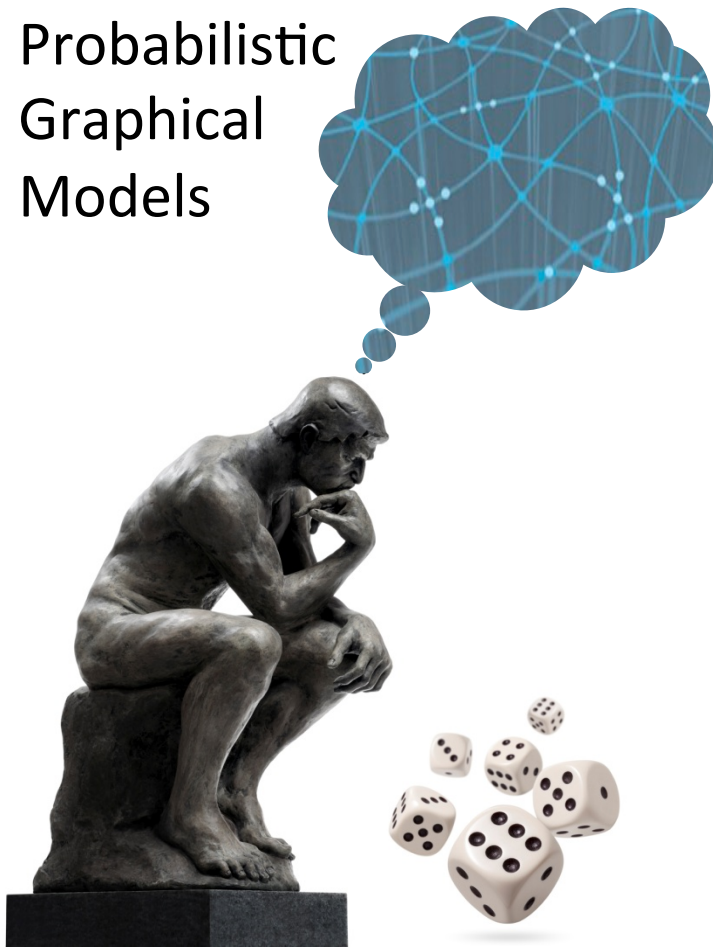


Probabilistic
Graphical
Models



Acting

Decision Making

Maximum
Expected
Utility

Simple Decision Making

A simple decision making situation \mathcal{D} :

- A set of possible actions $\text{Val}(\underline{A}) = \{\underline{a^1}, \dots, \underline{a^K}\}$
- A set of states $\text{Val}(\underline{X}) = \{\underline{x^1}, \dots, \underline{x^N}\}$
- A distribution $\underline{P(X \mid A)}$
- A utility function $\underline{U(X, A)}$

