CS 1571 Introduction to AI
Lecture 25

Decision making in the presence of uncertainty

Milos Hauskrecht

milos@cs.pitt.edu

5329 Sennott Square

Decision-making in the presence of uncertainty

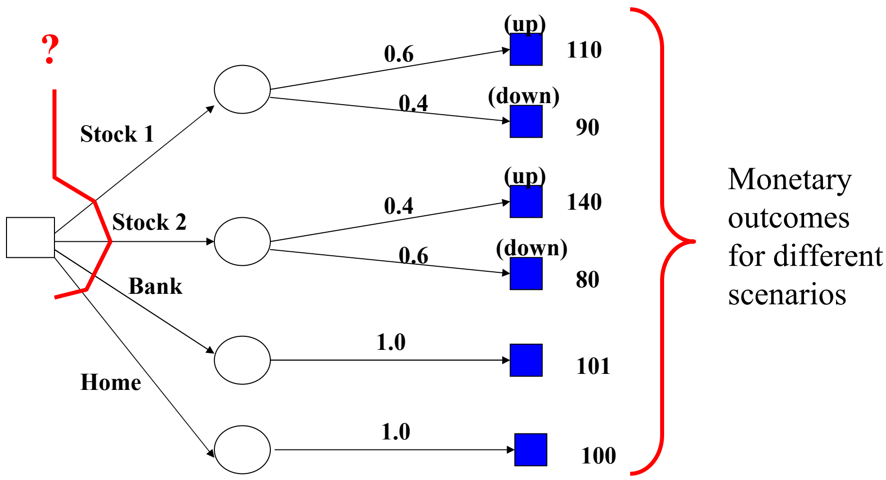
- Many real-world problems require **to choose future actions in the presence of uncertainty**
- **Examples:** patient management, investments

Main issues:

- **How to model the decision process in the computer ?**
- **How to make decisions about actions in the presence of uncertainty?**

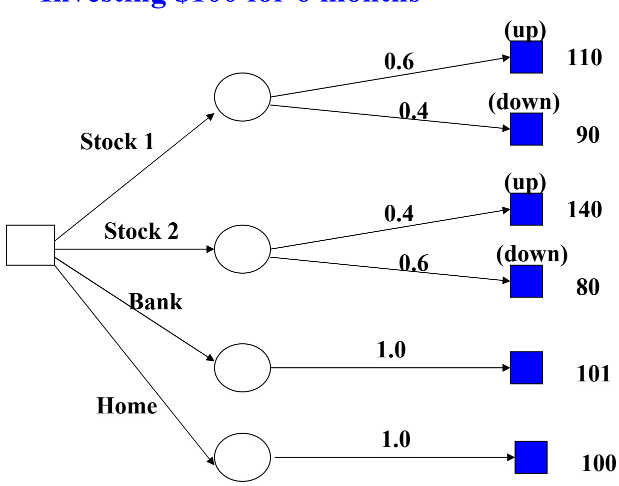
Decision making example.

We need to make a choice whether to invest in Stock 1 or 2, put money into bank or keep them at home. But how?



Decision tree representation of the problem

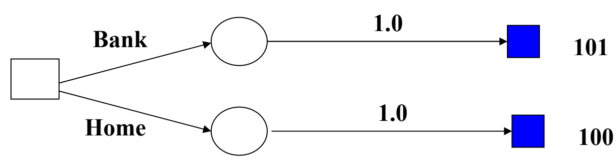
Investing $100 for 6 months



Decision making example.

Assume the simplified problem with the Bank and Home choices only.

The result is guaranteed – the outcome is deterministic

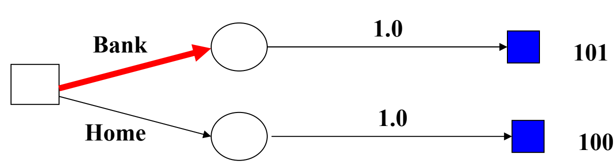


What is the rational choice assuming our goal is to make money?

Decision making. Deterministic outcome.

Assume the simplified problem with the Bank and Home choices only.

These choices are deterministic.



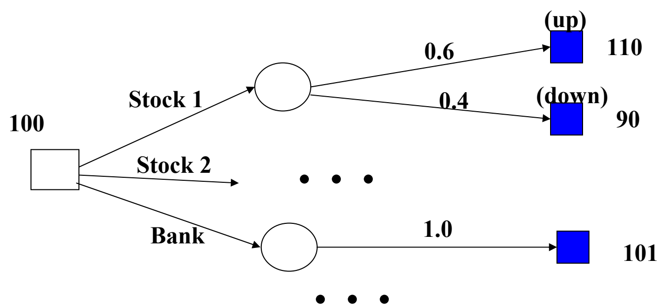
Our goal is to make money. What is the rational choice?

