Intelligent Autonomous Agents

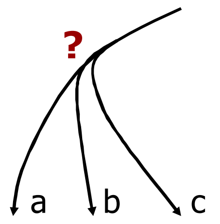
and Cognitive Robotics

Topic 7: Decision-Making under Uncertainty
Simple Decisions

Ralf Möller, Rainer Marrone
Hamburg University of Technology

1

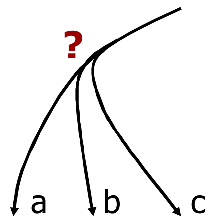
Non-Deterministic vs. Probabilistic Uncertainty



{a,b,c}

→ decision that is best for worst case

Non-deterministic model



$\{a(p_a),b(p_b),c(p_c)\}$

→ decision that maximizes expected utility value

Probabilistic model

~ Adversarial search

2

Expected Utility

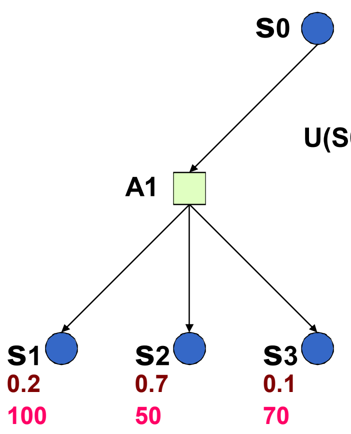
- Random variable *X* with *n* values $x_1,\ldots,x_n$ and distribution $(p_1,\ldots,p_n)$
  X is the state reached after doing an action **A** under uncertainty
- Function **U** of **X** : **U** is the utility of a state
- The expected utility of **A** is

$$EU[A] = \Sigma_{i=1,\ldots,n}\, p(x_i|A)U(x_i)$$

$$MEU = \underset{A}{\text{argmax}}\; EU[A]$$

3

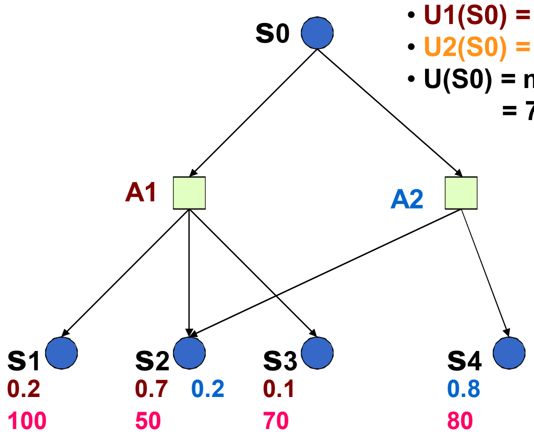
One State/One Action Example



U(S0) = 100 x 0.2 + 50 x 0.7 + 70 x 0.1
= 20 + 35 + 7
= 62

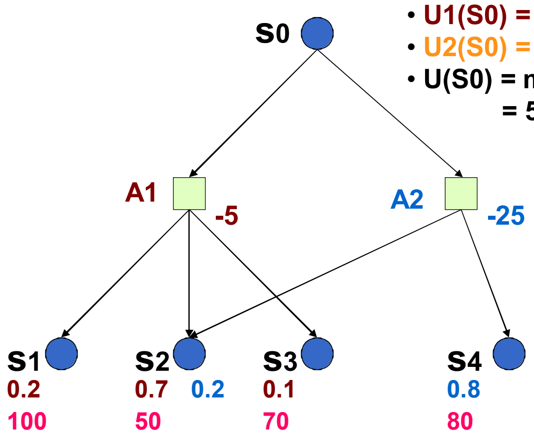
4

One State/Two Actions Example



- U1(S0) = 62
- U2(S0) = 0.2x50+0.8*80 = 74
- U(S0) = max{U1(S0),U2(S0)}
  = 74

Introducing Action Costs



- U1(S0) = 62 – 5 = 57
- U2(S0) = 74 – 25 = 49
- U(S0) = max{U1(S0),U2(S0)}
  = 57

MEU Principle

- A **rational agent** should choose the action that maximizes agent's expected utility
- This is the basis of the field of **decision theory**
- The MEU principle provides a **normative criterion** for rational choice of action

7

But …

- Must have **complete** model of:
  - Actions
  - Utilities
  - States
- Even if you have a complete model, it might be computationally **intractable**
- In fact, a truly rational agent takes into account the utility of reasoning as well---**bounded rationality**
- Nevertheless, great progress has been made in this area recently, and we are able to solve much more complex decision-theoretic problems than ever before

8

We’ll look at

- Decision-Theoretic Planning
  - Simple decision making (ch. 16)
  - Sequential decision making (ch. 17)

9

Rational preferences

Idea: preferences of a rational agent must obey constraints.
Rational preferences →
behavior describable as maximization of expected utility

MEU is not the only possible solution:
minimize worst case
only preferences without numeric values
…

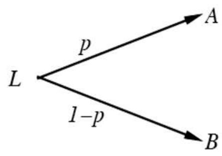
Why should a utility function with numerical values exist?

