Rational decisions

Chapter 13.1, 16.1-3

Adapted from slides kindly shared by Stuart Russell

Appreciations

- ♢ Appreciations :)
- ♢ Modern medicine
- ♢ Improvisation

Share some of yours?

Outline

- ♢ Uncertainty
- ♢ Rational preferences
- ♢ Maximizing expected utility
- ♢ Utilities
- ♢ Money

Uncertainty

Let action A_t = leave for airport t minutes before flight
Will A_t get me there on time?

Problems:

1) partial observability (road state, other drivers' plans, etc.)
2) noisy sensors (Google traffic map overlays)
3) uncertainty in action outcomes (flat tire, etc.)
4) immense complexity of modelling and predicting traffic

Logic, Reasoning, Planning (Ch 7-12): develop contingency plan

A purely logical approach either
1) risks falsehood: "A_{25} will get me there on time"
or 2) leads to conclusions that are too weak for decision making:
"A_{25} will get me there on time if there's no accident on the bridge and it doesn't rain and my tires remain intact etc etc."

(A_{1440} might be said to get me there on time, but ...)

Methods for handling uncertainty

Probability

Given the available evidence,

A_{25} will get me there on time with probability 0.04

Mahaviracarya (9th C.) - bets that can't lose

Cardano (1565) - theory of gambling

Money, a timeless motivator....

"Against the Gods - the Remarkable Story of Risk" by Bernstein.

Probability

Probabilistic assertions **summarize** effects of

laziness: failure to enumerate exceptions, qualifications, etc.

ignorance: lack of relevant facts, initial conditions, etc.

Subjective or Bayesian probability:

Probabilities relate propositions to one's own state of knowledge

e.g., $P(A_{25} | \text{no reported accidents}) = 0.06$

These are **not** claims of a "probabilistic tendency" in the current situation (but might be learned from past experience of similar situations)

Probabilities of propositions change with new evidence:

e.g., $P(A_{25} | \text{no reported accidents, 5 a.m.}) = 0.15$

Making decisions under uncertainty

Suppose I believe the following:

$$
\begin{aligned}
P(A_{25} \text{ gets me there on time} | \ldots) &= 0.04 \\
P(A_{90} \text{ gets me there on time} | \ldots) &= 0.70 \\
P(A_{120} \text{ gets me there on time} | \ldots) &= 0.95 \\
P(A_{1440} \text{ gets me there on time} | \ldots) &= 0.9999
\end{aligned}
$$

Which action to choose?

Depends on my preferences for missing flight vs. airport cuisine, etc.

Utility theory is used to represent and infer preferences

Decision theory $=$ utility theory $+$ probability theory

Maximizing Expected Utilities

Maximizing expected utilities is at the heart of this course

A rational agent should choose the action which maximizes its expected utility, given its knowledge

Breaking that down:

◊ Expectation given knowledge: **probabilitic inference** - middle of course

◊ To do that, need **experience, learning** - later in course

◊ **Maximization** - figure out which action to choose given what might happen as a result - current focus

Reducing preferences to tractable utilities

Can your preferences be summarized with utilities?

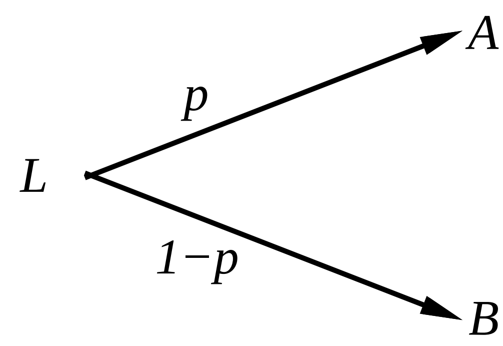
If you're "rational", yes.

Experiments suggest otherwise, humans not behaving rationally

Heads up: don't gamble or play the wrong sorts of games if you prefer not to adopt these constraints

Preferences

An agent chooses among prizes (A, B, etc.) and lotteries, i.e., situations with uncertain prizes



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

Notation:

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

Rational preferences

Idea: preferences of a rational agent must obey constraints.

