy:

$$MEU(W = \text{rain}) = \max_a EU(a|W = \text{rain}) = 70$$



W	P(W)
sun	0.7
rain	0.3

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

- What is VPI(weather)?

$$0.7 \cdot 100 + 0.3 * 70 - 70$$
$$= 91 - 70 = 21$$

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

- In other words, we'd be willing to give up 21 units of utility for perfect information about the weather.

Value of Sensor Information

- Goal: Find the value of information about the sensor.

MEU if we don’t know anything:

$$MEU(\emptyset) = \max_a EU(a) = 70$$

MEU if sensor says good

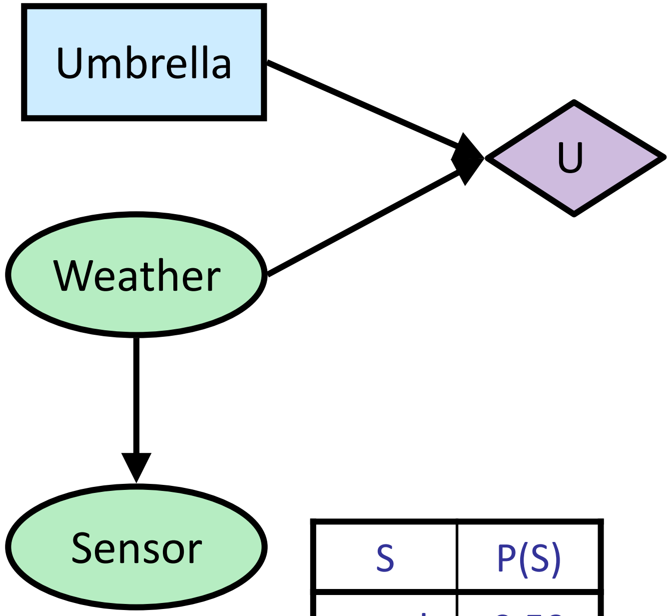
$$MEU(S = \text{good}) = \max_a EU(a|S = \text{good}) = 95$$

MEU if sensor says bad

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

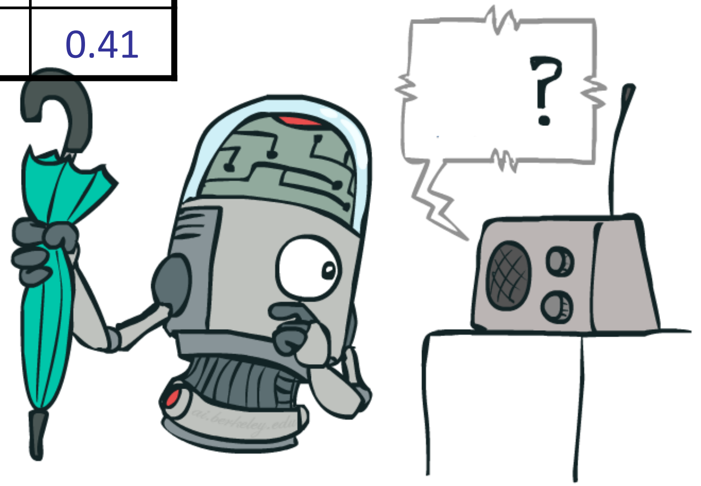
- What is VPI(sensor)?

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



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

S	P(S)
good	0.59
bad	0.41



Value of Sensor Information

- Goal: Find the value of information about the sensor.

MEU if we don't know anything:

$$MEU(\emptyset) = \max_a EU(a) = 70$$

MEU if sensor says good

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

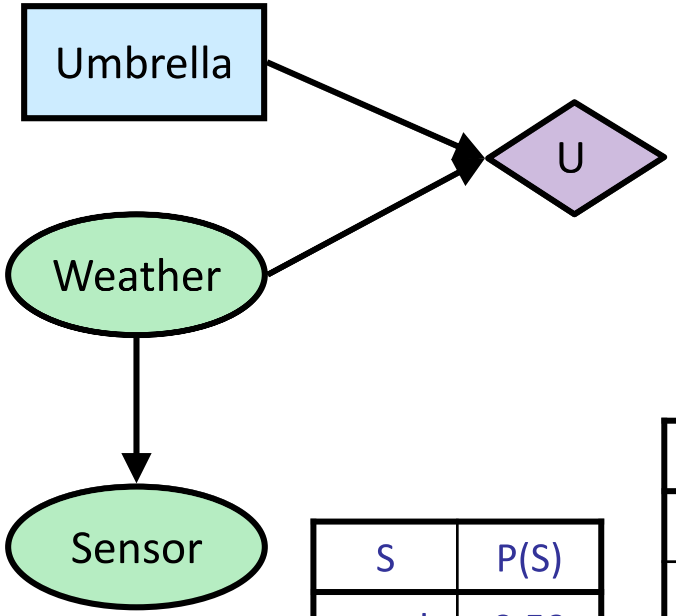
MEU if sensor says bad

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

- What is VPI(sensor)? $0.59 \cdot 95 + 0.41 * 53 - 70$
$= 77.8 - 70 = 7.8$

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

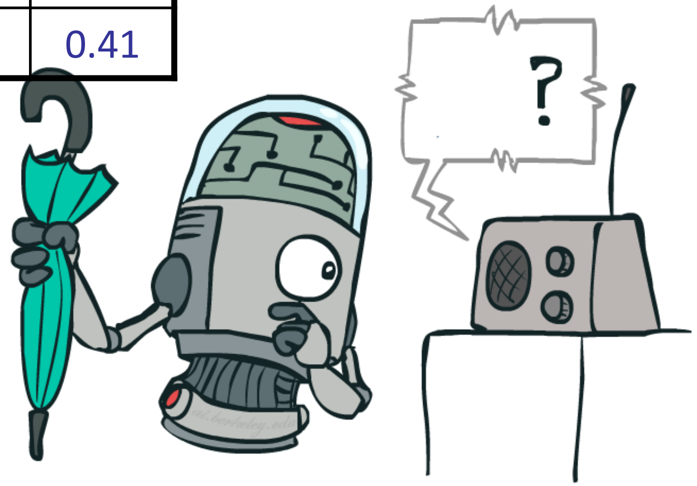
- We'd be willing to give up 7.8 units of utility to read the sensor.



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

W	P(W|S=good)
sun	0.95
rain	0.05

S	P(S)
good	0.59
bad	0.41



Conditional Value of Weather Information

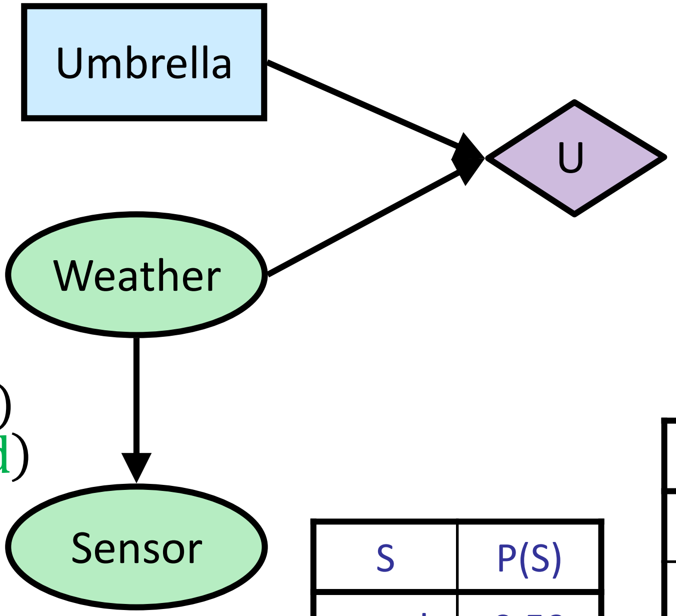
- Suppose we already know that $S = \text{good}$.

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

- What is $VPI(Weather|S = \text{good})$?
  - $= P(W = \text{sun}|S = \text{good}) \cdot MEU(W = \text{sun}|S = \text{good})$
    $+ P(W = \text{rain}|S = \text{good}) \cdot MEU(W = \text{rain}|S = \text{good})$
    $- MEU(S = \text{good})$
  - $= 0.95 \cdot 100 + 0.05 \cdot 70 - 95 = 98.5 - 95 = 3.5$

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

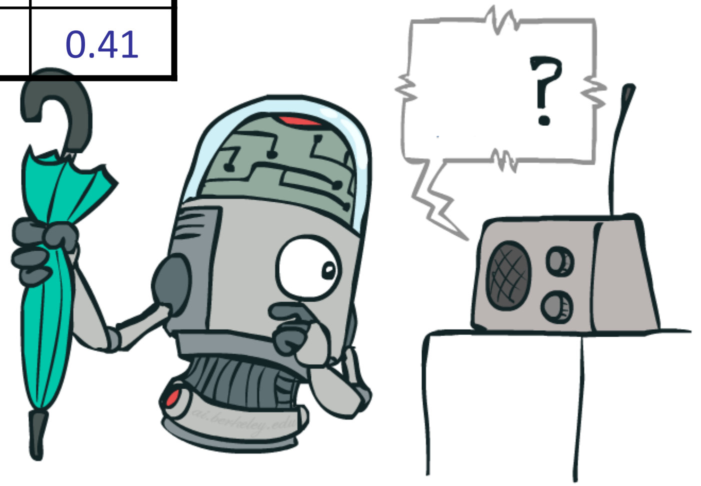
- If already observed sensor = good, only willing to pay 3.5 utility to see true weather.