10

Basis of utility theory: constrains on preferences

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

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



$A \succ B$ the agent prefers A over B.

$A \sim B$ the agent is indifferent between A and B.

A and B can be lotteries again: Prizes are special lotteries: [1, X; 0, not X]

11

Axioms of Utility Theory

- **Orderability:** Given any two states, the rational agent prefers one of them, else the two as equally preferable.

$$(A \succ B) \vee (B \succ A) \vee (A \sim B)$$

- **Transitivity:** Given any three states, if an agent prefers A to B and prefers B to C, agent must prefer A to C.

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

- **Continuity:** If some state B is between A and C in preference, then there is a p for which the rational agent will be indifferent between state B and the lottery in which A comes with probability p, C with probability (1-p).

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

12

Rational preferences contd.

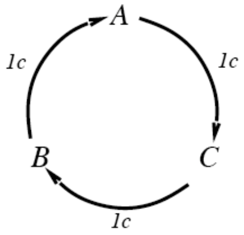
Violating the constraints leads to self-evident irrationality

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

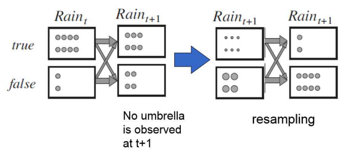
If $C \succ A$, then an agent who has A would pay (say) 1 cent to get C



13

Last time

- Kalman filters
- Failure models for DBN: transient, persistent
- Approximate inference in DBNs: Particle filtering



- Utility theory
  - Lotteries and axioms for preferences

14

Axioms of Utility Theory

- **Substitutability:** If an agent is indifferent between two lotteries, A and B, then there is a more complex lottery in which A can be substituted with B.

This also holds for $\succ$

$$(A \sim B) \Rightarrow [p : A; (1-p) : C] \sim [p : B; (1-p) : C]$$

- **Monotonicity:** If an agent prefers A to B, then the agent must prefer the lottery in which A occurs with a higher probability

$$(A \succ B) \Rightarrow (p > q \Leftrightarrow [p : A; (1-p) : B] \succ [q : A; (1-q) : B])$$

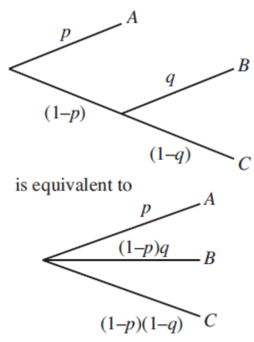
- **Decomposability:** Compound lotteries can be reduced to simpler lotteries using the laws of probability.

$$[p : A; (1-p) : [q : B; (1-q) : C]] \Rightarrow$$
$$[p : A; (1-p)q : B; (1-p)(1-q) : C]$$

No fun in gambling

15

Decomposabilty



16

And then there was utility

- Theorem by Neumann and Morgenstern, 1944
  Given preferences satisfying the constraints there exists a real-valued function *U* such that

$$U(A) \geq U(B) \;\Leftrightarrow\; A \succsim B$$
$$U([p_1, S_1;\ \ldots\ ;\ p_n, S_n]) = \Sigma_i\, p_i U(S_i)$$

MEU principle:
Choose the action that maximizes expected utility

Allais Paradox

A : 80% chance of \$4000
B : 100% chance of \$3000

When presented with a choice between A and B, most people would choose the sure thing B.

C : 20% chance of \$4000
D : 25% chance of \$3000

When presented with a choice between C and D, most people would choose the C, with higher expected utility (800 vs. 750).

These choices together are inconsistent

1*U(3000) > 0.8*U(4000)

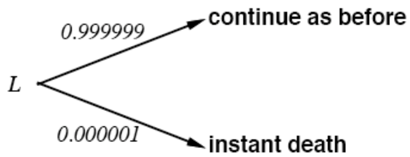
0.25*U(3000) < 0.2*U(4000)
1*U(3000) < 0.8*U(4000)

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" $u_\perp$ with probability $(1-p)$
adjust lottery probability p until $A \sim L_p$

pay \$30 ~ -and-continue -as-before



19

Utility scales

Normalized utilities: $u_\top = 1.0$, $u_\perp = 0.0$ U(pay \$30...) = 0.999999

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. +ve 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

