Chap. 16

Making Simple Decisions

How should an agent make decisions
to achieve what it wants – on average, at least?

Outline

- Rational Preferences
- Utilities
- Money
- Multiattribute Utilities
- Decision Networks
- Value of Information

Introduction

- ***Decision Theory*** (what an agent should do)

  ⊃ *Probability Theory* (what an agent should believe on the basis of evidence)

  \+ *Utility Theory* (what an agent wants)
- *DT*: Build a *Rational Agent* which makes the *best decision* to choose an action which leads to the *best expected outcome* by considering all possible actions.
- *UT*: an agent can be described as possessing a *utility function*, and selections actions as if maximizing its *expected utility*.
- *Multi-attribute UT* deals with utilities that depend on several distinct attributes of states – *Stochastic dominance*: useful for making unambiguous decisions.
- *Decision Network*: a formalism to express/solve decision problems.

  \- an extension of Bayesian networks.
- The *Value of Information*: the expected improvement in utility compared with making a decision w/o the information.
- *Expert System*: KB + inference system + decision making + the value of information

3

Introduction: Decision_Theoretic Agent

- Probability Theory + Utility Theory $\Rightarrow$ Decision-Theoretic Agent,

  i.e. Belief + Desire/Preference
- A continuous *measure of state quality*.
  - Utility function $U: State \rightarrow Number$ (the desirability/preference of a state)
- The maximization of expected utility: probability + utility of each action.
- The behavior of any rational agent can be captured by supposing a utility function that is being maximized.
- How to handle utility functions that depend on several quantities?
- *Decision networks*: a formalism to express/solve decision problems.

  \- an extension of Bayesian Networks with actions + utilities.

4

Beliefs + Desires under Uncertainty

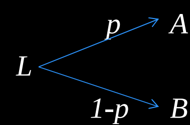
- Action:
  - Nondeterministic
  - a current state $\rightarrow^{A}$ possible outcome states ($Resut(A)=s_i$)
  - $P(Result(A)=s_i \mid A, E)$: probability of each outcome state s_i
    where E summarizes the agent's available evidence a/b the world and A is the proposition that action A is executed in the current state.
- The *Expected Utility* of the action given the evidence:
  $EU(A \mid E) = \sum_i P(Result(A)= s_i \mid A, E)\cdot U(s_i)$
- The Principle of Maximum Expected Utility (MEU):
  - A rational agent should choose an action that maximizes the agent's EU.
  - It's infeasible for long sequences of actions to enumerate all action sequences and choose the best.
  - $Action = argmax_A\ EU(A|E)$

The Principle of MEU

- The Principle of Maximum Expected Utility (MEU):
  - A rational agent should choose an action that maximizes the agent's EU.
  - It's infeasible for long sequences of actions to enumerate all action sequences and choose the best.
  - $Action = argmax_A\ EU(A|E) = argmax_A \sum_i P(Result(A)=s_i|A,E)\cdot U(s_i)$
- Problem:
  - Computing $P(Result(A)=s_i|A,E)$ requires a complete causal model of the world
  - i.e. the prohibited computations & NP-hard inference in a large BNs.
  - Computing $U(s_i)$ requires searching/planning because an agent doesn't know the quality of s_i until it know to where s_i yields the agent.
- If an agent acts so as to maximize a utility function that correctly reflects the performance measure, then the agent will achieve the highest possible performance score (averaged over all the possible environment.)

Preferences

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

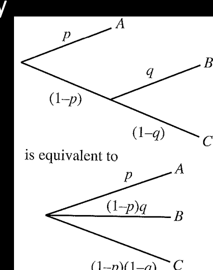


- Lottery $L = [p, A; (1\text{-}p), B]$ or $[p_1, C_1; p_2, C_2; \ldots p_n, C_n]$ in general:
  a probability distribution over a set of actual outcomes (the prizes of the lottery) where each outcome is either an atomic state or another lottery.
- Notation:
  - $A \succ B$ A is preferred to B
  - $A \sim B$ indifference between A and B
  - $A \succsim B$ B is not preferred to A, i.e. $(A \succ B \vee A \sim B)$

7

Rational Preferences

- Idea: preferences of a rational agent must obey constraints.
- Rational preferences $\Rightarrow$
  behavior describable as maximization of expected utility
- How preferences b/t complex lotteries are related to preferences b/t the underlying states in those lotteries?

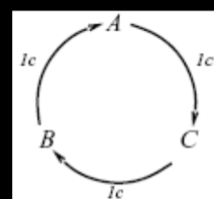


- Constraints: The Axioms of Utility Theory
  - Orderability: $(A \succ B) \vee (B \succ A) \vee (A \sim B)$
  - Transitivity $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$
  - Continuity $A \succ B \succ C \Rightarrow \exists p \; [p, A; \; 1-p, C] \sim B$
  - Substitutability $A \sim B \Rightarrow [p, A; \; 1-p, C] \sim [p, B; 1-p, C]$
  - Monotonicity $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A; \; 1-p, B] \succsim [q, A; \; 1-q, B])$
  - Decomposability $[p,A; 1\text{-}p,[q,B; 1\text{-}q, C]] \sim [p,A; (1\text{-}p)q, B; (1\text{-}p)(1\text{-}q), C]$

8

.. continued

- 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



9

Preferences lead to Utility

- Existence of Utility function

  If an agent's preferences obey the axioms of utility,
  then there exists a function U s.t.
  $U(A) > U(B) \Leftrightarrow A \succ B$, $U(A) = U(B) \Leftrightarrow A \sim B$

- Expected Utility of a Lottery

  The utility of a lottery is : $L = [p_1, S_1; p_2, S_2; \ldots p_n, S_n]$
  the sum of the probability of each outcome × the utility of that outcome,
  the sum of the expected utility of each outcome.

  $$U([p_1, S_1; p_2, S_2; \ldots p_n, S_n]) = \sum_i p_i \cdot U(S_i)$$

- Utility function doesn't establish its uniqueness: agent's behavior wouldn't change if U(S) were transformed according to

  $U'(S) = a \cdot U(S) + b$ where a, b are constants and $a > 0$.