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

W	P(W|S=good)
sun	0.95
rain	0.05

S	P(S)
good	0.59
bad	0.41



Value of Information: Summary

- Assume we have evidence E=e. Value if we act now, where s=parents of U:

$$MEU(e) = \max_a \sum_e P(s|e)U(s,a)$$

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

$$MEU(e,e') = \max_a \sum_e P(s|e,e')U(s,a)$$

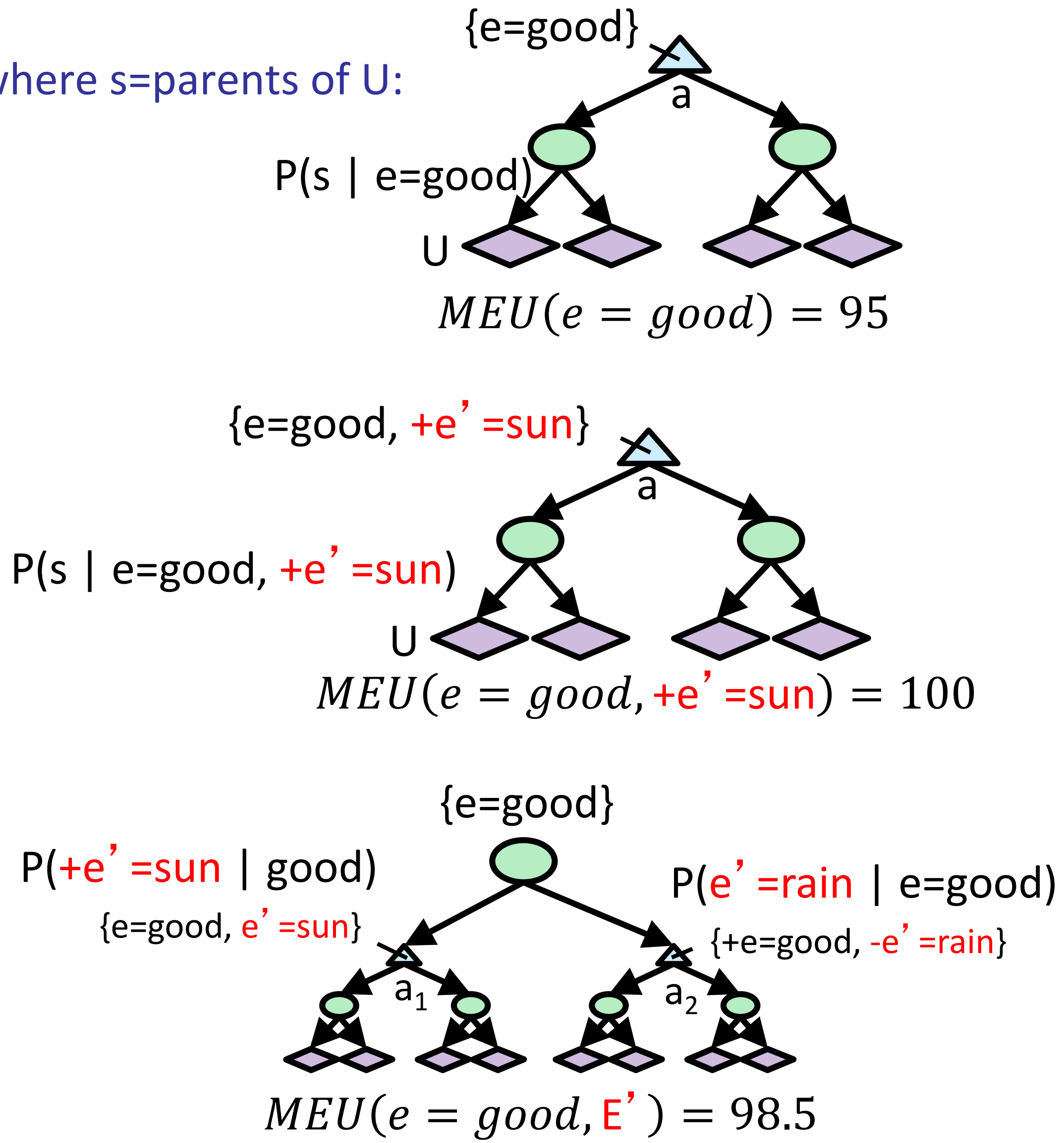
- BUT E' is a random variable whose value is unknown, so we don't know what e' will be

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

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

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

$$VPI(E'|e) = MEU(e,E') - MEU(e)$$



VPI Properties

- Nonnegative

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

- Nonadditive

  (think of observing E_j twice)

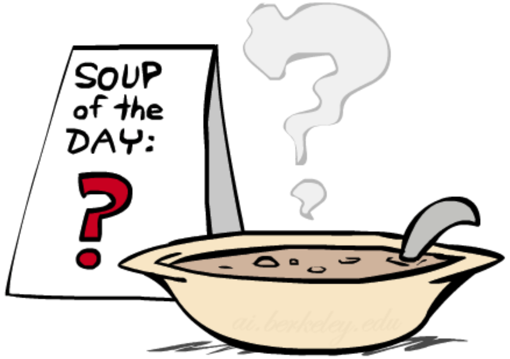
$$\text{VPI}(E_j, E_k|e) \neq \text{VPI}(E_j|e) + \text{VPI}(E_k|e)$$