20

Value Functions

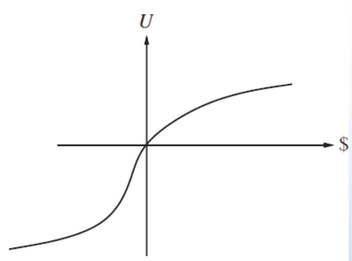
- Provides a ranking of alternatives, but not a meaningful metric scale
- Also known as an "ordinal utility function"

- Remember the expectiminimax example:
  - Sometimes, only relative judgments (value functions) are necessary
  - At other times, absolute judgments (utility functions) are required

21

Money Versus Utility

- Money <> Utility
  - More money is better, but not always in a linear relationship to the amount of money
- Expected Monetary Value
- Risk-averse – $U(L) < U(S_{EMV(L)})$
- Risk-seeking – $U(L) > U(S_{EMV(L)})$
- Risk-neutral – $U(L) = U(S_{EMV(L)})$



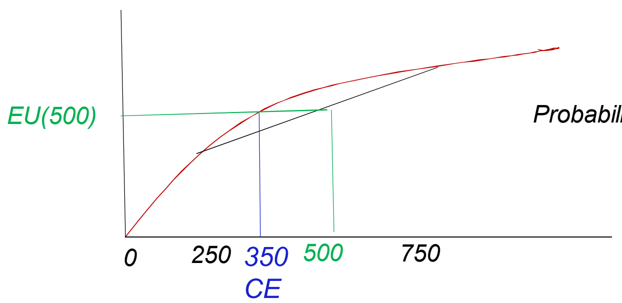
23

Two Concepts

- The ***certainty equivalent of a lottery:*** the sum of money, X, which, if received with certainty will yield the same utility as the gamble
  X is **CE** if $u(X) = EU = p_G \times u(c_G) + p_B \times u(c_B)$
- The ***risk premium associated with a lottery is*** the maximum amount a person is prepared to pay to avoid the gamble
  RP = EMV - CE

24

Risk averse



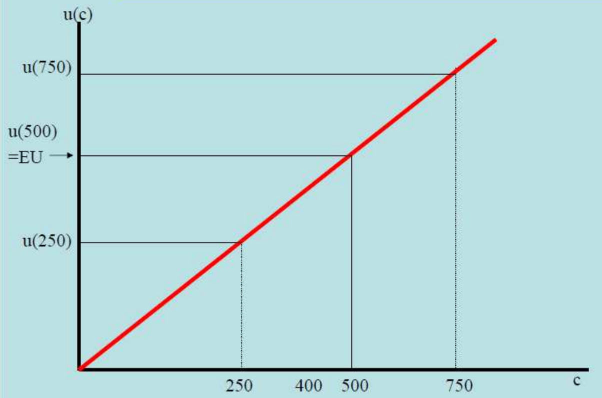
CE is the utility one get for sure when not choosing the lottery. In our case 350.
The RP is the money someone pays for not participating in the lottery and getting the sure thing.

The risk premium is 150=500-350.

25

Risk Neutral

A risk neutral person / with constant MU / with a linear utility function will be indifferent between accepting/rejecting a fair gamble

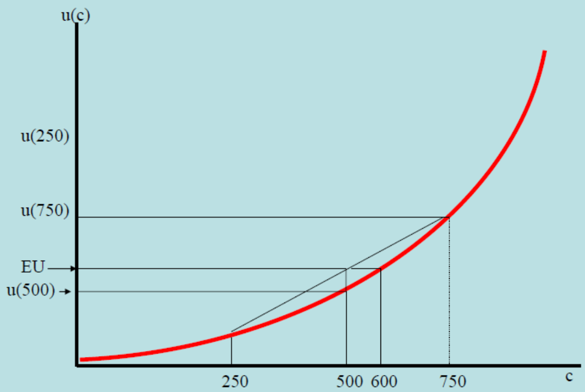


The certainty equivalent of the gamble is $500; the risk premium is $0

28

Risk Seeking

A risk loving person / with increasing MU / with a convex utility function will accept a fair gamble



The certainty equivalent of the gamble is $600; the risk premium is -$100

29

Multiattribute Utility Theory

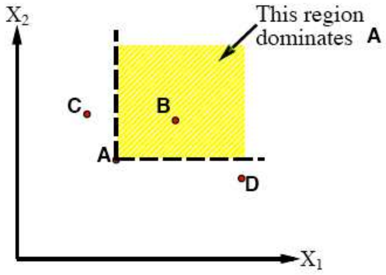
- A given state may have multiple utilities
  - ...because of multiple evaluation criteria
  - ...because of multiple agents (interested parties) with different utility functions