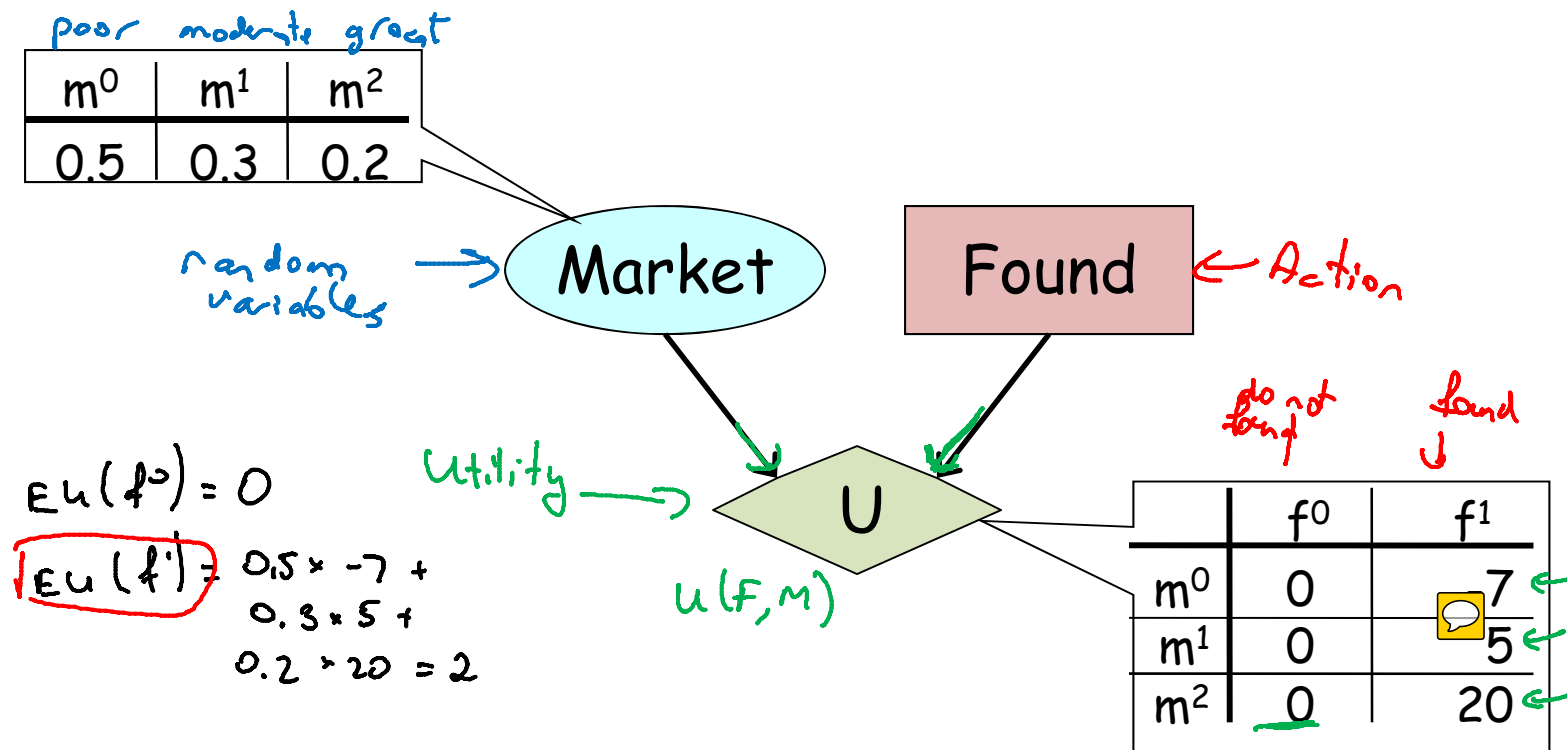
Expected Utility

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

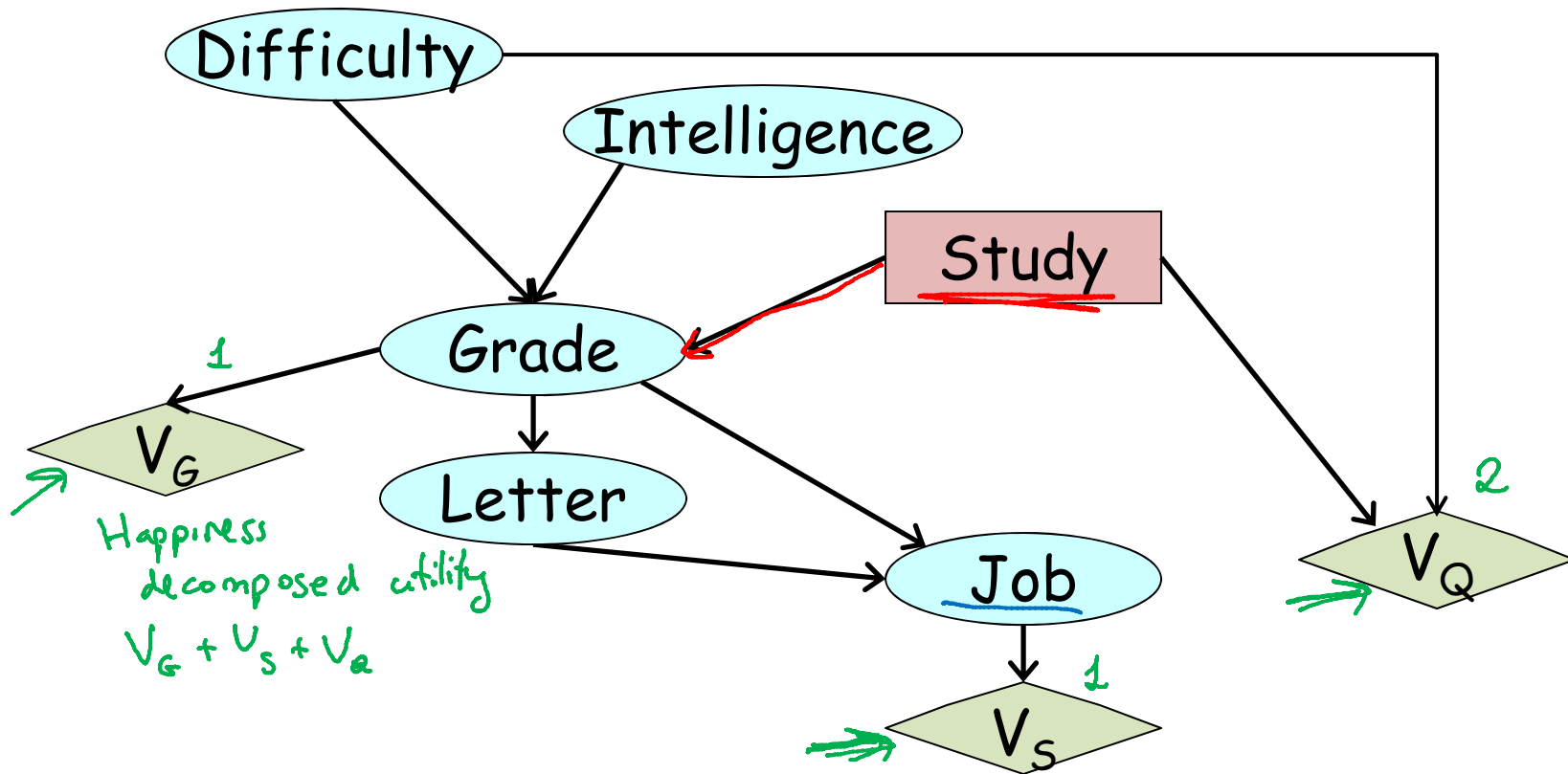
- Want to choose action \hat{a} that maximizes the expected utility *Max. expected utility*