- Order-independent

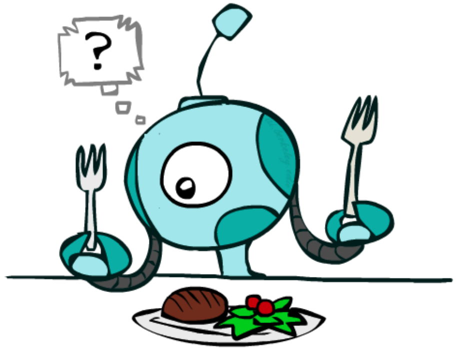
$$\text{VPI}(E_j, E_k|e) = \text{VPI}(E_j|e) + \text{VPI}(E_k|e, E_j)$$
$$= \text{VPI}(E_k|e) + \text{VPI}(E_j|e, E_k)$$

Quick VPI Questions

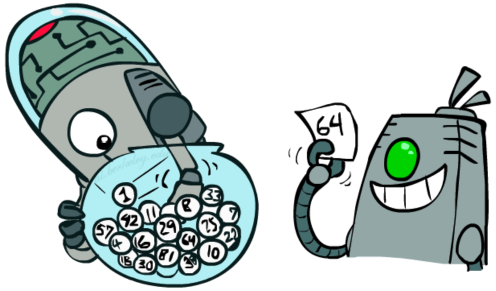
- The soup of the day is either clam chowder or split pea, but you wouldn’t order either one. What’s the value of knowing which it is?



- There are two kinds of plastic forks at a picnic. One kind is slightly sturdier. What’s the value of knowing which?



- You get one free lottery ticket. 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? Assume U(x dollars) = x.

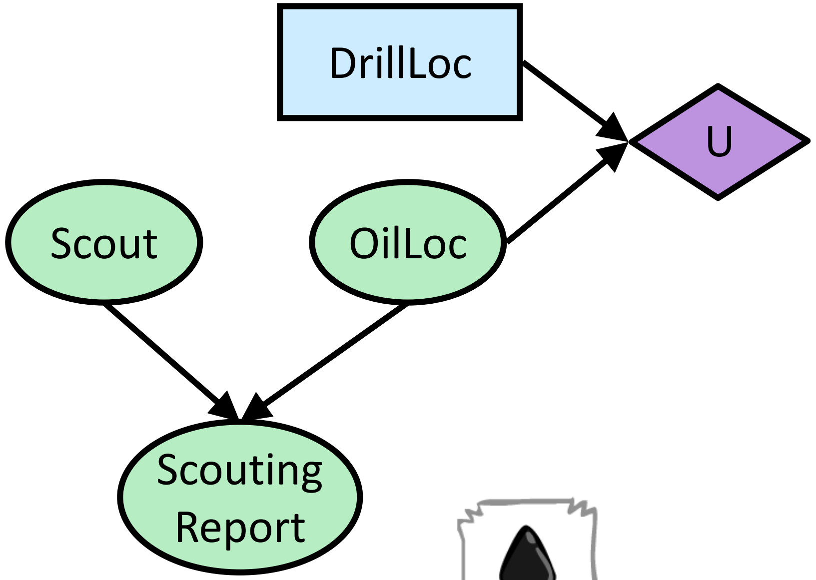


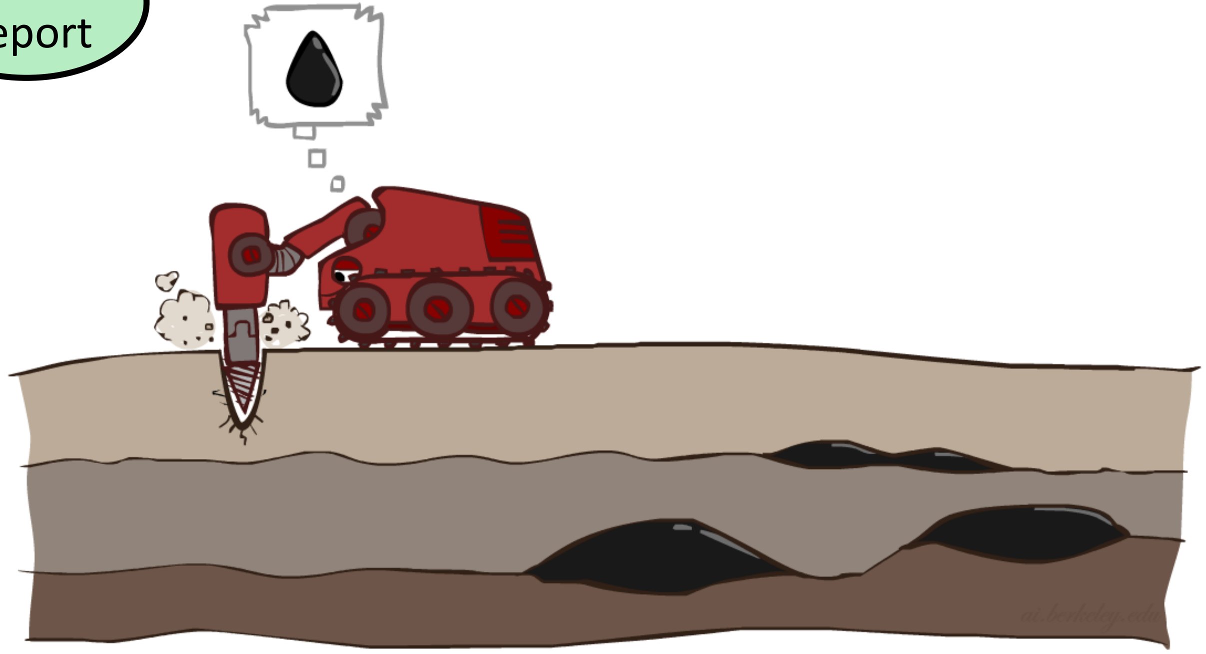
Value of Imperfect Information?

