
Artificial Intelligence

IV. Uncertain Knowledge and Reasoning

5. Rational Decisions

F. C. Langbein

School of Computer Science
and Informatics
Cardiff University





Overview

- Utility theory
 - Preferences and lotteries
 - Utility
 - Maximum expected utility
 - Decision networks
 - Multi-attribute utility
 - Dominance
 - Independence
- Human decisions and game theory examples
- Value of Information

Planning a Party

- Agent's problem: decide whether to have a party outside
 - Maximise rating of party based on utility ratings:

	rain	¬rain
outside	1 (poor)	4 (excellent)
inside	3 (good)	2 (satisfactory)

- Probability of rain is $1/3$
- Choice between two actions with uncertain outcomes
 - Expected utility of an action a with uncertain result r is

$$EU(a) = P(r|a) * U(r|a) + P(\neg r|a)U(\neg r|a)$$

$$EU(\text{outside}) = (1/3)1 + (2/3)4 = 3$$

$$EU(\text{inside}) = (1/3)3 + (2/3)2 \approx 2.33$$

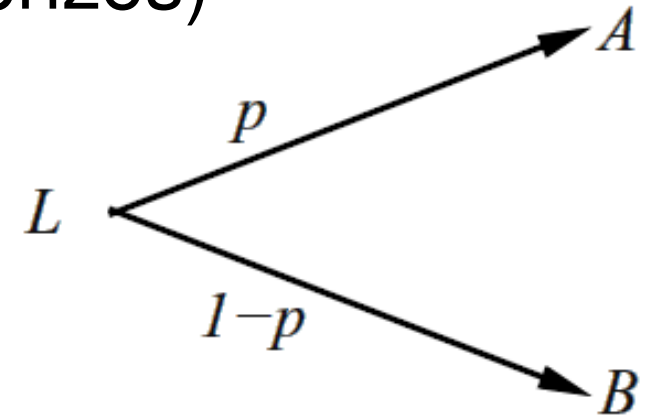
Lotteries and Preferences

➤ An agent chooses among **lotteries**

- Situations with uncertain outcomes (prizes)

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

- Notation: $L = [p_1, R_1; \dots; p_n, R_n]$



➤ An agent has preferences for **prizes**

- State which outcomes are preferred over other outcomes

- Notation for preferences:

- $A \succ B$ A preferred to B

- $A \sim B$ indifference between A and B

- $A \not\succeq B$ B not preferred to A

Compound Lotteries

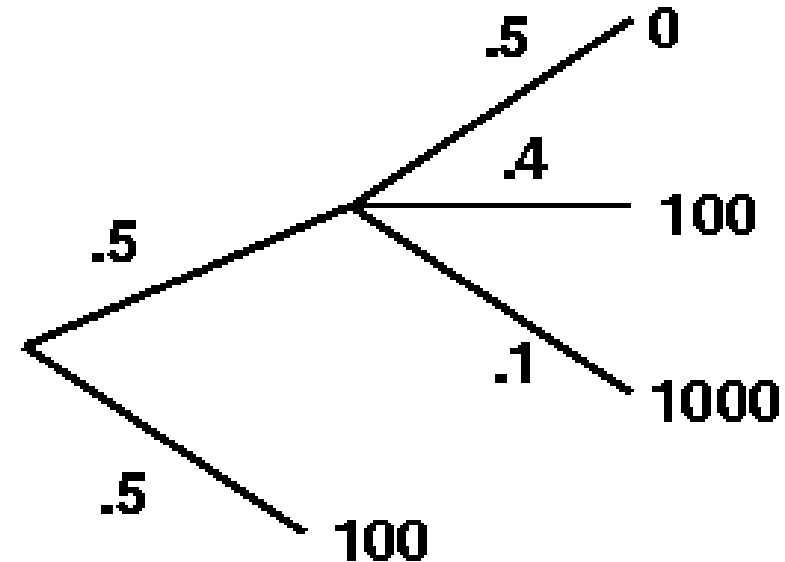
➤ Combine lotteries:

- $L_1 = [.5, L_2; .5, \text{win_100}]$
- $L_2 = [.5, \text{win_0}; .4, \text{win_100}; .1, \text{win_1000}]$
- Yields combined lottery:

$$\begin{aligned} L_{12} &= [.5 * .5, \text{win_0}; .5 * .4 + .5, \text{win_100}, .5 * .1, \text{win_1000}] \\ &= [.25, \text{win_0}; .7, \text{win_100}, .05, \text{win_1000}] \end{aligned}$$

➤ In general:

- Multiply probabilities of paths leading to result
- And add resulting probabilities for same result



Rational Preferences

- Preferences of a rational agent must obey constraints.
- **Rational preferences** \Rightarrow behaviour describable as **maximisation 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**:
 $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; 1 - p, B] \succeq [q, A; 1 - q, B])$

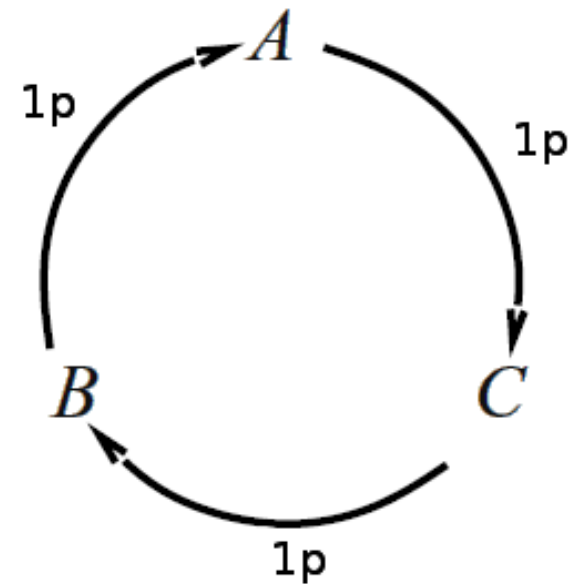
Rational Preferences

- 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 pence to get B

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

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



Maximising Expected Utility

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

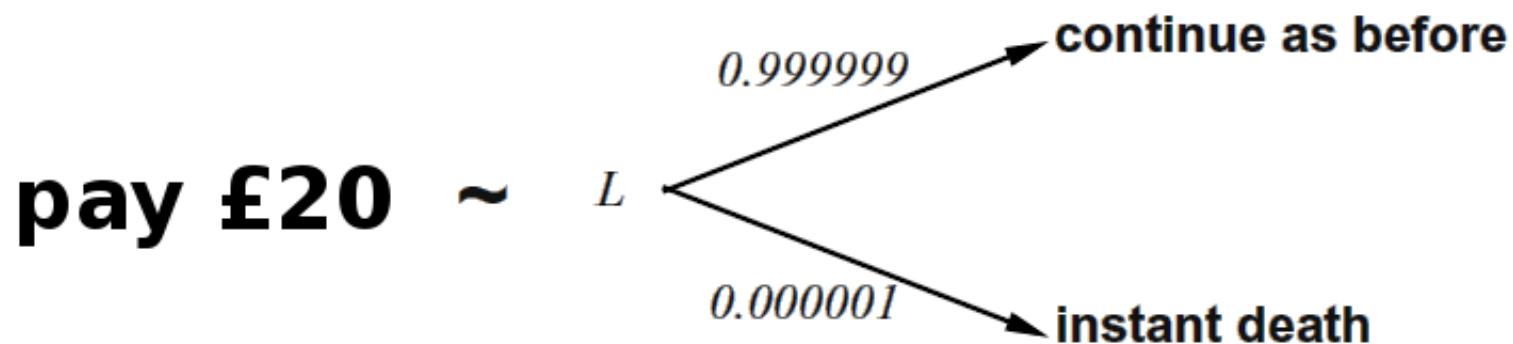
$$U(A) \geq U(B) \Leftrightarrow A \succsim B$$

$$U([p_1, S_1; \dots; p_n, S_n]) = \sum_{l=1}^n p_l U(S_l)$$

- **Maximum expected utility** (MEU) principle:
Choose the action that maximises 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 Tic Tac Toe