Answer: Put money into the bank. The choice is always strictly better in terms of the outcome

But what to do if we have uncertain outcomes?

Decision making. Stochastic outcome

- **How to quantify the goodness of the stochastic outcome?**
  We want to compare it to deterministic and other stochastic outcomes.



Idea: <u>**Use the expected value of the outcome**</u>

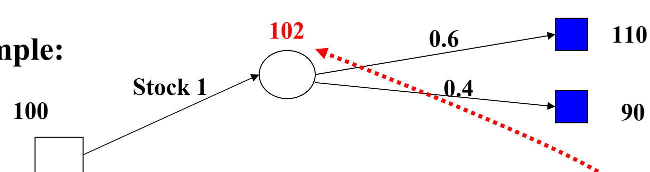
Expected value

- Let X be a random variable representing the monetary outcome with a discrete set of values Ω_X.
- **Expected value** of X is:

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

  - **Expected value** summarizes all stochastic outcomes into a single quantity

- **Example:**

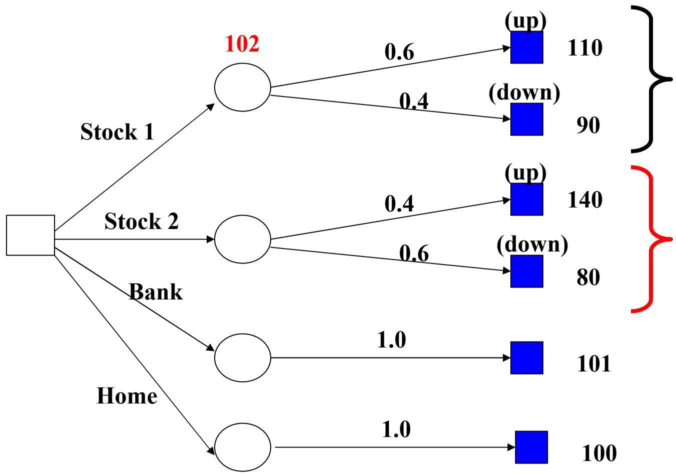


Expected value for the outcome of the Stock 1 option is:

$$0.6 \times 110 + 0.4 \times 90 = 66 + 36 = \boxed{102}$$

Expected values

Investing $100 for 6 months

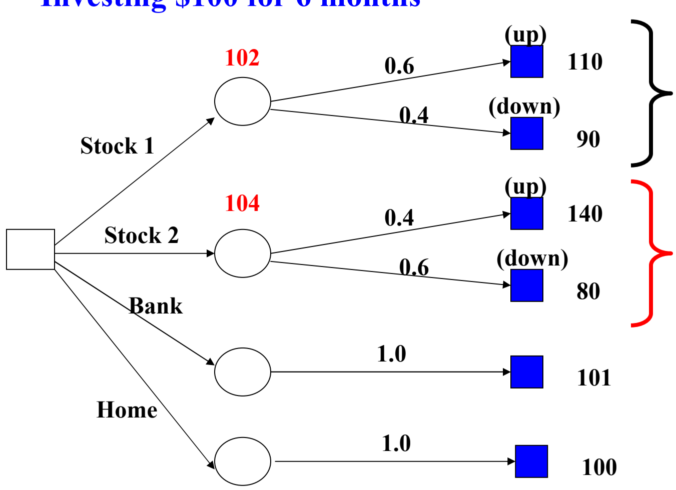


$0.6 \times 110 + 0.4 \times 90 =$
$66 + 36 = 102$

?

Expected values

Investing $100 for 6 months

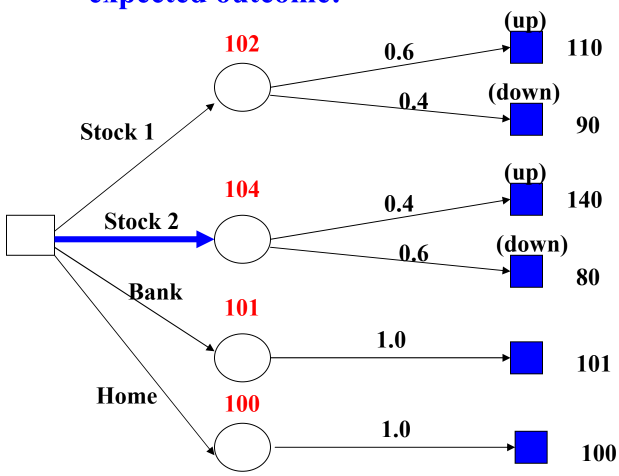


$0.6 \times 110 + 0.4 \times 90 =$
$66 + 36 = 102$

$0.4 \times 140 + 0.6 \times 80 =$
$56 + 48 = 104$

Selection based on expected values

The optimal action is the option that maximizes the expected outcome:



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?