- No such thing
- Information corresponds to the observation of a node in the decision network
- If data is “noisy” that just means we don’t observe the original variable, but another variable which is a noisy version of the original one

VPI Question

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



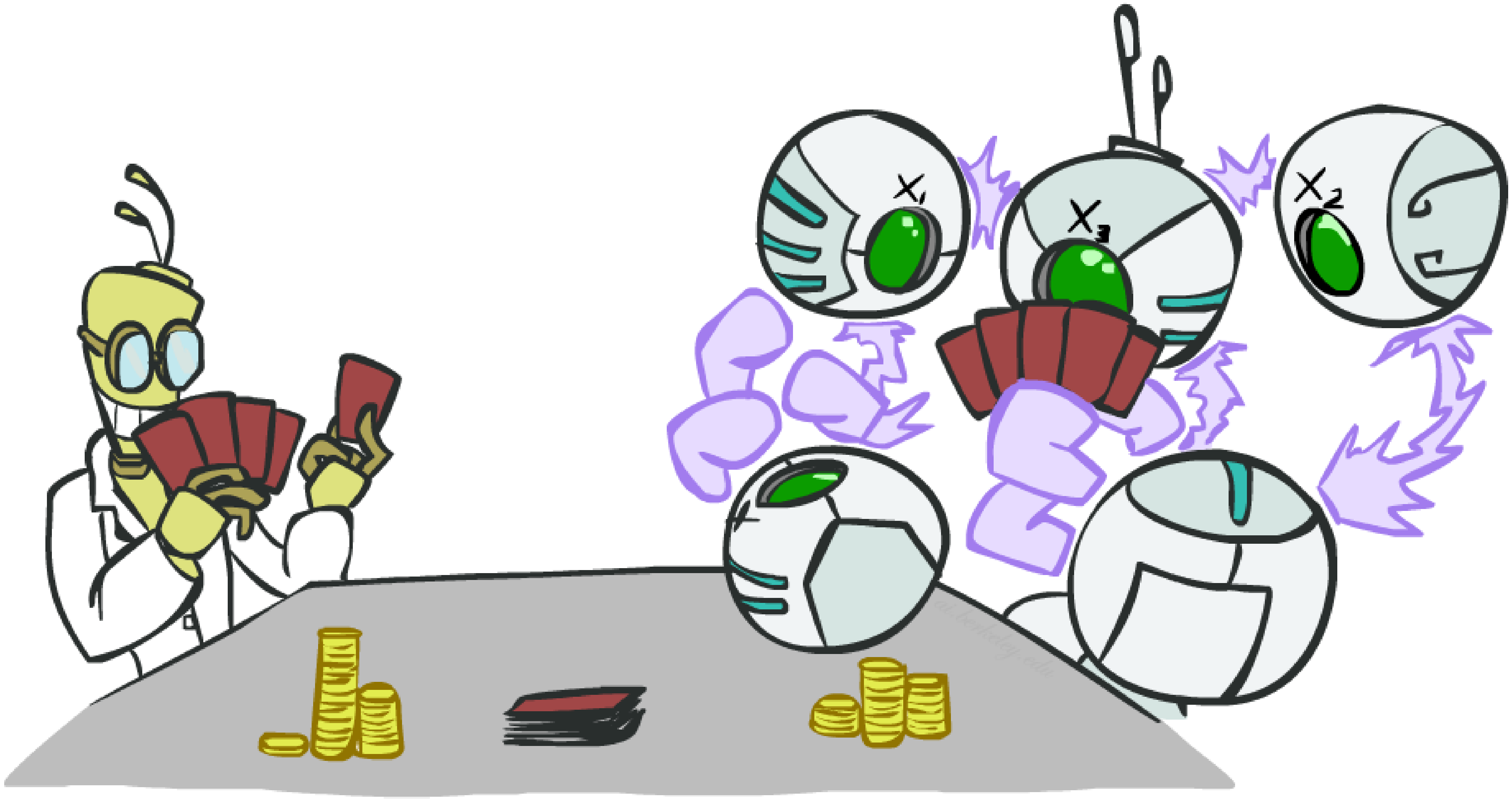


- Generally:

