Decision Making
in Robots and Autonomous Agents

Utility and Decision Theory:
How should a robot incorporate notions of choice?

Subramanian Ramamoorthy
School of Informatics

2 February, 2018

Types of Decisions

- Who makes it?
  - Individual
  - ‘Group’
- What are the conditions?
  - Certainty
  - Risk
  - Uncertainty

For much of this course, we’ll take the ‘individual’ viewpoint (potentially acting in conflict of interest scenarios), and we’ll be somewhere between risk and uncertainty

How to Model Decision under *Certainty*?

- Given a set of possible acts
- Choose one that maximizes some given index

If **a** is a generic act in a set of feasible acts **A**, f(**a**) is an index being maximized, then

Problem: Find **a*** in **A** such that f(**a***) > f(**a**) for all **a** in **A**.

The index f plays a key role, e.g., think of buying a painting.

Essential problem: How should the subject select an index function such that her choice reduces to finding maximizers?

Operational Way to Find *an* Index Function

- Observe subject's behaviour in restricted settings and predict purchase behaviour from that:
- Instruct the subject as follows:
  - Here are ten valuable reproductions
  - We will present these to you in pairs
  - You will tell us which one of the pair you prefer to own
  - After you have evaluated all pairs, we will pick a pair at random and present you with the choice you previously made (it is to your advantage to remember your true tastes)
- The subject's behaviour is **as though** there is a ranking over all paintings, so each painting can be summarized by a number

Some Properties of this Ranking

- *Transitivity*: Previous argument only makes sense if the rank is transitive – if A is preferred in (A, B) and B is preferred in (B, C) then A is preferred in (A, C); and this holds for all triples of alternatives A, B and C
- *Ordinal nature of index*: One is tempted to turn the ranking into a latent measure of 'satisfaction' but that is a mistake as utilities are non-unique.

  e.g., we could assign 3 utiles to A, 2 utiles to B and 1 utile to C to explain the choice behaviour

  Equally, 30, 20.24 and 3.14 would yield the same choice

While it is OK to compare indices, it is not OK to add or multiply

What Happens if we Relax Transitivity?

- Assume Pandora says (in the pairwise comparisons):
  - Apple < Orange
  - Orange < Fig
  - Fig < Apple
- Is this a problem for Pandora? Why?

- Assume a merchant who transacts with her as follows:
  - Pandora has an Apple at the start of the conversation
  - He offers to exchange Orange for Apple, if she gives him a penny
  - He then offers an exchange of Fig for Orange, at the price of a penny
  - Then, offers Apple for the Fig, for a penny
  - **Now, what is Pandora's net position?**

Decision Making under *Risk*

- Initially appeared as analysis of fair gambles, needed some notions of utility
- Gamble has n outcomes, each worth $a_1, \ldots, a_n$
- The probability of each outcome is $p_1, \ldots, p_n$
- How much is it worth to participate in this gamble?

$$b = a_1 p_1 + \ldots + a_n p_n$$

One may treat this monetary expected value as a fair price

Is this a sufficient description of choice behaviour under risk?

St. Petersburg Paradox of D. Bernoulli

- A fair coin is tossed until a head appears
- Gambler receives 2^n if the first head appears on trial n
- Probability of this event = probability of tail in first (n-1) trials and head on trial n, i.e., $(1/2)^n$

  Expected value $= 2.(1/2) + 4.(1/2)^2 + 8.(1/2)^8 + \ldots = \infty$

- Are you willing to bet in this way? Is anyone?

Defining Utility

- Bernoulli went on to argue that people do not act in this way
- The thing to average is the 'intrinsic worth' of the monetary values, not the absolute values

  e.g., intrinsic worth of money may increase with money but at a *diminishing rate*

- Let us say utility of m is $\log_{10} m$, then expected value is,

$$\log_{10} 2.(1/2) + \log_{10} 4.(1/2)^2 + \log_{10} 8.(1/2)^8 + \ldots = b < \infty$$

  Monetary fair price of the gamble is a where $\log_{10} a = b$.

