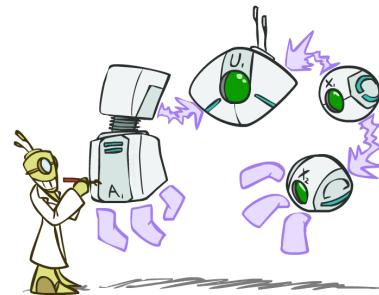


CS 188: Artificial Intelligence



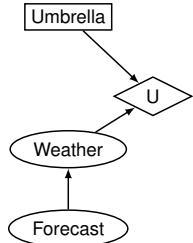
Decision Networks and Value of Perfect Information

Decision Networks



Decision Networks

MEU: choose action which maximizes expected utility given evidence

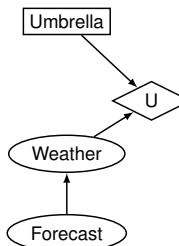


- Directly operationalize with decision networks
- Bayes nets with nodes for **utility** and **actions**
- Can calculate expected utility for actions

New node types:

- Chance nodes (just like BNs)
- Actions (rectangles, no parents, acts as observed evidence)
- Utility node (diamond, depends on action/chance nodes)

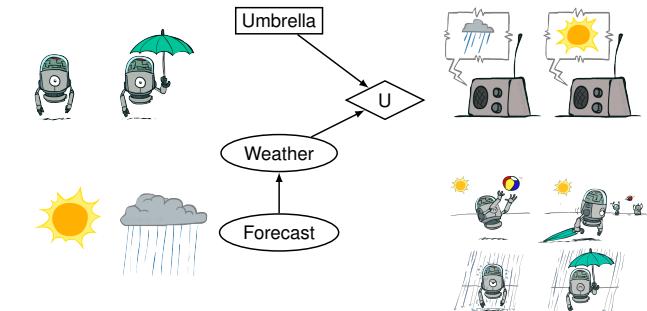
Decision Networks



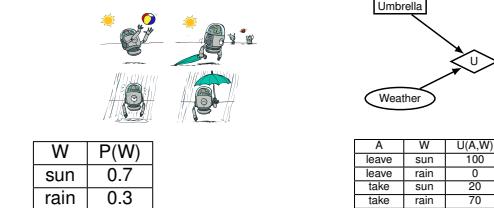
Action selection

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

Decision Networks



Decision Networks



Umbrella = leave

$$EU(\text{leave}) = \sum_w P(w) U(\text{leave}, w) = 0.7 \times 100 + 0.3 \times 0 = 70$$

Umbrella = take

$$EU(\text{take}) = \sum_w P(w) U(\text{take}, w) = 0.7 \times 20 + 0.3 \times 70 = 35$$

