Fundamentals of Decision Theory

Chapter 16

Mausam

(Based on slides from NPS)

Uncertainty in AI

- Uncertainty comes in many forms
  - uncertainty due to another agent's policy
  - uncertainty in outcome of my own action
  - uncertainty in my knowledge of the world
  - uncertainty in how the world evolves

Decision Theory

- is the study of agent's choices
- lays out principles for how an agent arrives at an *optimal* choice

  - Good decisions may occasionally have unexpected bad outcomes
    - it is still a good decision if made properly
  - Bad decisions may occasionally have good outcomes if you are lucky
    - it is still a bad decision

Steps in Decision Theory

1. List the possible actions (actions/decisions)

2. Identify the possible outcomes

3. List the payoff or profit or reward

4. Select one of the decision theory models

5. Apply the model and make your decision

Running Example

- Problem.
  - The FoxPhone India Co. must decide whether or not to expand its product line by manufacturing indigenous smartphones in India

- Step 1: List the possible alternatives/actions

  alternative: "a course of action or strategy that may be chosen by the decision maker"

  - (1) Construct a large plant to manufacture the phones
  - (2) Construct a small plant
  - (3) Do nothing

The FoxPhone India Co.

- Step 2: Identify the states of nature
  - Market for indigenous smartphones could be favorable
    - high demand
  - Market for indigenous smartphones could be unfavorable
    - low demand

state of nature: "an outcome over which the agent has little or no control"

e.g., lottery, coin-toss, whether it will rain today

The FoxPhone India Co.

- Step 3: List the possible rewards
  - A reward for all possible combinations of actions and states of nature
  - *Conditional values*: "reward depends upon the action and the state of nature"
    - with a favorable market:
      - a large plant produces a net profit of ₹200,000
      - a small plant produces a net profit of ₹100,000
      - no plant produces a net profit of ₹0
    - with an unfavorable market:
      - a large plant produces a net loss of ₹180,000
      - a small plant produces a net loss of ₹20,000
      - no plant produces a net profit of ₹0

Reward Tables

	States of Nature	
Actions		

- A means of organizing a decision situation, including the rewards from different situations given the possible states of nature

The FoxPhone India Co.

	States of Nature	
Actions	**Favorable Market**	**Unfavorable Market**
Large plant	**₹200,000**	**-₹180,000**
Small plant	**₹100,000**	**-₹20,000**
No plant	**₹0**	**₹0**

The FoxPhone India Co.

- Steps 4/5: Select an appropriate model and apply it
  - Model selection depends on the operating environment and degree of uncertainty

Future Uncertainty

- Nondeterministic
- Probabilistic

Non-deterministic Uncertainty

	States of Nature	
Actions	**Favorable Market**	**Unfavorable Market**
Large plant	**₹200,000**	**-₹180,000**
Small plant	**₹100,000**	**-₹20,000**
No plant	**₹0**	**₹0**

- What should we do?

Maximax Criterion

"Go for the Gold"

- Select the decision that results in the maximum of the maximum rewards
- A very optimistic decision criterion
  - Decision maker assumes that the most favorable state of nature for each action will occur
- Most risk prone agent

Maximax

Decision	States of Nature		Maximum in Row
	Favorable	Unfavorable	
Large plant	₹200,000	-₹180,000	₹200,000
Small plant	₹100,000	-₹20,000	₹100,000
No plant	₹0	₹0	₹0

- **FoxPhone India Co. assumes that the most favorable state of nature occurs for each decision action**
- **Select the maximum reward for each decision**
  - **All three maximums occur if a favorable economy prevails (a tie in case of no plant)**
- **Select the maximum of the maximums**
  - **Maximum is ₹200,000; corresponding decision is to build the large plant**
  - **Potential loss of ₹180,000 is completely ignored**

Maximin Criterion

"Best of the Worst"

- Select the decision that results in the maximum of the minimum rewards
- A very pessimistic decision criterion
  - Decision maker assumes that the minimum reward occurs for each decision action
  - Select the maximum of these minimum rewards
- Most risk averse agent

Maximin

Decision	States of Nature		Minimum in Row
	Favorable	Unfavorable	
Large plant	₹200,000	-₹180,000	-₹180,000
Small plant	₹100,000	-₹20,000	-₹20,000
No plant	₹0	₹0	₹0

- **FoxPhone India Co. assumes that the least favorable state of nature occurs for each decision action**
- **Select the minimum reward for each decision**
  - **All three minimums occur if an unfavorable economy prevails (a tie in case of no plant)**
- **Select the maximum of the minimums**
  - **Maximum is ₹0; corresponding decision is to do nothing**
  - **A conservative decision; largest possible gain, ₹0, is much less than maximax**