  If $\quad \text{Parents}(U) \perp\!\!\!\perp Z \mid \text{CurrentEvidence}$

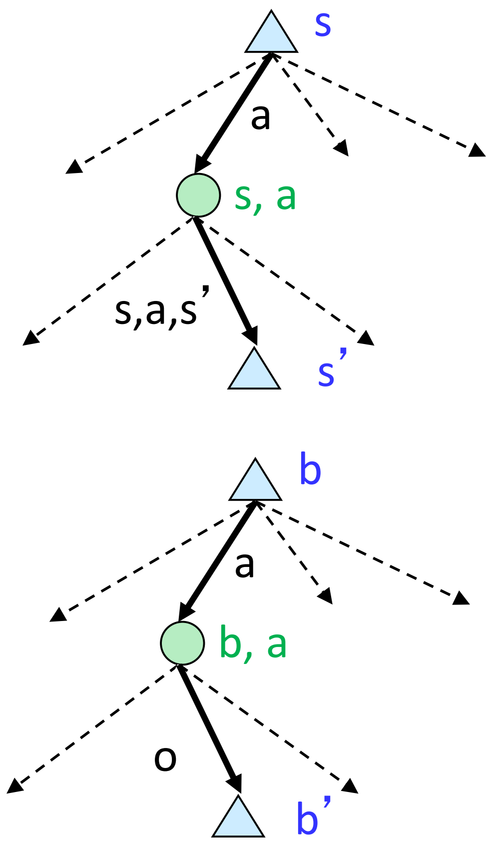
  Then $\quad \text{VPI}(Z \mid \text{CurrentEvidence}) = 0$

