Ch 16 - Making Simple Decisions
Ch 17 - Making Complex Decisions

Outline

- Combining Beliefs & Desires
- Basis of Utility Theory
- Utility Function
- Decision Networks
- Value of Information

- Markov Decision Processes
- Optimal policy
- Value Iteration

Example

I'm going to buy tickets for two performances. I have two options. I can either buy both of them now at a discount or I can buy them separately closer to the performance. The probability of finding the time for a performance is 0.4. A single ticket costs $20, and a combined ticket costs $30. The "value" of going to a performance is 20.

- Which ticket should I buy?

Example…

Option	F, F (p=0.16)	F, ¬F (p=0.24)	¬F, F (p=0.24)	¬F, ¬F (p=0.36)
Combined ticket	C=$30 V=$40 G=$10	C=$30 V=$20 G=-$10	C=$30 V=$20 G=-$10	C=$30 V=$0 G=-$30
Single ticket	C=$40 V=$40 G=$0	C=$20 V=$20 G=$0	C=$20 V=$20 G=$0	C=$0 V=$0 G=$0

The "expected value" of buying a combined ticket is
0.16*10 + 0.24*(-10) + 0.24*(-10) + 0.36 *(-30) = -$14.00

Example…

- Which ticket should I buy?
- Buying a combined ticket in advance is not a good idea when the probability of attending the performance is low.
- Now, change that probability to 0.9.
- The “expected value” of buying a combined ticket is $6.00.
- It is a rational decision in this case!

Combining Beliefs & Desires

- Rational Decision
  - based on Beliefs & Desires
  - where uncertainty & conflicting goals exist
- An agent's preferences are captured by a utility function U which maps a state S to a number U(S) describing the desirability of S.
- A nondeterministic action A may have several outcome states $Result_i(A)$ indexed by the different outcomes of A.
- Prior to executing an action A, the agent assigns a probability $P(Result_i(A) \mid Do(A),E))$ to each outcome.

Maximum Expected Utility

- Expected Utility

$$EU(A \mid E) = \sum_i P(Result_i(A) \mid E, Do(A)) U(Result_i(A))$$

- Maximum Expected Utility (MEU) principle
  - Choose an action which maximize agent's expected utility
- If the agent's utility function U correctly reflects its performance measure, then it will achieve the highest possible performance averaged over the environments in which it could be placed.
- Simple decision
  - concerned only with single or **one-shot** decisions

The basis of Utility Theory

- As a justification for the MEU principle, some constraints are imposed on the preferences that a rational agent should possess.
- In utility theory, different attainable outcomes (prizes) and the respective probabilities (chances) are formalized as lotteries:
  - A lottery L having outcomes $A_1,\ldots, A_n$ with probabilities $p_1+\ldots+p_n = 1$ is denoted $[p_1, A_1;\ldots; p_n, A_n]$.
  - A lottery [1,A] with a single outcome A is abbreviated as A.
- Preference relations for lotteries (or states) A and B:
  - $A > B \Leftrightarrow$ the agent prefers A to B
  - $A \sim B \Leftrightarrow$ the agent is indifferent between *A* and *B*
  - $A \geq B \Leftrightarrow$ the agent prefers A to B or is indifferent

Axioms of utility theory

- **Orderability**: $(A > B) \vee (A < B) \vee (A \sim B)$
- **Transitivity**: $(A > B) \wedge (B>C) \Rightarrow (A>C)$.
- **Continuity**: If $A>B>C$ then $\exists p[p,A;\ 1-p,C] \sim B$
- **Substitutability**:
  - If $A \sim B$ then $[p,\ A;\ 1-p,\ C] \sim [p,\ B;\ 1-p,\ C]$
- **Monotonicity**: If A is better than B, then a lottery that differs only in assigning higher prob to A is better: $p \geq q \Leftrightarrow [p,\ A;\ 1-p,\ B] \geq [q,\ A;\ 1-q,\ B]$
- **Decomposability**: Compound lotteries can be reduced using the laws of probability ("no fun in gambling" – you don't care whether you gamble once or twice, as long as the results are the same)

Utility principle

- Utility principle

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

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

- Maximum Expected Utility principle

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

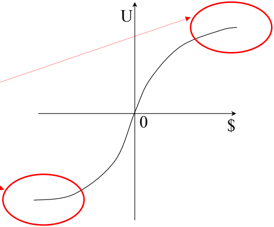
- Utility Function
  - Represents what the agent's actions are **trying to achieve**.
  - Can be **constructed by observing agent's preferences**.

The utility of money

- A choice example
  - either you take the $1,000,000 prize
  - or gamble
    - if the coin comes up H, you get nothing
    - if the coin comes up T, you get $3,000,000
  - Expected Monetary Value (EMV)
    - ½($0)+½($3,000,000) = $1,500,000 > $1,000,000
  - Is gamble a better decision?
- Money does not usually behave as a utility function
  - Empirical data suggests that the value of money is logarithmic
  - For most people getting $5 million is good, but getting $6 million is not 20% better

Utility Function

- Money =/= utility
  - Given a lottery ***L***
  - risk-averse
    $U(S_L) < U(S_{EMV(L)})$
  - risk-seeking
    $U(S_L) > U(S_{EMV(L)})$

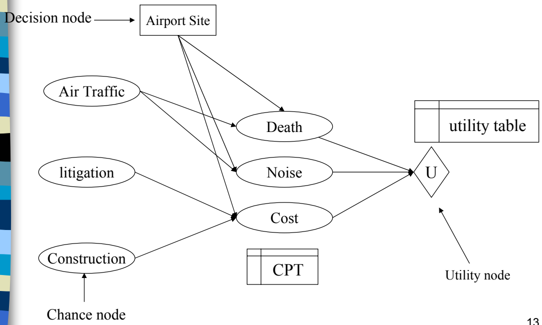


- The value an agent will accept in lieu of a lottery is called it **certainty equivalent** (CE).
- Insurance premium = EMV - CE

