Rational Decisions

Chapter 16

3rd ed.

Rational Decisions

◇ Decision-theoretic agent:

combines probability theory with the concept of utility
for making rational decisions even in the face of uncertainty.

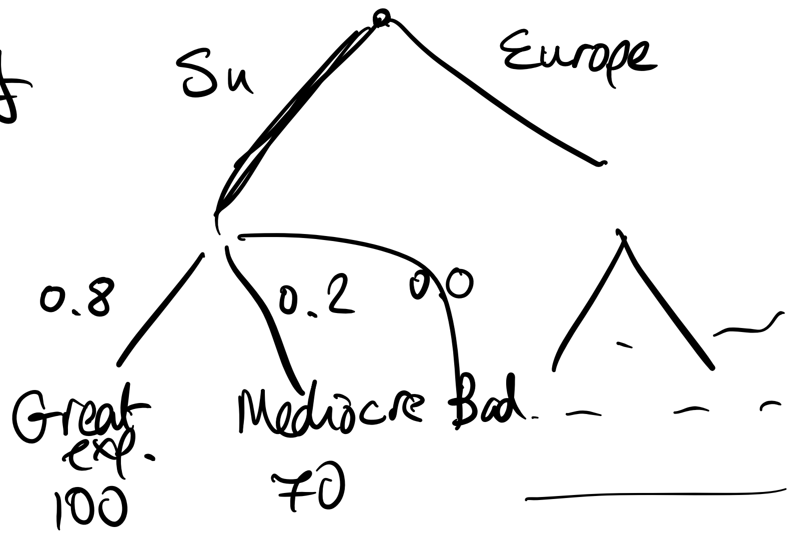
There are couple of important concepts in this chapter:

◇ a lottery versus a fixed "prize"

◇ expected utility of a lottery

state $\xrightarrow{\text{utility}}$ real # indicating the goodness of that state

lottery



: acts

: possible outcomes

Exp. Util. of SU: 0.8x 100 + 0.2x70 =

Utility and Expected Utility

◇ ◇ Utility function assigns a single number to express the desirability of a state: $U(S)$

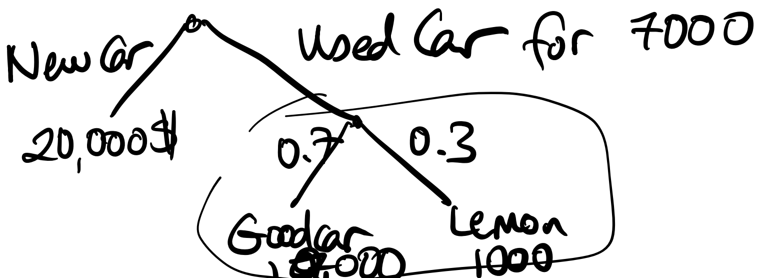
◇ Expected utility of a non-deterministic action A, with possible outcomes $Result_i(A)$ is:

Good Mediocre Bad

$$EU(A|E) = \Sigma_i \underbrace{P(Result_i(A)|E, A)} \times \underbrace{U(Result_i(A))}$$

A: Choosing SU

E.g. Action buy the car. One result may be that it was a good car, another one may be that it was a bad car that will have problems...



◇ The principle of Maximum Expected Utility says that a rational agent should choose an action that maximizes the agent's expected utility.

Decision Theory

Decision Theory provides a framework incorporating many components of AI:

- ♢ Knowing the initial state: representation, perception,...
- ♢ $U(Result_i(A))$: search (to find where you can get to from that state, as done in adversarial search)
- ♢ $P(Result_i(A)|E, A)$: inference, learning...

Basis of Utility Theory

Why maximizing the average utility?

Why not minimizing the worst possible loss?

Can an agent compare states without assigning them a value?

$\Rightarrow$ Constraints on the preferences that an agent should have answer these questions and derives MEU.

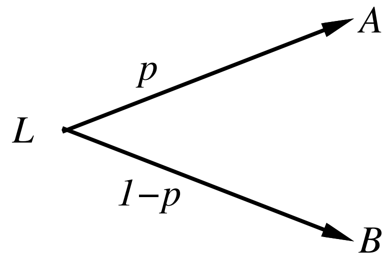
Preferences

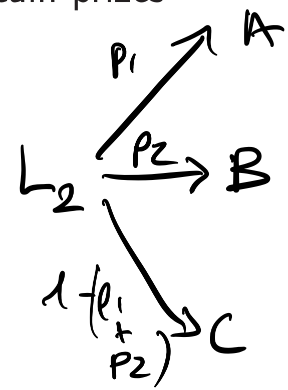
Utility theory deals with preferences. Preferences about what?