$$a^* = \operatorname{argmax}_a EU[\mathcal{D}[a]]$$

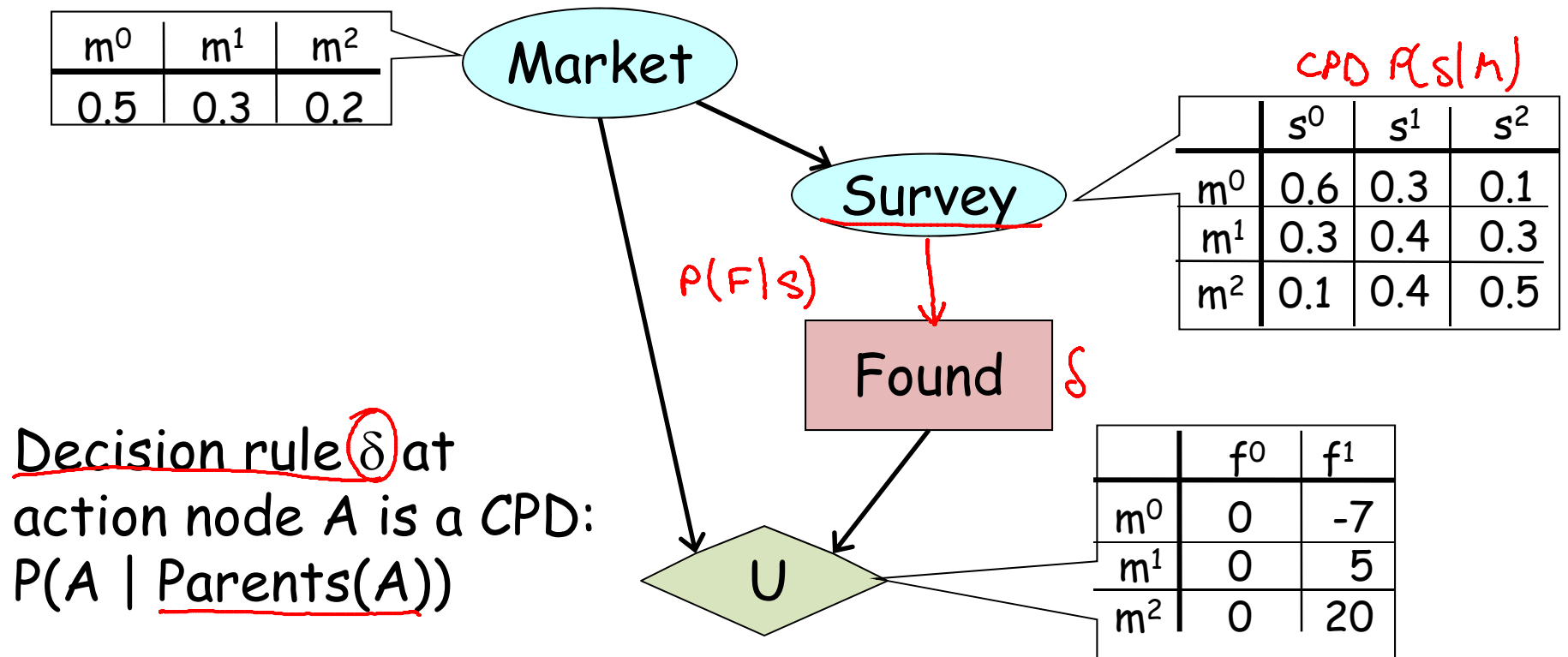
Simple Influence Diagram



More Complex Influence Diagram



Information Edges



Expected Utility with Information

$$\text{EU}[\mathcal{D}[\delta_A]] = \sum_{\mathbf{x}, a} \overbrace{P_{\delta_A}(\mathbf{x}, a)}^{\text{joint prob. dist. over } \overline{X} \cup \{A\}} \underbrace{U(\mathbf{x}, a)}$$

- Want to choose the decision rule δ_A that maximizes the expected utility

$$\operatorname{argmax}_{\delta_A} \text{EU}[\mathcal{D}[\delta_A]]$$

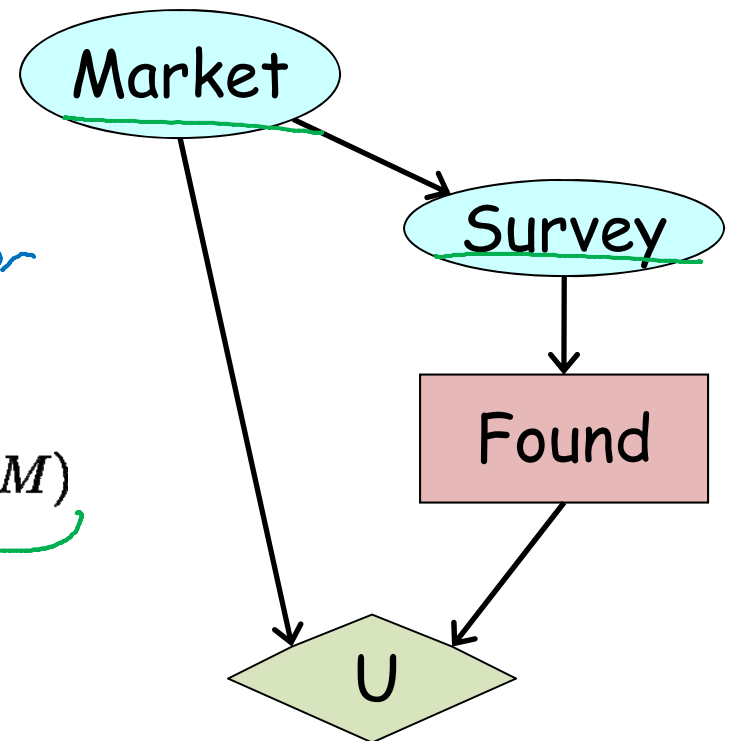
$$\text{MEU}(\mathcal{D}) = \max_{\delta_A} \text{EU}[\mathcal{D}[\delta_A]]$$

Finding MEU Decision Rules

$$\text{EU}[\mathcal{D}[\delta_A]] = \sum_{\mathbf{x}, a} P_{\delta_A}(\mathbf{x}, a) \underline{U(\mathbf{x}, a)}$$

optimize (pointing to $\mathcal{D}[\delta_A]$)

$$\begin{aligned} \sum_{M, S, F} \underline{P(M)} \underline{P(S | M)} \underline{\delta_F(F | S)} \underline{U(F, M)} &= \\ &= \sum_{S, F} \underline{\delta_F(F | S)} \underbrace{\sum_M P(M) P(S | M) U(F, M)}_{\text{factor}} \\ &= \sum_{S, F} \underline{\delta_F(F | S)} \underline{\mu(F, S)} \end{aligned}$$



Finding MEU Decision Rules

$$\sum_{S,F} \delta_F(F | S) \left[\sum_M P(M) P(S | M) U(F, M) \right]$$

$$= \sum_{S,F} \delta_F(F | S) \mu(F, S)$$

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

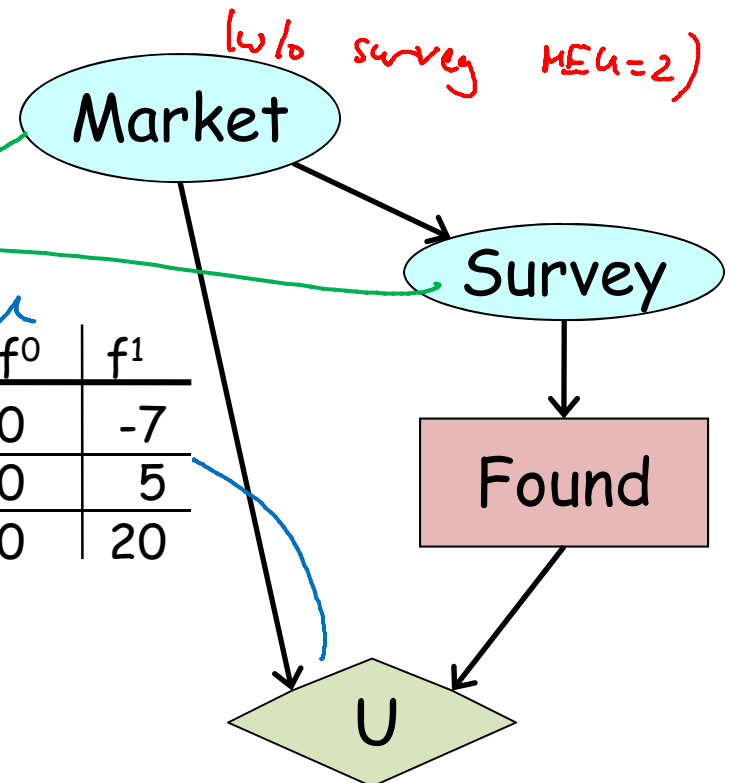
	s^0	s^1	s^2
m^0	0.6	0.3	0.1
m^1	0.3	0.4	0.3
m^2	0.1	0.4	0.5