- Value function, Ordinal Utility function: Agent just needs a preference ranking on states - the numbers don't matter.

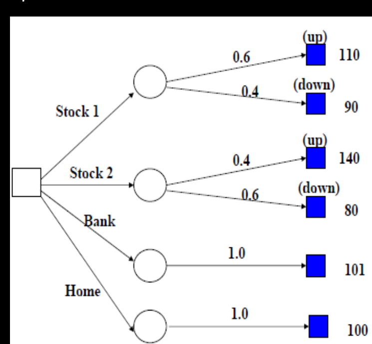
Maximizing Expected Utility (MEU)

- **Theorem** (Ramsey, 1931; von Neumann and Morgenstern, 1944):
  Given preferences satisfying the constraints
  there exists a real-valued function U such that
  $U(A) \geq U(B) \Leftrightarrow A \succsim B$
  $U([p_1, S_1; \ldots ; p_n, S_n]) = \sum_i p_i U(S_i)\ (\Rightarrow EU(A \mid E))$: utility of a lottery
- **MEU principle:**
  Choose the action that maximizes expected utility
  $argmax_A\ EU(A|E) = argmax_A \sum_i P(S_i\ |A,E)U(S_i)$
- 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
- Having a utility function that describes an agent's preference doesn't necessarily mean that agent uses MEU explicitly in its own deliberation. Knowing the rational agent's preferences helps to construct the utility function that represents what the agent is actually trying to achieve.

11

Example: Investment

- Assume we want to invest $100 for 6 months.
- 4 choices:
  1. Invest in *Stock 1*
  2. Invest in *Stock 2*
  3. Save it in *Bank*
  4. Keep it at *Home*



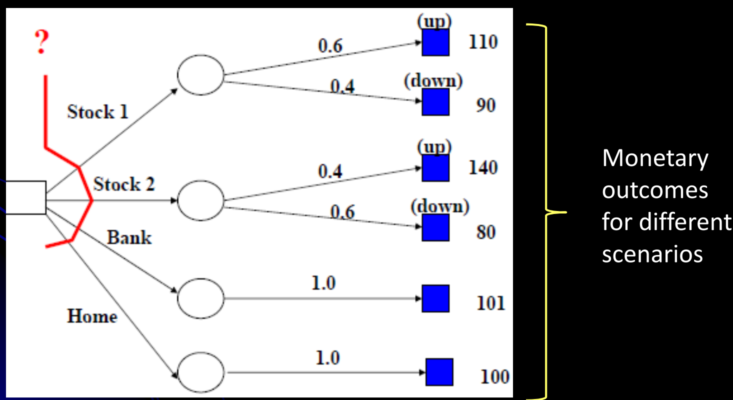
Monetary outcomes for different states

Stock 1 value can go *Up* or *Down*: [.6, Up; .4, Down]
Up: with probability 0.6
Down: with probability 0.4

12

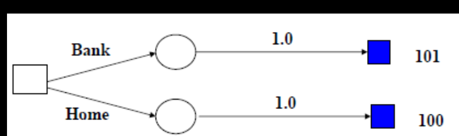
Cont.

- Assume we want to invest $100 for 6 months.
- A need of making a choice among *{Stock1, Stock2, Bank, Home}*

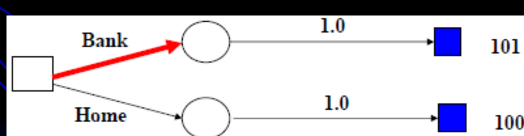


Cont.

- Simple Cases: *Bank* and *Home*
- The result is guaranteed – the outcome is deterministic.

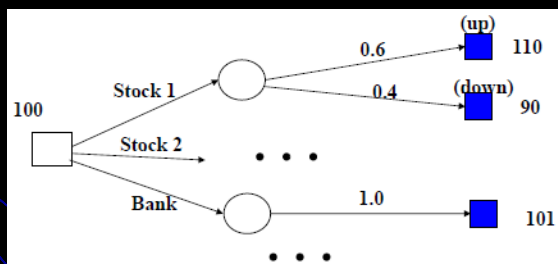


- These choices are *deterministic.*
- Goal: To make money. What is the rational choice?



Cont.

- Uncertain cases: *Stock1* and *Stock2*
  - What to do with uncertain outcomes?
- How to quantify the value of the stochastic outcome?
  - deterministic outcome vs. stochastic outcome



- Idea: Use *Expected Value* of the Outcome
- Notice that a monetary value of *x* is used as a utility of X for convenience:
  i.e. $U(x) = x$.

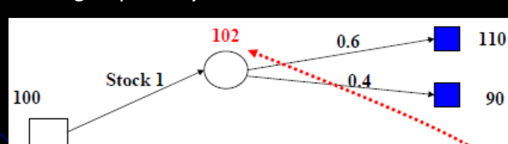
15

Cont.

- Let X be a random variable representing the monetary outcome with a discrete set of values D_X.
- *Expected (Monetary) Value* of X is:

$$E(X) = \sum_{x \in D_X} x P(X = x)$$

- Intuition: Expected value summarizes all stochastic outcomes into a *single quantity*.
- Example:

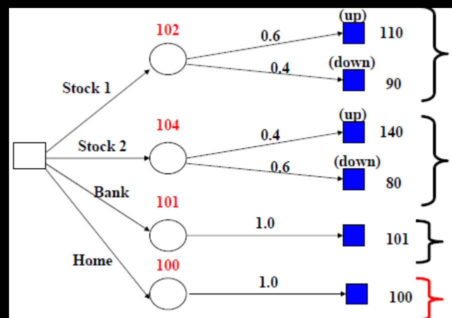


- What is the expected value of the outcome of Stock1 option?
  $0.6 \times 110 + 0.4 \times 90 = 66 + 36 = 102$

16

Cont.

