Decision Theory

CPS 170
Ronald Parr

Decision Theory

What does it mean to make an optimal decision?

- Asked by economists to study consumer behavior
- Asked by MBAs to maximize profit
- Asked by leaders to allocate resources
- Asked in OR to maximize efficiency of operations
- Asked in AI to model intelligence
- Asked (sort of) by any intelligent person every day

Utility Functions

- A *utility function* is a mapping from world states to real numbers
- Sometimes called a *value function*
- Rational or optimal behavior is typically viewed as maximizing expected utility:

$$\max_a \sum_s P(s \mid a) U(s)$$

a = actions, s = states

Are Utility Functions Natural?

- Some have argued that people don't really have utility functions
  - What is the utility of the current state?
  - What was your utility at 8:00pm last night?
  - *Utility elicitation* is difficult problem
- It's easy to communicate *preferences*
- Theorem (sorta): Given a plausible set of assumptions about your preferences, there must exist a consistent utility function

Axioms of Utility Theory

- Orderability: $(A \succ B) \vee (A \prec B) \vee (A \sim B)$
- 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: $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; 1-p, B] \geq [q, A; 1-q, B])$
- Decomposability:

$$[p, A; (1-p), [q, B; (1-q), C]] \sim [p, A; (1-p)q, B; (1-p)(1-q), C]$$

Consequences of Preference Axioms

- Utility Principle
  - There exists a real-valued function U:

$$U(A) > U(B) \Leftrightarrow A \succ B$$
$$U(A) = U(B) \Leftrightarrow A \sim B$$

- Expected Utility Principle
  - The utility of a lottery can be calculated as:

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

More Consequences

- Scale invariance
- Shift invariance

Maximizing Utility

- Suppose you want to be famous
- You can be either (N,M,C)
  - Nobody
  - Modestly Famous
  - Celebrity
- Your utility function:
  - U(N) = 20
  - U(M) = 50
  - U(C) = 100
- You have to decide between going to grad school and becoming a professor (G) or going to Hollywood and becoming an actor (A)

Outcome Probabilities

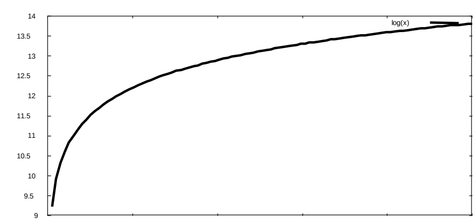
- P(N|G)=0.5, P(M|G)=0.4, P(C|G)=0.1
- P(N|H)=0.6, P(M|H)=0.2, P(C|H)=0.2
- Maximize expected utility:
  - U(N) = 20, U(M) = 50, U(C) = 100

$$EU_G = 0.5(20) + 0.4(50) + 0.1(100) = 40$$
$$EU_H = 0.6(20) + 0.2(50) + 0.2(100) = 42$$

Hollywood wins!

Utility of Money

- How much happier are you with an extra $1M?
- How much happier is Bill Gates with an extra $1M?
- Some have proposed:

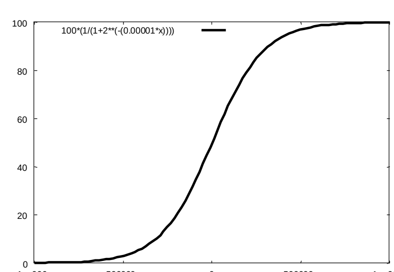


Utility of Money

- U(money) should drop slowly in negative region too
- If you're solvent, losing $1M is pretty bad
- If already $10M in debt, losing another $1M isn't that bad
- Utility of money is probably sigmoidal

A Sigmoidal Utility Function

$$U(\$X) = 100\frac{1}{1+2^{-0.00001X}}$$



Utility & Gambling

- Suppose U($X)=X, would you spend $1 for a 1 in a million chance of winning $1M?
- Suppose you start with c dollars:
  - EU(gamble)=1/1000000(1000000-1+c)+(1-1/1000000)(c-1)=c
  - EU(do_nothing)=c
- Starting amount doesn't matter
- You have no expected benefit from gambling

Sigmoidal Utility & Gambling

- Suppose: $U(\$X) = 100\frac{1}{1+2^{-0.00001X}}$
- Suppose you start with $1M
  - EU(gamble)-EU(do_nothing)=-5.7*10^{-7}
  - Winning is worthless
- Suppose you start with -$1M
  - EU(gamble)-EU(do_nothing)=+4.9*10^{-5}
  - Gambling is rational because losing doesn't hurt

Additive Independence

- Suppose it makes me happy to have my car clean
- Suppose it makes me happy to have coffee
- U=U(coffee)+U(clean)
- It seems that these don't interact
- However, suppose there's a tea variable
- U=U(coffee)+U(tea)+U(clean)???
- Probably not. I'd need U(coffee,tea)+U(clean)
- Often implicit!

Value of Information

- Expected utility of action a with evidence E:

$$\mathrm{EU}_{\mathrm{E}}(A \mid E) = \max_a \sum_i P(S_i \mid E, a) U(S_i)$$

- Expected utility given new evidence E'

$$\mathrm{EU}_{E,E'}(A \mid E, E') = \max_a \sum_i P(S_i \mid E, E', a) U(S_i)$$

- Value of knowing E' (value of perfect information)

$$\mathrm{VPI}_E(E') = \underbrace{\left(\sum_{E'} P(E' \mid E) \mathrm{EU}_{E,E'}(a \mid E, E') \right)}_{\text{Expected utility given New information (weighted)}} - \underbrace{\mathrm{EU}(a_E \mid E)}_{\text{Previous Expected utility}}$$

Properties of VPI

- VPI is non-negative!
- VPI is not additive
- VPI is order independent
- VPI is easy to compute and is often used to determine how much you should pay for one extra piece of information. Why is this myopic?

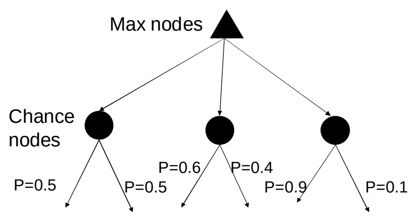
For example, knowing X AND Y together may useful, while knowing just one alone may be useless.

More Properties of VOI

- Acquiring information optimally is very difficult
- Need to construct a conditional plan for every possible outcome before you ask for even the first piece of information
  - Suppose you're a doctor planning to treat a patient
  - Picking the optimal test to do first requires that you consider all subsequent tests and all possible treatments as a result of these tests
- General versions of this problem are intractable!

Decision Theory as Search

- Can view DT probs as search probs
- States = atomic events



DT as Search

- Attach costs to arcs, leaves
- Path(s) w/lowest expected cost = optimal
- Minimizing expect cost = maximizing expected utility
- Expectimax:

$$\mathrm{V}(n_{\max}) = \max_{s \in \text{succesors}(n)} \mathrm{V}(s)$$

$$\mathrm{V}(n_{\text{chance}}) = \sum_{s \in \text{succesors}(n)} \mathrm{V}(s) p(s)$$

Optimal Solutions and Computation Cost

- Optimal solutions to decision theoretic problems are necessarily paths
- Why?
- What does this say about cost of decision theoretic reasoning?

Conclusions

- Decision theory provides a framework for optimal decision making
- Principle: Maximize Expected Utility
- Easy to describe in principle
- Application to complex problems can require advanced planning and probabilistic reasoning techniques