Utilities

- Utilities map states (prizes) 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_{\perp} with probability $(1-p)$
 - Adjust lottery probability p until $A \sim L_p$

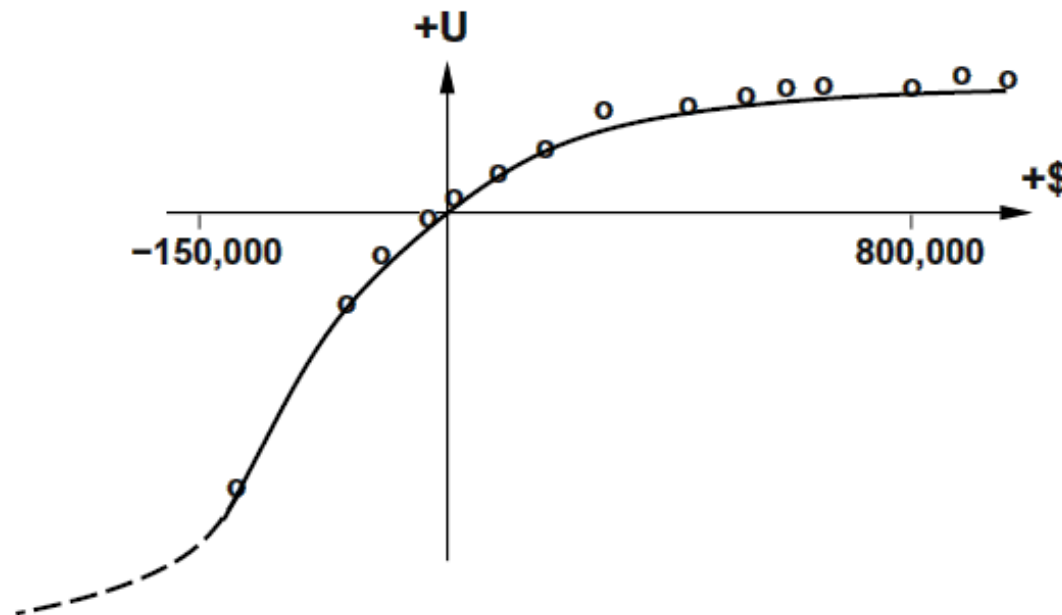


Utility Scales

- **Normalised 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: behaviour is *invariant* w.r.t. linear transformation
$$U'(x) = k_1 U(x) + k_2 \quad \text{where } k_1 > 0$$
- Only when all prizes are deterministic (no lottery choices):
 - *Ordinal utility* can be determined
(i.e. total order on prizes)