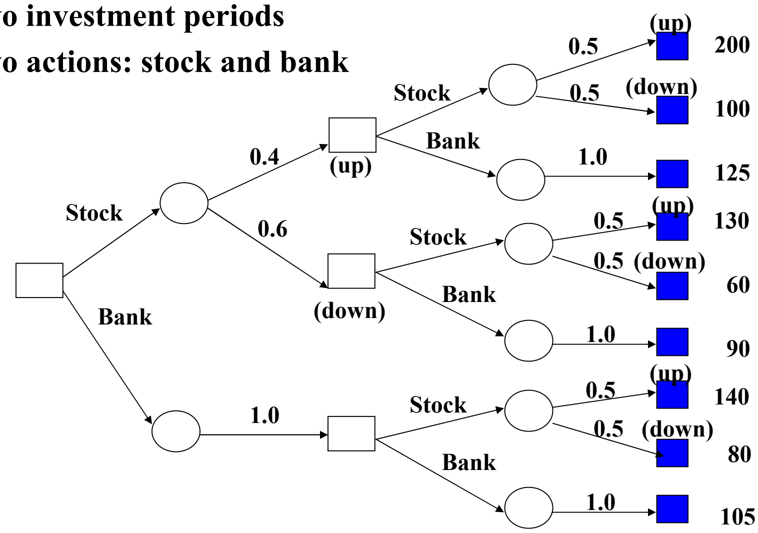
- Start from the leaves of the decision tree (outcome nodes)
- Compute expectations at chance nodes
- Maximize at the decision nodes

Algorithm is sometimes called **expectimax**

Multi-step problem example

Assume:

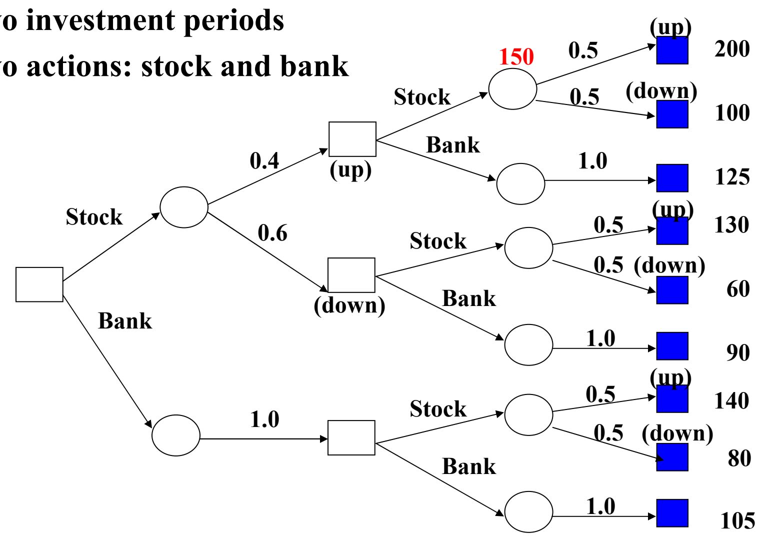
- Two investment periods
- Two actions: stock and bank



Multi-step problem example

Assume:

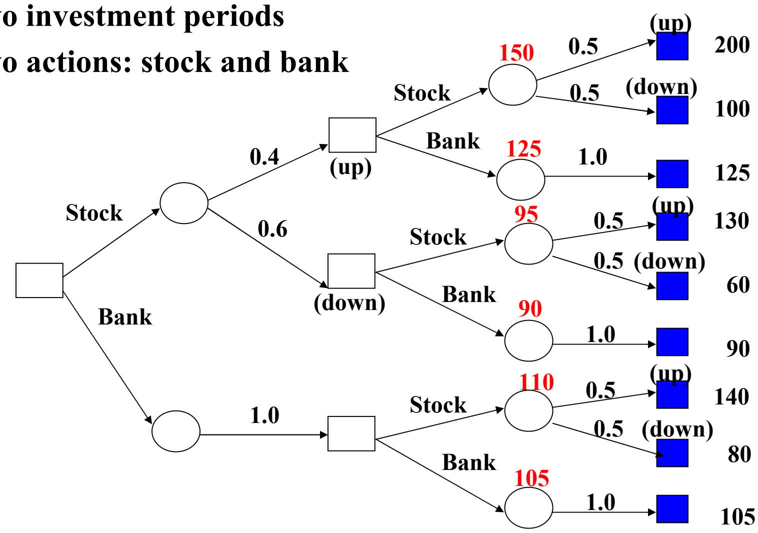
- Two investment periods
- Two actions: stock and bank



