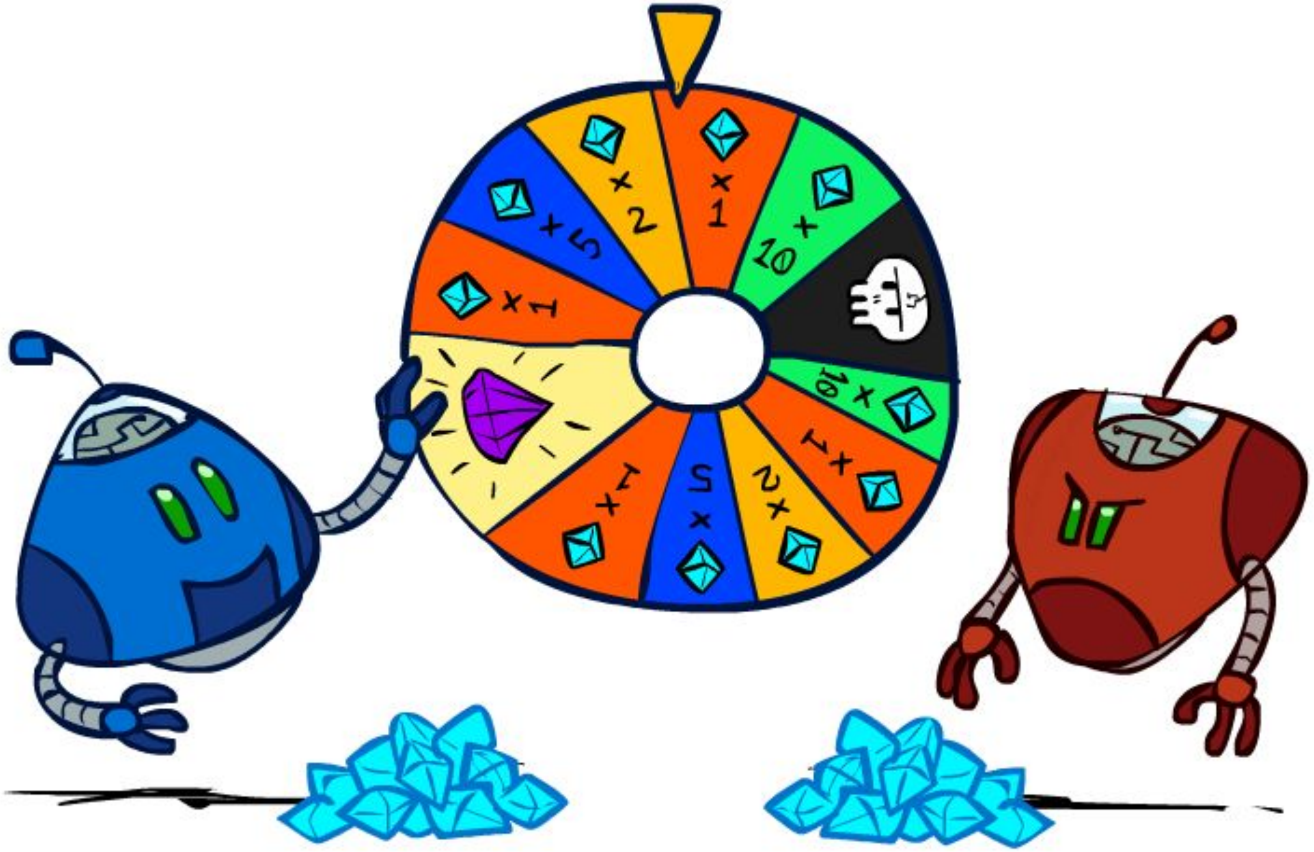
CS 188: Artificial Intelligence

Decision Networks and VPI



Instructor: Stuart Russell and Dawn Song

University of California, Berkeley

Rational Preferences

The Axioms of Rationality

?

Rational Preferences

The Axioms of Rationality

Orderability:

$$(A > B) \vee (B > A) \vee (A \sim B)$$

Transitivity:

$$(A > B) \wedge (B > C) \Rightarrow (A > C)$$

Continuity:

$$(A > B > C) \Rightarrow \exists p\ [p, A;\ 1\text{-}p, C] \sim B$$

Substitutability:

$$(A \sim B) \Rightarrow [p, A;\ 1\text{-}p, C] \sim [p, B;\ 1\text{-}p, C]$$

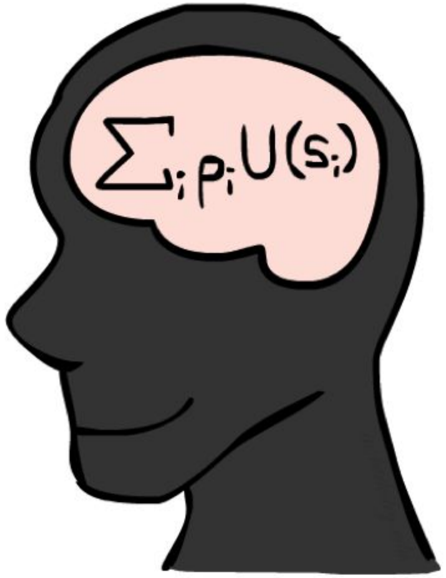
Monotonicity:

$$(A > B) \Rightarrow$$
$$(p \geq q) \Leftrightarrow [p, A;\ 1\text{-}p, B] \geq [q, A;\ 1\text{-}q, B]$$

What's the implication?