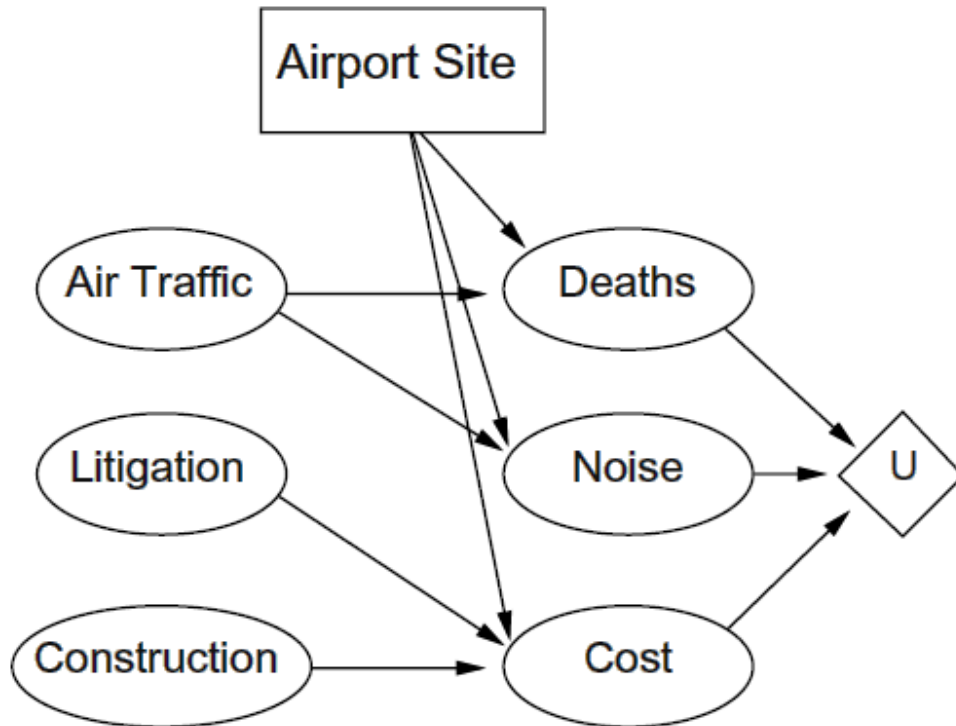
Money

- Money does *not* behave as a utility function
- Given a lottery L with expected monetary value $EMV(L)$: usually $U(L) < U(EMV(L))$ (people are *risk-averse*)
- *Utility curve*: for what probability p am I indifferent between fixed prize x and a lottery $[p, \$M; (1 - p), \$0]$ for large M ?
- Typ. empirical data, extrapolated with *risk-prone* behaviour



Decision Networks

- Add **action nodes** and **utility** nodes to Bayesian network (**chance** nodes) to enable rational decision making



$$EU(a|e) = \sum_s U(s)P(s|e, a)$$

➤ Algorithm

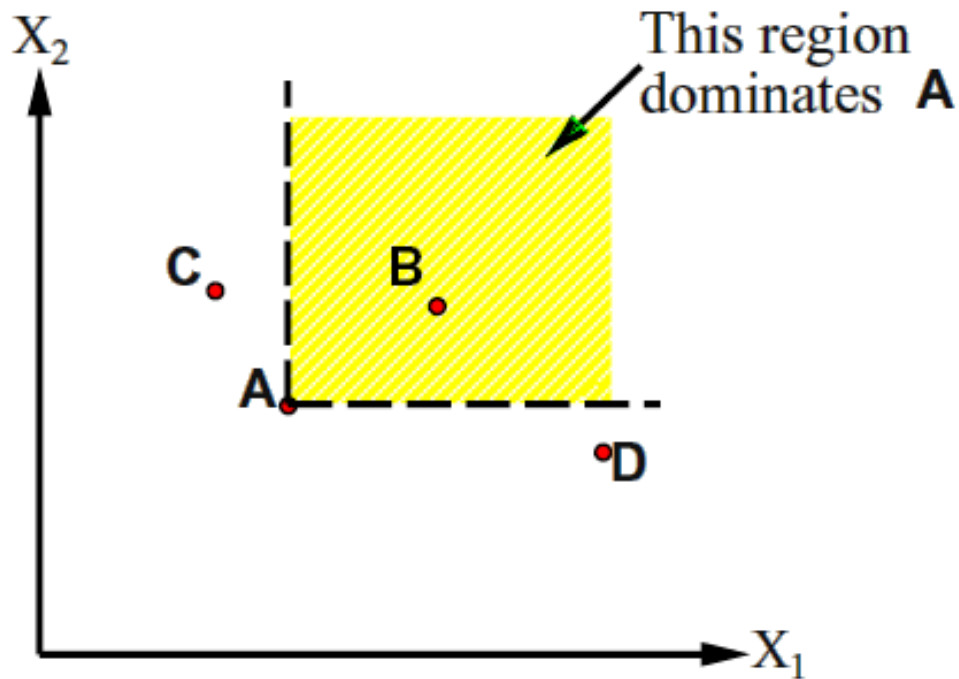
```
For  $a \in net.action\_node$   
     $EU[a] \leftarrow$  EXPECTEDUTILITY ( $net.utility\_node,$   
                                      $\{a\} \cup net.evidence$ )  
return argmax ( $EU[a] : a \in net.action\_node$ )
```