30

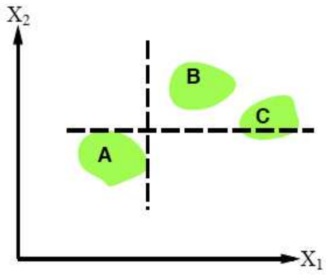
Strict dominance

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



Deterministic attributes



Uncertain attributes

31

Stochastic Dominance

- Introduced by Rothschild and Stiglitz (1970)
- When distribution F(.) yields unambiguously higher returns than G(.)?
  - When **every** expected utility maximizer (who values more money over less) prefers F(.) to G(.)
  - When for every amount of money *x* the probability of getting at least *x* is higher under F(.) than under G(.)
- Fortunately, these two definitions are equivalent

32

Stochastic dominance

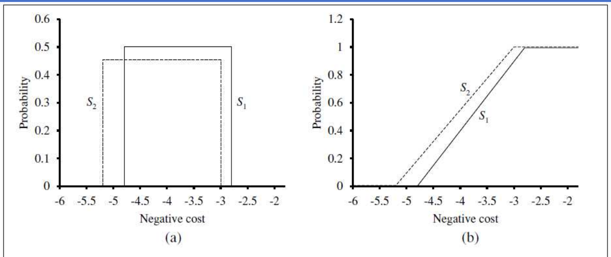


Figure 16.5 Stochastic dominance. (a) S_1 stochastically dominates S_2 on cost. (b) Cumulative distributions for the negative cost of S_1 and S_2.

If two actions S1 and S2 lead to probability distributions $p_1(x)$ and $p_2(x)$ on attribute X, then S1 stochastically dominates S2 on X if:

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

For any monotonically non-decreasing utility function U(x), the expected utility of S1 is at least as high as the expected utility of S2. Hence, S2 can be discarded.

33

Stochastic dominance contd.

Stochastic dominance can often be determined without exact distributions using **qualitative** reasoning

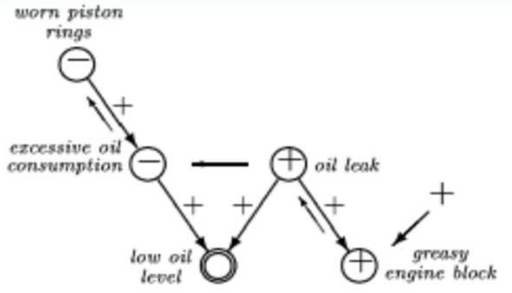
E.g., construction cost increases with distance from city
S_1 is closer to the city than S_2
$\Rightarrow$ S_1 stochastically dominates S_2 on cost

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 $\mathbf{P}(Y|x_2, \mathbf{z})$

36

Example



Qualitative influence of greasy engine block on worn piston rings:

Greasy engine block is evidence of oil leak.
Oil leak and excessive oil consumption can each cause low oil level.
Oil leak explains low oil level and so is evidence against excessive oil consumption.
Decreased likelihood of excessive oil consumption is evidence against worn piston rings.
Therefore, greasy engine block is evidence against worn piston rings.

37

Preference structure: Deterministic

X_1 and X_2 preferentially independent 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., $\langle Noise, Cost, Safety\rangle$:
$\langle$20,000 suffer, \$4.6 billion, 0.06 deaths/mpm$\rangle$ vs.
$\langle$70,000 suffer, \$4.2 billion, 0.06 deaths/mpm$\rangle$

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..

Theorem (Debreu, 1960): mutual P.I. $\Rightarrow$ $\exists$ additive value function:

$$V(S) = \Sigma_i V_i(X_i(S))$$

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

38

Multi-attribute utility functions

- Multi-dimensional or multi-attribute utility theory deals with expressing such utilities
- Example: you are made a set of job offers, how do you decide?
  u(job-offer) = *u*(salary) + *u*(location) +
  u(pension package) + *u*(career opportunities)

  u(job-offer) = 0.4*u*(salary) + 0.1*u*(location) +
  0.3*u*(pension package) + 0.2*u*(career opportunities)

But if there are interdependencies between attributes, then additive utility functions do not suffice. Multiplicative utility function:
$u(x,y)=w_x u(x)+w_y u(y)+w_x w_y u(x)u(y)$

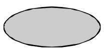
39

Decision Networks/ Influence diagrams

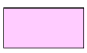
- Extend BNs to handle actions and utilities
- Also called *influence diagrams*
- Use BN inference methods
- Perform *Value of Information* calculations

40

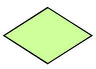
Decision Networks cont.