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

	f^0	f^1
s^0	0	-1.25
s^1	0	1.15
s^2	0	2.1

$$\begin{array}{r} 0 \\ + 1.15 \\ + 2.1 \\ \hline 3.25 \end{array}$$

$s \rightarrow f$
 $s_1 \rightarrow f_1$
 $s_2 \rightarrow f_1$



More Generally

$$\begin{aligned}
 \text{EU}[\mathcal{D}[\delta_A]] &= \sum_{\mathbf{x}, a} \overbrace{P_{\delta_A}(\mathbf{x}, a)}^{\text{joint dist.}} \underbrace{U(\mathbf{x}, a)} \\
 &= \sum_{X_1, \dots, X_n, A} \left(\underbrace{\left(\prod_i P(X_i \mid \mathbf{Pa}_{X_i}) \right)}_{\text{prob}} U(\mathbf{Pa}_U) \underbrace{\delta_A(A \mid \mathbf{Z})}_{\text{prior to } A} \right) \\
 &= \sum_{\mathbf{Z}, A} \underbrace{\delta_A(A \mid \mathbf{Z})}_{\text{prior to } A} \sum_{\mathbf{W}} \underbrace{\left(\left(\prod_i P(X_i \mid \mathbf{Pa}_{X_i}) \right) U(\mathbf{Pa}_U) \right)}_{\text{prob}} \\
 &= \sum_{\mathbf{Z}, A} \delta_A(A \mid \mathbf{Z}) \underbrace{\mu(A, \mathbf{Z})}_{\text{prob}} \quad \delta_A^*(a \mid \mathbf{z}) = \begin{cases} 1 & a = \operatorname{argmax}_A \mu(A, \mathbf{z}) \\ 0 & \text{otherwise} \end{cases}
 \end{aligned}$$

$\mathbf{Z} = \mathbf{Pa}_A$ *observations prior to A*
 $\mathbf{W} = \{X_1, \dots, X_n\} - \mathbf{Z}$

MEU Algorithm Summary

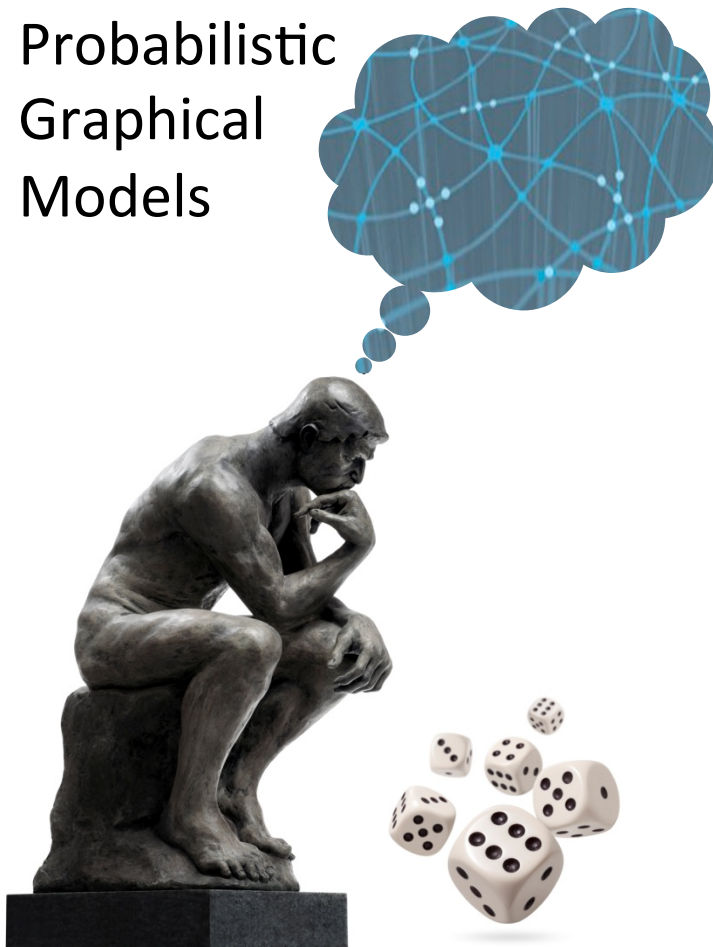
- To compute MEU & optimize decision at A :
 - Treat A as random variable with arbitrary CPD
 - Introduce utility factor with scope Pa_U
 - Eliminate all variables except A, Z (A 's parents) to produce factor $\mu(A, Z)$
 - For each \mathbf{z} , set:

$$\delta_A^*(a | \mathbf{z}) = \begin{cases} 1 & a = \operatorname{argmax}_A \mu(A, \mathbf{z}) \\ 0 & \text{otherwise} \end{cases}$$

Decision Making under Uncertainty

- MEU principle provides rigorous foundation
- PGMs provide structured representation for probabilities, actions, and utilities
- PGM inference methods (VE) can be used for
 - Finding the optimal strategy
 - Determining overall value of the decision situation
- Efficient methods also exist for:
 - Multiple utility components
 - Multiple decisions