Equal Likelihood Criterion

- Assumes that all states of nature are equally likely to occur
  - Maximax criterion assumed the most favorable state of nature occurs for each decision
  - Maximin criterion assumed the least favorable state of nature occurs for each decision

- Calculate the *average reward* for each action and select the action with the maximum number
  - *Average reward*: the sum of all rewards divided by the number of states of nature

- Select the decision that gives the highest average reward

Equal Likelihood

Decision	States of Nature		Row Average
	Favorable	Unfavorable	
Large plant	₹200,000	-₹180,000	₹10,000
Small plant	₹100,000	-₹20,000	₹40,000
No plant	₹0	₹0	₹0

Row Averages

$$\text{Large Plant} = \frac{\$200{,}000 - \$180{,}000}{2} = \$10{,}000$$

$$\text{Small Plant} = \frac{\$100{,}000 - \$20{,}000}{2} = \$40{,}000$$

$$\text{Do Nothing} = \frac{\$0 + \$0}{2} = \$0$$

- Select the decision with the highest weighted value
  - Maximum is ₹40,000; corresponding decision is to build the small plant

Criterion of Realism

- Also known as the weighted average or Hurwicz criterion
  - A compromise between an optimistic and pessimistic decision

- A coefficient of realism, α, is selected by the decision maker to indicate optimism or pessimism about the future

$$0 \leq \alpha \leq 1$$

When α is close to 1, the decision maker is optimistic.

When α is close to 0, the decision maker is pessimistic.

- *Criterion of realism* = α(row maximum) + (1-α)(row minimum)
  - A weighted average where maximum and minimum rewards are weighted by α and (1 - α) respectively

Criterion of Realism

- Assume a coefficient of realism equal to 0.8

Decision	States of Nature		Criterion of Realism
	Favorable	Unfavorable	
Large plant	**₹200,000**	**-₹180,000**	**₹124,000**
Small plant	**₹100,000**	**-₹20,000**	**₹76,000**
No plant	**₹0**	**₹0**	**₹0**

Weighted Averages

Large Plant = (0.8)(₹200,000) + (0.2)(-₹180,000) = ₹124,000

Small Plant = **(0.8)(₹100,000) + (0.2)(-₹20,000) = ₹76,000**

Do Nothing = **(0.8)(₹0) + (0.2)(₹0) = ₹0**

Select the decision with the highest weighted value

Maximum is ₹124,000; corresponding decision is to build the large plant

Minimax Regret

- Regret/Opportunity Loss: "the difference between the optimal reward and the actual reward received"

- Choose the action that minimizes the maximum regret associated with each action
  - Start by determining the max regret for each action
  - Pick the action with the minimum number

Regret Table

- If I knew the future, how much I'd regret my decision…

- Regret for any state of nature is calculated by subtracting each outcome in the column from the best outcome in the same column

Minimax Regret

Decision	States of Nature				
	Favorable		Unfavorable		Row
	Payoff	Regret	Payoff	Regret	Maximum
Large plant	₹200,000	₹0	-₹180,000	₹180,000	₹180,000
Small plant	₹100,000	₹100,000	-₹20,000	₹20,000	₹100,000
No plant	₹0	₹200,000	₹0	₹0	₹200,000
Best payoff	₹200,000		₹0		

- Select the action with the lowest maximum regret

Minimum is ₹100,000; corresponding decision is to build a small plant

Summary of Results

Criterion	Decision
Maximax	**Build a large plant**
Maximin	**Do nothing**
Equal likelihood	**Build a small plant**
Realism	**Build a large plant**
Minimax regret	**Build a small plant**

Future Uncertainty

- Non deterministic
- Probabilistic

Probabilistic Uncertainty

- Decision makers know the probability of occurrence for each possible outcome
  - Attempt to maximize the expected reward

- Criteria for decision models in this environment:
  - Maximization of expected reward
  - Minimization of expected regret
    - Minimize expected regret = maximizing expected reward!

Expected Reward (Q)

- called Expected Monetary Value (EMV) in DT literature
- "the probability weighted sum of possible rewards for each action"
  - Requires a reward table with conditional rewards and probability assessments for all states of nature

Q(action a) = (reward of 1st state of nature)

X (probability of 1st state of nature)

+ (reward of 2nd state of nature)

X (probability of 2nd state of nature)

+ . . . + (reward of last state of nature)

X (probability of last state of nature)

The FoxPhone India Co.

- Suppose that the probability of a favorable market is exactly the same as the probability of an unfavorable market. Which action would give the greatest Q?