Some Critiques of Bernoulli's Formulation

von Neumann and Morgenstern (vNM), who 'started' game theory, raised the following questions:

- The assignment of utility to money is arbitrary and *ad hoc*
  - There are an infinity of functions that capture 'diminishing rate', how should we choose?
  - The association may vary from person to person
- Why is the definition of the decision based upon expected value of this notion of utility?
  - Is this actually descriptive of a single gambler, in "one-shot" choice?

von Neumann & Morgenstern Formulation

- If a person is able to express preferences between every possible pair of gambles

  where gambles are taken over some basic set of alternatives

- Then one *can* introduce utility associations to the basic alternatives in such a manner that

- If the person is guided solely by the utility expected value, *he is acting in accord with his true tastes*.
  - provided his tastes are consistent in some way

Constructing Utility Functions

- Suppose we know the following preference order:
  - A < b ~ c < d < e

- The following are utility functions that capture this:

	a	b	c	d	E
U	0	1/2	1/2	3/4	1
V	-1	1	1	2	3
W	-8	0	0	1	8

  - So, in situations like St Petersburg paradox, the revealed preference of any realistic player may differ from the case of infinite expected value
  - Satisfaction at some large value, risk tolerance, time preference, etc.

Certainty Equivalents and Indifference

- The previous statement applies equally well to certain events and gambles or lotteries
- So, even attitudes regarding tradeoffs between the two ought to be captured
- Basic issue – how to compare?
- Imagine the following choice (A > B > C pref.) : (a) you get B for certain, (b) you get A with probability p and C otherwise
- If p is near 1, option b is better; if p is near 0, then option a: there is a single point where we switch
- Indifference is described as something like

$$(2/3)\ (1) + (1 - 2/3)\ (0) = 2/3$$

A C B

Caveats

- As before, we need to remember that the utility values should not be mis-interpreted
- The number 2/3 is determines by choices among risky alternatives and reflect attitude to 'gambling'
- For instance, imagine a subject who would be indifferent to paying \$9 and a 50-50 chance of paying \$10 or nothing;
- This suggests utilities for \$0, -\$9, -\$10 are 1, ½, 0.
- However, we can't say it is just as enjoyable for him to go from -\$10 to -\$9 as it is to go from -\$9 to \$0!
- Subject's preferences among alternatives or lotteries come prior to numerical characterization of them

Axiomatic Treatment of Utility

vNM and others formalize the above to define axioms for utility:

1) Any two alternatives shall be comparable, i.e., given any two, subject will prefer one over the other of be indifferent
2) Both preference and indifference relations for lotteries are transitive
3) In case a lottery has as one of its alternatives another lottery, then the first lottery is decomposable into the more basic alternatives through the use of the probability calculus
4) If two lotteries are indifferent to the subject then they are interchangeable as alternatives in any compound lottery

Axiomatic Treatment of Utility, contd.

vNM and others formalize the above to define axioms for utility:

5) If two lotteries involve the same two alternatives, then the one in which the more preferred alternative has a higher probability of occurring is itself preferred
6) If A is preferred to B and B to C, then there exists a lottery involving A and C (with appropriate probabilities) which is indifferent to B

Decision Making under *Uncertainty*

- A choice must be made from among a set of acts, $A_1, \ldots, A_m$.
- The relative desirability of these acts depends on which state of nature prevails, either $s_1, \ldots, s_n$.
- As decision maker we know that one of several things is true and this influences our choice but we **do not** have a probabilistic characterization of these alternatives

- Savage's omelet problem: Your friend has broken 5 good eggs into a bowl when you come in to volunteer and finish the omelet. A sixth egg lies unbroken (you must use it or waste it altogether). Your three acts: break it into bowl, break it into saucer – inspect and pour into bowl, throw it uninspected

Decision Making under *Uncertainty*

Table 1. Savage's example illustrating acts, states, and consequences

Act	State	
	Good	Rotten
Break into bowl	six-egg omelet	no omelet, and five good eggs destroyed
Break into saucer	six-egg omelet, and a saucer to wash	five-egg omelet, and a saucer to wash
Throw away	five-egg omelet, and one good egg destroyed	five-egg omelet

- To each outcome, we could assign a utility and maximize it
- What do we know about the state of nature?
  - We may act *as though* there is one true state and we just don't know it
  - If we assume a probability over s, this is decision under risk

- What criteria do we have for a decision problem under ***uncertainty*** (d.p.u.u.)?