Decision Networks

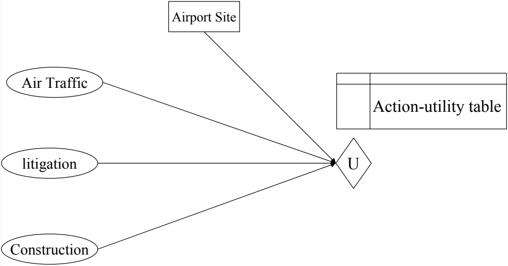
- Represents information about the agent's current state, its possible actions, the state that will result from the agent's action, and the utility of that state.
- A.k.a. influence diagram
- Belief network + decision & utility node
- Nodes
  - Chance node : represent random variables
  - Decision node : represent points where the decision-maker has a choice of actions
  - Utility node : represent the agent's utility function

Decision Networks



Decision Networks

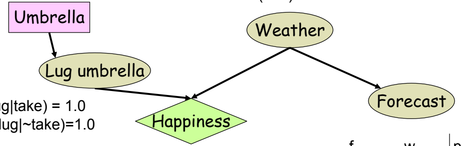
- Simplified representation



Example: Umbrella network

take/don't take

P(rain) = 0.4



P(lug|take) = 1.0
P(~lug|~take)=1.0

U(lug, rain) = 25
U(lug, ~rain) = -20
U(~lug, rain) = -100
U(~lug, ~rain) = 100

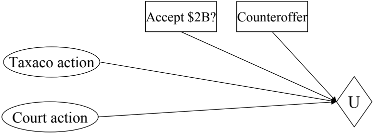
f	w	p(f|w)
sunny	rain	0.3
rainy	rain	0.7
sunny	no rain	0.8
rainy	no rain	0.2

Another example: Pennzoil

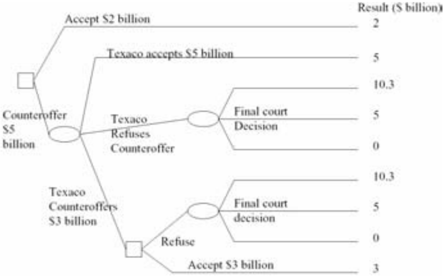
In early 1984, Pennzoil and Getty Oil agreed to the terms of a merger. But before any formal documents could be signed, Texaco offered Getty Oil a substantially better price, and Gordon Getty, who controlled most of the Getty stock, reneged on the Pennzoil deal and sold to Texaco. Naturally, Pennzoil felt as if it had been dealt with unfairly and filed a lawsuit against Texaco alleging that Texaco had interfered illegally in Pennzoil-Getty negotiations. Pennzoil won the case; in late 1985, it was awarded \$11.1 billion, the largest judgment ever in the United States. A Texas appeals court reduced the judgment by \$2 billion, but interest and penalties drove the total back up to \$10.3 billion. James Kinnear, Texaco's chief executive officer, had said that Texaco would file for bankruptcy if Pennzoil obtained court permission to secure the judgment by filing liens against Texaco's assets. . . .

Another example: Pennzoil…

. . . Furthermore Kinnear had promised to fight the case all the way to the U.S. Supreme Court if necessary, arguing in part that Pennzoil had not followed Security and Exchange Commission regulations in Its negotiations with Getty. In April 1987, just before Pennzoil began to file the liens, Texaco offered Pennzoil $2 billion to settle the entire case. Hugh Liedtke, chairman of Pennzoil, indicated that his advisors were telling him that a settlement of between $3 billion and $5 billion would be fair.



Another example: Pennzoil…



Decision Networks

- Evaluating decision networks

  1. set the evidence variables for the current state
  2. For each possible value of the decision node
     (a) Set the decision node to that value
     (b) Calculate the posterior probabilities for the parent nodes of the utility node, using a standard probabilistic inference algorithm.
     (c) Calculate the resulting utility for the action
  3. Return the action with the highest utility

The value of information

- Not all available information is provided to the agent before it makes its decision
- One of the most important parts of decision making is knowing what questions to ask.
- To conduct expensive and critical tests or not depends on two factors:
  - Whether the different possible outcomes would make a significant difference to the optimal course of action
  - The likelihood of the various outcomes
- Information value theory enables an agent to choose what information to acquire.

Value of Information

- Not all available information is provided to the agent before it makes its decision
- Example
  - An oil company is hoping to buy one of n indistinguishable blocks of ocean drilling rights
  - Exactly one of the blocks contain oil worth C dollars, and that the price of each block is C/n dollars.
  - A seismologist offers the results of a survey of block number 3, which indicates definitively whether the block contains oil. How much should the company pay for this information?

Value of Information

- Solution

  1) block 3 contains oil → the company will buy this block

  profit : C – C/n = (n-1)C/n

  2) block 3 contains no oil → the company will buy a different block

  profit : C/(n-1) – C/n = C/n(n-1)

expected profit : $\frac{1}{n} \times \frac{(n-1)C}{n} + \frac{n-1}{n} \times \frac{C}{n(n-1)} = \frac{C}{n}$

Value of Information

- General formula
  - Value of perfect information(VPI)

$$EU(\alpha \mid E) = \max_{A} \sum_i U(S_i) P(S_i \mid E, Do(A))$$

$$EU(\alpha_{E_j} \mid E, E_j) = \max_{A} \sum_i U(S_i) P(S_i \mid E, Do(A), E_j)$$

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

- E : current evidence, α : current best choice
- E_j : newly obtained evidence