◇ States ⇒ prizes (A, B, etc.)

◇ Scenario with possible outcomes ⇒ lotteries with uncertain prizes

Lottery $L = [p, A;\ (1-p), B]$





$L_2 = [p_1, A;\ p_2, B;\ 1-(p_1+p_2), C]$

Notation:

$A \succ B$	A preferred to B
$A \sim B$	indifference between A and B
$A \succsim B$	B not preferred to A

Primary issue for utility theory: understand how preferences between lotteries are related to preferences between the underlying states in those lotteries.

Rational preferences

Idea: preferences of a rational agent must obey constraints.
Rational preferences $\Rightarrow$
behavior describable as maximization of expected utility

Constraints:

Orderability
$$(A \succ B) \vee (B \succ A) \vee (A \sim B)$$
Transitivity
$$(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$$
Continuity
$$A \succ B \succ C \Rightarrow \exists\, p \;\; [p, A;\; 1-p, C] \sim B$$
Substitutability
$$A \sim B \Rightarrow [p, A;\, 1-p, C] \sim [p, B; 1-p, C]$$
Monotonicity
$$A \succ B \Rightarrow (p \geq q \;\Leftrightarrow\; [p, A;\; 1-p, B] \succsim [q, A;\; 1-q, B])$$

A B C
$100 \succ 75 \succ 50$
$B \sim [p, A; 1-p, C]$
$0.5, A; 0.5, C]$

Rational preferences contd.

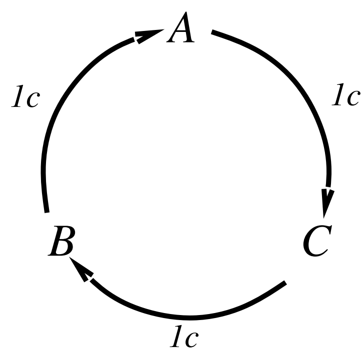
Violating the constraints leads to self-evident irrationality

For example: an agent with intransitive preferences can be induced to give away all its money

If $B \succ C$, then an agent who has C would pay (say) 1 cent to get B

If $A \succ B$, then an agent who has B would pay (say) 1 cent to get A

If $C \succ A$, then an agent who has A would pay (say) 1 cent to get C



Utility Theory

Axioms of utility theory talk about preferences. The existence of a utility function follows from the axioms of utility:

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944): If an agent's preferences obeys the axioms of utility, then there exists a real-valued function U such that:

$$U(A) \geq U(B) \quad \Leftrightarrow \quad A \succ B$$
$$U(A) = U(B) \quad \Leftrightarrow \quad A = B$$

Expected utility of a lottery: The utility of a lottery is the sum of the utilities of each outcome weighted by the corresponding probabilities:

$$U([p_1, S_1; \; \ldots \; ; \; p_n, S_n]) = \Sigma_i \; p_i U(S_i)$$

$\Rightarrow$ justifies the MEU decision rule

Maximizing expected utility

An agent can be entirely rational (consistent with MEU)
without ever representing or manipulating utilities and probabilitie

E.g., a lookup table for perfect tictactoe

Utilities

Utilities map states to real numbers.

◇ Which numbers?

An agent/person can have any preferences it wants (e.g. preferring dented cars over shiny cars etc.)

◇ But we can design utility functions such that when installed in an agent, they will generate the desired behaviors.

Utility of Money

Utility theory has its roots in economics.

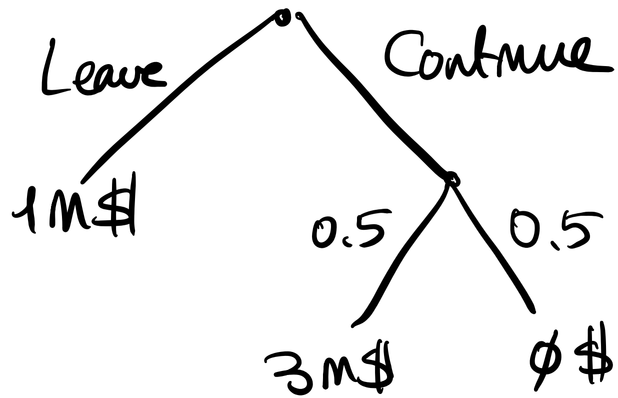
Can money be used as a utility measure?

Game: You gained 1,000,000 $ at a game show. To continue the game, you must flip a coin

Heads: Get 3,000,000$

Tails: Loose everything

What will you do?