Multi-step problem example

Assume:

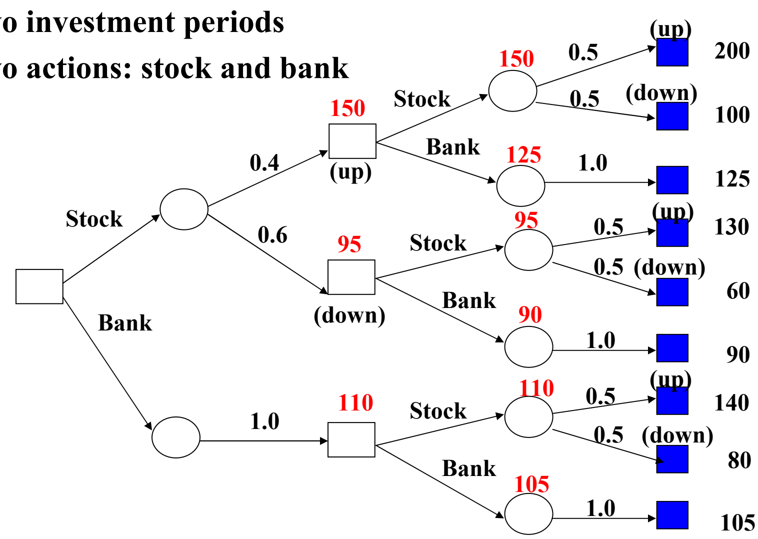
- Two investment periods
- Two actions: stock and bank



Multi-step problem example

Assume:

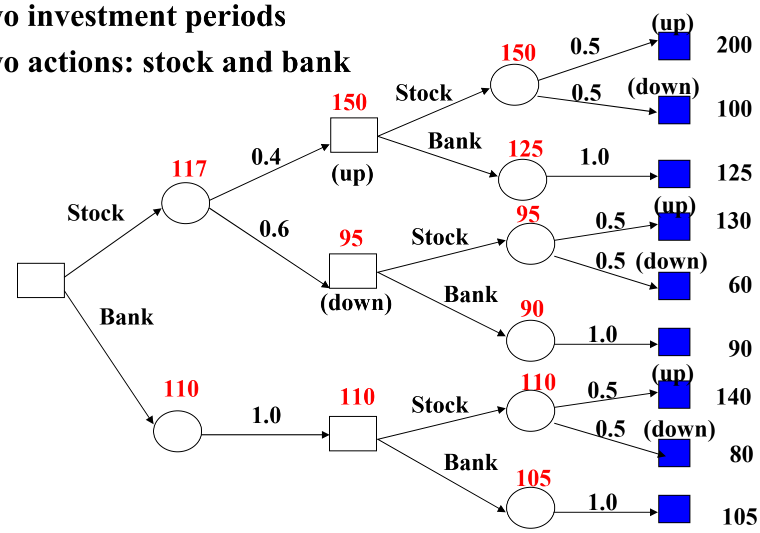
- Two investment periods
- Two actions: stock and bank



Multi-step problem example

Assume:

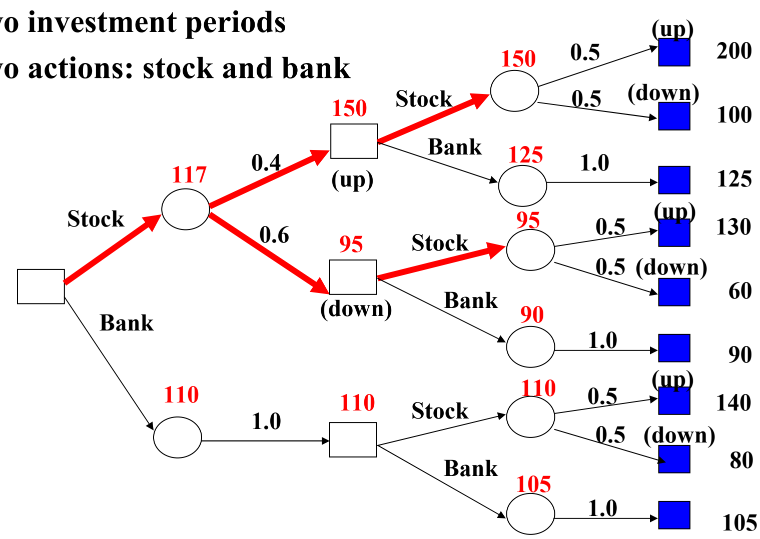
- Two investment periods
- Two actions: stock and bank