Multi-attribute Utility

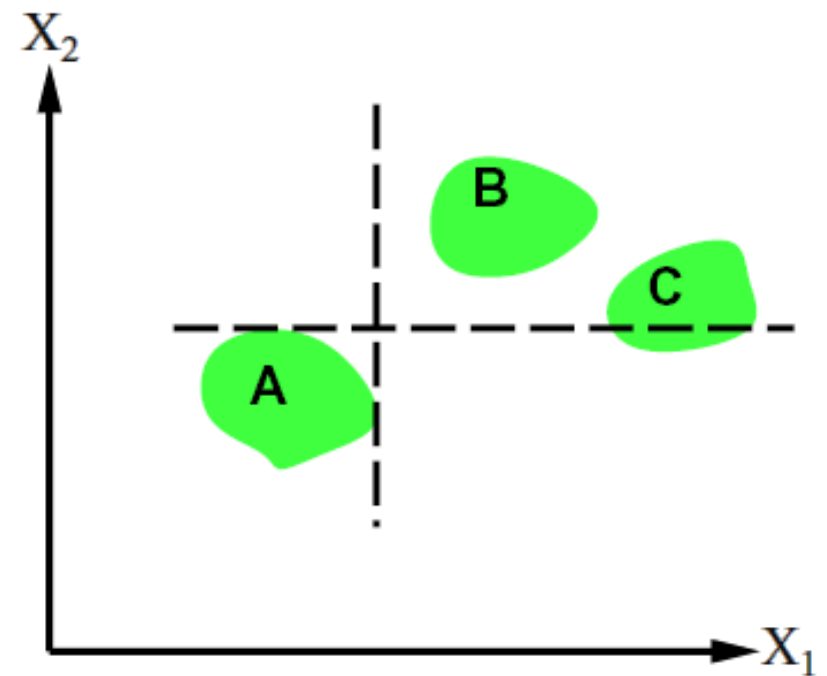
- How can we handle utility functions of *many* variables X_1, \dots, X_n ?
 - E.g. what is $U(\text{Deaths}, \text{Noise}, \text{Cost})$?
- How can *complex* utility functions be assessed from preference behaviour?
- Idea 1: **Dominance**
 - Identify conditions under which decisions can be made without complete identification of $U(X_1, \dots, X_n)$
- Idea 2: **Independence**
 - Identify various types of independence in preferences and derive consequent canonical forms for $U(X_1, \dots, X_n)$

Strict Dominance

- Typically define attributes such that U is **monotonic** in each
- **Strict dominance**: choice B strictly dominates choice A iff
$$\forall l \ X_1(B) \geq X_1(A) \quad (\text{and hence } U(B) \geq U(A))$$