- The Expected Value (EV) of each outcome:

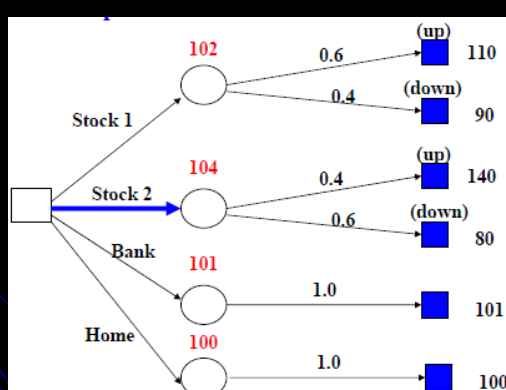


  - $EV(S_{Stock1}) = 0.6 \times 110 + 0.4 \times 90 = 66 + 36 = 102$
  - $EV(S_{Stock2}) = 0.4 \times 140 + 0.6 \times 80 = 56 + 48 = 104$
  - $EV(S_{Bank}) = 1 \times 101 = 101$
  - $EV(S_{Home}) = 1 \times 100 = 100$

17

Cont.

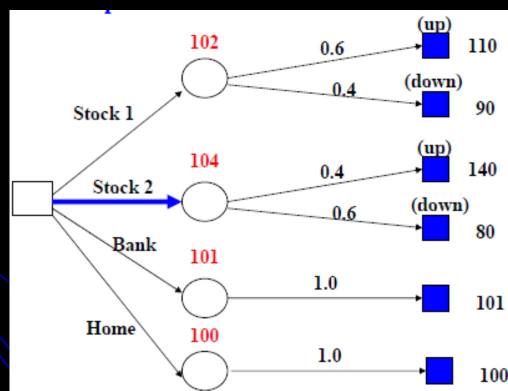
- The optimal action is the choice that maximizes the expected outcome.



18

Cont.

- The optimal action is the choice that maximizes the expected outcome.



- Decision tree: decision node( ), chance node( ), outcome node( ) 19

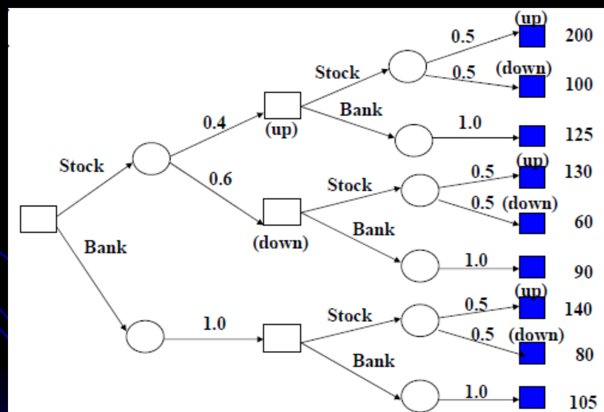
Example: Investment 2

- The Sequential (Multi-Step) Problems:
  - The decision tree can be build to capture multi-step decision problems:
    - Choose an action
    - Observe the stochastic outcome
    - And, repeat
  - How to make decisions for multi-step problems?
    - Start from the leaves of the decision tree (outcome nodes)
    - Compute the values of expectation at chance nodes.
    - Maximize at the decision nodes.
- Example:
  - Assumption: 2 investment periods and 2 actions *{Stock, Bank}*

20

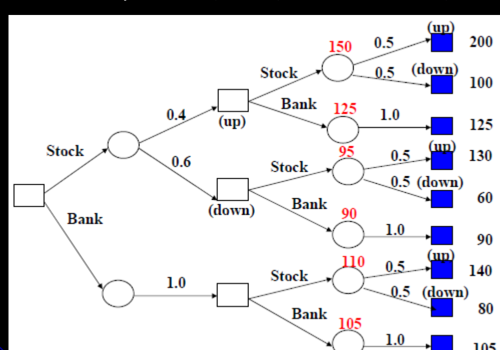
Cont.

- Example: 2 investment periods {t1, t2} and 2 actions *{Stock, Bank}*



Cont.

- Example: 2 investment periods {t1, t2} and 2 actions *{Stock, Bank}*

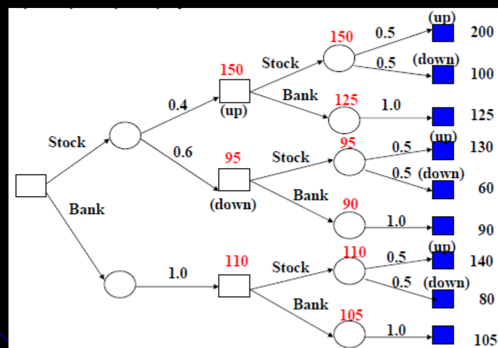


Expected Values of chance nodes at period 2:

- $EV(S_{Stock_{t2}} | S_{Stock_{t1}}=up) = 0.5\times200+0.5\times100=150$; $EV(S_{Bank_{t2}} | S_{Stock_{t1}}=up) = 1.0\times125=125$
- $EV(S_{Stock_{t2}} | S_{Stock_{t1}}=down) = 0.5\times130+0.5\times60=95$; $EV(S_{Bank_{t2}} | S_{Stock_{t1}}=down) = 1.0\times90=90$
- $EV(S_{Stock_{t2}} | S_{bank_{t1}}=true) = 0.5\times140+0.5\times80=110$; $EV(S_{Bank_{t2}} | S_{bank_{t1}}=true) = 1.0\times105=105$

Cont.

- Example: 2 investment periods {t1, t2} and 2 actions {*Stock, Bank*}



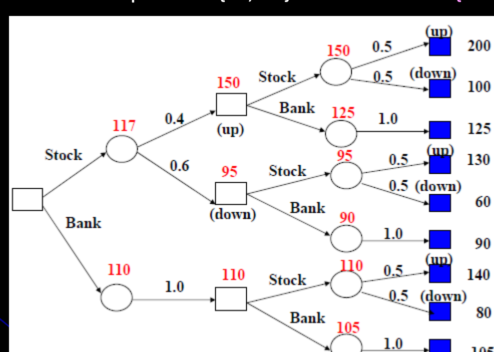
Maximum Expected Values for decision nodes at period 2:

- MEV(D_{up}) = Max (EV($S_{Stock(up,t2)}$), EV($S_{Bank(up,t2)}$)) = 150;
- MEV(D_{down}) = Max (EV($S_{Stock(down,t2)}$), EV($S_{Bank(down,\ t2)}$) = 95
- MEV(D_{bank}) = Max(EV($S_{Stock(bank,t2)}$), EV($S_{Bank(bank,\ t2)}$) = 110

23

Cont.