Multi-step problem example

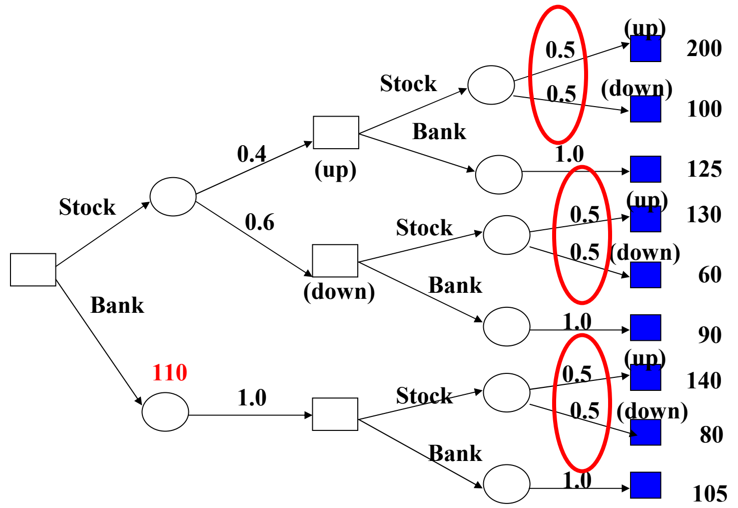
Assume:

- Two investment periods
- Two actions: stock and bank



Multi-step problems. Conditioning.

- Notice that the probability of stock going up and down in the 2nd step is independent of the 1st step (=0.5)



Conditioning in the decision tree

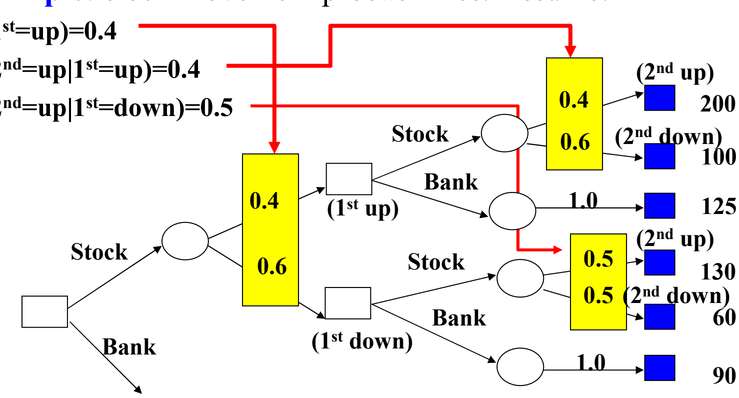
- **But this may not be the case. In decision trees:**
  - **Later outcomes can be conditioned on the earlier stochastic outcomes and actions**

Example: stock movement probabilities. Assume:

P(1st=up)=0.4

P(2nd=up|1st=up)=0.4

P(2nd=up|1st=down)=0.5



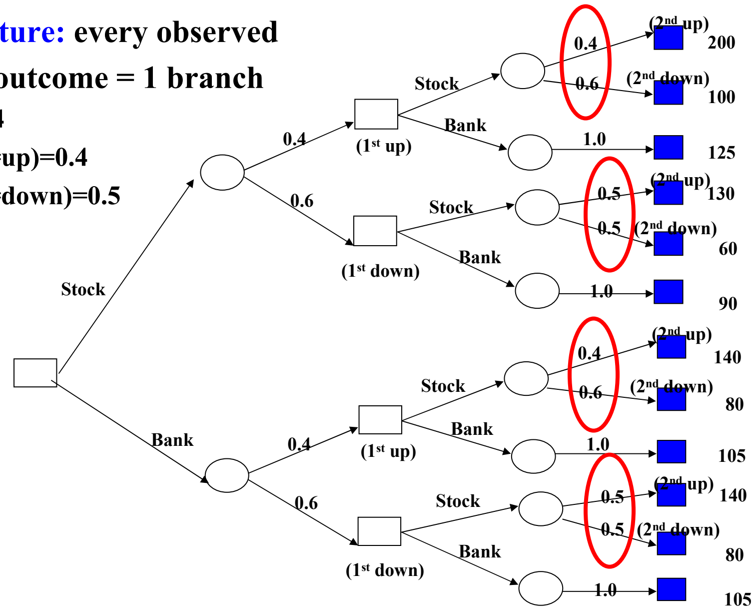
Multi-step problems. Conditioning.