Some Criteria for d.p.u.u.

Maximin criterion: To each act, assign its security level as an index. Index of A_i is the minimum of the utilities $u_{i1}, \ldots, u_{in}$

Choose the act whose associated index is maximum.

	s1	s2
A1	0	1
A2	1	0

- What is the security level for each act?
- What happens if we allow for mixed strategies (i.e., akin to a compound lottery, e.g., p = 0.5 for a1 and p = 0.5 for a2) ?
- Interpretation as game against nature: best response against nature's minimax strategy (least favourable a priori strategy)

Point to Ponder about Maximin

- Is nature a *conscious* adversary?!
- Consider:

	s1	s2
A1	0	100
A2	1	1

- What are the safety values for the actions?
  - If mixed strategies are allowed?

- What if 100 went up to 10^6 and 1 came down to 0.0001?

Some Criteria for d.p.u.u.

- **Minimax risk criterion** (Savage): Consider a setup as follows:

	s1	s2
A1	0	100
A2	1	1

If s1 is the true state, choosing A2 poses no 'risk' whereas if s2 is the true state then considerable 'risk' in A2.

Savage's procedure: (i) Create new risk payoffs which are amounts to be added to utility to match maximum column utility, (ii) Choose act which minimizes maximum risk index

Minimax Risk Criterion

- Transform Utility Payoff to Risk Payoff:

	s1	s2
A1	0	100
A2	1	1

	s1	s2
A1	1	0
A2	0	99

- Take the u_{ij} and define r_{ij} so that it is the amount that has to be added to u_{ij} to equal maximum utility payoff in column j.
- Critique (due to Chernoff):
  - "Regret" may not be measured by utility difference
  - Different states of nature may not be traded off properly
  - **Taking away an irrelevant (obviously bad) action may change optimal decision!**

More Criteria for d.p.u.u.

- **Pessimism-optimism index criterion** of Hurwicz:

  Let m_i and M_i be minimum and maximum utility. Assume a fixed pessimism-optimism index, α. To each act, associate an α-index $\alpha m_i + (1 - \alpha) M_i$.

  Of two acts, the one with higher α-index is preferred.

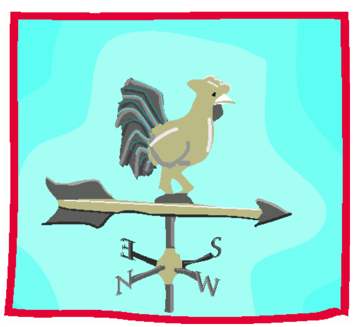
- "**Principle of insufficient reason**": If one is completely ignorant, one should act as though all states are equally likely; so choice should be based on a utility index which is the average of utility for all possible states for any act

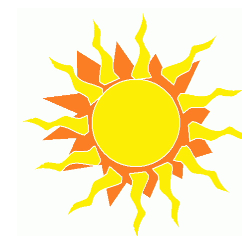
What is the effect of the way we enumerate possible states of nature?

Use of Bayesian Principles for Decisions: Simple Example

Bob observes the weather forecast before deciding whether to carry an umbrella to work. Bob wishes to stay dry, but carrying an umbrella around is annoying.

Forecast





Setup of Decision Theory

- Set ***A*** of actions
  - **Umbrella**=*{true, false}*
- Set ***E*** of (unobserved) events
  - **Weather**=*{rain, sun}*
- Set ***O*** of observations
  - **Forecast**=*{rain, sun}*
- Probability distribution over
  - events ***P(E)***
  - observations given events ***P(O | E)***
- Utility function from actions and events to real numbers.