- Example: 2 investment periods {t1, t2} and 2 actions {*Stock, Bank*}



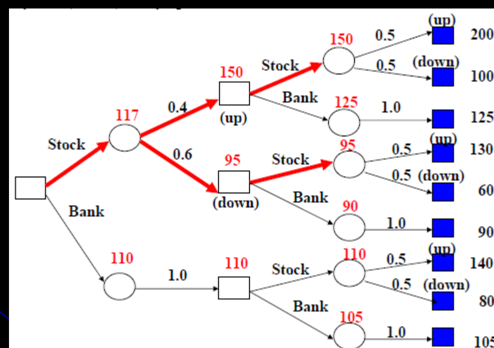
Expected Values for chance nodes at period 1:

- EV($S_{Stock_{t1}}$) = 0.4 × 150 + 0.6 × 95 = 117;
- EV($S_{Bank_{t1}}$) = 1 × 110 = 110;

24

Cont.

- Example: 2 investment periods {t1, t2} and 2 actions *{Stock, Bank}*

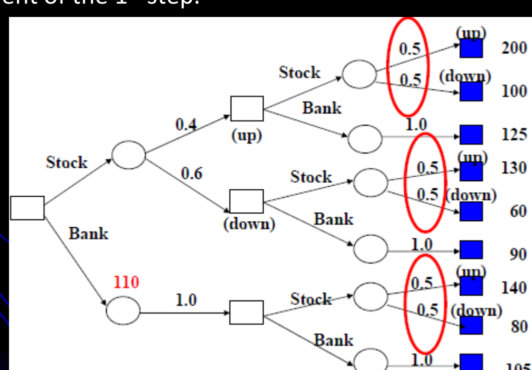


Optimal Decision on the Actions(i.e. choices):

- Stock_1 at period 1 + Stock_2 at period 2
  with $\text{EV}(SStock_2 \mid SStock_1 = up) = 150$ and $\text{EV}(SStock_2 \mid SStock_1 = down) = 95$;

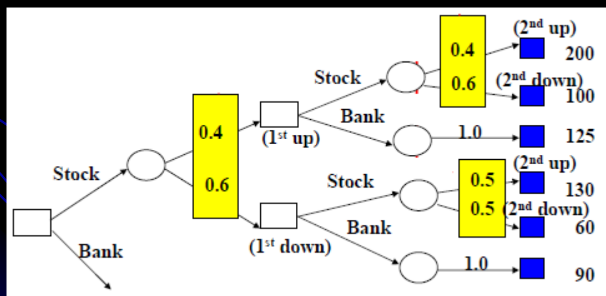
Cont.

- Multi-Step Problems: Conditioning.
  The probability of stock going up and down in the 2^{nd} step is independent of the 1^{st} step.



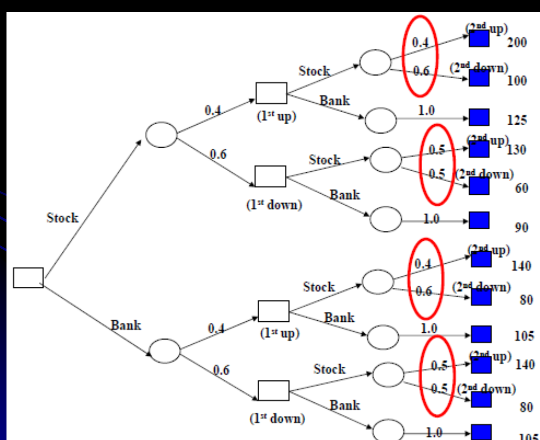
Cont.

- Multi-Step Problems: Conditioning in the Decision Tree
  Later outcomes can be conditioned on the earlier stochastic outcomes and actions.
- Example: Stock movement probabilities.
  Assume $P(Sstock_1=up) = 0.4;$
  $P(Sstock_2=up|Sstock_1=up) = 0.4;$ $P(Sstock_2=up|Sstock_1=down) = 0.5;$



Cont.

- Multi-Step Problems: Conditioning in the Decision Tree
  - Tree Structure: every observed stochastic outcome = 1 branch.



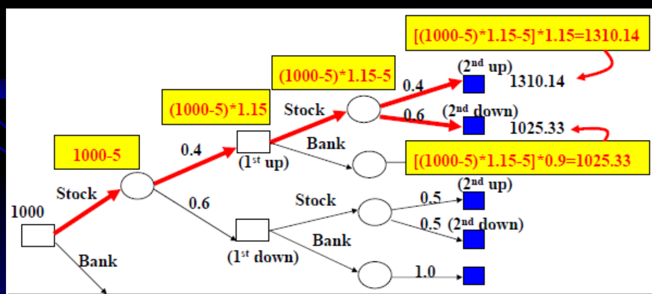
$P(Sstock_1=up) = 0.4;$
$P(Sstock_2=up|Sstock_1=up) = 0.4;$
$P(Sstock_2=up|Sstock_1=down) = 0.5;$

Cont.

- Trajectory payoffs
  - Outcome values at leaf nodes (e.g. monetary values)
    - Rewards and costs for the path trajectory
  - Example: Stock fees and gains.

    Assumption: Net Asset = $1,000, Fee/period = $5 at the beginning,

    Gain for up = 15%, Loss for down = 10%



Example: Investment

- Construction of a Decision Tree
  - The decision tree is rarely given to you directly
    – its construction is a part of the problem.
  - Example: stocks, bonds, bank for *k* period.
    - Stock:
      - Probability of stocks going up in the 1st period = 0.3, i.e. $P(Sstock_1=up)=0.3$
      - Probability of stocks going up in subsequent periods:

        $P(Sstock_k=up|\ Sstock_{k-1}=up)=0.4;\quad P(Sstock_k=up|\ Sstock_{k-1}=down)=0.5$
      - Return if stock goes up: 15%; if down: 10%
      - Fixed fee/period = $5.
    - Bonds:
      - Probability of value up: 0.5, down: 0.5
      - Return if bond value is going up: 7%; if down: 3%
      - Fixed fee/period = $2
    - Bank:
      - Guaranteed return of 3%/period, NO fee.