The axioms of rationality:

<u>Orderability</u>

Exactly one of $(A \succ B) \lor (B \succ A) \lor (A \sim B)$ holds

<u>Transitivity</u>

$(A \succ B) \land (B \succ C) \Rightarrow (A \succ C)$

<u>Continuity</u>

$A \succ B \succ C \Rightarrow \exists p \ [p, A;\ 1-p, C] \sim B$

<u>Substitutability</u>

$A \sim B \Rightarrow [p, A;\ 1-p, C] \sim [p, B; 1-p, C]$

<u>Monotonicity</u>

$A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A;\ 1-p, B] \succsim [q, A;\ 1-q, B])$

Rational preferences $\Rightarrow$

behavior describable as maximization of expected utility

Rational preferences contd.

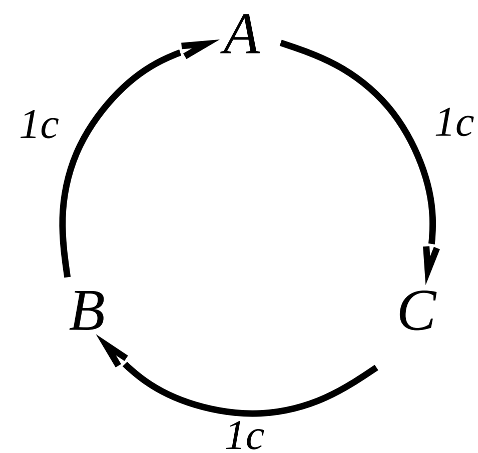
Violating the constraints leads to self-evident irrationality

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



Theorem: Rational preferences imply behavior describable as maximazation of expected utility

Maximizing expected utility

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944):
Given preferences satisfying the constraints
there exists a real-valued function U such that

Greater utility means preference

$U(A) \geq U(B) \quad \Leftrightarrow \quad A \succsim B$

Utility of a lottery is the expectation of the utilities

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

MEU principle:

Choose the action that maximizes expected utility

Note: an agent can be entirely rational (consistent with MEU)
without ever representing or manipulating utilities and probabilities

E.g., a lookup table for perfect tictactoe

Utility scales

Normalized utilities: $u_\top = 1.0$, $u_\perp = 0.0$

Micromorts: one-millionth chance of death
useful for Russian roulette, paying to reduce product risks, etc.

QALYs: quality-adjusted life years
useful for medical decisions involving substantial risk

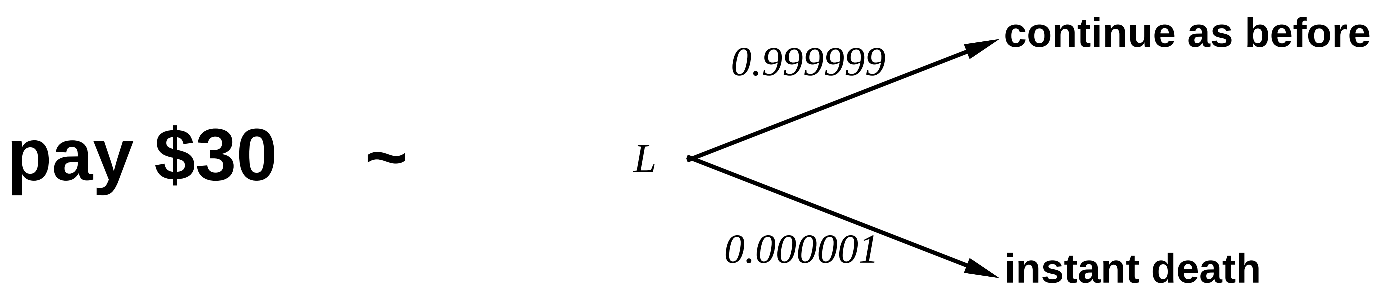
Note: behavior is **invariant** w.r.t. positive linear transformation

$$U'(x) = k_1 U(x) + k_2 \quad \text{where } k_1 > 0$$

Human Utilities

Utilities map states to real numbers. Which numbers?