- Chance nodes: random variables, as in BNs: X={x1, …, xn}



- Decision nodes: actions that decision maker can take: A={a1, …, an}



- Utility function nodes: the utility of the outcome state: U(X,A)

41

Expected Utility in DN/ID

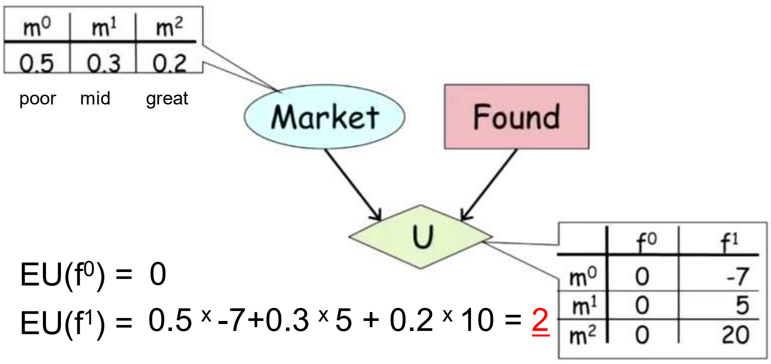
$$EU[D(a)] = \sum_{x} P(x|a)U(x,a)$$

- Want to choose action *a* that maximizes the expected utility

$$a^* = argmax_a EU[D(a)]$$

42

Simple example

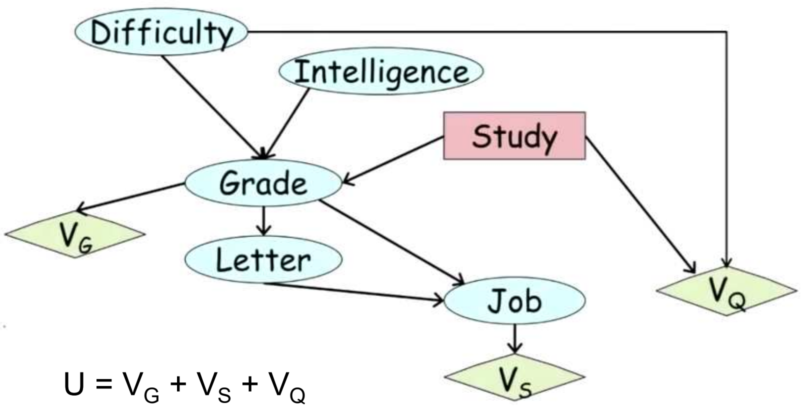


EU(f^0) = 0

EU(f^1) = 0.5 × -7+0.3 × 5 + 0.2 × 10 = 2

43

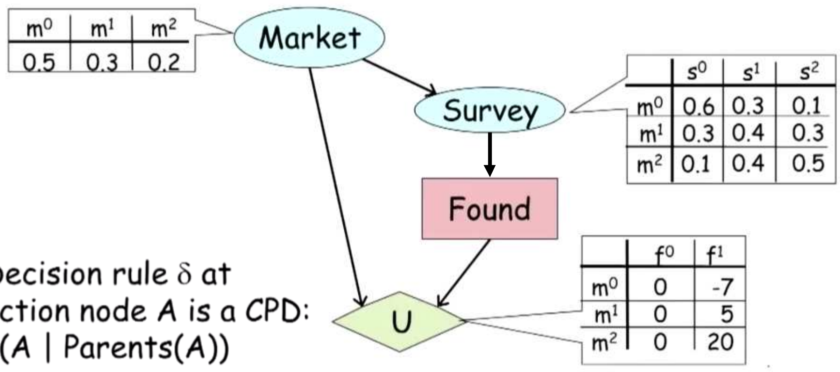
A more complex network



$U = V_G + V_S + V_Q$

44

Information edges



m^0	m^1	m^2
0.5	0.3	0.2

	s^0	s^1	s^2
m^0	0.6	0.3	0.1
m^1	0.3	0.4	0.3
m^2	0.1	0.4	0.5

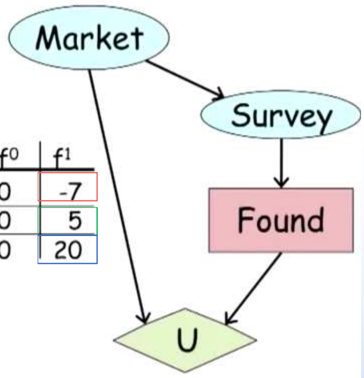
	f^0	f^1
m^0	0	-7
m^1	0	5
m^2	0	20

Decision rule δ at action node A is a CPD: $P(A \mid Parents(A))$

45

Finding MEU Decision rules

$$\sum_{S,F} \delta_F(F \mid S) \sum_M P(M)P(S \mid M)U(F,M)$$
$$= \sum_{S,F} \delta_F(F \mid S)\mu(F,S)$$