Maximizing Expected Utility (MEU)

- **Theorem** (Ramsey, 1931; von Neumann and Morgenstern, 1944):
  Given preferences satisfying the constraints
  there exists a real-valued function U such that
  $U(A) \geq U(B) \Leftrightarrow A \succeq B$
  $U([p_1, S_1; \ldots ; p_n, S_n]) = \sum_i p_i U(S_i)$ ($\Rightarrow$**EU(A | E)**) : utility of a lottery
- **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

31

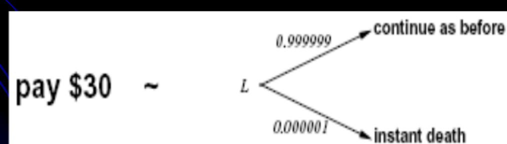
Utility Assessment

- Utilities map states to real numbers. Which numbers?
- *Preference elicitation* involves presenting choices to the agent and using the observed preferences to decide the underlying utility function.
- Standard approach to assessment of human utilities:
  - Compare a given state A to a standard lottery L_p $[p, u_\top ; (1-p), u_\perp]$ that has
    "best possible prize" $u_\top$ with probability p
    "worst possible catastrophe" $u_\perp$ with probability $(1-p)$
  - Adjust lottery probability p until $A \sim L_p$
- *Normalized utilities*: $U(S) = u_\top = 1.0$ (best possible prize),
  $U(S) = u_\perp = 0.0$ (worst possible catastrophe)

32

Utility scales

- **Micromorts**: one-millionth chance of death (i.e. micro probability of death)
  - A scale used in medical and safety analysis.
  - 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. linear transformation
  $U'(x) = k_1 U(x) + k_2$ where $k_1 > 0$
- With deterministic prizes only (no lottery choices), only ordinal utility can be determined, i.e. total order on prizes

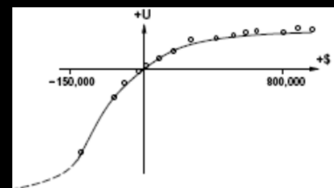


The Utility of Money

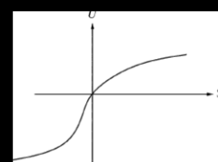
- In economics, a utility measure is money: an agent's total net asset.
- The agent exhibits *a monotonic preference* for more money; but money does not behave as a utility function – it says nothing a/b preferences b/t lotteries involving money.
- *certainty_equivalent(L)* of lottery: the value an agent accepts in lieu of a lottery
- *Insurance Premium(IP)* = *U(EMV(L)) – certainty_equivalent(L)*
  where *EMV* is the expected monetary value
- CE(L) = $1,000,000 while a lottery L = [½, 0; ½, $2,500,000].
  Then, U(EMV(L) = ½*0 + ½*2,500,000 = $1,250,000.
- Let S_n be the state of possessing total wealth $*n*, the current wealth is $*k*.
  - EU(Accept L) = ½*U(S_k) + ½*U($S_{k+2,500,000}$).
  - EU(Reject L, but get CE) = U($S_{k+1,000,000}$).
- To decide an action, we need to assign utility to the outcome state.
- Note that utility is not directly proportional to monetary value.
  $U(S_{n+k}) = -263.31 + 22.09 \log_2 (n+150{,}000)$ for $n \in [-\$150{,}000, \$800{,}000]$

The Utility of Money

- Utility curve: for what probability *p* am I indifferent between a prize *x* and a lottery $[p, \$M; (1-p), \$0]$ for large *M*?
  - proportional to the logarithm of the amount:
  - $U(S_n)=a \log_2 n+b$ where S_n is the state of having \$*n*.
  - e.g.) $U(S_{n+k}) = -263.31 + 22.09 \log_2 (n+150{,}000)$
  - for $n \in [-\$150{,}000, \$800{,}000]$



- It's likely/mostly to have a utility function that is concave for positive wealth, but preferences b/t different levels of debt can display a reversal of the concavity.
- In the positive part of curves where a slope is decreasing, for given a lottery L with expected monetary value *EMV(L)*, *u*sually
  - $U(L) < U(EMV(L))$, i.e. people are risk-averse: IP > 0
  - -- prefer a sure thing with a payoff < EMV(L)
- Typical empirical data, extrapolated with risk-seeking behavior.



35

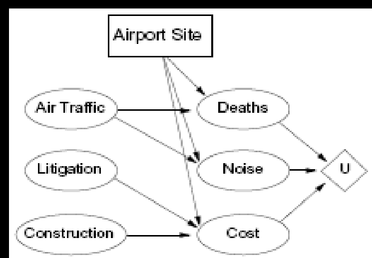
The Utility of Money

- *certainty_equivalent(L)* of lottery: the value an agent accepts in lieu of a lottery
- *Insurance Premium(IP)* = $U(EMV(L)) - certainty\_equivalent(L)$
  where *EMV* is the expected monetary value
- Given a lottery L with expected monetary value *EMV(L)*, usually $U(L) < U(EMV(L))$, i.e. people are risk-averse: IP > 0
- CE(L) = \$1,000,000 while a lottery L = [½, 0; ½, \$2,500,000].
  Then, U(EMV(L) = ½*0 + ½*2,500,000 = \$1,250,000.
- Let S_n be the state of possessing total wealth \$*n*, the current wealth is \$*k*.
  - EU(Accept L) = ½*U(S_k) + ½*U($S_{k+2,500,000}$).
  - EU(Reject L, but get CE) = U($S_{k+1,000,000}$).
- To decide an action, we need to assign utility to the outcome state.
- Note that utility is not directly proportional to monetary value.

36

Decision Networks

- DN 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.