	Weather
sun	0.7
rain	0.3

	Forecast	
Weather	*sun*	*rain*
sun	0.6	0.4
rain	0.4	0.6

Weather	Umbrella	Utility
sun	*TRUE*	-10
sun	*FALSE*	100
rain	*TRUE*	100
rain	*FALSE*	-10

Choosing the Best Action

Let $U^a(\text{Bob} \mid e)$ be Bob's reward for taking action $a \in \mathbf{A}$ after event $e \in \mathbf{E}$ has occurred.

The expected utility for Bob after observing $o \in \mathbf{O}$ is

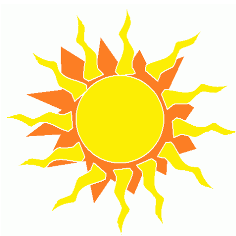
$$EU^a(\text{Bob} \mid o) = \sum_{e \in \mathbf{E}} P(e \mid o) \cdot U^a(\text{Bob} \mid e)$$

Optimal behavior — Given observation o choose the action that leads to maximal expected utility.

$$a^* = \text{argmax}_{a \in \mathbf{A}} EU^a(\text{Bob} \mid o)$$

Computing an Optimal Strategy for Bob

- A strategy for Bob must specify whether to take an umbrella for any possible value of the forecast.
- Suppose forecast predicts sun. What is Bob's expected utility for taking an umbrella ?



Computing Expected Utility for Bob for taking Umbrella

$$EU^{\mathsf{UM}}(\mathrm{Bob} \mid \mathsf{F} = sun) = P(\mathsf{W} = sun \mid \mathsf{F} = sun) \cdot U^{\mathsf{UM}}(\mathrm{Bob} \mid \mathsf{W} = sun) + P(\mathsf{W} = rain \mid \mathsf{F} = sun) \cdot U^{\mathsf{UM}}(\mathrm{Bob} \mid \mathsf{W} = rain)$$

Weather	Umbrella	Utility
sun	*TRUE*	-10
sun	*FALSE*	100
rain	*TRUE*	100
rain	*FALSE*	-10

Marginal probability

$$
\begin{aligned}
P(\mathsf{F} = sun) =& P(\mathsf{F} = sun \mid \mathsf{W} = sun) \cdot P(\mathsf{W} = sun) + \\
& P(\mathsf{F} = sun \mid \mathsf{W} = rain) \cdot P(\mathsf{W} = rain) \\
=& 0.6 \cdot 0.7 + 0.4 \cdot 0.3 = 0.54
\end{aligned}
$$

Bayes Rule

$$
\begin{aligned}
P(\mathsf{W} = sun \mid \mathsf{F} = sun) =& \frac{P(\mathsf{F} = sun \mid \mathsf{W} = sun) \cdot P(\mathsf{W} = sun)}{P(\mathsf{F} = sun)} \\
=& \frac{0.6 \cdot 0.7}{0.54} = 0.77
\end{aligned}
$$

Computing Expected Cost

$$
\begin{aligned}
EU^{\mathsf{UM}}(\mathrm{Bob} \mid \mathsf{F} = sun) = & P(\mathsf{W} = sun \mid \mathsf{F} = sun) \cdot U^{\mathsf{UM}}(\mathrm{Bob} \mid \mathsf{W} = sun) + \\
& P(\mathsf{W} = rain \mid \mathsf{F} = sun) \cdot U^{\mathsf{UM}}(\mathrm{Bob} \mid \mathsf{W} = rain) \\
= & 0.77 \cdot (-10) + 0.23 \cdot 100 = 15.3
\end{aligned}
$$