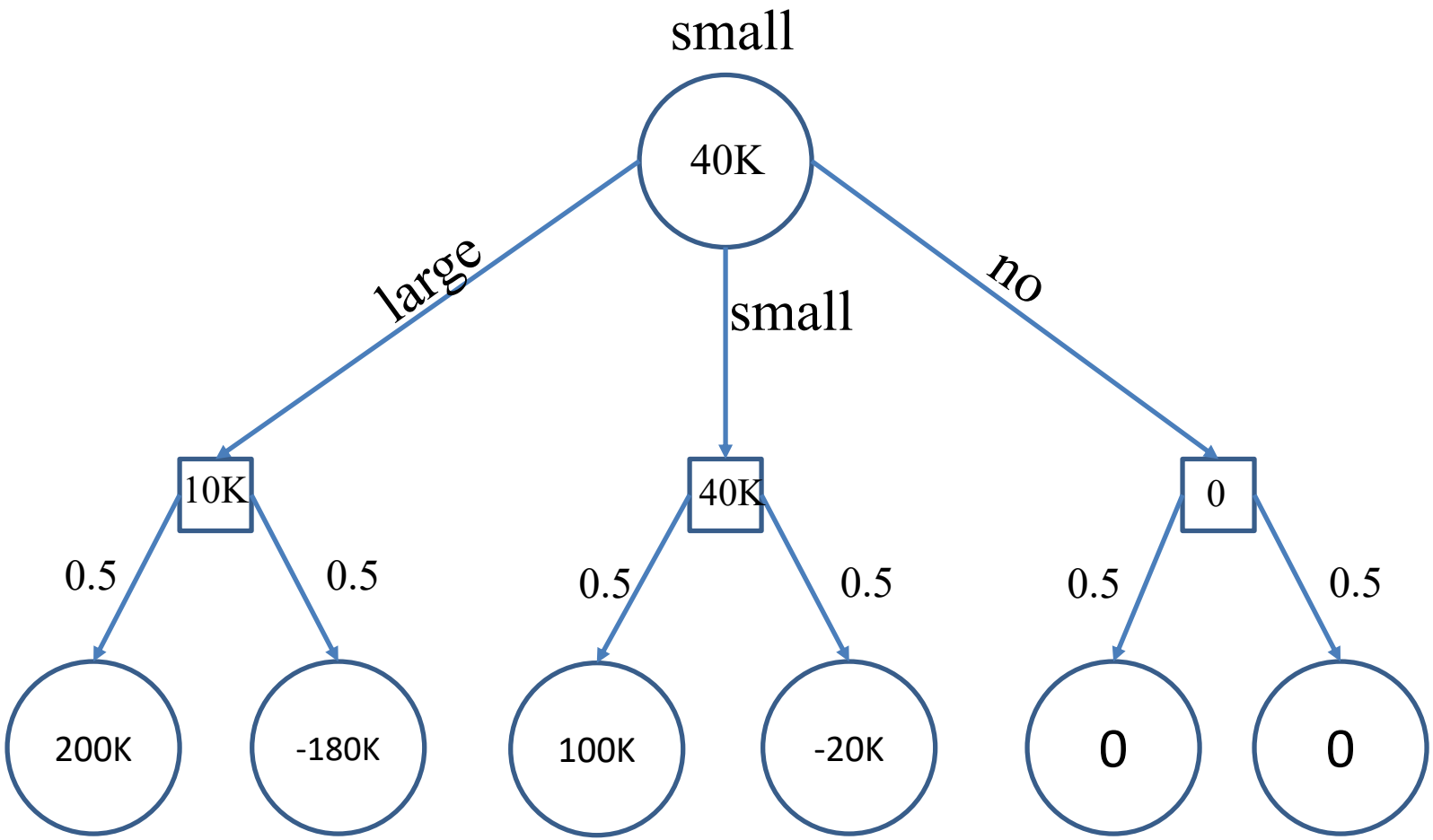
	States of Nature		
Decision	Favorable Mkt p = 0.5	Unfavorable Mkt p = 0.5	EMV
Large plant	₹200,000	-₹180,000	₹10,000
Small plant	₹100,000	-₹20,000	₹40,000
No plant	₹0	₹0	₹0