MEU Principle

- Theorem [Ramsey, 1931; von Neumann & Morgenstern, 1944]
  - Given any preferences satisfying the previous constraints, there exists a real-valued function *U* such that:

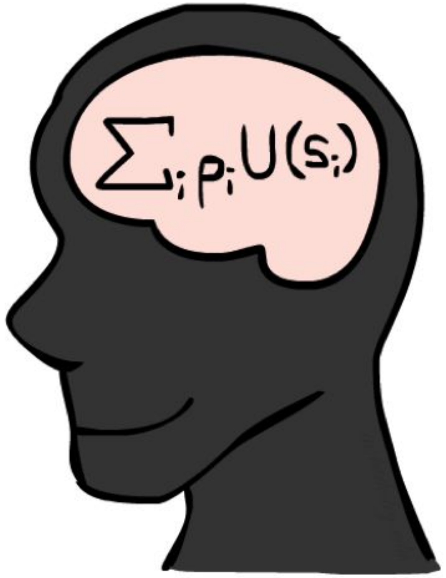


MEU Principle

- Theorem [Ramsey, 1931; von Neumann & Morgenstern, 1944]
  - Given any preferences satisfying the previous constraints, there exists a real-valued function U such that:

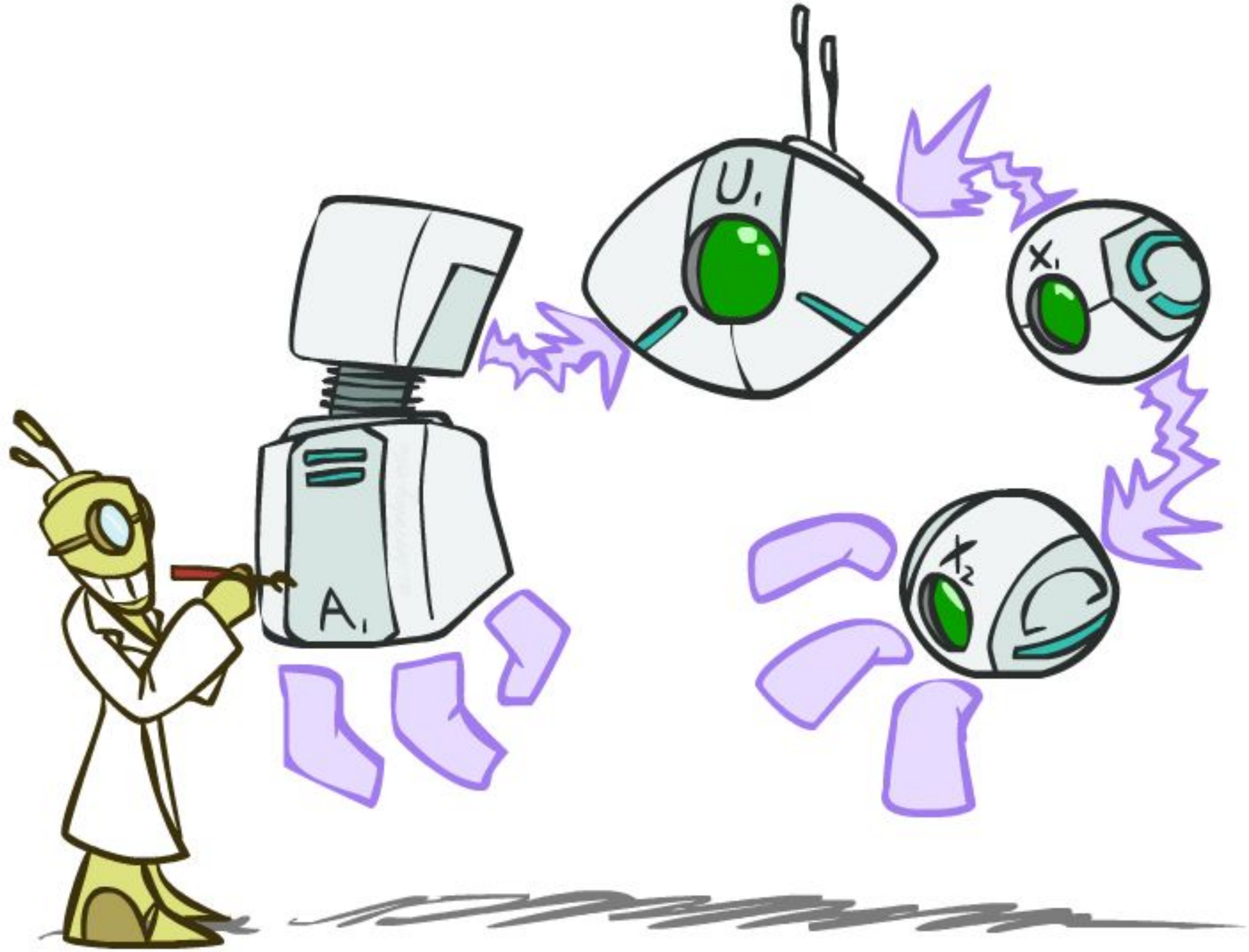
$$U(A) > U(B) \Leftrightarrow A > B \text{ and } U(A) \sim U(B) \Leftrightarrow A \sim B$$

$$U([p_1, S_1; \ldots ; p_n, S_n]) = p_1 U(S_1) + \ldots + p_n U(S_n)$$



- Maximum expected utility (MEU) principle:
  - A rational agent chooses the action that maximizes expected utility