Probabilistic
Graphical
Models



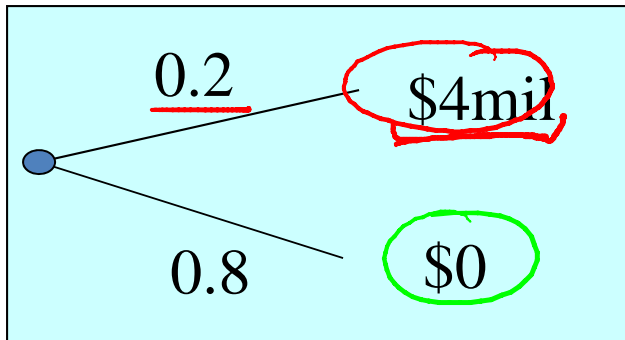
Acting

Decision Making

Utility
Functions

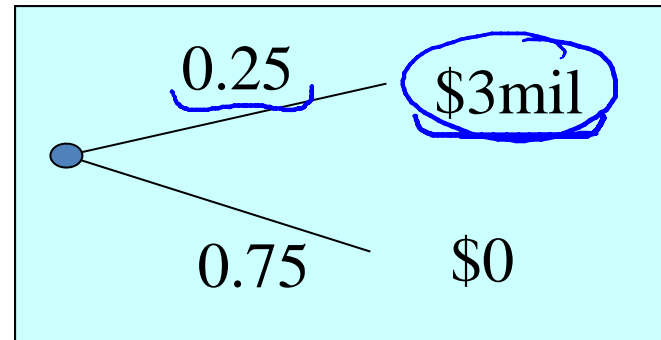
Utilities and Preferences

lotteries



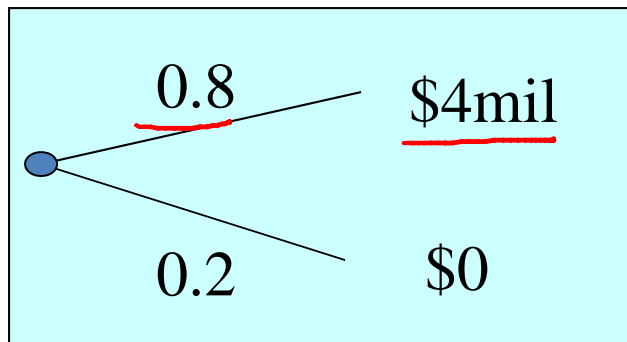
$$0.2 \times u(4) + 0.8 u(0)$$

\succsim



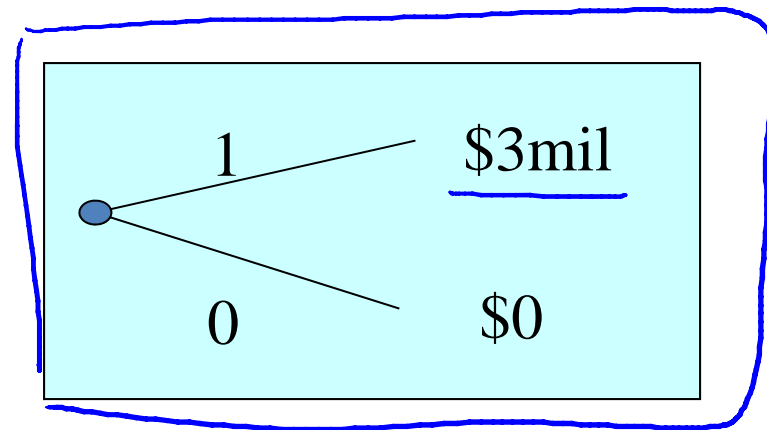
$$0.25 u(3) + 0.75 u(0)$$

Utility = Payoff?



$$\begin{aligned} &\$4\text{mil} \times 0.8 = \\ &\quad \underline{\underline{\$3.2\text{mil}}} \end{aligned}$$