Q(large plant) = **(0.5)(₹200,000) + (0.5)(-₹180,000) = ₹10,000**

Q(small plant) = **(0.5)(₹100,000) + (0.5)(-₹-20,000) = ₹40,000**

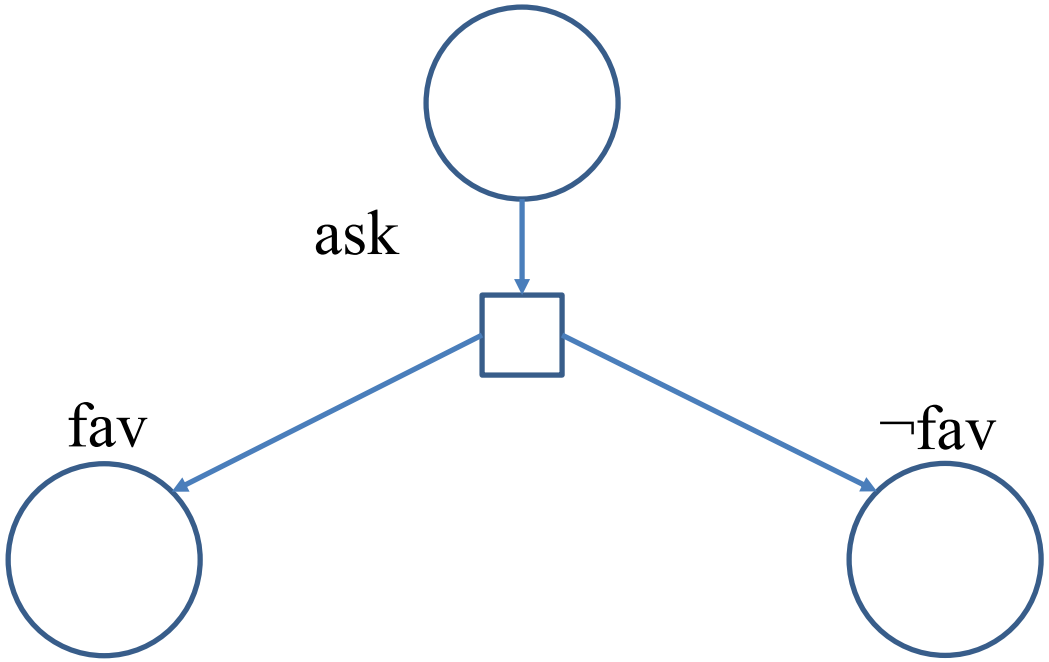
Q(no plant) = **(0.5)(₹0) + (0.5)(₹0) = ₹0**

Build the small plant


small
40K
large
small
no
10K
40K
0
0.5
0.5
0.5
0.5
0.5
0.5
200K
-180K
100K
-20K
0
0

Expected Value of Perfect Information (EVPI)

- It may be possible to purchase additional information about future events and thus make a better decision
  - FoxPhone India Co. could hire an economist to analyze the economy in order to more accurately determine which economic condition will occur in the future
    - How valuable would this information be?