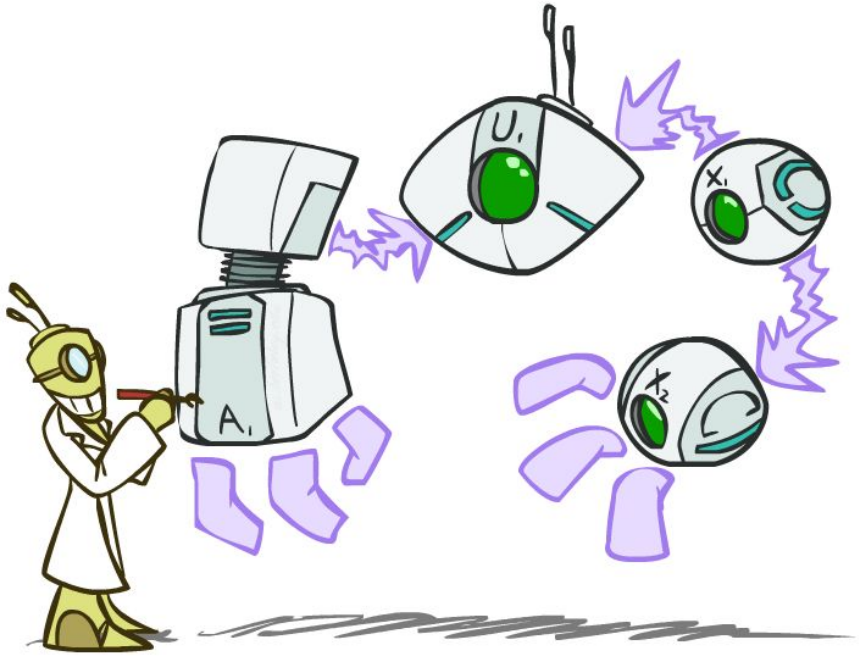
Decision Networks



Decision Networks

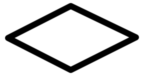
In its most general form, a decision network represents information about

- Its current state
- Its possible actions
- The state that will result from its actions
- The utility of that state



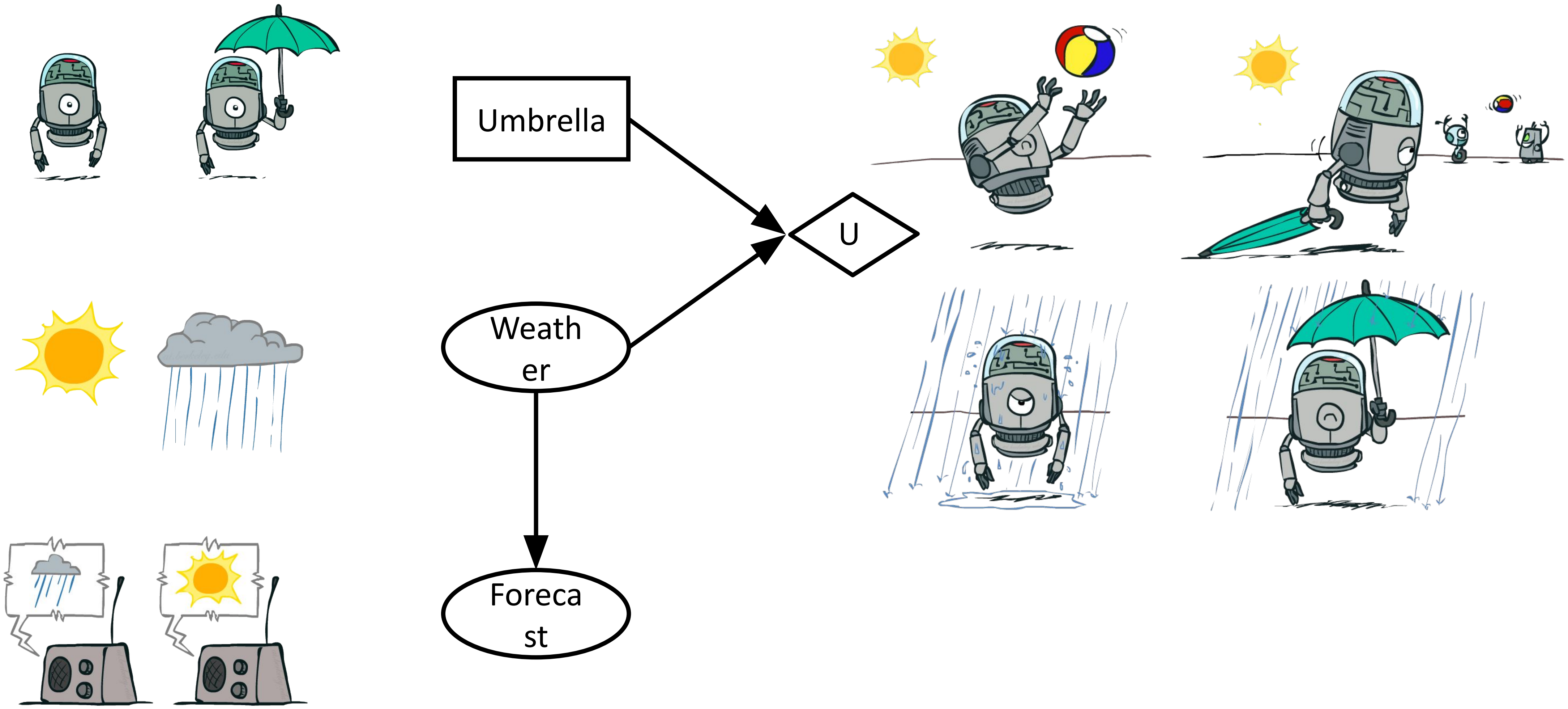
Decision network = Bayes net + Actions + Utilities

- ***Action nodes*** (rectangles, cannot have parents, will have value fixed by algorithm)