$EU[\text{Continue}] = 1.5M\$$

Utility of Money

You can represent the game as:

$L : [0.5, Heads(triple); 0.5, Tails(loose)]$

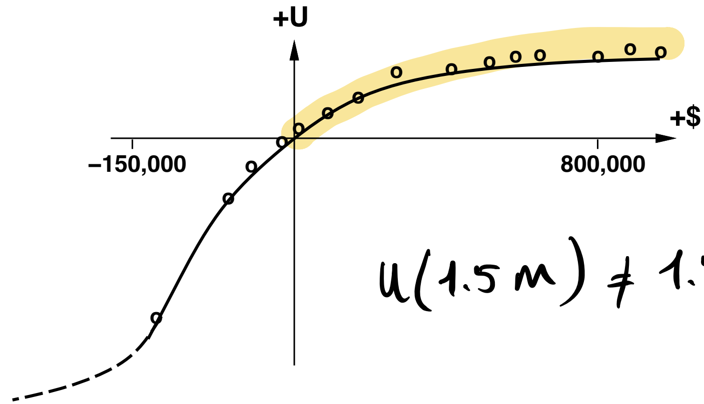
$EU(L) = 0.5 \times 3,000,000 + 0.5 \times 0 = 1,500,000$

Utility of Money

Yet, most people prefer to leave with the money. Denote S_k the state of possessing total wealth of k \$.

◇ $U(S_{k+\$})$ is not linear (almost exactly logarithmic for positive values)

◇ depends on k

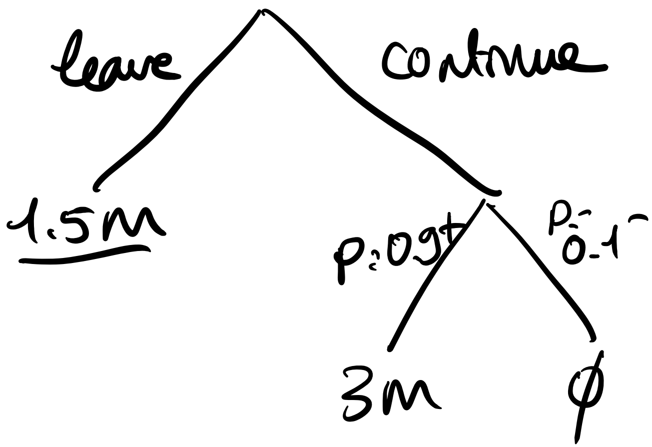


$U(1.5M) \neq 1.5 \times U(1M)$

Typical empirical data, extrapolated with risk-prone behavior.

Assessing Utilities

Preference elicitation: Money is not always appropriate to use as utility. We need to work out the utility function for an agent, so that it can make decisions.



Assessing Utilities

One procedure for assessing utilities is to establish a scale with a "best possible prize" with U(S) $= u_\top$ and a "worst possible catastrophe" with U(S) $= u_\perp$.

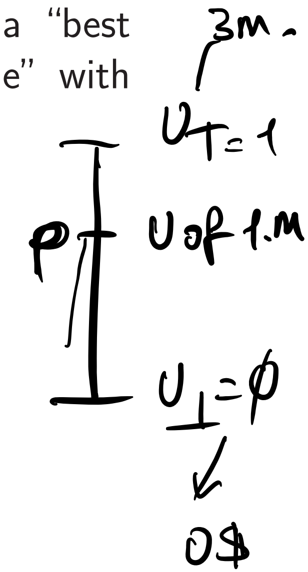
Normalized utilities: $u_\top = 1$ and $u_\perp = 0$.

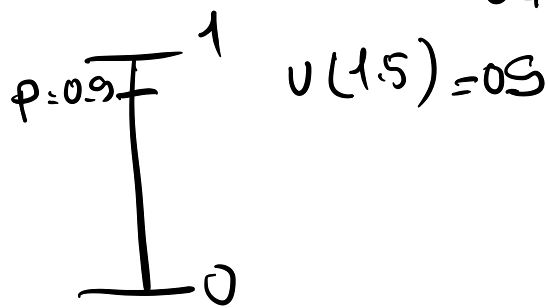
Utilities of an intermediate outcomes state S is assessed by asking the agent to indicate a preference between:

S (*fixed prize*) and a <u>standard lottery</u> $L_p = [p, u_\top; (1-p), u_\perp]$,

adjusting p until the agent is indifferent.