Standard approach to assessment of human utilities:
compare a given state A to a standard lottery L_p that has
"best possible prize" $u_\top$ with probability p
"worst possible catastrophe" $u_\bot$ with probability $(1-p)$
adjust lottery probability p until $A \sim L_p$



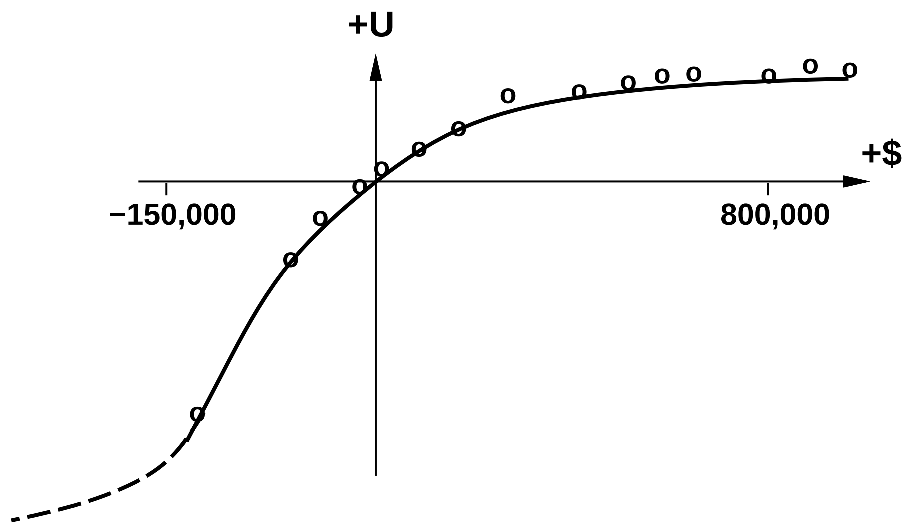
Money

Money does **not** behave as a utility function

Given a lottery L with expected monetary value $EMV(L)$,
usually $U(L) < U(EMV(L))$, i.e., people are risk-averse

Utility curve: for what probability p am I indifferent between a prize x and a lottery $[p, \$M;\ (1-p), \$0]$ for large M?

Typical empirical data, extrapolated with risk-prone behavior:



Insurance

Consider the lottery [0.5, $1000; 0.5 $0]

♢ What it the expected monetary value? ($500)

♢ What is its certainty equivalent? (monetary value acceptable in lieu of lottery)

$400 is a typical value

Difference of $100 is the insurance premium

Insurance company has much larger assets, nearly linear utility curve there, less risk-averse.

Not zero sum - everyone happy

Human Rationality?

Example of Allais (1953). Two lotteries:

◇ A: [0.8, $4000; 0.2, $0]

◇ B: [1.0, $3000; 0.0, $0]

◇ C: [0.2, $4000; 0.8, $0]

◇ D: [0.25, $3000; 0.75, $0]

Human Rationality?

Example of Allais (1953). Two lotteries:

◇ A: [0.8, $4000; 0.2, $0]

◇ B: [1.0, $3000; 0.0, $0]

◇ C: [0.2, $4000; 0.8, $0]

◇ D: [0.25, $3000; 0.75, $0]

Most people prefer $B \succ A, C \succ D$

But if U($0) = 0, then

◇ $B \succ A \Rightarrow U(\$3000) > 0.8U(\$4000)$

◇ $C \succ D \Rightarrow 0.8U(\$4000) > U(\$3000)$

Human Rationality?