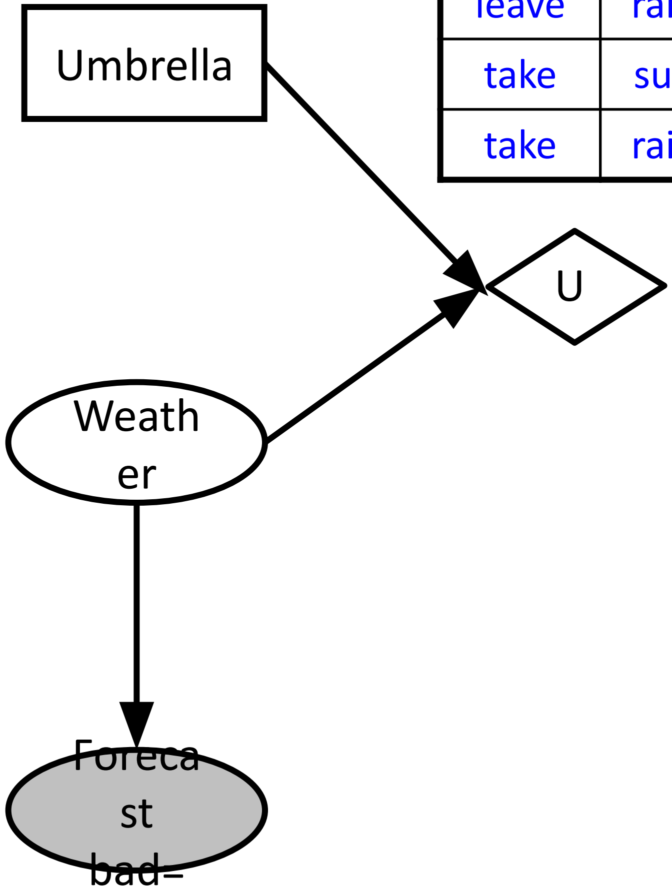
- ***Utility nodes*** (diamond, depends on action and chance nodes)

Decision Networks



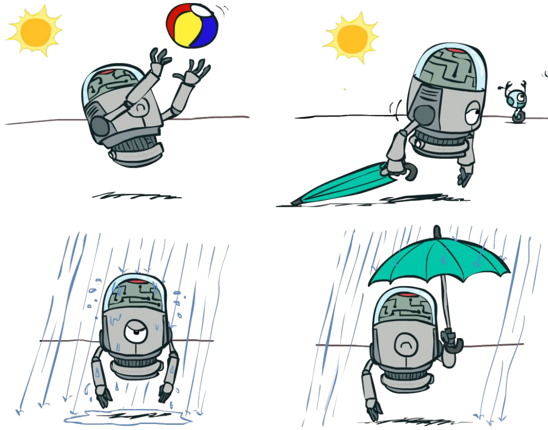
Example: Take an umbrella?

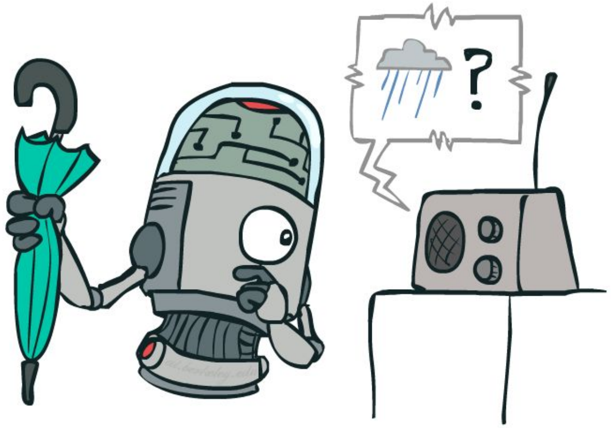
A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70



W	P(W)
sun	0.7

W	P(F=bad|W)
sun	0.17
rain	0.77

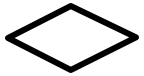




Decision Networks

- Decision network = Bayes net + Actions + Utilities

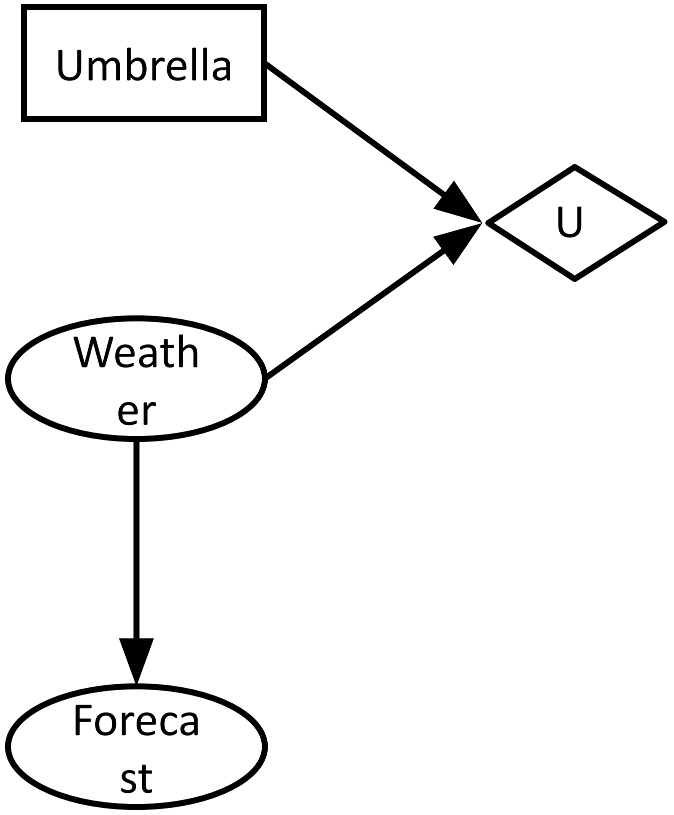
- ***Action nodes*** (rectangles, cannot have parents, will have value fixed by algorithm)



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

- Decision algorithm:
  - Fix evidence $\mathbf{e}$
  - For each possible action a
    - Fix action node to a
    - Compute posterior $P(\mathbf{W}|\mathbf{e},a)$ for parents $\mathbf{W}$ of U
    - Compute expected utility $\sum_w P(\mathbf{w}|\mathbf{e},a)\ U(a,\mathbf{w})$
  - Return action with highest expected utility

Bayes net inference!