Tree Structure: every observed stochastic outcome = 1 branch

P(1st=up)=0.4

P(2nd=up|1st=up)=0.4

P(2nd=up|1st=down)=0.5



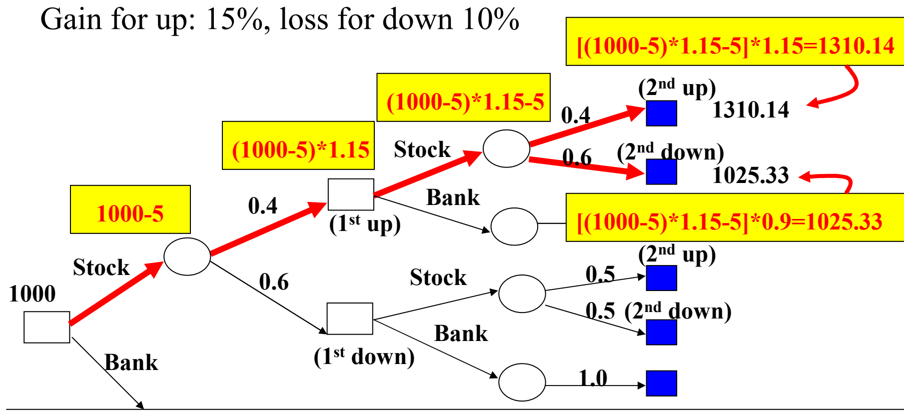
Trajectory payoffs

- **Outcome values at leaf nodes (e.g. monetary values)**
  - **Rewards and costs for the path trajectory**

Example: stock fees and gains. **Assume:**

Fee per period: $5 paid at the beginning

Gain for up: 15%, loss for down 10%



Constructing a decision tree

- **The decision tree is rarely given to you directly.**
  - Part of the problem is to construct the tree.

Example: stocks, bonds, bank for k periods

Stock:
- Probability of stocks going up in the first period: 0.3
- Probability of stocks going up in subsequent periods:
  - P(kth step=Up| (k -1)th step =Up)=0.4
  - P(kth step =Up| (k -1)th step=Down)=0.5
- Return if stock goes up: 15 % if down: 10%
- Fixed fee per investment period: $5

Bonds:
- Probability of value up: 0.5, down: 0.5
- Return if bond value is going up: 7%, if down: 3%
- Fee per investment period: $2

Bank:
- Guaranteed return of 3% per period, no fee

Information-gathering actions

- **Many actions and their outcomes irreversibly change the world**
- **Information-gathering (exploratory) actions:**
  - **make an inquiry about the world**
  - **Key benefit:** reduction in the uncertainty
- **Example: medicine**
  - Assume a patient is admitted to the hospital with some set of initial complaints
  - We are uncertain about the underlying problem and consider a surgery, or a medication to treat them
  - But there are often lab tests or observations that can help us to determine more closely the disease the patient suffers from
  - **Goal of lab tests:** Reduce the uncertainty of outcomes of treatments so that better treatment option can be chosen

Decision-making with exploratory actions

In decision trees:

- **Exploratory actions can be represented and reasoned about the same way as other actions**.

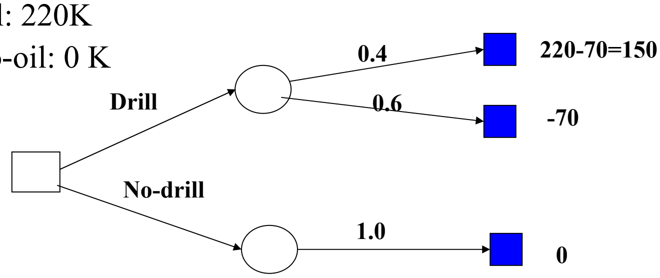
How do we capture the effect of exploratory actions in the decision tree model?

- Information obtained through exploratory actions may affect the probabilities of later outcomes
  - Recall that the probabilities on later outcomes can be conditioned on past observed outcomes and past actions
  - Sequence of past actions and outcomes is "remembered" within the decision tree branch

Oil wildcatter problem.