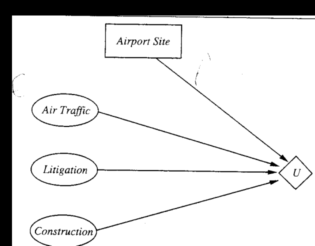
- Chance nodes (ovals): represent random variables.
  - Each node has associated with it a conditional distribution being indexed by the state of the parent nodes
  - Parent nodes can include decision nodes as well as chance nodes.
- *Decision nodes* (rectangles): where a decision maker has a choice of actions.
- *Utility nodes* (diamonds): a description of the agent's utility as a function of the parent attributes.

37

Decision networks

- A simplified DN with the chance node of current state, decision node and utility node.

  -- the omission of outcome states.



- Utility node: EU(A|E) – expected utility associated with each action.
- Algorithm:

  1. Set the evidence variables for the current state
  2. For each value of action node (i.e. decision node)

     Set the decision node to that value

     compute expected value of utility node given an action, evidence

     i.e. $EU(A|E) = \sum_i P(Result_i(A)|Do(A), E) \cdot U(Result_i(A))$ where $Result\_i(A)$ is outcome states and E is evidences.
  3. Return MEU action

38

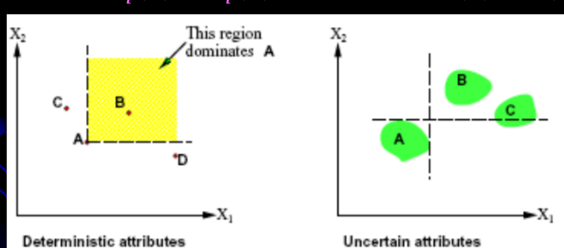
Multiattribute Utility

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

39

Strict dominance

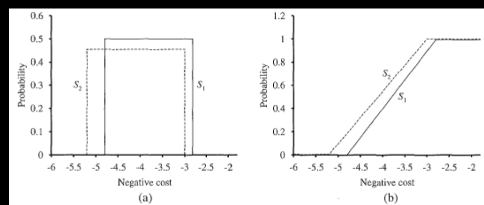
- Typically define attributes such that U is monotonic in each.
- **Strict dominance**: choice *B* strictly dominates choice *A* iff
  $\forall i \;\; X_i(B) \geq X_i(A)$ (and hence $U(B) \geq U(A)$)



- Strict dominance seldom holds in practice

40

Stochastic dominance

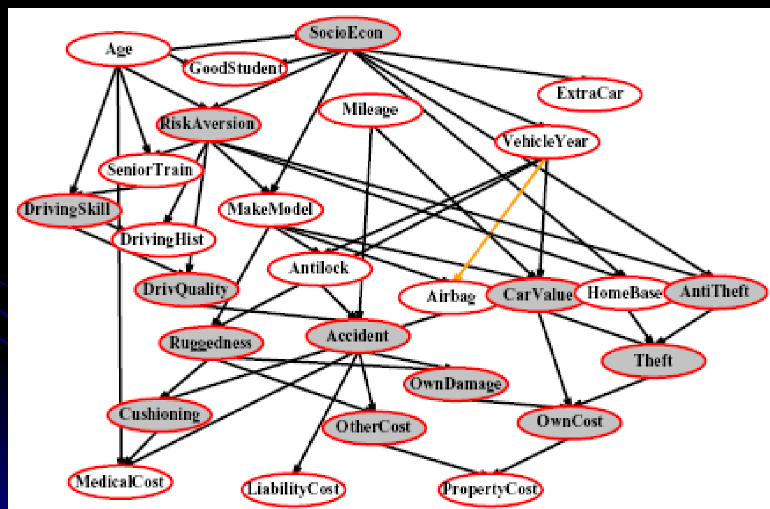


- Distribution p_1 stochastically dominates distribution p_2 iff
  $\forall t \;\; \int_{-\infty}^{t} p_1(x)dx \leq \int_{-\infty}^{t} p_2(t)dt$: S_1 is cheaper than S_2.
- If U is monotonic in x, then A_1 with outcome distribution p_1 stochastically dominates A_2 with outcome distribution p_2:
  $\int_{-\infty}^{\infty} p_2(x)U(x)dx \leq \int_{-\infty}^{\infty} p_1(x)U(x)dx$: Expected Utility (i.e. EU(A_2) ≤ EU(A_1))
- Thus, an action(A_2) which is stochastically dominated by another action(A_1) on all attributes can be *discarded.*

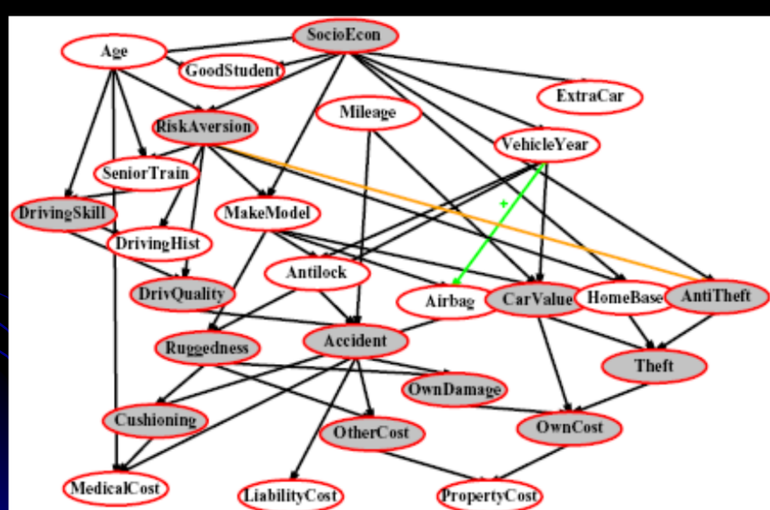
Stochastic dominance contd.