POMDPs



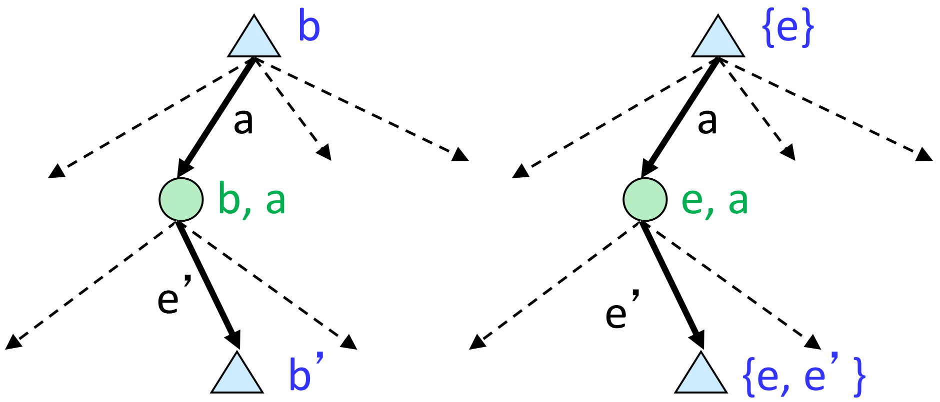
POMDPs

- 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

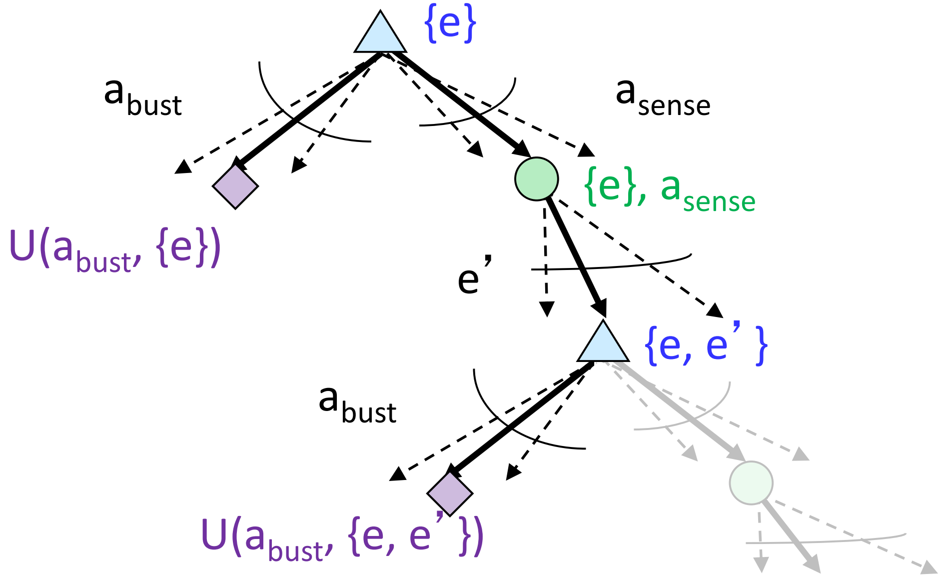


Example: Ghostbusters

- 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



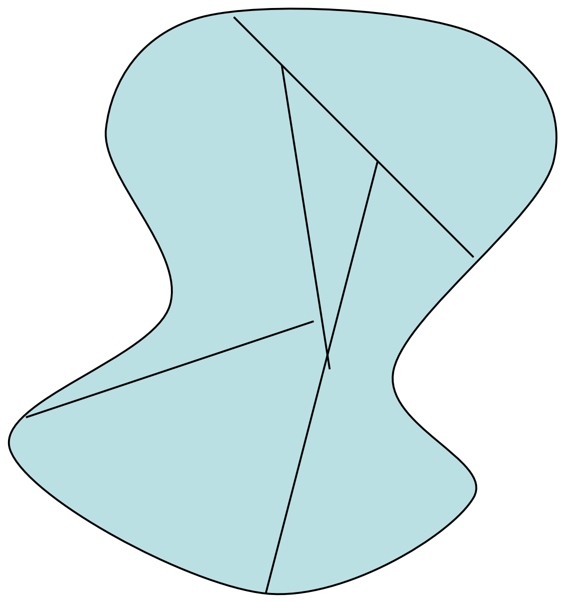
- Solving POMDPs
  - One way: use truncated expectimax to compute approximate value of actions
  - Example: What if you only considered busting or one sense followed by a bust?
  - You get a VPI-based agent!



Video of Demo Ghostbusters with VPI

More Generally*

- General solutions map belief functions to actions
  - Can divide regions of belief space (set of belief functions) into policy regions (gets complex quickly)
  - Can build approximate policies using discretization methods
  - Can factor belief functions in various ways
- Overall, POMDPs are very (actually PSPACE-) hard
- Most real problems are POMDPs, but we can rarely solve then in general!



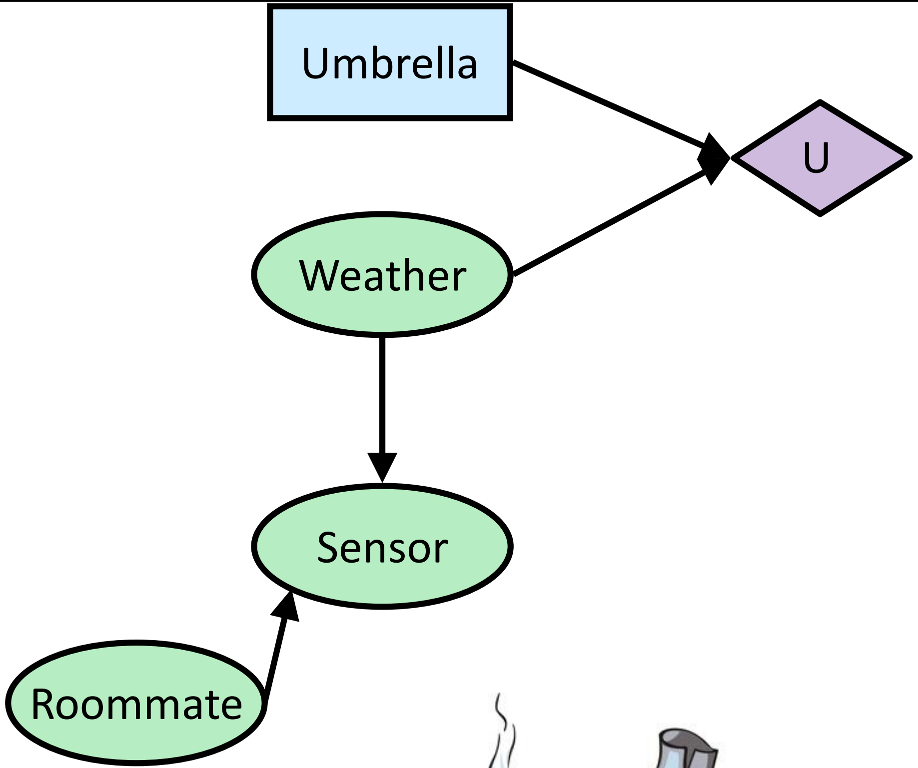
Value of Roommate Information

- Suppose we have our roommate read the sensor.
  - There is a chance they might mess up.
  - Calculating VPI(Roommate) is just the same as before.

- Suppose we know that $\mathrm{R} = \mathrm{said\_bad}$
  - Also, suppose that $MEU(R = \mathrm{said\_bad}) = 49$ (can calculate, but tedious).

- What is $VPI(Sensor|\mathrm{R} = \mathrm{said\_bad})$?
  - $= P(S = \mathrm{good}|R = \mathrm{said\_bad}) \cdot MEU(S = \mathrm{good}, R = \mathrm{said\_bad})$
    $+ P(S = \mathrm{bad}|R = \mathrm{said\_bad}) \cdot MEU(S = \mathrm{bad}, R = \mathrm{said\_bad})$
    $- MEU(\mathrm{said\_bad})$