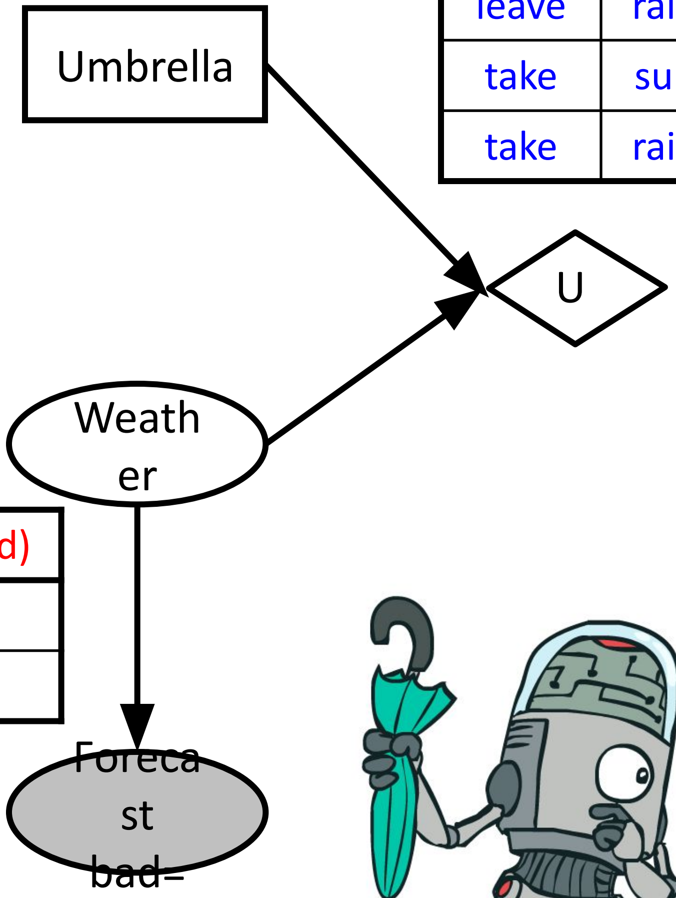


Example: Take an umbrella?

- Decision algorithm:
  - Fix evidence ***e***
  - For each possible action *a*
    - Fix action node to *a*
    - Compute posterior $P(\boldsymbol{W}|\boldsymbol{e},a)$ for parents $\boldsymbol{W}$ of U
    - Compute expected utility of action a: $\sum_w P(\boldsymbol{w}|\boldsymbol{e},a)\, U(a,\boldsymbol{w})$
  - Return action with highest expected utility

Bayes net inference!

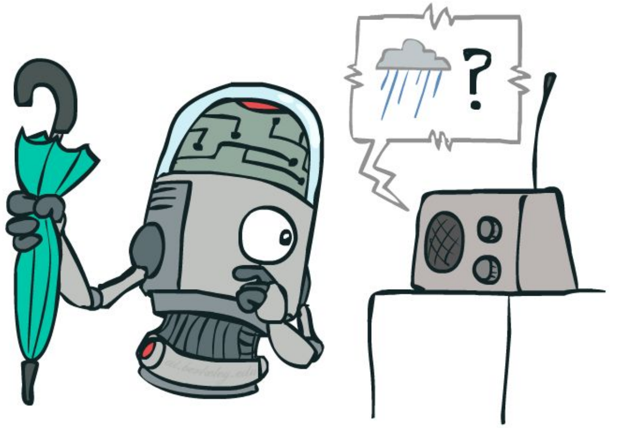
A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70