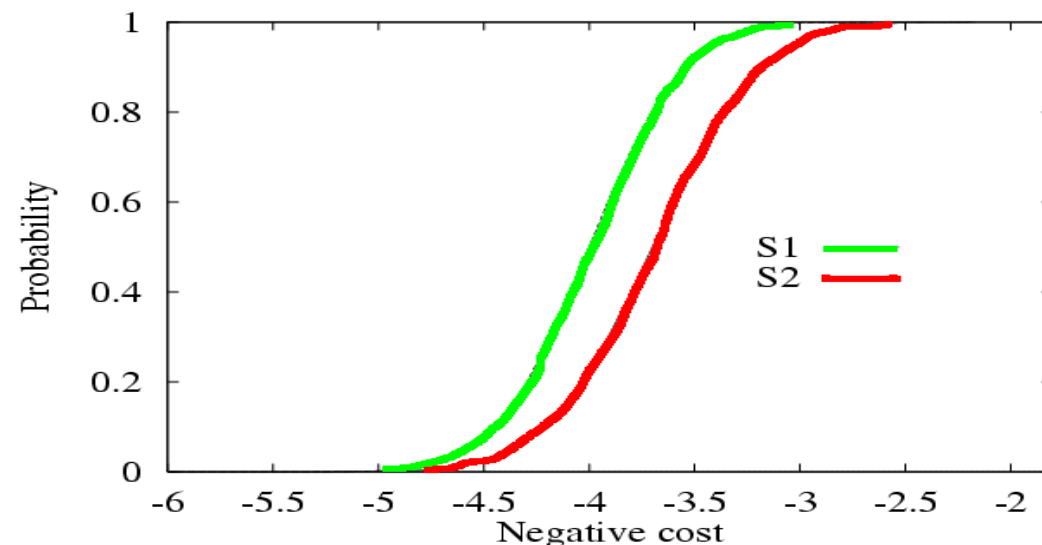
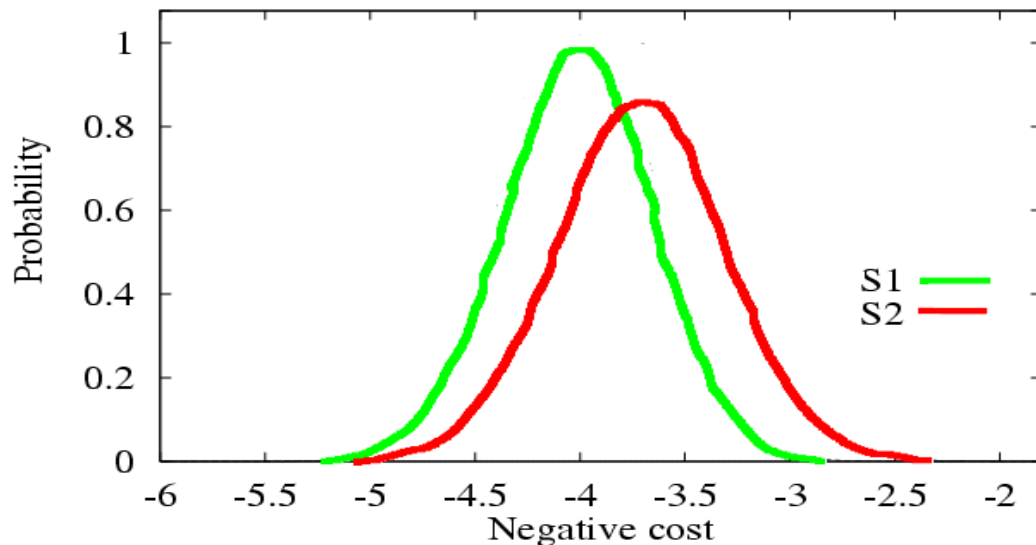
Deterministic attributes



Uncertain attributes

- Strict dominance seldom holds in practice

Stochastic Dominance



- Action S_1 **stochastically dominates** distribution S_2 on X iff

$$\forall t \int_{-\infty}^t p_2(x) dx \leq \int_{-\infty}^t p_1(x) dx$$

- If S_1 stochastically dominates S_2 , then for any monotonically non-decreasing utility U , the expected utility of S_2 is at least as high as the expected utility of S_1
- Multi-attribute case: *stochastic dominance* on *all* attributes
⇒ *optimal*