ask
fav
¬fav

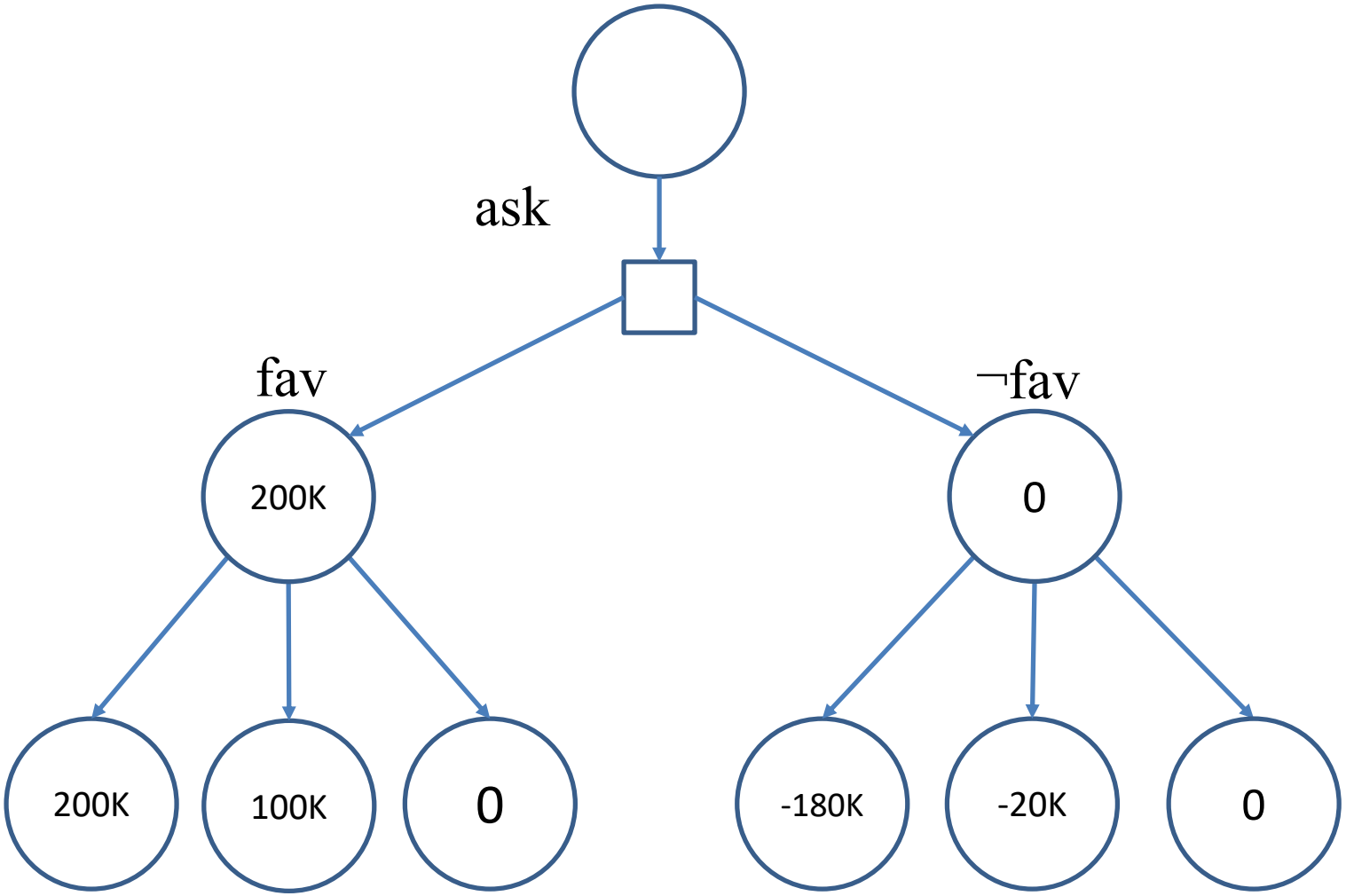
EVPI Computation

- Look first at the decisions under each state of nature
  - If information was available that perfectly predicted which state of nature was going to occur, the best decision for that state of nature could be made
    - *expected value with perfect information* (EV w/ PI): "the expected or average return if we have perfect information before a decision has to be made"

EVPI Computation

- Perfect information changes environment from decision making under risk to decision making with certainty
  - Build the large plant if you know for sure that a favorable market will prevail
  - Do nothing if you know for sure that an unfavorable market will prevail

Decision	States of Nature	
	Favorable p = 0.5	Unfavorable p = 0.5
Large plant	₹200,000	-₹180,000
Small plant	₹100,000	-₹20,000
No plant	₹0	₹0


ask
fav
¬fav
200K
0
200K
100K
0
-180K
-20K
0

EVPI Computation

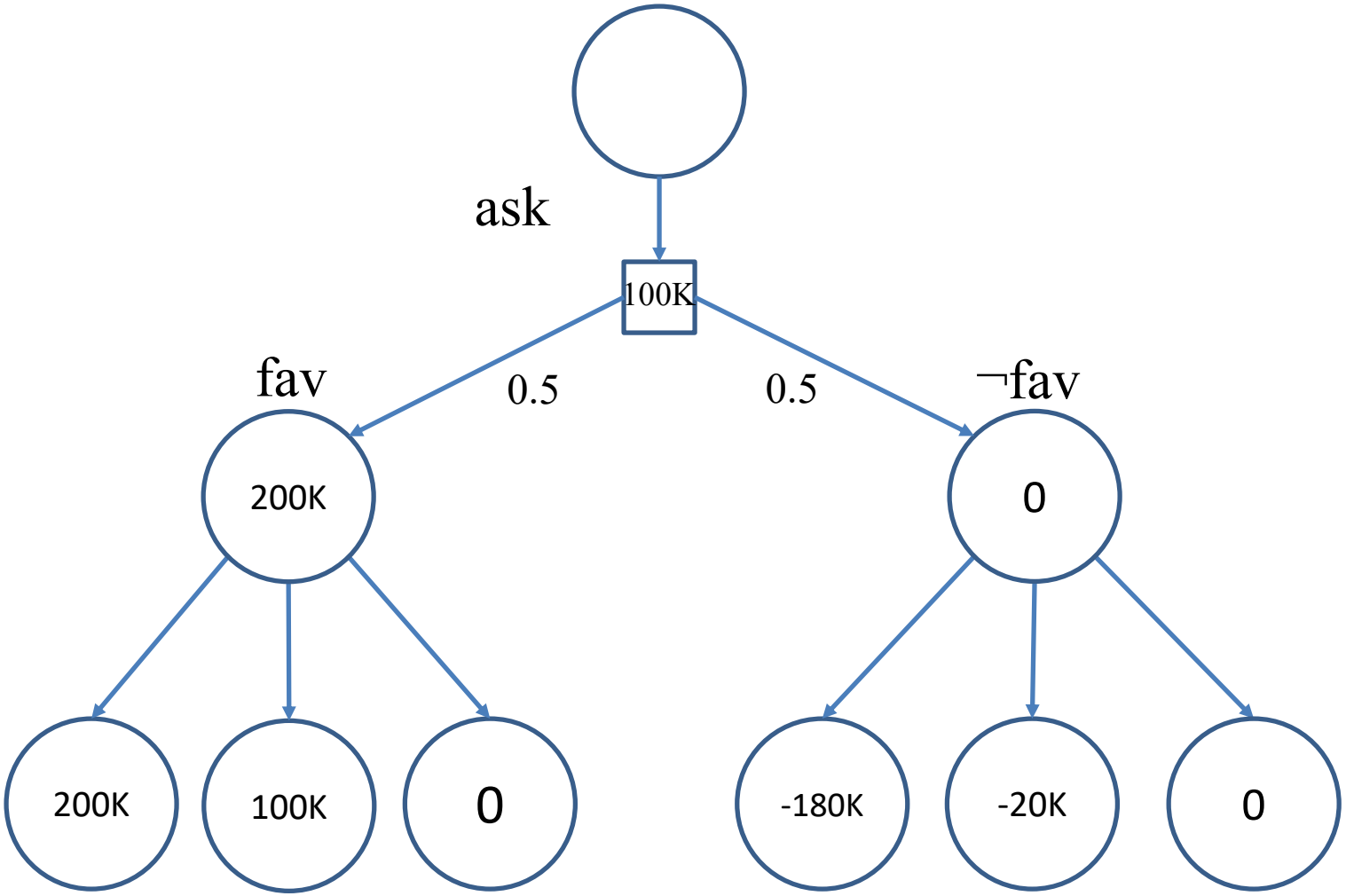
- Even though perfect information enables FoxPhone India Co. to make the correct investment decision, each state of nature occurs only a certain portion of the time
  - A favorable market occurs 50% of the time and an unfavorable market occurs 50% of the time
  - EV w/ PI calculated by choosing the best action for each state of nature and multiplying its reward times the probability of occurrence of the state of nature