W	P(W)
sun	0.7

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

W	P(F=bad|W)
sun	0.17
rain	0.77



Umbrella = leave

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

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

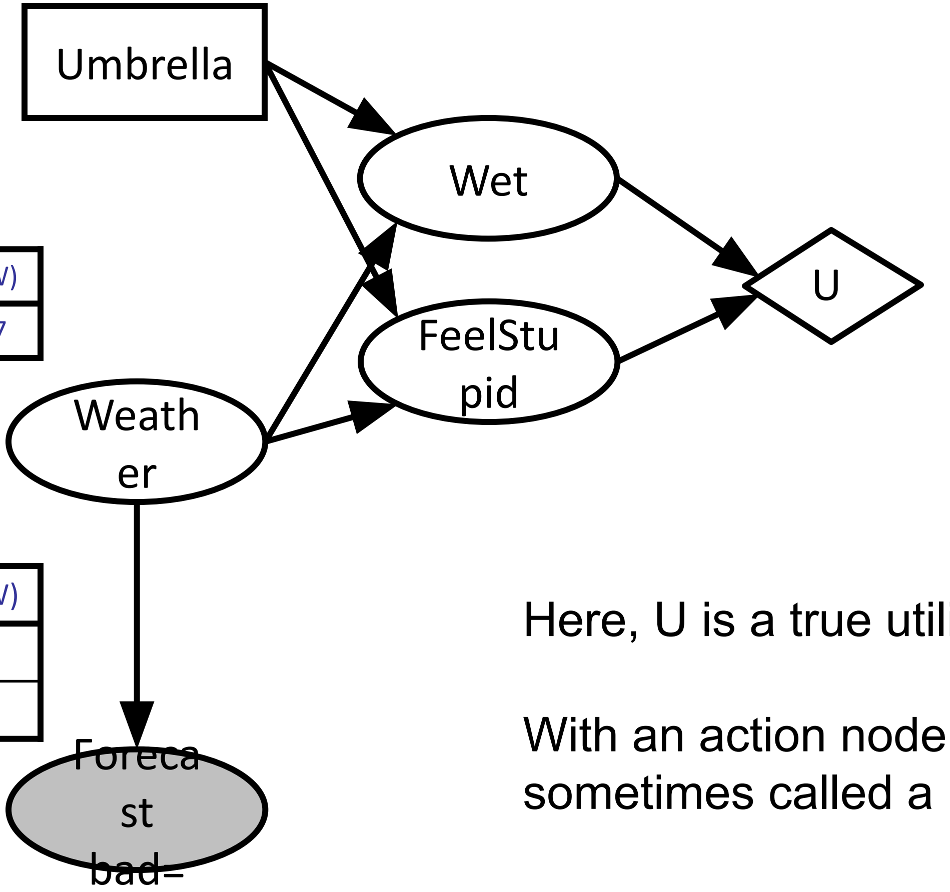
Umbrella = take

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

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

Optimal decision = take!

Decision network with utilities on outcome states



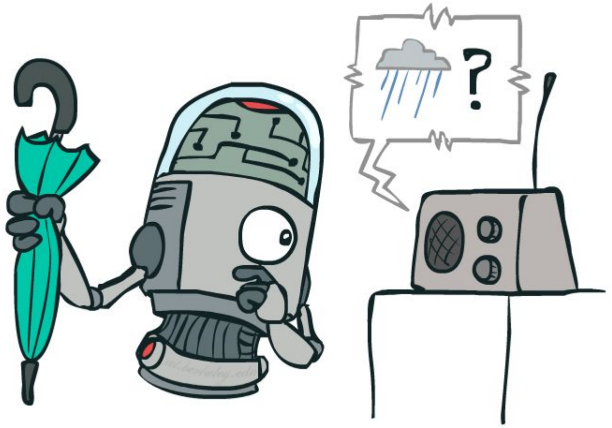
W	P(W)
sun	0.7

W	P(F=bad|W)
sun	0.17
rain	0.77

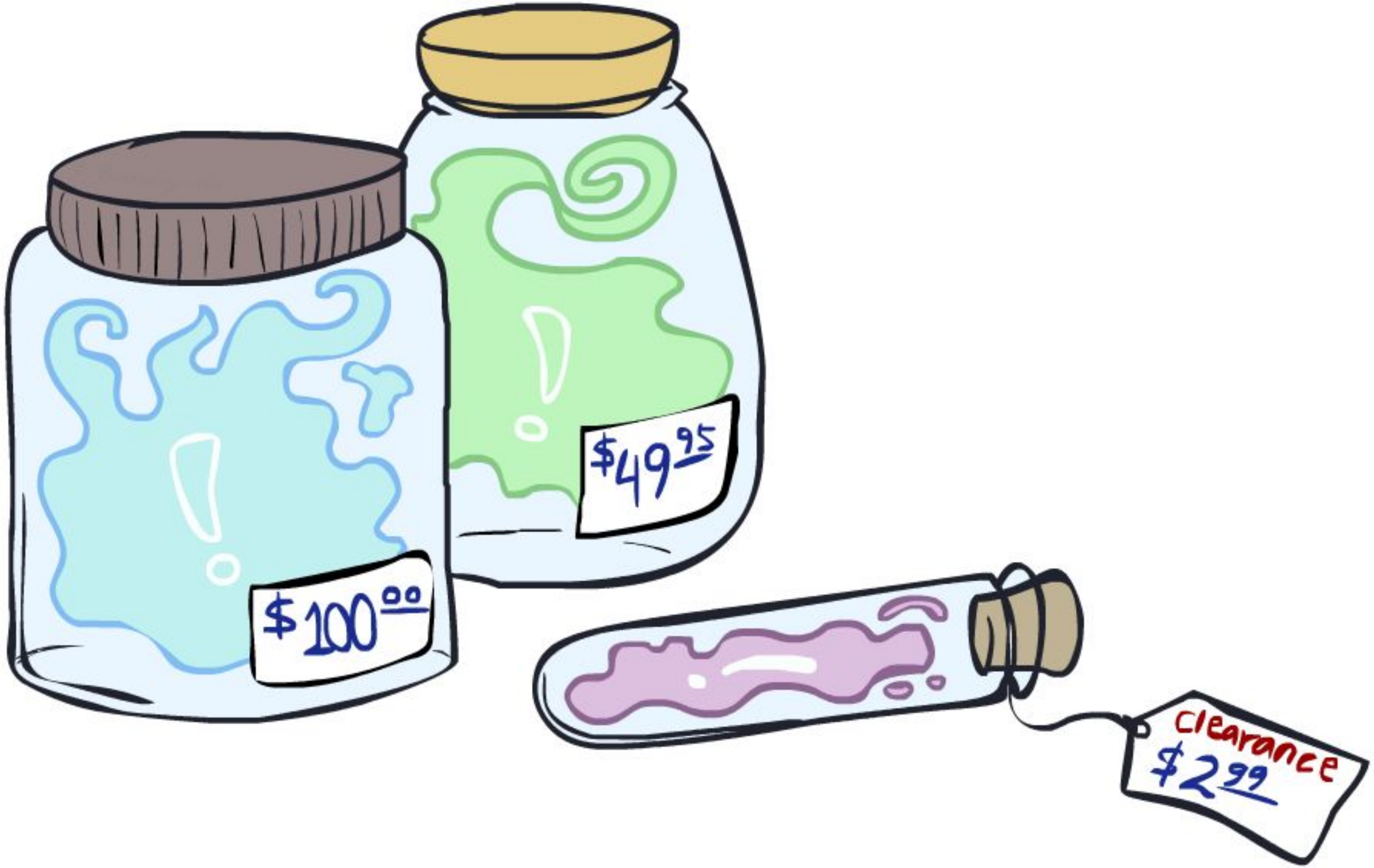
Wet	FeelStupid	U
dry	false	100
wet	true	0
dry	true	20
damp	false	70

Here, U is a true utility.