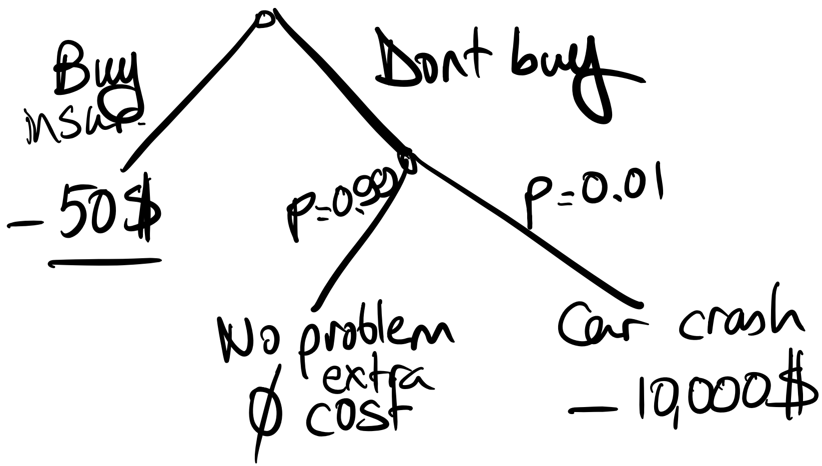
Then, $U(S) = p$ assuming normalized utilities.





Assessing Utilities

Insurance premium: the difference between the expected monetary value of a lottery and its certainty equivalent (what you would prefer to the lottery).



Insurance Premium ~ Lottery of not buying insu.

Multiattribute utility - Mostly skipped

How can we handle utility functions of many variables $X_1 \ldots X_n$?
E.g., what is $U(Deaths, Noise, Cost)$?

How can complex utility functions be assessed from preference behaviour?

Idea 1: identify conditions under which decisions can be made without complete identification of $U(x_1, \ldots, x_n)$ (without combining attribute values into a single utility value)

Idea 2: identify various types of independence in preferences and derive consequent canonical forms for $U(x_1, \ldots, x_n)$

x x

AL_1 AL_2

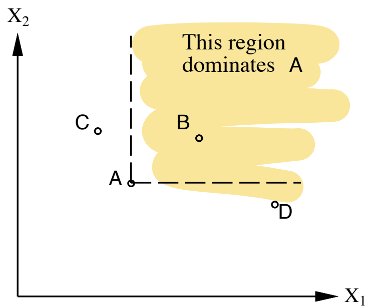
Strict dominance

Typically define attributes such that U is monotonic in each

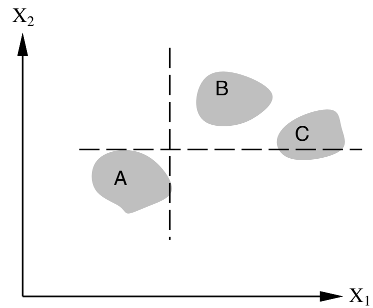
Strict dominance: choice B strictly dominates choice A iff

$$\forall i \;\; X_i(B) \geq X_i(A) \qquad \text{(and hence } U(B) \geq U(A))$$

e.g. Site 1 costs less, causes less pollution and is safer, than Site 2.



Deterministic attributes



Uncertain attributes

Strict dominance seldom holds in practice. Other approaches (e.g. stochastic dominance) is skipped.

Value of information

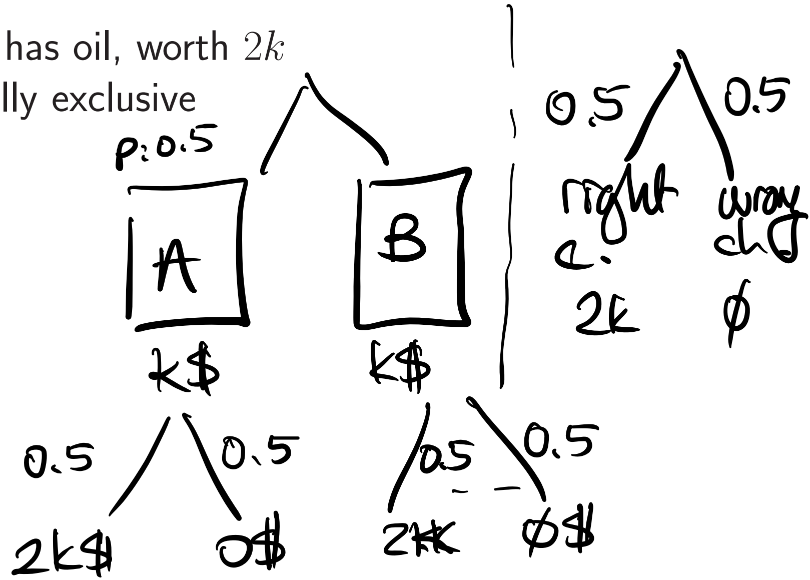
Example: buying oil drilling rights

Two blocks A and B, exactly one has oil, worth $2k$

Prior probabilities 0.5 each, mutually exclusive

Current price of each block is k

What should you do?