An oil wildcatter has to make a decision of whether to drill or not to drill on a specific site

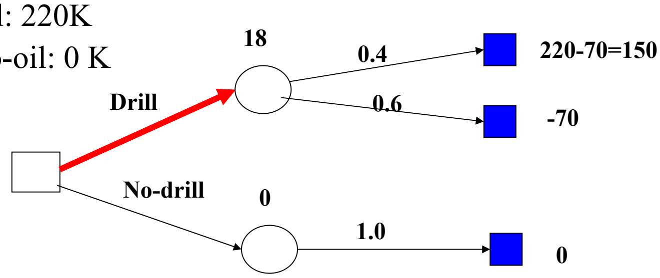
- **Chance of hitting an oil deposit:**
  - Oil: 40% $P(Oil = T) = 0.4$
  - No-oil: 60% $P(Oil = F) = 0.6$
- **Cost of drilling:** 70K
- **Payoffs:**
  - Oil: 220K
  - No-oil: 0 K



Oil wildcatter problem.

An oil wildcatter has to make a decision of whether to drill or not to drill on a specific site

- **Chance of hitting an oil deposit:**
  - Oil: 40% $P(Oil = T) = 0.4$
  - No-oil: 60% $P(Oil = F) = 0.6$
- **Cost of drilling:** 70K
- **Payoffs:**
  - Oil: 220K
  - No-oil: 0 K



Oil wildcatter problem

- Assume that in addition to the drill/no-drill choices we have an option to run the **seismic resonance test**
- **Seismic resonance test results:**
  - **Closed pattern** (more likely when the hole holds the oil)
  - **Diffuse pattern** (more likely when empty)