With an action node as parent, it is sometimes called a ***Q-value***



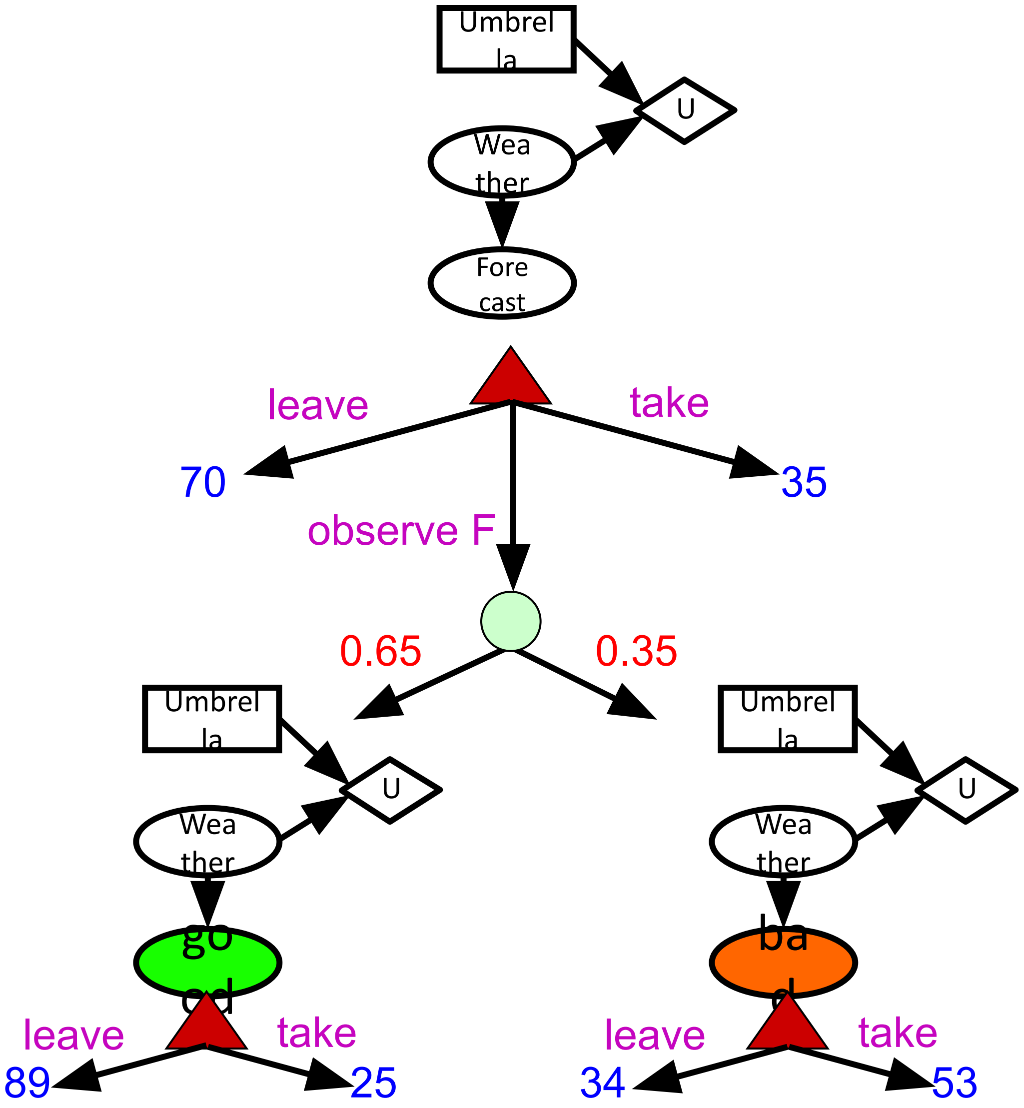
Value of Information



Value of information

- Suppose you haven't yet seen the forecast
  - EU(leave |) = 0.7x100 + 0.3x0 = 70
  - EU(take |) = 0.7x20 + 0.3x70 = 35
- ***What if you look at the forecast?***
- If Forecast=good
  - EU(leave | F=good) = 0.89x100 + 0.11x0 = 89
  - EU(take | F=good) = 0.89x20 + 0.11x70 = 25
- If Forecast=bad
  - EU(leave | F=bad) = 0.34x100 + 0.66x0 = 34
  - EU(take | F=bad) = 0.34x20 + 0.66x70 = 53
- P(Forecast) = <0.65,0.35>
- Expected utility given forecast
  - = 0.65x89 + 0.35x53 = 76.4
- ***Value of information*** = 76.4-70 = 6.4

Bayes net inference!



Value of information contd.