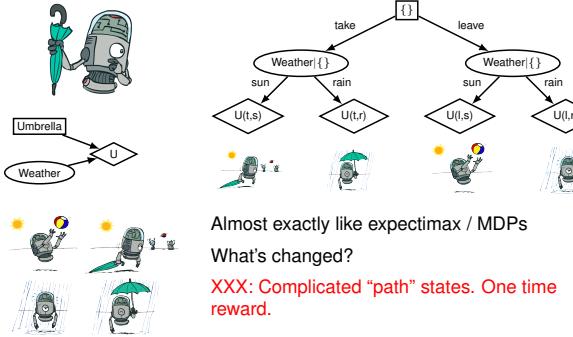
Optimal decision = leave

$$MEU(\phi) == \max_a EU(a).$$

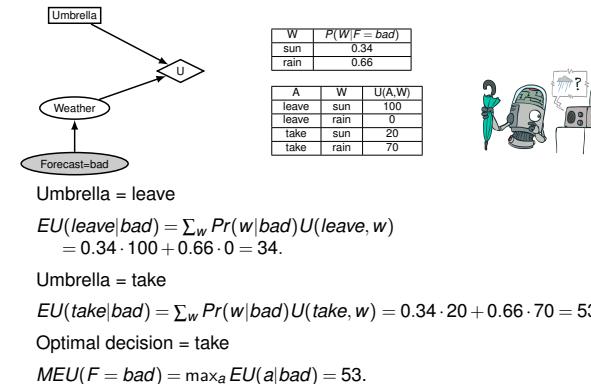
W	P(W)
sun	0.7
rain	0.3

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

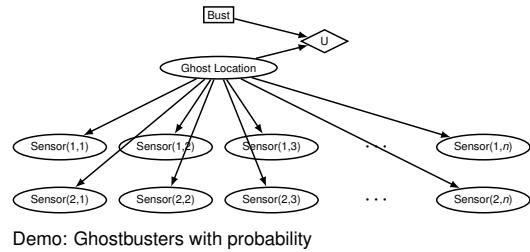
Decisions as Outcome Trees



Example: Decision Networks



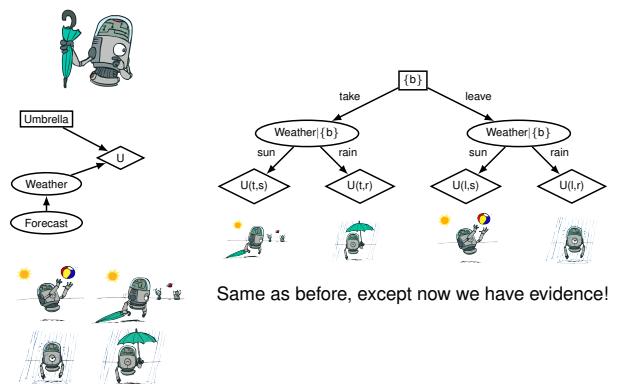
Ghostbusters Decision Network



Video of Demo Ghostbusters with Probability



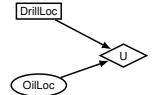
Decisions as Outcome Trees



Value of Information

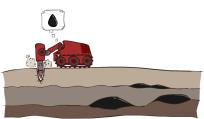


Value of Information



D	O	U
a	1/2	k
b	0	0
b	0	k

O	P
a	1/2
b	1/2



- Idea: compute value of acquiring evidence
- Can be done directly from decision network
- Example: buying oil drilling rights
- Two blocks, oil in one of A or B. Equally likely.
 - Can drill in one location
 - Drilling in A or B: has $EU = k/2$, $MEU = k/2$
- Question: what's value of information of O?
- Value of knowing whether A or B has oil.
 - Value is expected gain in MEU from info.
 - Survey may say "oil in a" or "oil in b."
 - If know OilLoc, **MEU is k** (either way).
 - Gain in MEU from knowing OilLoc?
 - $VPI(OilLoc) = k/2$
 - Fair price of information: **k/2**

VPI Properties



Nonnegative:

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

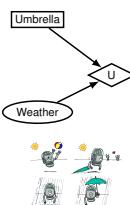
Nonadditive (think of observing E_j twice)

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

Order-independent

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

VPI Example: Weather



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

F	P(F)
good	0.59
bad	0.41

MEU with no evidence: 70

MEU if forecast is bad:

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

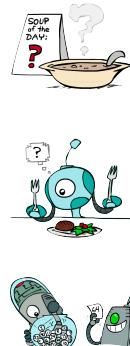
MEU if forecast is good:

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

Forecast distribution:

$$\begin{aligned} VPI(E'|e) &= (\sum_{e'} P(e' | e) MEU(e, e')) - MEU(e) \\ VPI &= 0.59 \cdot (95) + 0.41 \cdot (53) - 70 = 77.8 - 70 = 7.8 \end{aligned}$$

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'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?

Value of Information

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

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

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

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

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:

$$MEU(e, E') = \sum_e P(e'|e) MEU(e, e')$$

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

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

Value of Imperfect Information?

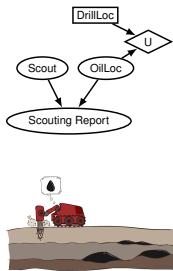
No such thing

Information corresponds to the observation of a node in the decision network

"Noisy" means don't observe another variable: noisy version of original variable



VPI Question

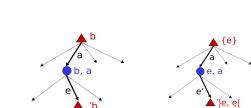


- VPI(OilLoc) ?
- VPI(ScoutingReport) ?
- VPI(Scout) ?
- VPI(Scout | ScoutingReport) ?
- Generally:
 - If Parents(U) ⊥ Z | CurrentEvidence
 - Then VPI(Z | CurrentEvidence) = 0

POMDPs



Example: Ghostbusters



In (static) Ghostbusters:

- Belief state determined by evidence e.
- Tree really over evidence sets.
- Probabilistic reasoning predicts new evidence given past evidence

Solving POMDPs

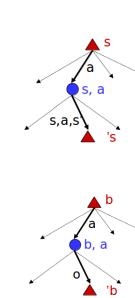
- One way: use truncated expectimax to compute approximate value of actions
- What if you only considered busting or one sense followed by a bust?
- You get a VPI-based agent!

Demo: Ghostbusters with VPI

Video of Demo Ghostbusters with VPI



POMDPs



MDPs have:

- States S
- Actions A
- Transition function $P(s'| s,a)$ (or $T(s,a,s')$)
- Rewards $R(s,a,s')$

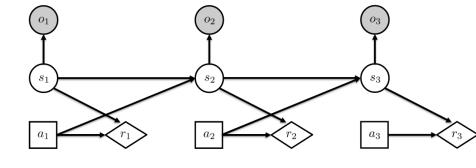
POMDPs add:

- Observations O
- Observation function $P(o | s)$ (or $O(s,o)$)

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

This is like belief states in Hidden Markov Models!

POMDPs as Decision Networks



MDP:

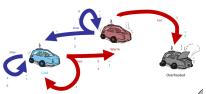
- States: S
- Actions: A
- Rewards: $R(s,a,s')$

POMCPs add:

- Observations: O
- Observation: $P(o|s)$

POMDPs More Generally*

How can we solve POMDPs?



POMDP is MDP over *belief* b .

$s \in \{\text{cool, warm, overheated}\}$.

$b \in [0, 1]^3 \leftarrow \text{vector of three continuous numbers!}$

Use: Value iteration, policy iteration, etc. over b .

b is continuous (discretize or use functions.)

b is very big (one number for each state.)

POMDPs More Generally*

General solutions map belief functions to actions

- 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 them in general! E.g. depth limited in ghostbusters.