EVPI Computation

EV w/ PI = **(best reward for 1st state of nature)**
X (probability of 1st state of nature)
+ (best reward for 2nd state of nature)
X (probability of 2nd state of nature)

EV w/ PI = **(₹200,000)(0.5) + (₹0)(0.5) = ₹100,000**

Decision	States of Nature	
	Favorable p = 0.5	Unfavorable p = 0.5
Large plant	₹200,000	-₹180,000
Small plant	₹100,000	-₹20,000
No plant	₹0	₹0

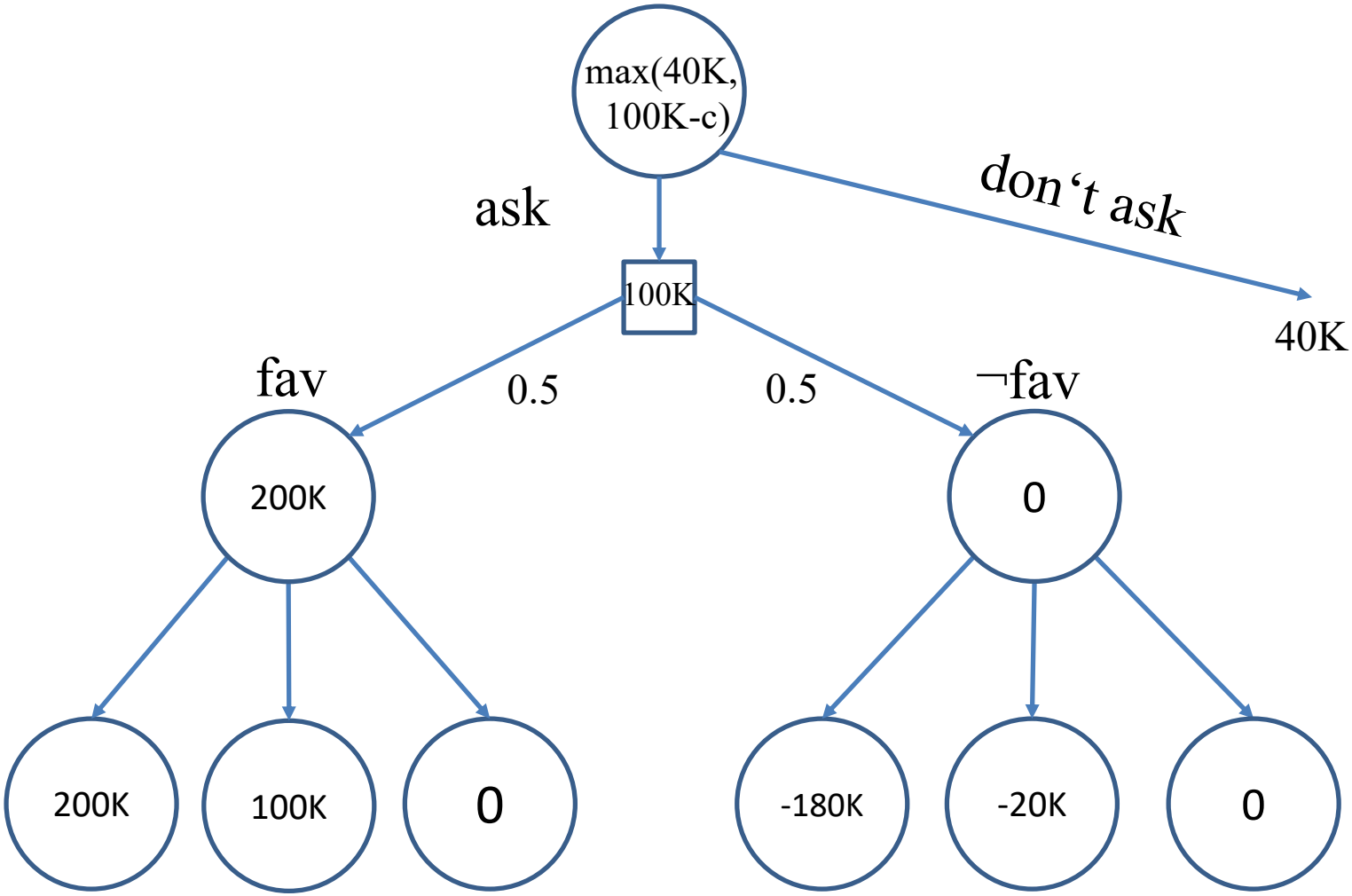

ask
100K
fav
0.5
0.5
¬fav
200K
0
200K
100K
0
-180K
-20K
0

EVPI Computation

- FoxPhone India Co. would be foolish to pay more for this information than the extra profit that would be gained from having it
  - *EVPI*: "the maximum amount a decision maker would pay for additional information resulting in a decision better than one made *without perfect information* "
    - EVPI is the expected outcome with perfect information minus the expected outcome without perfect information

EVPI = EV w/ PI - Q

EVPI = ₹100,000 - ₹40,000 = ₹60,000


max(40K,
100K-c)
ask
don‘t ask
100K
40K
fav
0.5
0.5
¬fav