- General idea: value of information = ***expected improvement in decision quality*** from observing value of a variable
  - E.g., oil company deciding on seismic exploration and test drilling
  - E.g., doctor deciding whether to order a blood test
  - E.g., person deciding on whether to look before crossing the road
- Key point: decision network contains everything needed to compute it!
- $\mathrm{VPI}(E_i \mid \mathrm{e}) = \left[\sum_{e_i} P(e_i \mid e) \max_a \mathrm{EU}(a \mid e_i, e) \right] - \max_a \mathrm{EU}(a \mid e)$

VPI Properties

VPI is non-negative! $VPI(E_i \mid e) \geq 0$

VPI is not (usually) additive: $VPI(E_i, E_j \mid e) \neq VPI(E_i \mid e) + VPI(E_j \mid e)$

VPI is order-independent: $VPI(E_i, E_j \mid e) = VPI(E_j, E_i \mid e)$

Decisions with unknown preferences

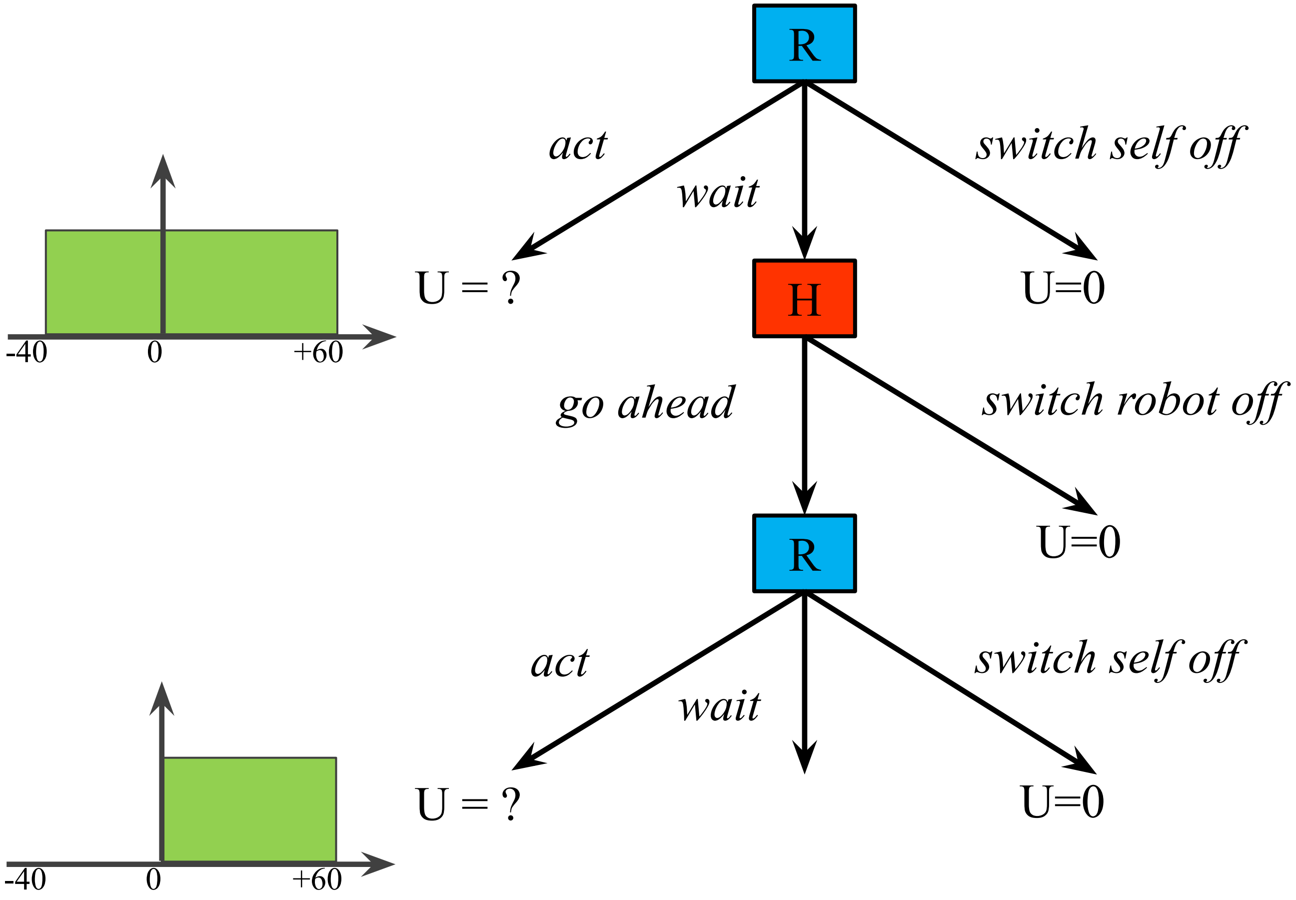
- In reality the assumption that we can write down our exact preferences for the machine to optimize is false
- A machine optimizing the wrong preferences causes problems

I’m sorry, Dave, I’m afraid I can’t do that

Decisions with unknown preferences

- In reality the assumption that we can write down our exact preferences for the machine to optimize is false
- A machine optimizing the wrong preferences causes problems
- A machine that is explicitly uncertain about the human's preferences will defer to the human (e.g., allow itself to be switched off)

Off-switch problem (example)



EU(act) = +10

EU(wait) = (0.4 * 0) + (0.6 * 30) = +18

Off-switch problem (general proof)

- $EU(act) = \int_{-\infty}^{+\infty} P(u) \cdot u \, du = \int_{-\infty}^{0} P(u) \cdot u \, du + \int_{0}^{+\infty} P(u) \cdot u \, du$
- $EU(wait) = \int_{-\infty}^{0} P(u) \cdot 0 \, du + \int_{0}^{+\infty} P(u) \cdot u \, du$
- Obviously $\int_{-\infty}^{0} P(u) \cdot u \, du \leq \int_{-\infty}^{0} P(u) \cdot 0 \, du$
- Hence $EU(act) \leq EU(wait)$