\approx



$\$3\text{mil}$

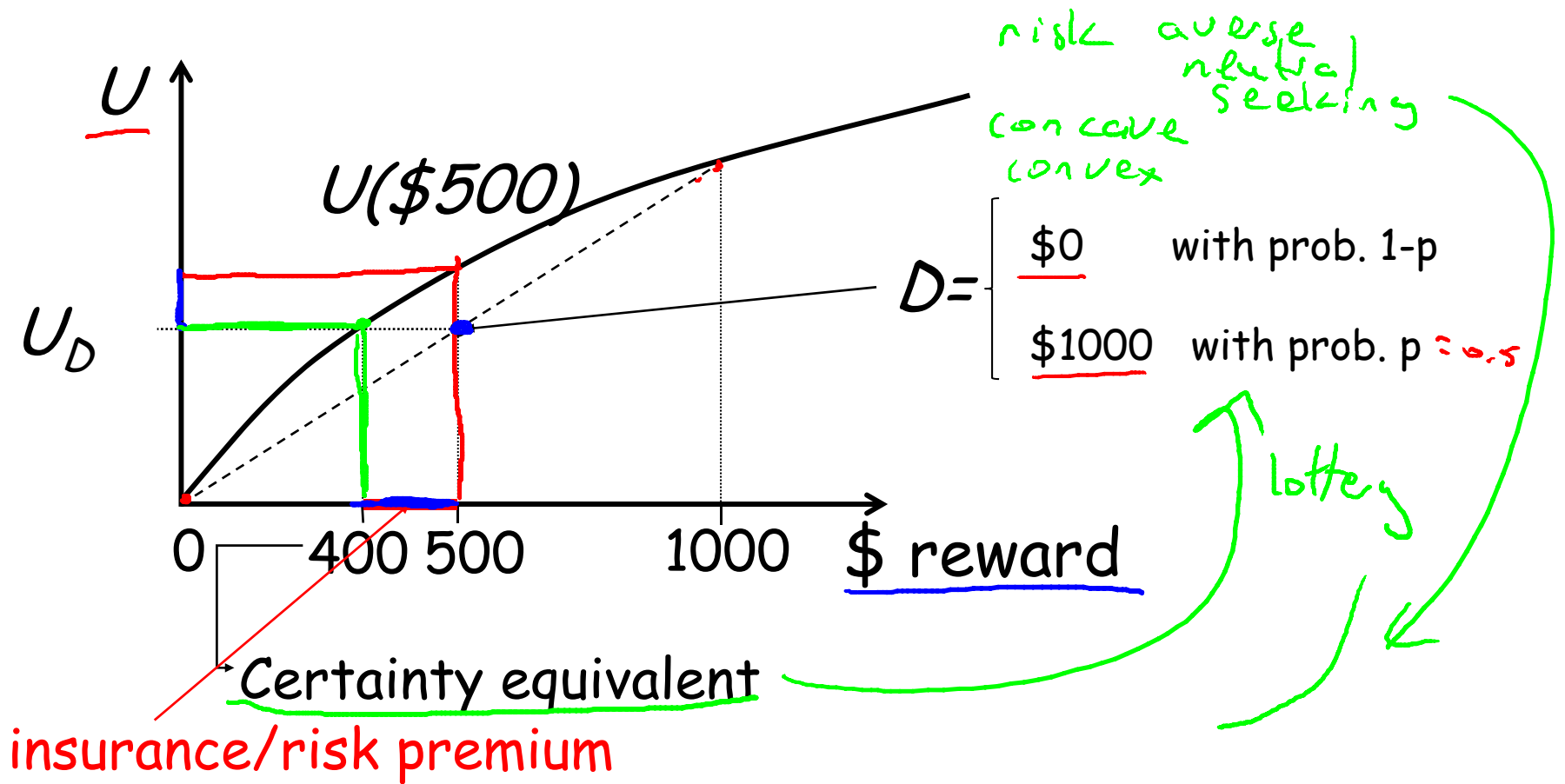
St. Petersburg Paradox



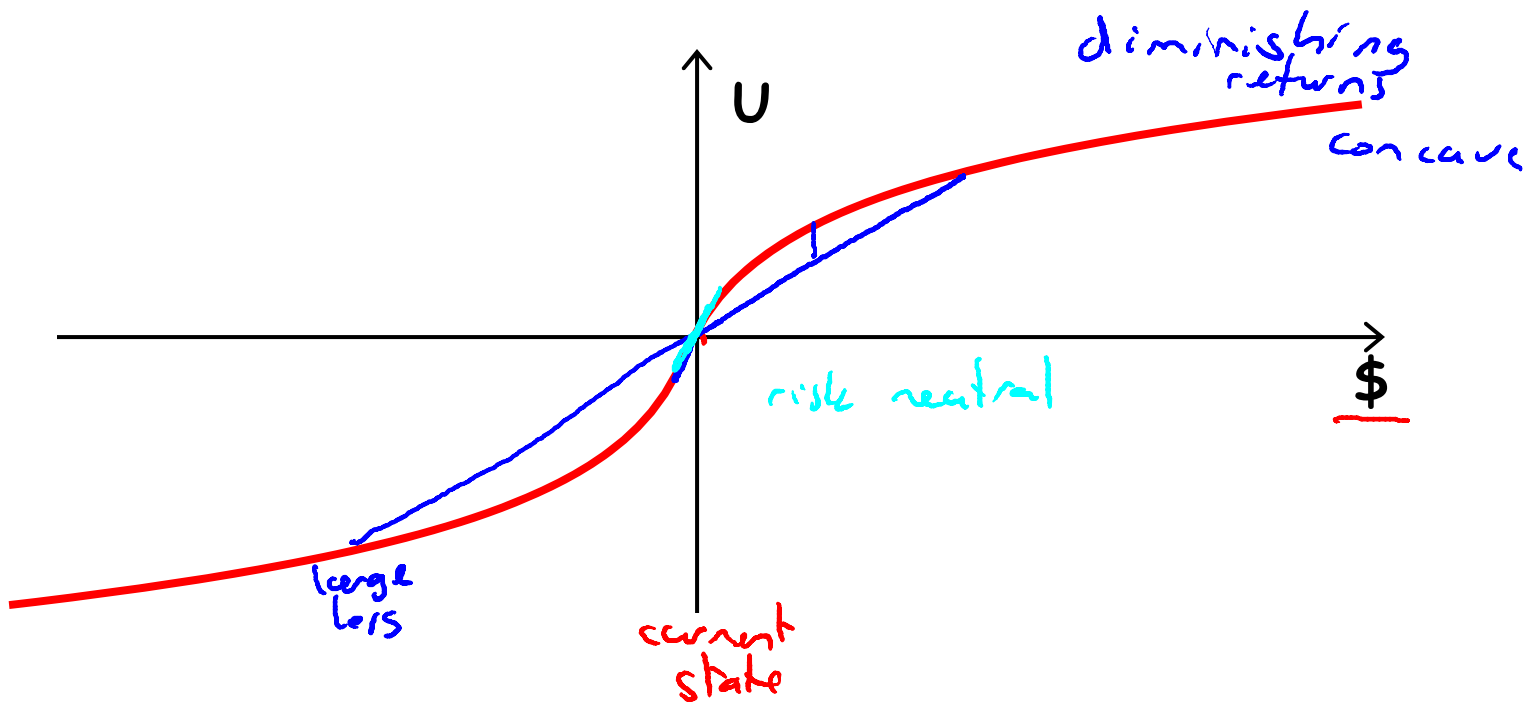
- Fair coin is tossed repeatedly until it comes up heads, say on the n^{th} toss
- Payoff = $\$2^n$

$$\frac{1}{2} \times 2 + \frac{1}{4} \times 4 + \frac{1}{8} \times 8 + \dots = \infty$$

most people value $\approx \$2$



Typical Utility Curve



Multi-Attribute Utilities

- All attributes affecting preferences must be integrated into one utility function

money, time, pleasure, ...

- Human life

- Micromorts *$1/1000000$ chance of death \approx \$20 1980*
- QALY (quality-adjusted life year)

Example: Prenatal diagnosis

$$\underline{U_1(T)} + \underline{U_2(K)} + \underline{U_3(D,L)} + \underline{U_4(L,F)}$$

Testing

Knowledge

Down's
syndrome

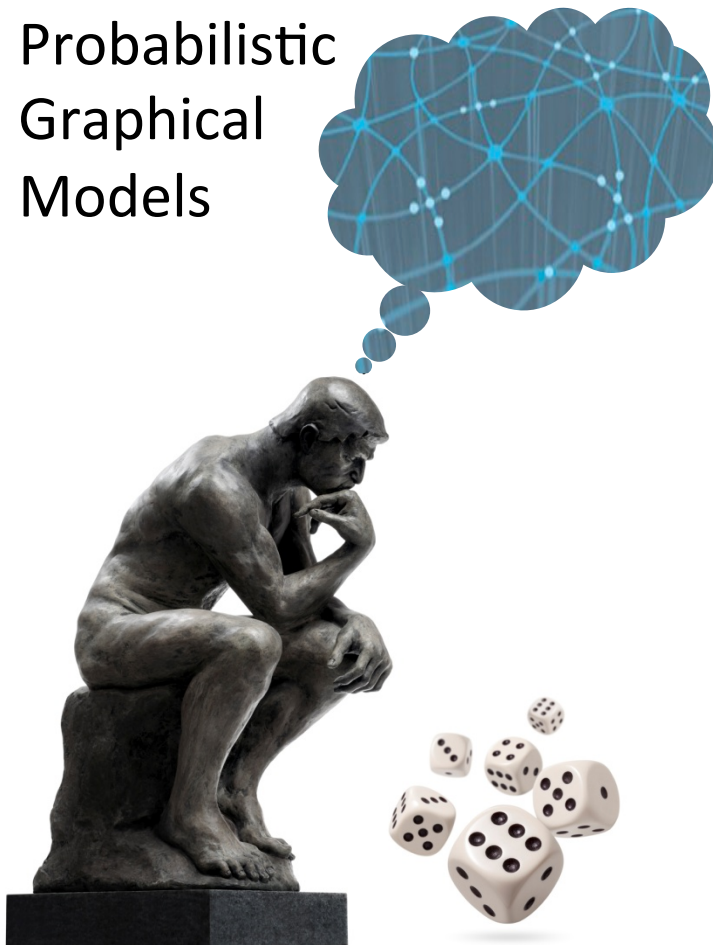
Loss of
fetus

Future
pregnancy

Summary

- Our utility function determines our preferences about decisions that involve uncertainty
- Utility generally depends on multiple factors
 - Money, time, chances of death, ...
- Relationship is usually nonlinear
 - Shape of utility curve determines attitude to risk
- Multi-attribute utilities can help decompose high-dimensional function into tractable pieces