- Stochastic dominance can often be determined without exact distributions using qualitative reasoning.
- E.g., construction cost increases with distance from city
  S_1 is closer to the city than S_2
  $\Rightarrow S_1$ stochastically dominates S_2 on cost
- E.g., injury increases with collision speed
- Can annotate belief networks with stochastic dominance information:
  $X \xrightarrow{+} Y$ (X positively influences Y) means that
  For every value $\mathbf{z}$ of Y's other parents $\mathbf{Z}$
  $\forall\, x_1, x_2,\; x_1 \geq x_2 \Rightarrow \mathbf{P}(Y|x_1, \mathbf{z})$ stochastically dominates $\mathbf{P}(Y|x_2, \mathbf{z})$

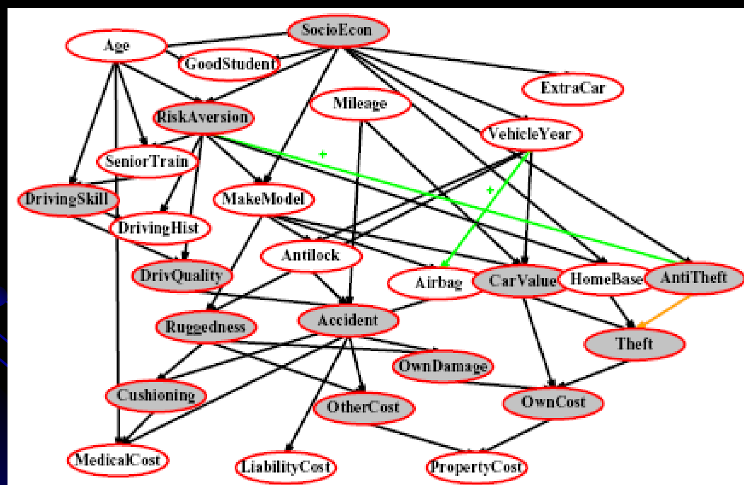
Label the arcs + and -



Label the arcs + and -

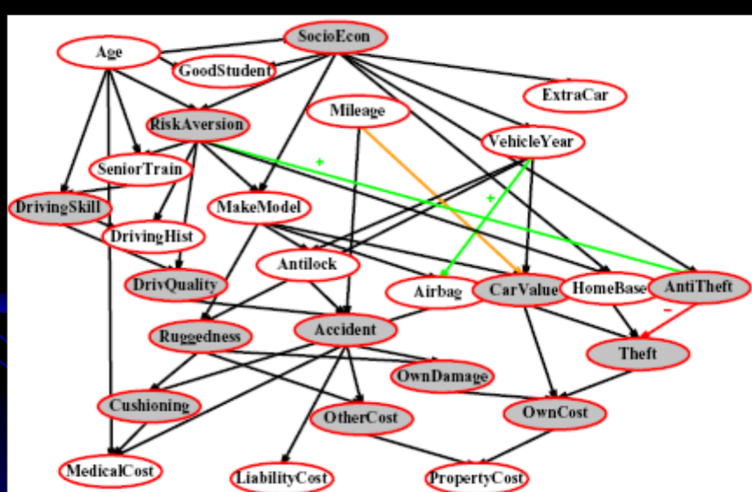


Label the arcs + and -



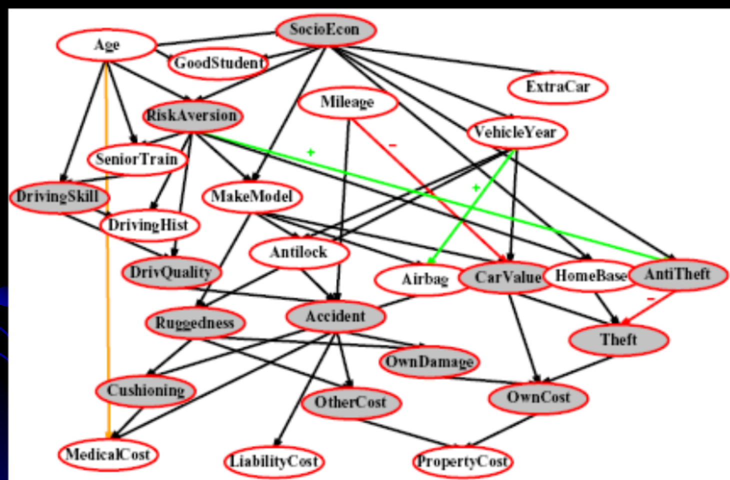
45

Label the arcs + and -



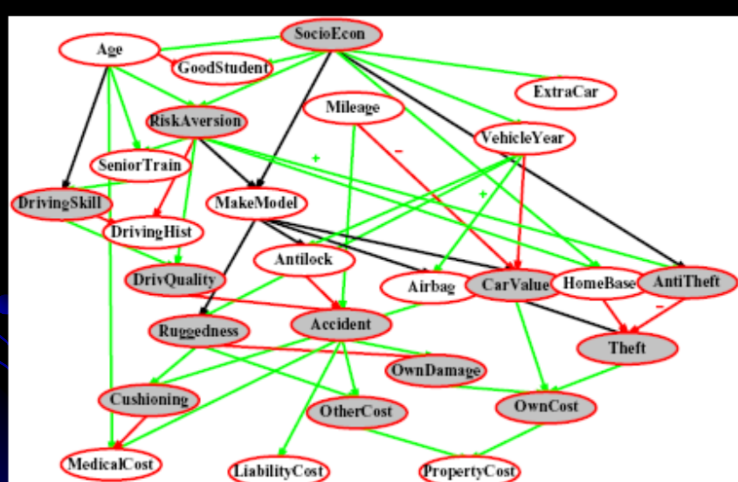
46

Label the arcs + and -



47

Label the arcs + and -



48

Preference structure: Deterministic

- Preferences of typical agents have much more structure than no regularity.
- Identify regularities in the preference behavior and use a representation theorem to show that an agent with a certain kind of preference structure has a utility function

  $U(x_1, \dots x_n) = F[f_1(x_1), \dots f_n(x_n)]$ where F is a simple function.

- X_1 and X_2 ***preferentially independent*** of X_3 iff

  preference between $<x_1, x_2, x_3>$ and $<x'_1, x'_2, x_3>$

  does not depend on x_3

- E.g., *<Noise, Cost, Death>*:

  <20,000 suffer, $4.6 billion, 0.06 deaths/mpm> vs.