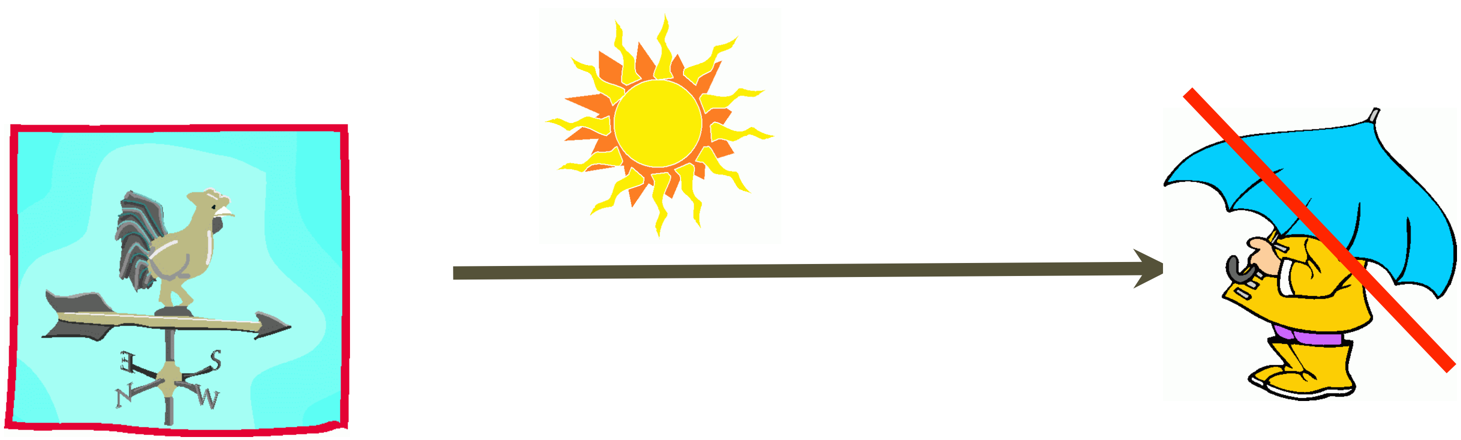
We now compute the expected utility for Bob for the case where Bob does not take an umbrella.

$$
\begin{aligned}
EU^{\overline{\mathsf{UM}}}(\mathrm{Bob} \mid \mathsf{F} = sun) = & P(\mathsf{W} = sun \mid \mathsf{F} = sun) \cdot U^{\overline{\mathsf{UM}}}(\mathrm{Bob} \mid \mathsf{W} = sun) + \\
& P(\mathsf{W} = rain \mid \mathsf{F} = sun) \cdot U^{\overline{\mathsf{UM}}}(\mathrm{Bob} \mid \mathsf{W} = rain) \\
= & 0.77 \cdot 100 + 0.23 \cdot (-10) = 74.7
\end{aligned}
$$

Computing Bob's Best Action

$$\overset{(15.3)}{EU^{\mathsf{UM}}(\text{Bob} \mid \mathsf{F} = sun)} < \overset{(74.7)}{EU^{\overline{\mathsf{UM}}}(\text{Bob} \mid \mathsf{F} = sun)}$$

If the forecast predicts sun, then Bob should not take the umbrella



Computing Bob's Best Action

We now compute Bob's decision for the case where the forecast predicts rain. We have that

$$\underset{(34)}{EU^{\mathsf{UM}}(\text{Bob} \mid \mathsf{F} = rain)} < \underset{(56)}{EU^{\overline{\mathsf{UM}}}(\text{Bob} \mid \mathsf{F} = rain)}$$

We get the following strategy for Bob

	Forecast	
	rain	*sun*
Umbrella	*FALSE*	FALSE

Making Sequential Decisions

The newspaper forecast is more reliable, but costs money, decreasing Bob's utility by 10 units. There are now two decisions:

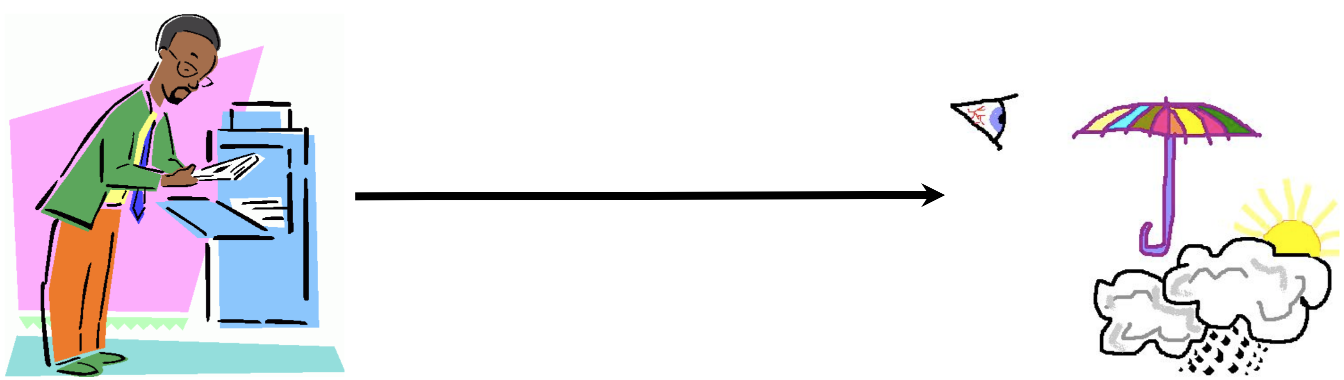
- Buying a newspaper
- Carrying an umbrella

	Forecast	
Weather	*sun*	*rain*
sun	0.8	0.2
rain	0.2	0.8

Weather	NP	Umbrella	Utility
sun	TRUE	*TRUE*	-20
sun	TRUE	*FALSE*	90
rain	TRUE	*TRUE*	90
rain	TRUE	*FALSE*	-20
…	….	….	….

Making Sequential Decisions

- Choosing the best action for one decision depends on the action for the other decision.
- How to weigh the tradeoff between these two decisions ?



Marginal probability

$$
\begin{aligned}
P^{\mathsf{NP}}(\mathsf{F}=sun) =& P^{\mathsf{NP}}(\mathsf{F}=sun \mid \mathsf{W}=sun)\cdot P(\mathsf{W}=sun) + \\
& P^{\mathsf{NP}}(\mathsf{F}=sun \mid \mathsf{W}=rain)\cdot P(\mathsf{W}=rain) \\
=& 0.8\cdot 0.7 + 0.2\cdot 0.3 = 0.62
\end{aligned}
$$