m^0	m^1	m^2
0.5	0.3	0.2

	s^0	s^1	s^2
m^0	0.6	0.3	0.1
m^1	0.3	0.4	0.3
m^2	0.1	0.4	0.5

	f^0	f^1
m^0	0	-7
m^1	0	5
m^2	0	20

*0.5*0.6*-7 = -2.1*
*0.3*0.3*5=0.45*
*0.2*0.1*20=0.4*

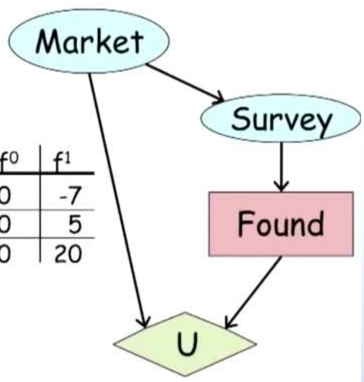
	f^0	f^1
s^0	0	
s^1	0	
s^2	0	

Summing leads to -1.25

46

Finding MEU Decision rules

$$\sum_{S,F} \delta_F(F \mid S) \sum_M P(M)P(S \mid M)U(F,M)$$
$$= \sum_{S,F} \delta_F(F \mid S)\mu(F,S)$$



m^0	m^1	m^2
0.5	0.3	0.2

	s^0	s^1	s^2
m^0	0.6	0.3	0.1
m^1	0.3	0.4	0.3
m^2	0.1	0.4	0.5

	f^0	f^1
m^0	0	-7
m^1	0	5
m^2	0	20

	f^0	f^1
s^0	$\boxed{0}$	-1.25
s^1	0	$\boxed{1.15}$
s^2	0	$\boxed{2.1}$

MEU= 0 + 1.15 + 2.1 = 3.25

47

More Generally

$$\mathrm{EU}[\mathcal{D}[\delta_A]] = \sum_{\boldsymbol{x},a} P_{\delta_A}(\boldsymbol{x},a)U(\boldsymbol{x},a)$$

$$\boldsymbol{Z} = \mathbf{Pa}_A$$

$$\boldsymbol{W} = \{X_1,\ldots,X_n\} - \boldsymbol{Z}$$

$$= \sum_{X_1,\ldots,X_n,A} \left(\left(\prod_i P(X_i \mid \mathbf{Pa}_{X_i})\right) U(\mathbf{Pa}_U)\delta_A(A \mid \boldsymbol{Z})\right)$$

$$= \sum_{\boldsymbol{Z},A} \delta_A(A \mid \boldsymbol{Z}) \sum_{\boldsymbol{W}} \left(\left(\prod_i P(X_i \mid \mathbf{Pa}_{X_i})\right) U(\mathbf{Pa}_U)\right)$$

$$= \sum_{\boldsymbol{Z},A} \delta_A(A \mid \boldsymbol{Z})\mu(A,\boldsymbol{Z})$$

$$\delta^*_A(a \mid \boldsymbol{z}) = \begin{cases} 1 & a = \mathrm{argmax}_A \mu(A,\boldsymbol{z}) \\ 0 & \text{otherwise} \end{cases}$$

48

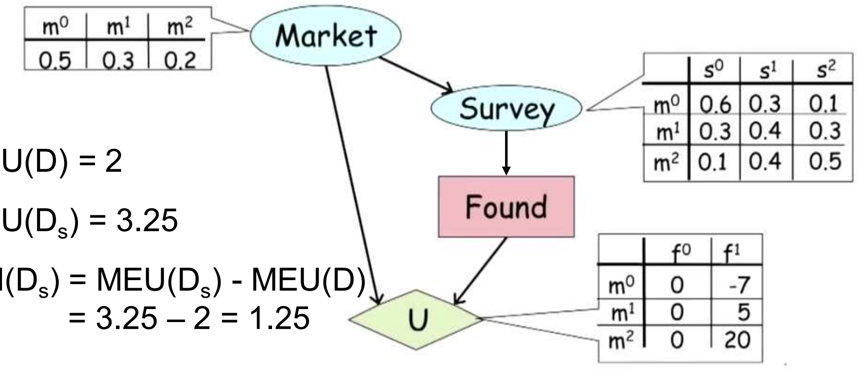
MEU Summary

- To compute MEU & optimize decision at A:
  - Treat A as random variable with arbitrary CPD
  - Introduce utility factor with scope Pa_U
  - Eliminate all variables except A, **Z** (A's parents) to produce factor μ(A, **Z**)
  - For each **z**, set:

$$\delta^*_A(a \mid \boldsymbol{z}) = \begin{cases} 1 & a = \mathrm{argmax}_A \mu(A,\boldsymbol{z}) \\ 0 & \text{otherwise} \end{cases}$$