  <70,000 suffer, $4.2 billion, 0.06 deaths/mpm>

49

Preference structure: Deterministic

- **Theorem** (Leontief, 1947):

  if every pair of attributes is P.I. of its complement, then every subset of attributes is P.I of its complement: ***mutual P.I***.

While each attribute may be important, it doesn't affect the way in which one trades off the other attributes against each other.

- **Theorem** (Debreu, 1960):

  If each X_i is mutual P.I.

  $\Rightarrow$ $\exists$ ***additive*** value function for the agent's preference behavior:

  $V(S) = \sum_i V_i(X_i(S))$ where V_i is a *value f^n* referring only to X_i.

  where $S = (X_1, X_2, \dots, X_n)$, V is an additive value f^n .

  Hence, assess *n* single-attribute functions; often a good approximation.

  e.g.) $V(noise, cost, deaths) = -noise \times 10^4 - cost - death \times 10^{12}$

Preference structure: Stochastic

- Need to consider preferences over lotteries:
  X is ***utility-independent*** of Y iff
  preference over lotteries in X do not depend on y of Y.
- Mutual U.I.: each subset is U.I. of its complement
  $\Rightarrow \exists$ ***multiplicative*** utility function:
  $$U = k_1U_1 + k_2U_2 + k_3U_3$$
  $$+ k_1k_2U_1U_2 + k_2k_3U_2U_3 + k_3k_1U_3U_1$$
  $$+ k_1k_2k_3U_1U_2U_3 \quad \text{where } U_i = U_i(x_i)$$
- An n-attribute problem exhibiting MUI can be modeled using n single-attribute utilities and n constants; each of single-attribute utility function can be developed independently of other attributes, then combine them for the overall preferences.

51

Value of Information

- Information Value Theory:
  Compute value of acquiring each possible piece of evidence by sensing actions.
  -- can be done directly from decision network
- With the information, one's course of action can be changed to suit the actual situation. Cf) w/o information, one has to do what's best on average over the possible situation.
- 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. Fair price?
- Solution: compute *expected value of information*
  = expected value of best action given the information
  – expected value of best action without information
- Survey may say "oil in A" or "no oil in A", prob. 0.5 each (given!)
  = [$0.5 \times$ value of "buy A" given "oil in A"
  + $0.5 \times$ value of "buy A" given "no oil in A"] - 0
  = $(0.5 \times k/2) + (0.5 \times k/2) - 0 = k/2$

52

Value of Information: Example

- Suppose an oil company is hoping to buy one of n indistinguishable blocks of ocean-drilling rights. Assume that exactly one of the blocks contains oil worth \$C, while the others are worthless. The asking price of each block is \$$C/n$. If the company is risk-neutral, then it will be indifferent between buying a block and not buying one.
- Suppose that a seismologist offers the company the results of a survey of block #3, which indicates definitively whether the block contains oil.
- How much should the company be willing to pay for the information?
- With p=1/n, the survey will indicate oil in block #3.
  → the company will buy block #3 for \$$C/n$ and make a profit of \$$C$ – \$$C/n$ = \$$(n-1)C/n$.
- With p=$(n-1)/n$, the survey will show that the block contains no oil → the company will buy a different block.
  So, p(finding oil in one of the other blocks) = $1/(n-1)$ from $1/n$,
  the expected profit is \$$C/(n-1)$ – \$$C/n$ = \$ $C/n(n-1)$.
- So, the expected profit, given the survey information is:

$$\frac{1}{n} \times \frac{(n-1)C}{n} + \frac{n-1}{n} \times \frac{C}{n(n-1)} = C/n\,.$$

- Thus, the company should be willing to pay the seismologist upto \$ C/n for the information. → the information is worth as much as the block itself.

General Formula: Value of Inf.

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

  $EU(\alpha \mid E) = max_a \sum_i U(S_i) \cdot P(S_i \mid a, E)$ where $S_i = Result_i(a)$

  : the value of the current best action α.
- Suppose we knew $E_j = e_{jk}$, then we would choose the new best action $\alpha_{e_{jk}}$, s.t.

  $EU(\alpha_{e_{jk}} \mid E, E_j = e_{jk}) = max_a \sum_i U(S_i) \cdot P(S_i \mid a, E, E_j = e_{jk})$
- E_j is a random variable whose value is *currently* unknown

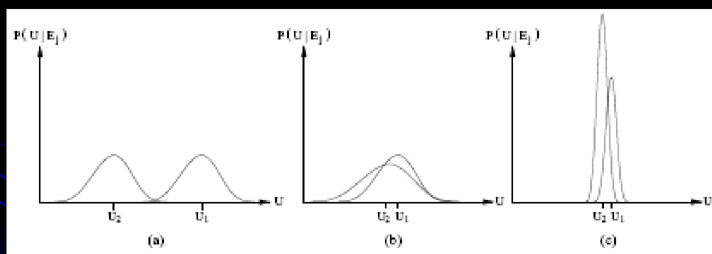
  $\Rightarrow$ must compute expected gain over all possible values e_{jk} :

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

  (VPI = *Value of Perfect Information*)

Qualitative behaviors

1. Choice is obvious, information worth little
2. Choice is nonobvious, information worth a lot
3. Choice is nonobvious, information worth little



Information has value to the extent that it is likely to cause a change of plan
and
to the extent that the new plan will be significantly better than the old plan.

55

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**
  $VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_{E,Ej}(E_k)$
  $= VPI_E(E_k) + VPI_{E,Ek}(E_j)$
- 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*

56

Implementing an Information-gathering Agent

- Gather information prior to acting.
- With each observable evidence E_j with its associated cost Cost(E_j), request the most efficient evidence in terms of utility gain per unit cost.

function INFORMATION-GATHERING-AGENT(*percept*) **returns** an *action*
  persistent: D, a decision network

  integrate *percept* into D
  $j \leftarrow$ the value that maximizes $VPI(E_j)\ /\ Cost(E_j)$
  if $VPI(E_j) > Cost(E_j)$
    return REQUEST(E_j)
  else return the best action from D

Figure 16.9 Design of a simple information-gathering agent. The agent works by repeatedly selecting the observation with the highest information value, until the cost of the next observation is greater than its expected benefit.