Stochastic Dominance

- Stochastic dominance can often be determined without exact distributions using **qualitative** reasoning
 - E.g. construction cost increases with distance from city
 - S_1 is further from the city than S_2
 - $\Rightarrow S_1$ stochastically dominates S_2 on cost
 - E.g. injury increases with collision speed
- This qualitative information can be handled by *qualitative probabilistic networks*
 - Annotate belief networks with qualitative information:
 $X \xrightarrow{+/-} Y$ (X *positively/negatively influences* Y)

Independence

- X_1 and X_2 **preferentially independent** (PI) of X_3 iff preference between $\langle x_1, x_2, x_3 \rangle$ and $\langle x'_1, x'_2, x_3 \rangle$ does not depend on x_3
 - E.g. $\langle 20\ 000\ \text{suffer}, \$4.6\ \text{billion}, 0.06\ \text{deaths/mpm} \rangle$ vs. $\langle 70\ 000\ \text{suffer}, \$4.2\ \text{billion}, 0.06\ \text{deaths/mpm} \rangle$
- Mutual PI $\Rightarrow \exists$ **additive** utility function $U(S) = \sum_i U_i(X_i(S))$
 - Use single-attribute utilities
 - Often good approximation
- Preferences over lotteries: X is **utility independent** (UI) of Y iff preferences over lotteries X do not depend on Y
 - **Mutual UI** $\Rightarrow \exists$ **multiplicative** utility function
$$U = k_1U_1 + k_2U_2 + k_3U_3 + k_1k_2U_1U_2 + k_2k_3U_2U_3 + k_3k_1U_3U_1 + k_1k_2k_3U_1U_2U_3$$

Utility Theory and Human Decisions

- D. Kahneman, A. Tversky study of human decision making:
 - Choice (1):
 - (a) *save 200 lives for sure*
 - (b) 1/3 chance to save 600 lives and 2/3 chance to save no one
 - ➔ Humans are risk-averse in choices between sure gains and favourable gambles
 - Choice (2):
 - (a) 400 people dying for sure
 - (b) *2/3 chance of 600 people dying and 1/3 chance no one dying*
 - ➔ Humans are risk-seeking in choices between sure losses and unfavourable gambles
- Humans are not immune to words, utility theory is



Utility Theory and Human Decisions

- Problems with the theory of expected utility
 - Human preferences are not rational preferences
 - Violations of constraints / axioms (e.g. transitivity)
 - Violations of invariance (e.g. reference point dependency, loss aversion)
 - Assumption that there are no other rational agents
 - ➔ *Non-cooperative game theory*

Game Theory Examples

- “Friends” with asymmetric preferences: John likes Betty, but Betty does not like John that much

		John Home	Pool
Betty	Home	(2,0)	(2,1)
	Pool	(3,0)	(1,2)

- *Pareto optimal*: no agent can be better off without making another agent worse off
 - Nash equilibrium: strategies are tied to each other – no one can gain by change of strategy unless someone else also changes strategy
- Prisoner’s dilemma: defect against each other or not?

		Leland Solidarity	Defection
Stan	Solidarity	(3,3)	(1,4)
	Defection	(4,1)	(1,1)

- Social dilemma: self-utility leads to inefficient outcome



Value of Information

- Which information to get next?
- Compute value of acquiring each possible piece of evidence
 - ➔ Can be done *directly from decision network*
- Example: buying oil drilling rights
 - Two blocks A and B, exactly one has oil, worth k
 - Prior probabilities 0.5 each, mutually exclusive
 - Current price of each block is $k/2$
 - Consultant offers accurate survey of A.
 - ➔ What is a fair price for this information?