200K
0
200K
100K
0
-180K
-20K
0

Using EVPI

- EVPI of ₹60,000 is the maximum amount that FoxPhone India Co. should pay to purchase perfect information from a source such as an economist
  - "Perfect" information is extremely rare
  - An investor typically would be willing to pay some amount less than ₹60,000, depending on how reliable the information is perceived to be

Is Expected Value sufficient?

- Lottery 1
  - returns ₹0 always
- Lottery 2
  - return ₹100 and -₹100 with prob 0.5

- Which is better?

Is Expected Value sufficient?

- Lottery 1
  - returns ₹100 always
- Lottery 2
  - return ₹10000 (prob 0.01) and ₹0 with prob 0.99

- Which is better?
  - depends

Is Expected Value sufficient?

- Lottery 1
  - returns ₹3125 always
- Lottery 2
  - return ₹4000 (prob 0.75) and -₹500 with prob 0.25

- Which is better?

Is Expected Value sufficient?

- Lottery 1
  - returns ₹0 always
- Lottery 2
  - return ₹1,000,000 (prob 0.5) and -₹1,000,000 with prob 0.5

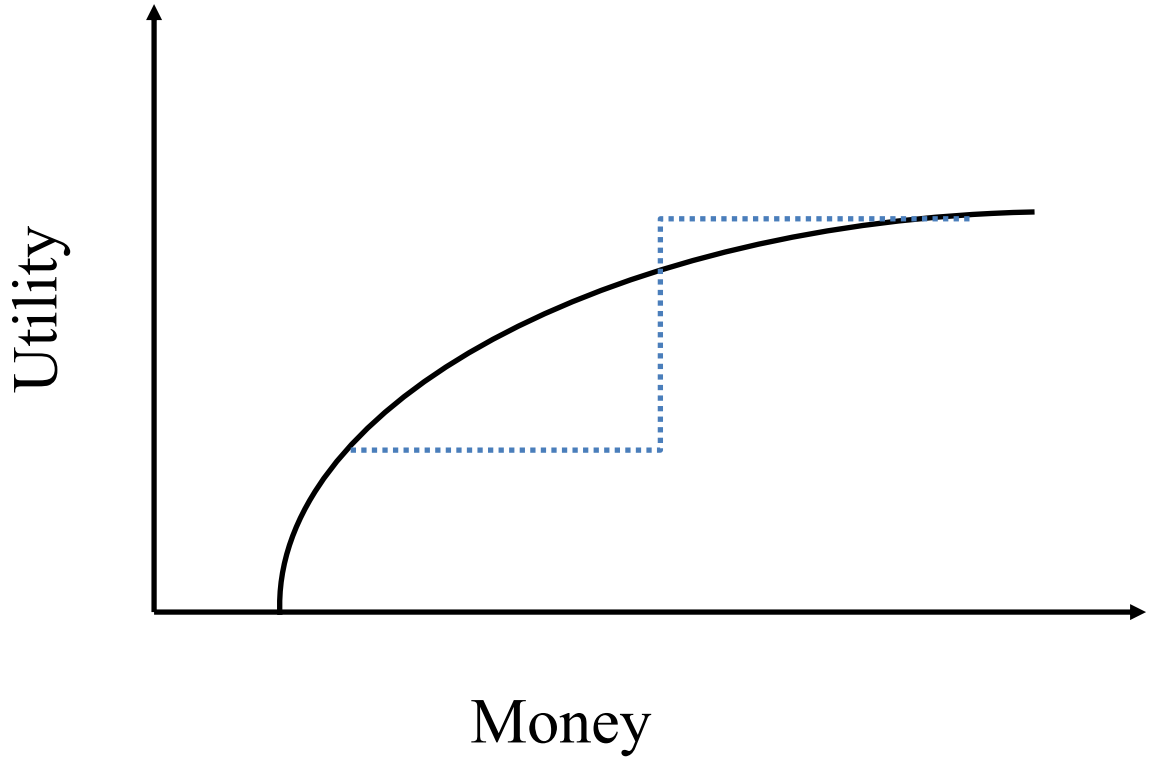
- Which is better?