49

Value of Perfect Information



MEU(D) = 2

MEU(D_s) = 3.25

VPI(D_s) = MEU(D_s) - MEU(D)
= 3.25 – 2 = 1.25

50

Value of Perfect Information

Current evidence **E**, current best action a
Possible actions outcomes $\mathbf{S_i}$, potential new evidence $\mathbf{E_j}$

$$MEU(a|E) = max_a \sum_i U(S_i)P(S_i|E,a)$$

Suppose we knew $\mathbf{E_j}$, we would choose $a_{\mathbf{ejk}}$

$$MEU\left(a_{e_{jk}}\middle|E, E_j = e_{jk}\right) = max_a \sum_i U(S_i)P(S_i|E, a, E_j = e_{jk})$$

$\mathbf{E_j}$ is not known. Must compute expected gain.

$$VPI(E_j) = \left(\sum_k P(E_j|E) MEU\left(a_{e_{jk}}\middle|E, E_j = e_{jk}\right)\right) - MEU(a|E)$$

51

Properties of VPI

Non negative

$$\forall j, E \; VPI_E(E_j) \geq 0$$

Non additive

$$VPI_E(E_j, E_k) \neq VPI_E(E_j) + VPI_E\,(E_k)$$

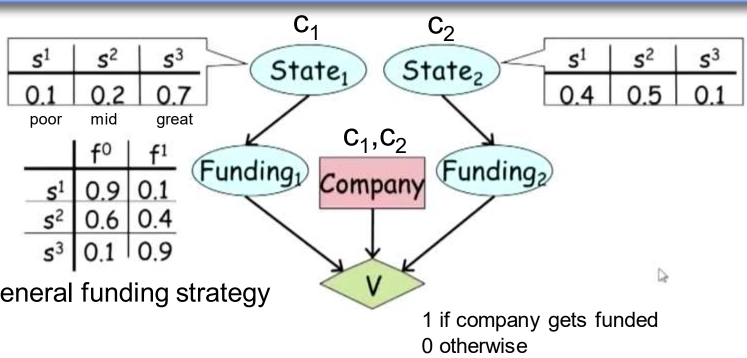
Order-independent

$$VPI_E(E_j, E_k) = VPI_E(E_j) + VPI_{E,E_j}(E_k) = VPI_E(E_k) + VPI_{E,E_k}(E_j)$$

When is information useful?

52

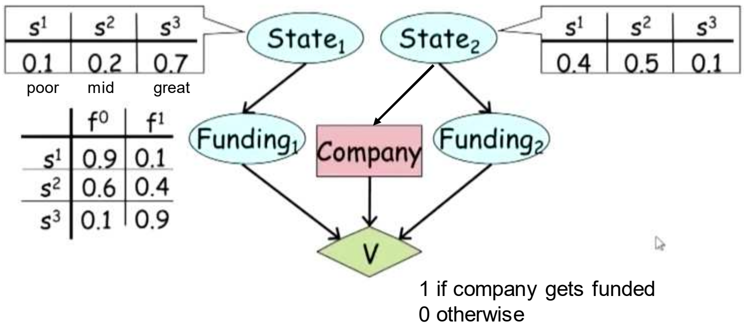
Example 1



EU(D[c_1]) = 0.1*0.1+0.2*0.4+0.7*0.9=0.72
EU(D[c_2]) = 0.4*0.1+0.5*0.4+0.1*0.9=0.33

53

Example 1

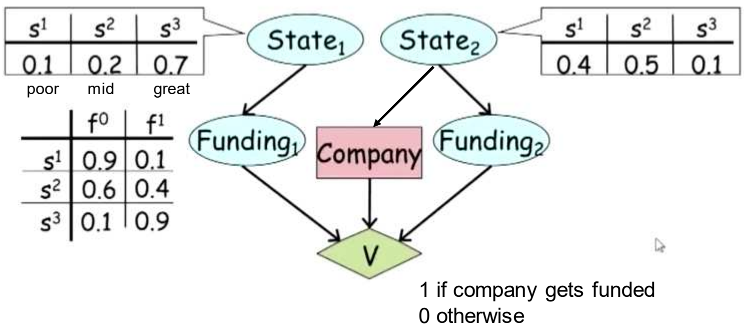


$EU(D[c_1]) = 0.72$
$EU(D[c_2]) = 0.33$

If c2 is in state s1, the utility is 0.1
If c2 is in state s2, the utility is 0.4
If c2 is in state s3, the utility is 0.9