Probabilistic
Graphical
Models



Acting

Decision Making

Value of
Perfect
Information

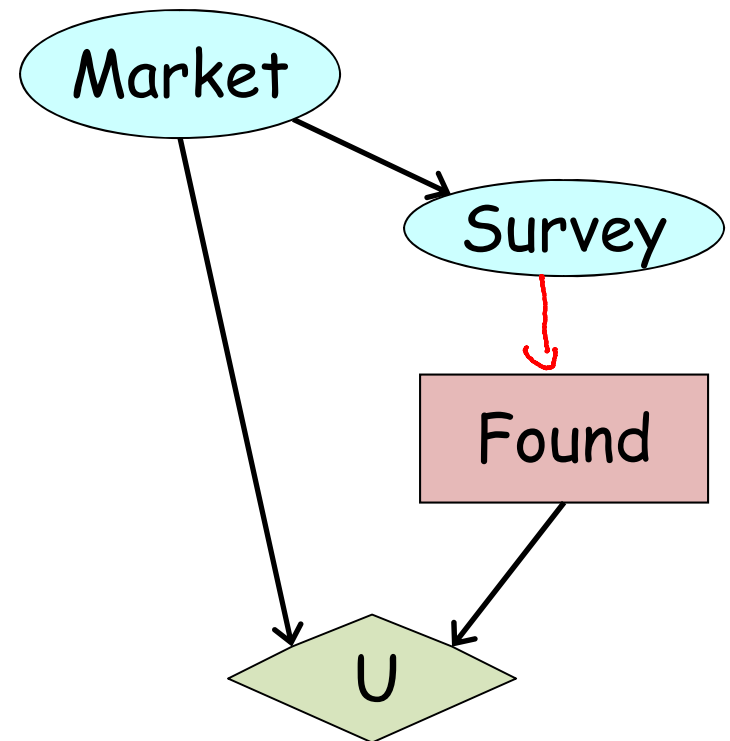
Value of Information

- ^{value of perfect information} VPI(A | X) is the value of observing X before choosing an action at A
- \mathcal{D} = original influence diagram
- $\mathcal{D}_{X \rightarrow A}$ = influence diagram with edge $X \rightarrow A$

$$\text{VPI}(A \mid X) := \text{MEU}(\mathcal{D}_{X \rightarrow A}) - \text{MEU}(\mathcal{D})$$

Finding MEU Decision Rules

$$\begin{array}{rcccl} \text{MEU}(D_{S \rightarrow F}) & - & \text{MEU}(D) & & \\ 3.25 & & 2 & = & 1.25 \end{array}$$



Value of Information

$$\text{VPI}(A \mid X) := \text{MEU}(\mathcal{D}_{X \rightarrow A}) - \text{MEU}(\mathcal{D})$$

- Theorem:

- $\text{VPI}(A \mid X) \geq 0$

- $\text{VPI}(A \mid X) = 0$ if and only if the optimal decision rule for \mathcal{D} is still optimal for $\mathcal{D}_{X \rightarrow A}$