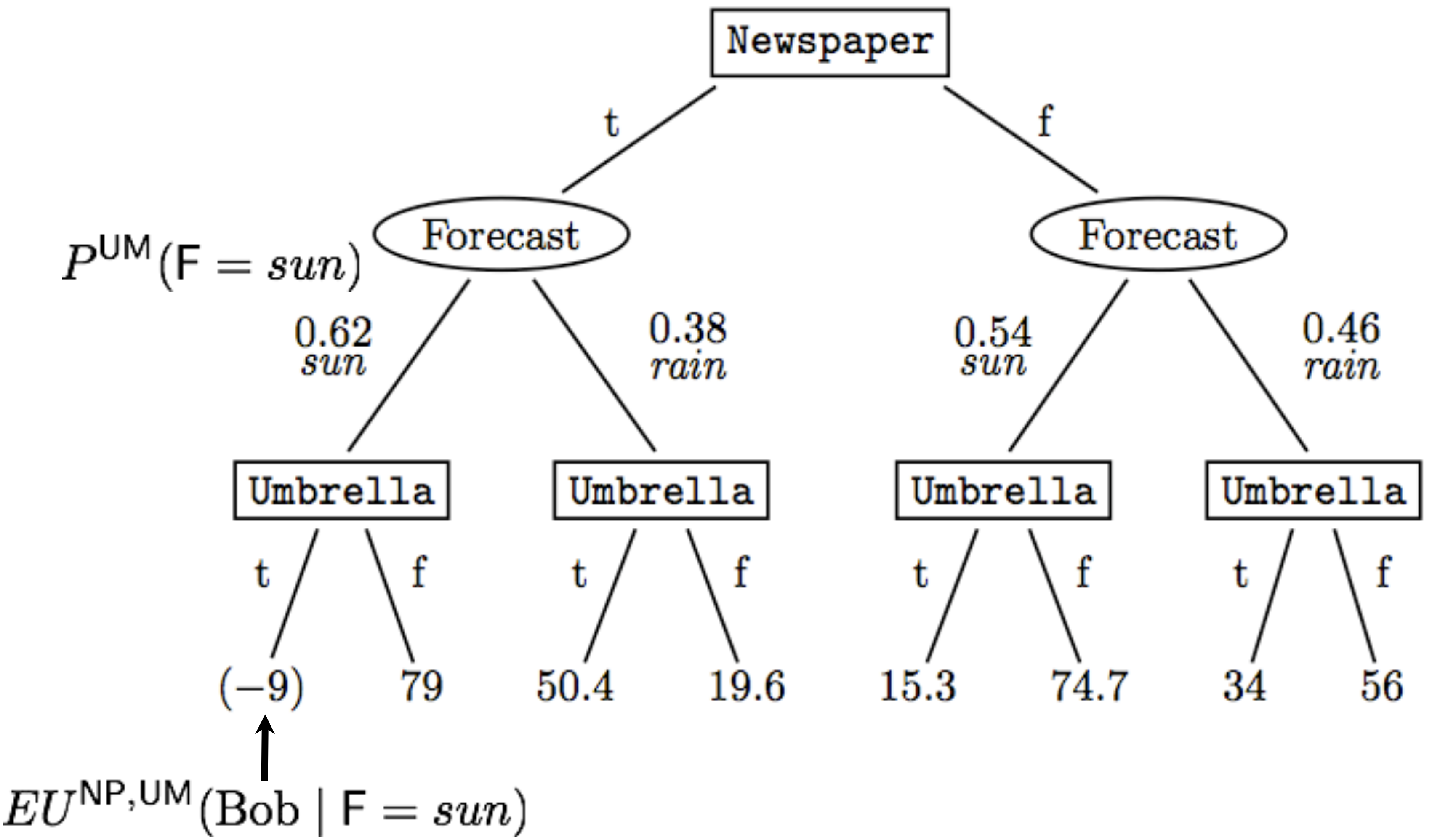
Bayes Rule

$$
\begin{aligned}
P^{\mathsf{NP}}(\mathsf{W}=sun \mid \mathsf{F}=sun) =& \frac{P^{\mathsf{NP}}(\mathsf{F}=sun \mid \mathsf{W}=sun)\cdot P(\mathsf{W}=sun)}{P^{\mathsf{NP}}(\mathsf{F}=sun)} \\
=& \frac{0.8\cdot 0.7}{0.62} = 0.90
\end{aligned}
$$

Expected utility

$$
\begin{aligned}
EU^{\mathsf{NP,UM}}(\mathrm{Bob} \mid \mathsf{F}=sun) =& P^{\mathsf{NP}}(\mathsf{W}=sun \mid \mathsf{F}=sun)\cdot U^{\mathsf{UM}}(\mathrm{Bob} \mid \mathsf{W}=sun) + \\
& P^{\mathsf{NP}}(\mathsf{W}=rain \mid \mathsf{F}=sun)\cdot U^{\mathsf{UM}}(\mathrm{Bob} \mid \mathsf{W}=rain) \\
=& 0.90\cdot(-20) + 0.10\cdot 90 = (-9)
\end{aligned}
$$