Simple

p=0.5 that I'll pick the right one + p=0.5 × I'll pick the wrong one

Expected gain = 0.5 × 2k$ + 0.5, 0$ = k

For expected value of k$, I don't want to spend k$. :(

Value of information

Idea: compute value of acquiring each possible piece of evidence

Example: buying oil drilling rights
Two blocks A and B, exactly one has oil, worth $2k$
Prior probabilities 0.5 each, mutually exclusive
Current price of each block is k

Consultant offers accurate survey of A (Q=oil in A?). What is the fair price for this expertise?

Solution: compute expected value of information
= expected value of best action given the information **minus** expected value of best action without information

Note: In each case, I take best action (choose land that he says) & not an expectation!

Value of information

Survey may say “oil in A” or “no oil in A”, prob. 0.5 each

$= 0.5 \times$ value of the best action (“buy A”) given that expert says oil is in A

$+ 0.5 \times$ value of he best action (“buy B”) given that expert says oil is in B

– 0

profit profit

$= (0.5 \times k) + (0.5 \times k) - 0 = k$

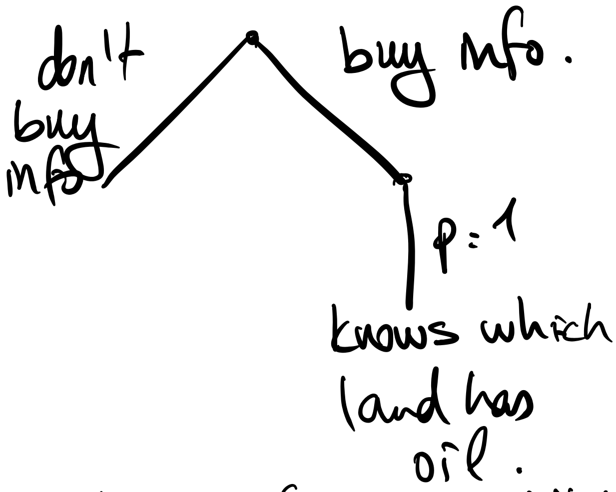
prob. he may say A ~ prob. he may say B

◇ We can pay up to k dollars for this information.

If we calculate expected value as in the prev. slide:

$0.5 \cdot 2k + 0.5 \times 2k = 2k$

he will say A he will say B



How much to pay that expert.

up to 2k expected value − k cost

up to = k $.

0.5

if he says {A — I'll buy A = 2k total value.
B — " B = 2k

General formula

As in the oil example, we need to calculate EU over each expert outcome:

Current evidence E, current best action α
Possible action outcomes S_i, potential new evidence E_j

$$EU(\alpha|E) = \max_a \Sigma_i\, U(S_i) \times P(S_i|E,a)$$

Suppose we knew $E_j = e_{jk}$, then we would choose $\alpha_{e_{jk}}$ s.t.

$$EU(\alpha_{e_{jk}}|E, E_j = e_{jk}) = \max_a \Sigma_i\, U(S_i) \times P(S_i|E,a,E_j = e_{jk})$$

$E_{j1}=A$ $E_{j2}=B$

But E_j is a random variable whose value is *currently* unknown
$\Rightarrow$ must compute expected gain over all possible values:

$$VPI_E(E_j) = \left(\Sigma_k\, P(E_j = e_{jk}|E) \times EU(\alpha_{e_{jk}}|E, E_j = e_{jk})\right) - EU(\alpha|E)$$

(VPI = value of perfect information)

Properties of VPI

Nonnegative—in **expectation**, not **post hoc**

$$\forall\, j, E \;\; VPI_E(E_j) \geq 0$$

Nonadditive—consider, e.g., obtaining E_j twice

$$VPI_E(E_j, E_k) \neq VPI_E(E_j) + VPI_E(E_k)$$

Order-independent

$$VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_{E,E_j}(E_k) = VPI_E(E_k) + VPI_{E,E_k}(E_j)$$

Note: when more than one piece of evidence can be gathered,
maximizing VPI for each to select one is not always optimal
$\Rightarrow$ evidence-gathering becomes a **sequential** decision problem

Qualitative behaviors

Sometimes the value of information is large, sometimes small:

a) Choice is obvious, information worth little
b) Choice is nonobvious, information worth a lot
c) Choice is nonobvious, information worth little