  - $= 0.1 \cdot 95 + 0.9 \cdot 53 - 49 = 57.2 - 49 = 8.2$

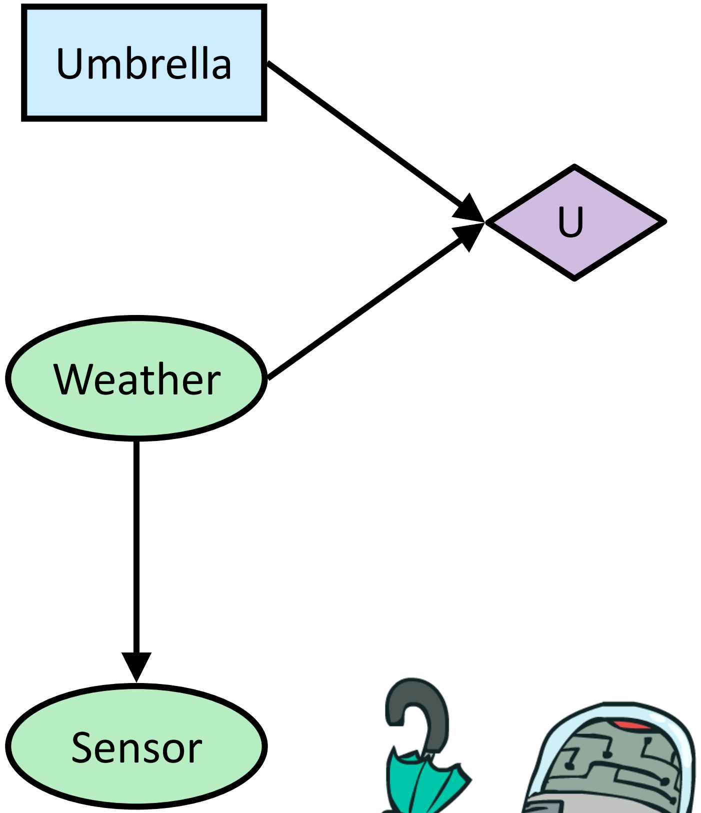


S	R	P(R|S)
good	said_good	0.9
good	said_bad	0.1
bad	said_good	0.1
bad	said_bad	0.9

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

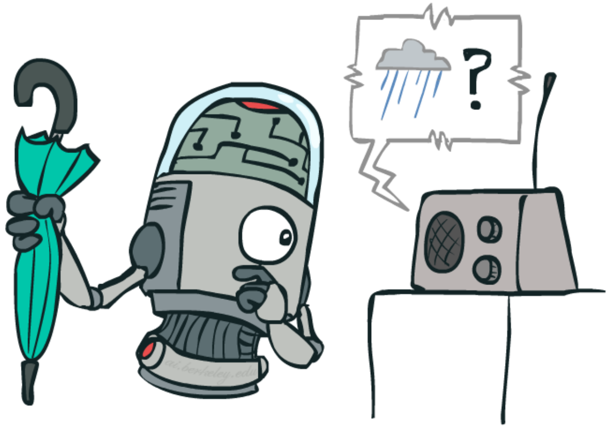
Appendix:

- If we have a sensor reading of good:
  - The maximum expected utility is 95, and we should leave.
  - Why 95? Can calculate from tables below.



W	P(W|S=good)
sun	0.95
rain	0.05

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



Conditional Value of Weather Information

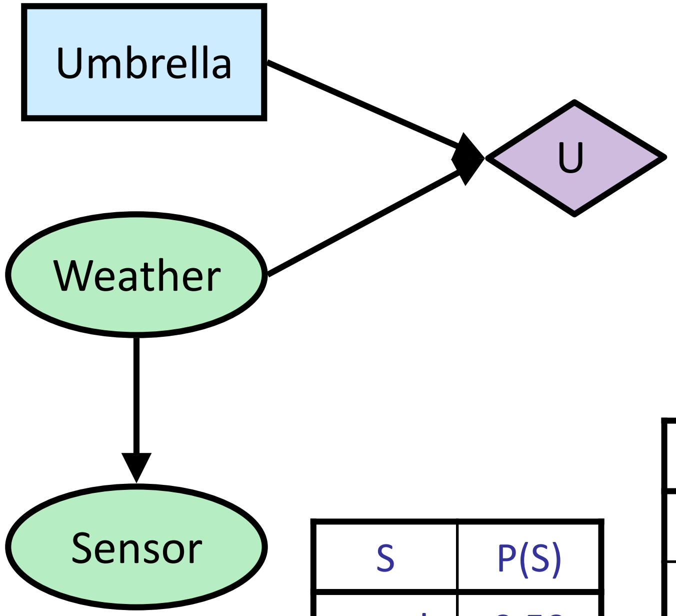
- Suppose we already know that S = bad.

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

- What is $VPI(Weather|S = \text{bad})$?
  - $= P(W = \text{sun}|S = \text{bad}) \cdot MEU(W = \text{sun}|S = \text{bad})$
    $+ P(W = \text{rain}|S = \text{bad}) \cdot MEU(W = \text{rain}|S = \text{bad})$
    $- MEU(S = \text{bad})$
  - $= 0.34 \cdot 100 + 0.66 \cdot 70 - 53 = 80.2 - 53 = 27.2$

- Observation:
  - Willing to pay 21 for direct observation of weather.
  - But given a bad forecast, willing to pay 27.2 for true weather value.

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



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

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

S	P(S)
good	0.59
bad	0.41