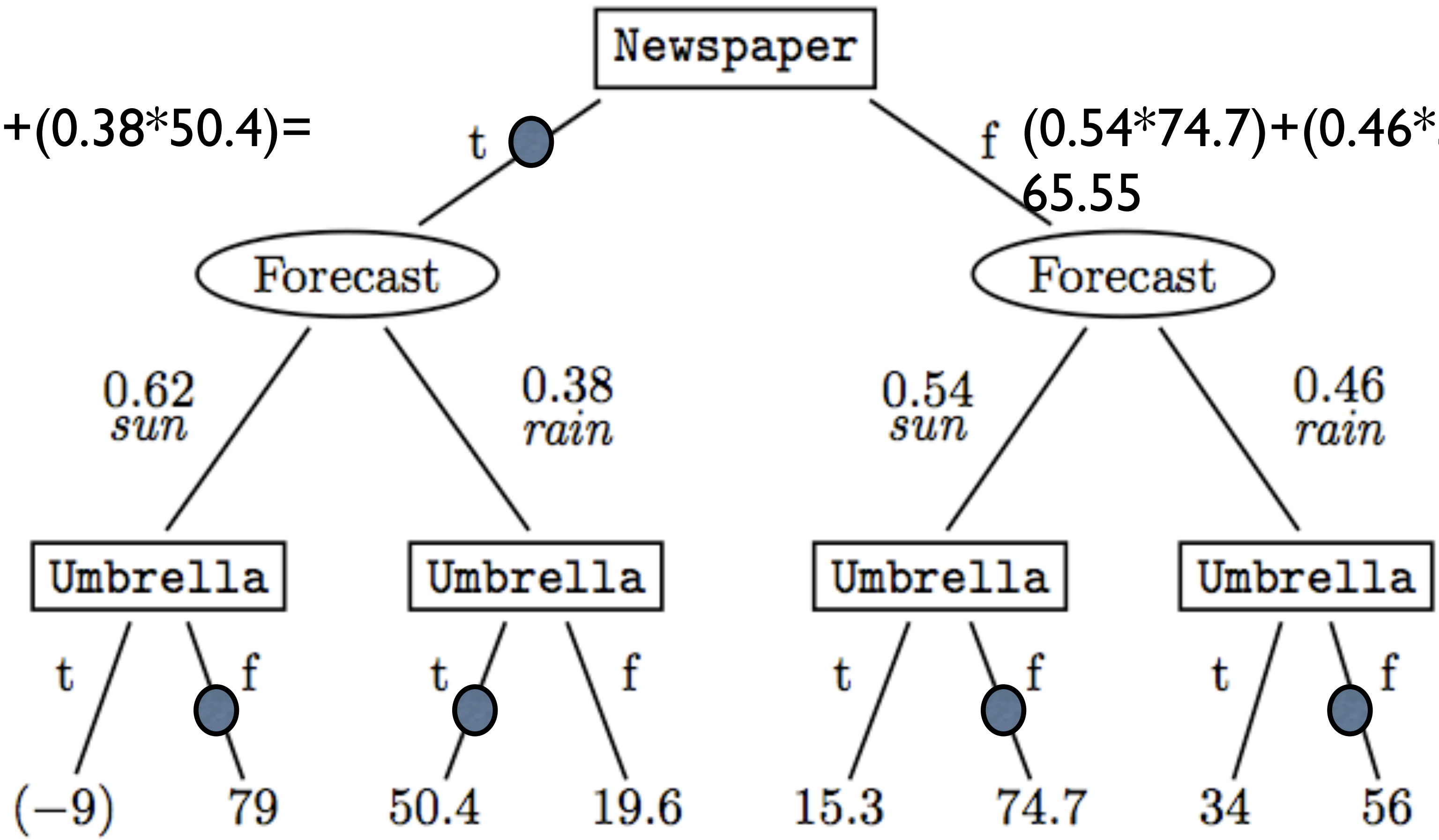
Decision Trees



Solving Decision Trees

(0.62*79)+(0.38*50.4)=
68.132

(0.54*74.7)+(0.46*56)=
65.55



Acknowledgements

- Much of the decision theory discussion is a paraphrased and condensed version of chapters in Luce and Raiffa's excellent book on Games and Decisions
- The example in the latter section is taken from a tutorial by Gal and Pfeffer at AAAI '08.