Any CPD $\delta(A \mid \bar{z})$ is also a CPD $\delta(A \mid \bar{z}, x)$

Clear notion of when information matters

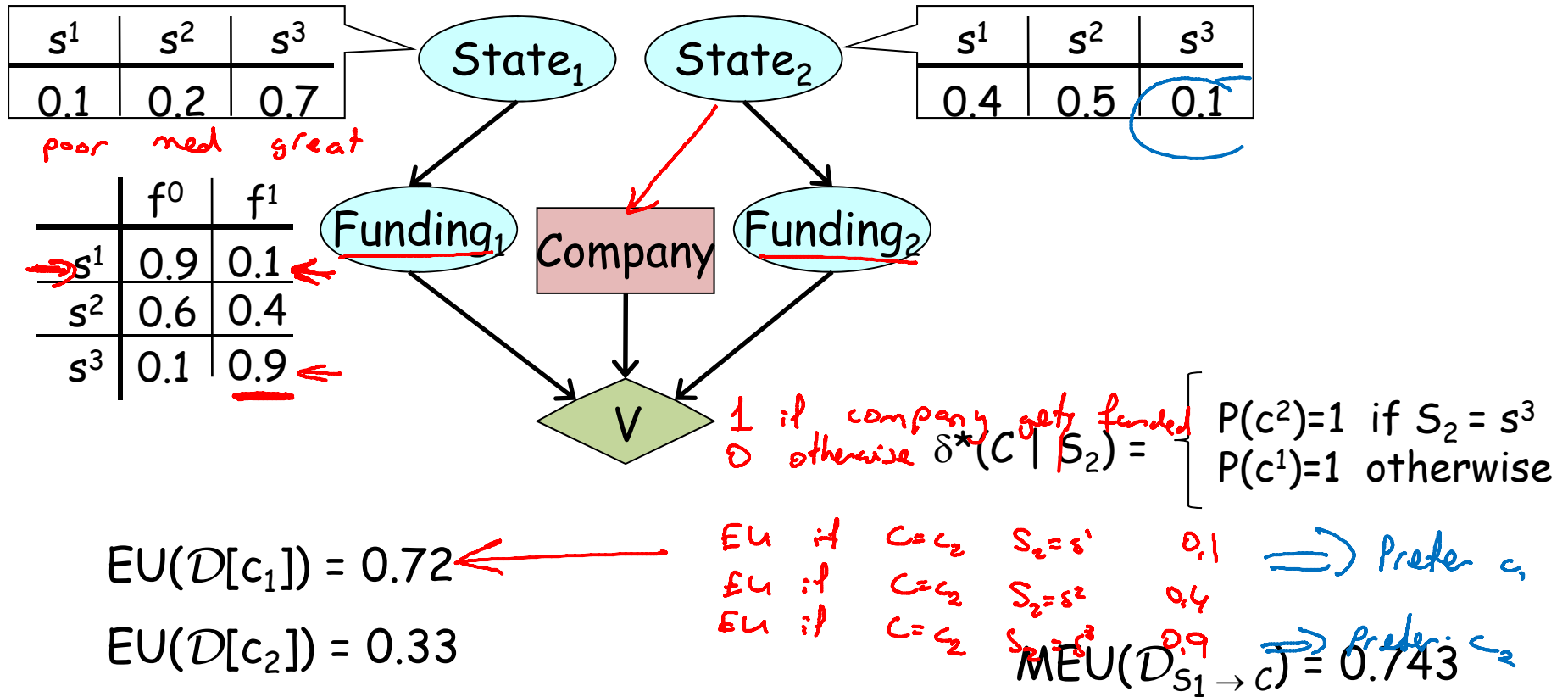


it changes my decision

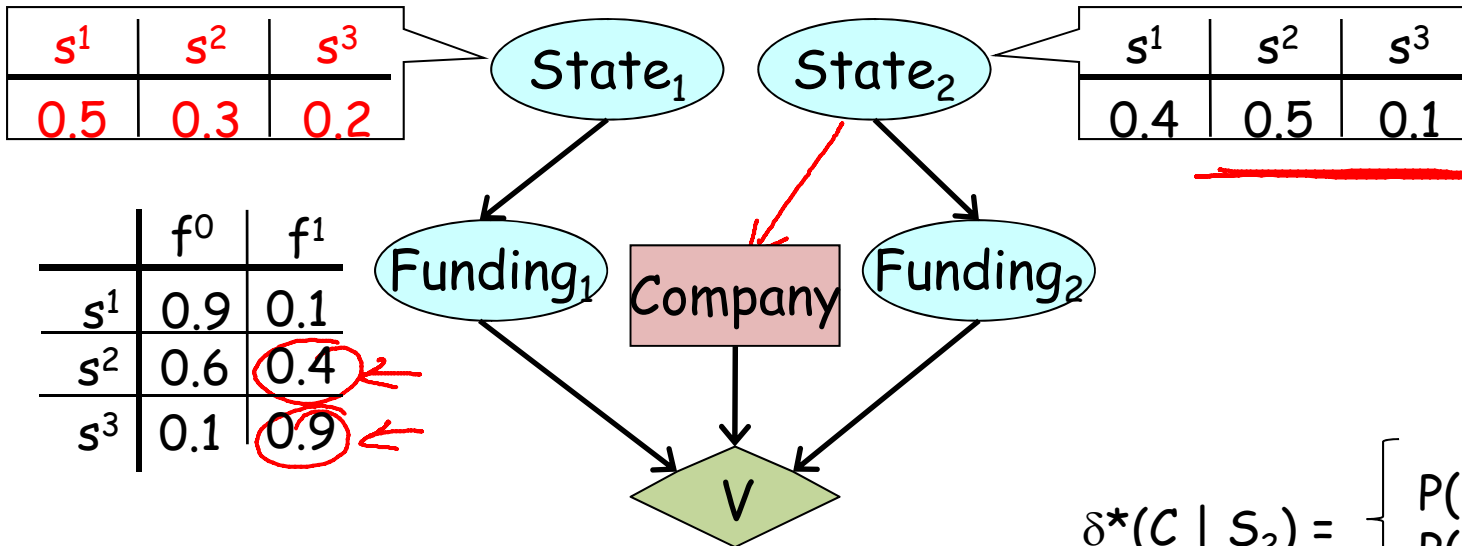
optimizing $\delta(A \mid \bar{z}, x)$

optimizing $\delta(A \mid \bar{z})$

Value of Information Example



Value of Information Example



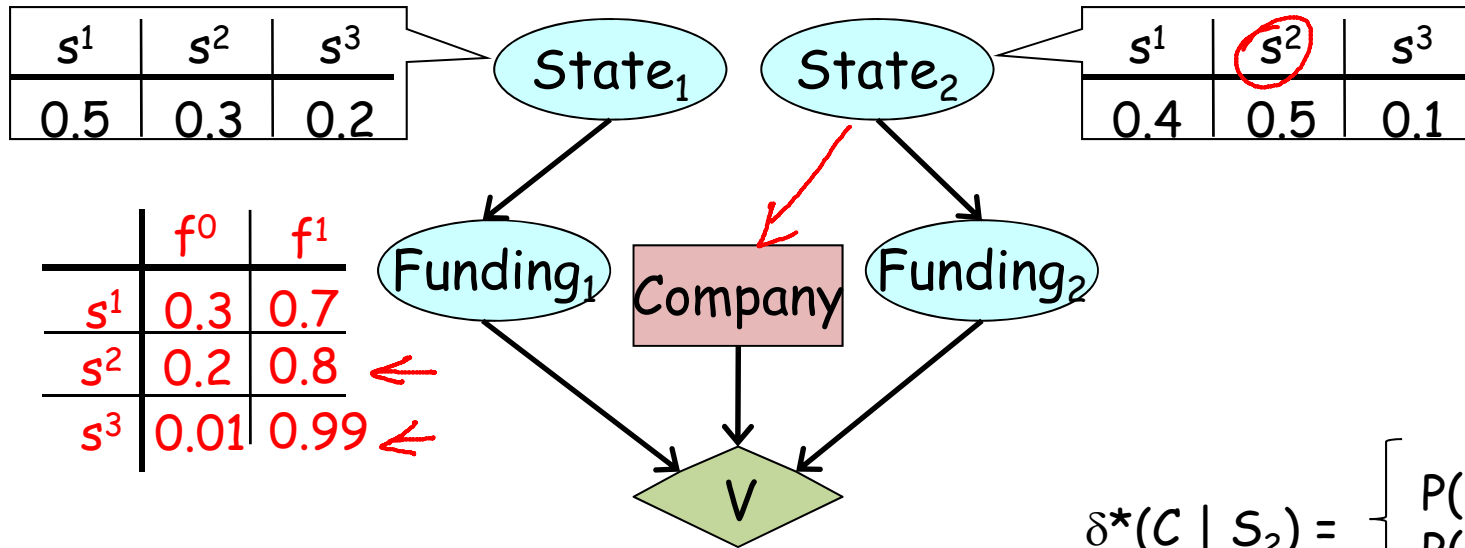
$$\delta^*(C \mid S_2) = \begin{cases} P(c^2)=1 & \text{if } S_2 = s^2, s^3 \\ P(c^1)=1 & \text{otherwise} \end{cases}$$

$$EU(D[c_1]) = 0.35$$

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

$$MEU(D_{S_2 \rightarrow c}) = \underline{0.43}$$

Value of Information Example



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

$$EU(D[c_1]) = 0.788$$

$$EU(D[c_2]) = 0.779$$

$$MEU(D_{s_1 \rightarrow c}) = \underline{0.8142}$$

Summary

- Influence diagrams provide clear and coherent semantics for the value of making an observation
 - Difference between values of two IDs
- Information is valuable if and only if it induces a change in action in at least one context