54

Example 1



$MEU(D_{State2}) = 0.4 * 0.72 = 0.288$
$0.5 * 0.72 = 0.36$
$0.1 * 0.9 = 0.09$
$\sum 0.738$

$EU(D[c_1]) = 0.72$
$EU(D[c_2]) = 0.33$

55

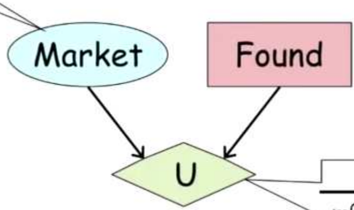
Last time

- Existence of a utility function
  - Additive vs multiplicative utility function
  - Stochastic dominance
- Risk profiles
  - Risk averse
  - Risk neutral
  - Risk seeking

56

Last time: Decision networks

m^0	m^1	m^2
0.5	0.3	0.2
poor	mid	great



	f^0	f^1
m^0	0	-7
m^1	0	5
m^2	0	20

$EU(f^0) = 0$

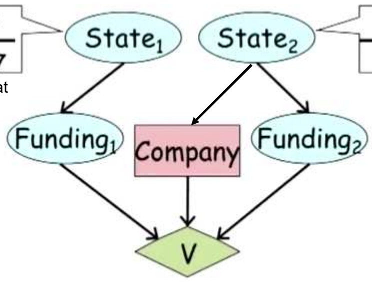
$EU(f^1) = 0.5 \times -7 + 0.3 \times 5 + 0.2 \times 10 = \underline{2}$

57

Example 1

s^1	s^2	s^3
0.1	0.2	0.7
poor	mid	great

	f^0	f^1
s^1	0.9	0.1
s^2	0.6	0.4
s^3	0.1	0.9



s^1	s^2	s^3
0.4	0.5	0.1

1 if company gets funded
0 otherwise
Select c_2 if $State_2 = s_3$, c_1 otherwise

$EU(D[c_1]) = 0.72$
$EU(D[c_2]) = 0.33$

$MEU(D_{State2}) = 0.738$

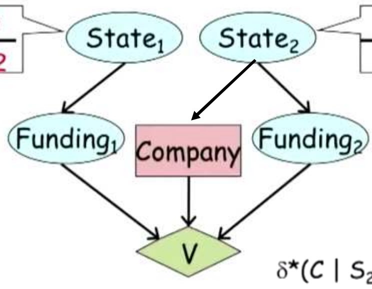
$VPI(D_{State2}) = 0.738 - 0.72 = 0.018$

58

Example 2

s^1	s^2	s^3
0.5	0.3	0.2

	f^0	f^1
s^1	0.9	0.1
s^2	0.6	0.4
s^3	0.1	0.9



s^1	s^2	s^3
0.4	0.5	0.1

$$\delta^*(C \mid S_2) = \begin{cases} P(c^2)=1 & \text{if } S_2 = s^2, s^3 \\ P(c^1)=1 & \text{otherwise} \end{cases}$$

$EU(\mathcal{D}[c_1]) = 0.35$

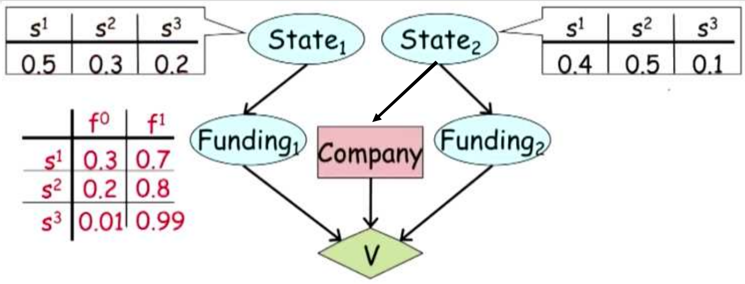
$EU(\mathcal{D}[c_2]) = 0.33$

$MEU(\mathcal{D}_{S_2 \to C}) = 0.43$

$VPI = 0.43 - 0.35 = 0.08$

59

Example 3



$EU(\mathcal{D}[c_1]) = 0.788$

$EU(\mathcal{D}[c_2]) = 0.779$

$$\delta^*(C \mid S_2) = \begin{cases} P(c^2)=1 & \text{if } S_2 = s^2, s^3 \\ P(c^1)=1 & \text{otherwise} \end{cases}$$

$MEU(\mathcal{D}_{S_2 \to C}) = 0.8142$

VPI = 0.8142 – 0.788 = 0.0262

60

Summary

- Influence diagrams provide clear and coherent semantics for the value of making an observation
  - VPI = P(new observation) * MEU(new observation) – MEU (with current observations)
- Information is valuable if and only if it induces a change in action in at least one context, and with (significant) higher MEU.

62