Two views:

♢ Humans broken

♢ Wrong / incomplete model

Need to consider looking at other factors

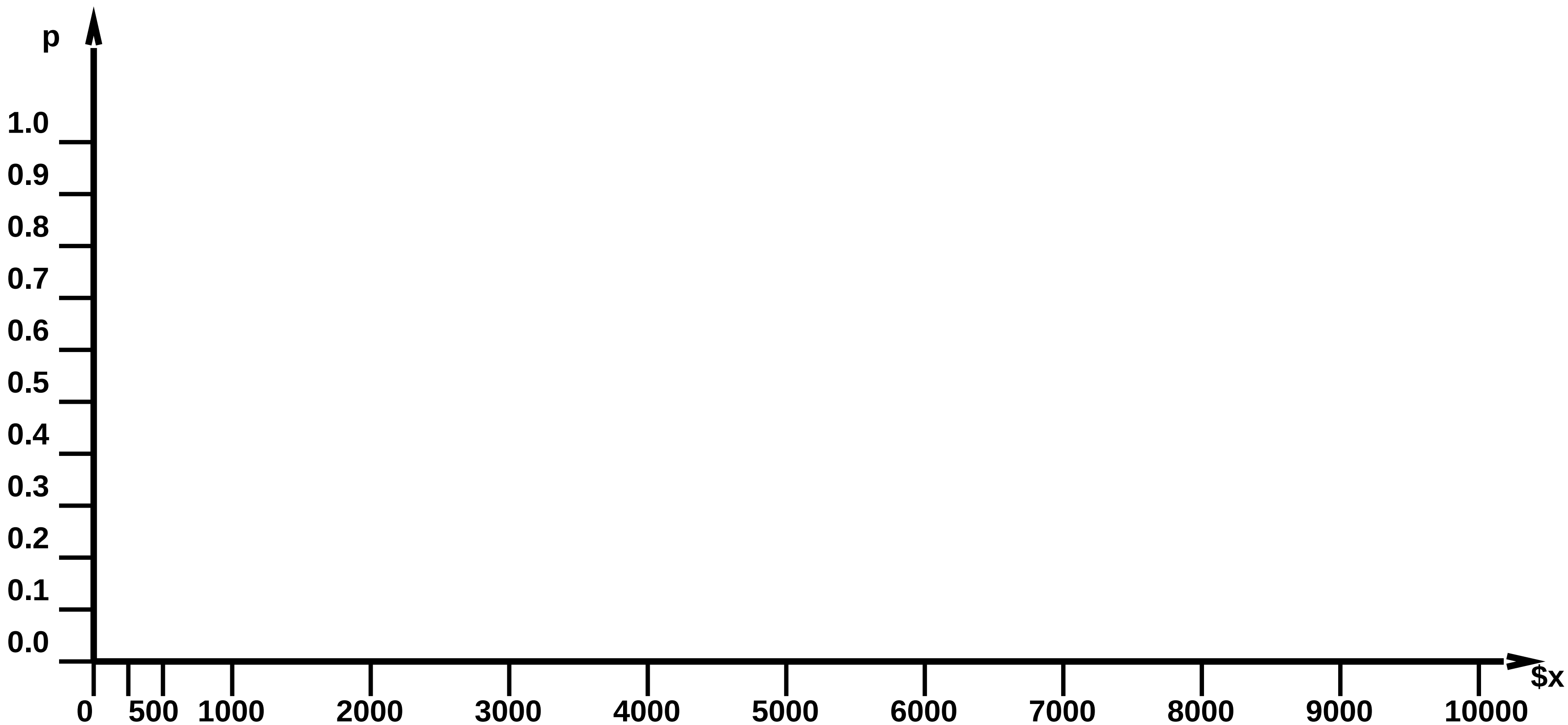
Get abstraction right - this will come up over and over

St. Petersburg Paradox

See "Expectimax, MDP, Utility" web page

Student group utility

For each x, adjust p until half the class votes for lottery (M=10,000)



Summary

Probability is a rigorous formalism for uncertain knowledge

Rational preferences imply behavior describable as maximazation of expected utility

Need to get the abstraction right

Humans can be hard to model, or irrational