Value of Information

- Compute **expected value of information**
 - Expected value of best action given the information minus expected value of best action without information
- Survey may say “oil in A” or “no oil in A” with probability 0.5 each
 - *Expected value of information* is
 - [0.5 * value of “buy A” given “oil in A”
+ 0.5 * value of “buy B” given “no oil in A”]
– 0
 - = $\left[0.5\frac{k}{2} + 0.5\frac{k}{2}\right] - 0 = k/2$

General Formulas

- With evidence e and actions A with possible outcomes S , choose action α such that expected utility

$$EU(\alpha|e) = \max_{\alpha \in A} \sum_{s \in S} U(s) P(s|e, \alpha)$$

- Suppose we knew $E^* = e_l^*$, then we would choose α_l s.t.

$$EU(\alpha_l|e, e_l^*) = \max_{\alpha \in A} \sum_{s \in S} U(s) P(s|e, e_l^*, \alpha)$$

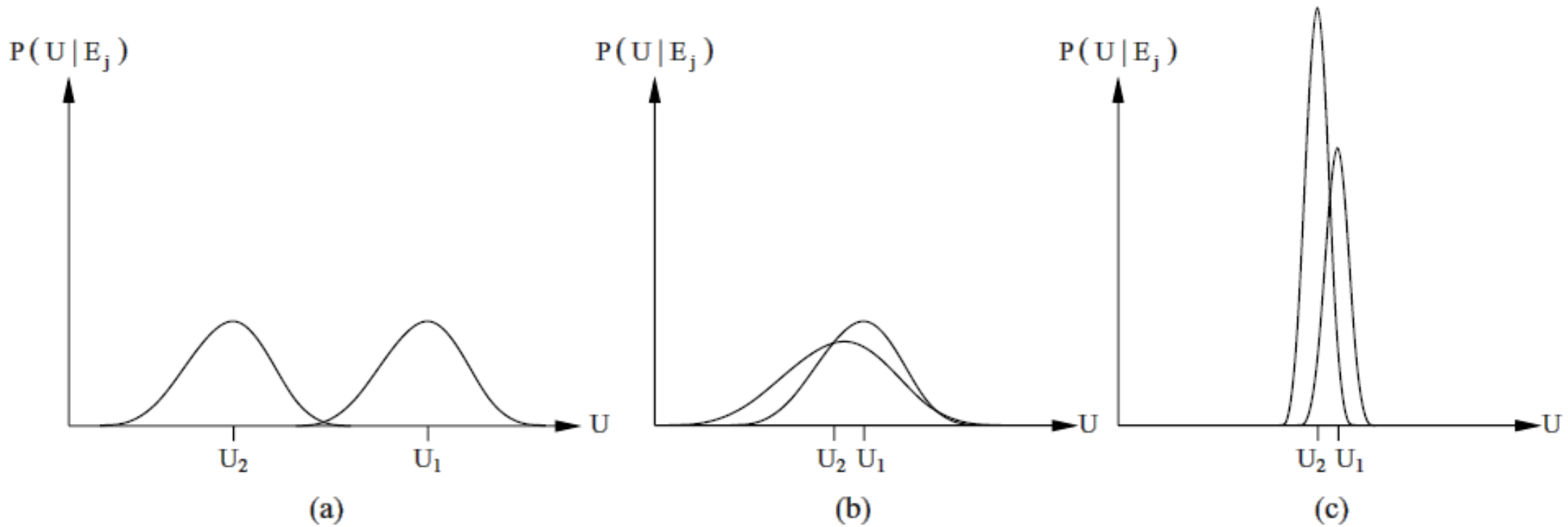
- E^* is a random variable whose value is *currently* unknown, so we must compute expected gain over all possible values:

$$VPI(E^*) = (\sum_l P(e_l^*|e) EU(\alpha_l|e, e_l^*)) - EU(\alpha|e)$$

(VPI = *value of perfect information*)

- When more than one piece of evidence can be gathered, maximising VPI for each individually is not always optimal
 - ➔ Evidence-gathering is a *sequential* decision problem

Qualitative Behaviours



- a) Choice is obvious, information not needed
- b) Choice is unclear, information crucial
- c) Choice is unclear, information less valuable