Utility Theory

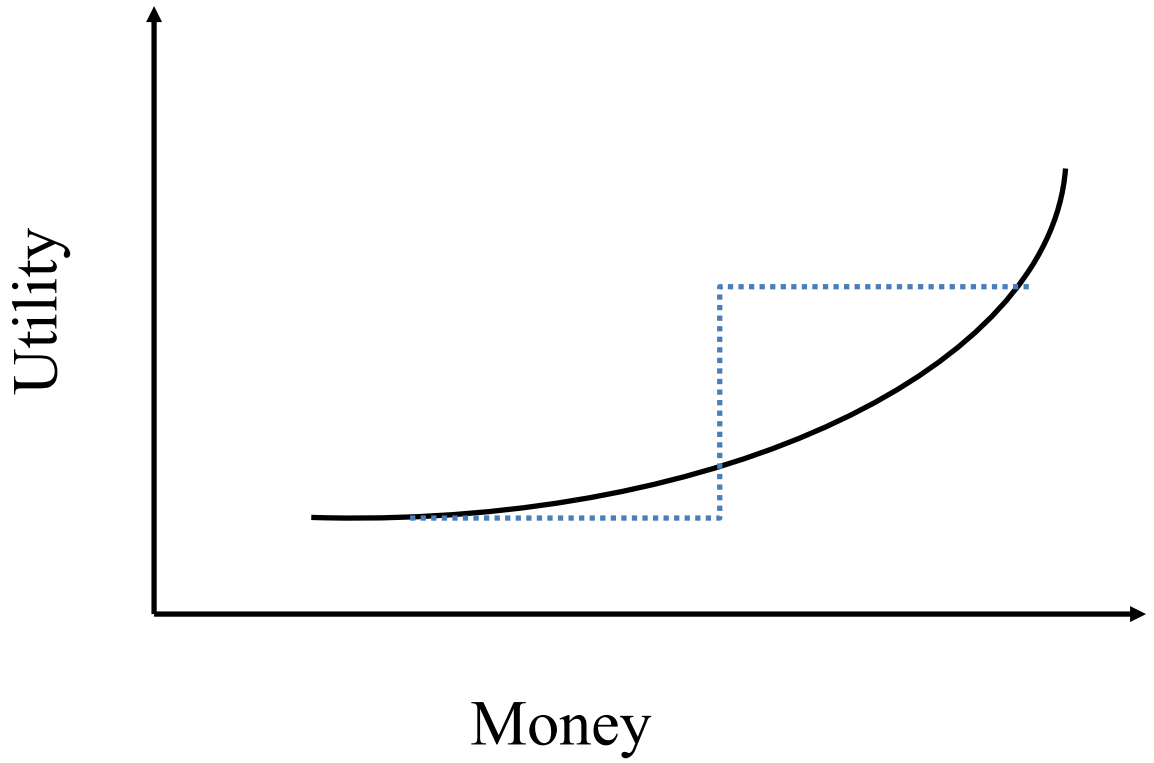
- Adds a layer of utility over rewards
- Risk averse
  - |Utility| of high negative money is much MORE than utility of high positive money
- Risk prone
  - Reverse

- Use expected utility criteria…

Utility function of risk-averse agent



Utility function of a risk-prone agent



Utility function of a risk-neutral agent