  $\mathbf{P}(Oil \mid Seismic\ resonance\ test)$

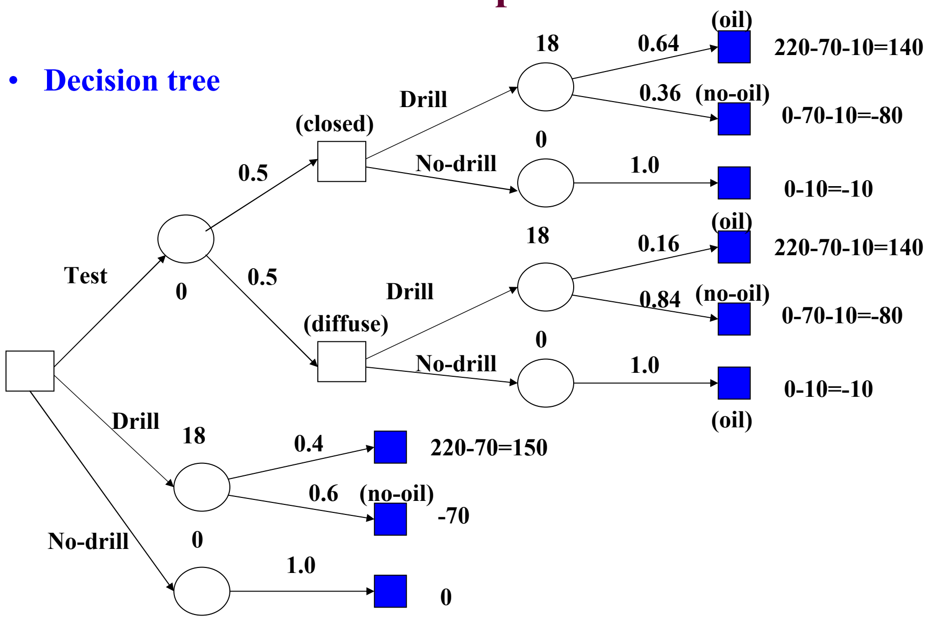
Seismic resonance test pattern

		closed	*diffuse*
Oil	*True*	0.8	0.2
	False	0.3	0.7

- **Test cost:** 10K

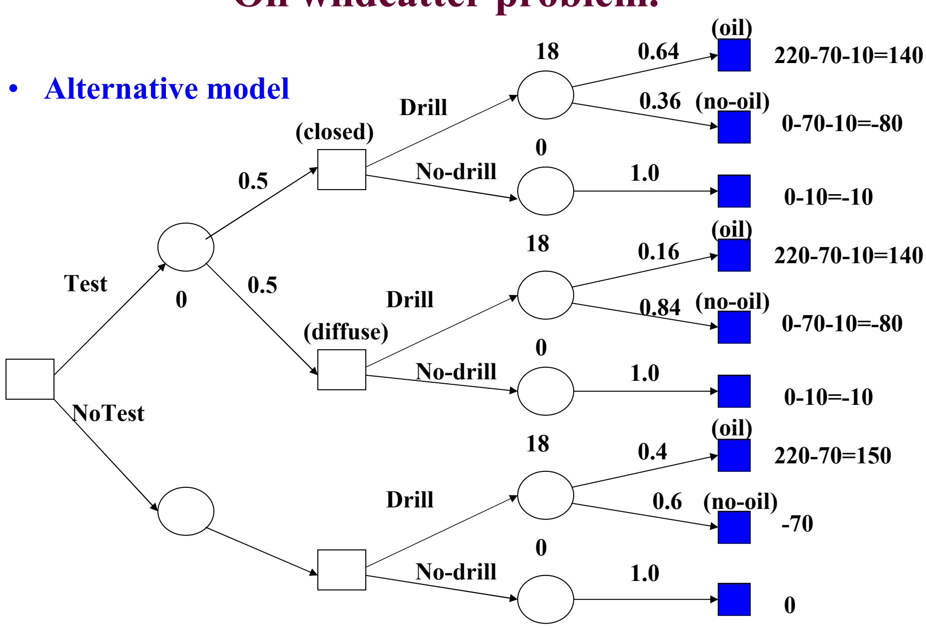
Oil wildcatter problem.

- **Decision tree**



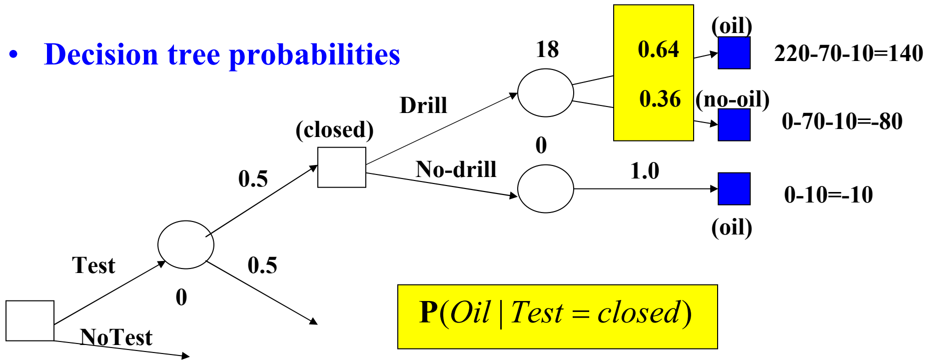
Oil wildcatter problem.

- **Alternative model**



Oil wildcatter problem.

- **Decision tree probabilities**



$\mathbf{P}(Oil \mid Test = closed)$

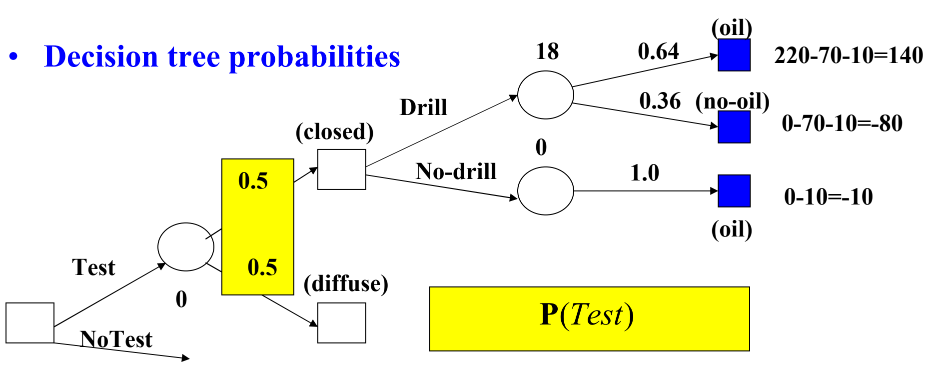
$$P(Oil = T \mid Test = closed) = \frac{P(Test = closed \mid Oil = T)P(Oil = T)}{P(Test = closed)}$$

$$P(Oil = F \mid Test = closed) = \frac{P(Test = closed \mid Oil = F)P(Oil = F)}{P(T = closed)}$$

$$P(Test = closed) = P(Test = closed \mid Oil = F)P(Oil = F) + P(Test = closed \mid Oil = T)P(Oil = T)$$

Oil wildcatter problem.

- **Decision tree probabilities**



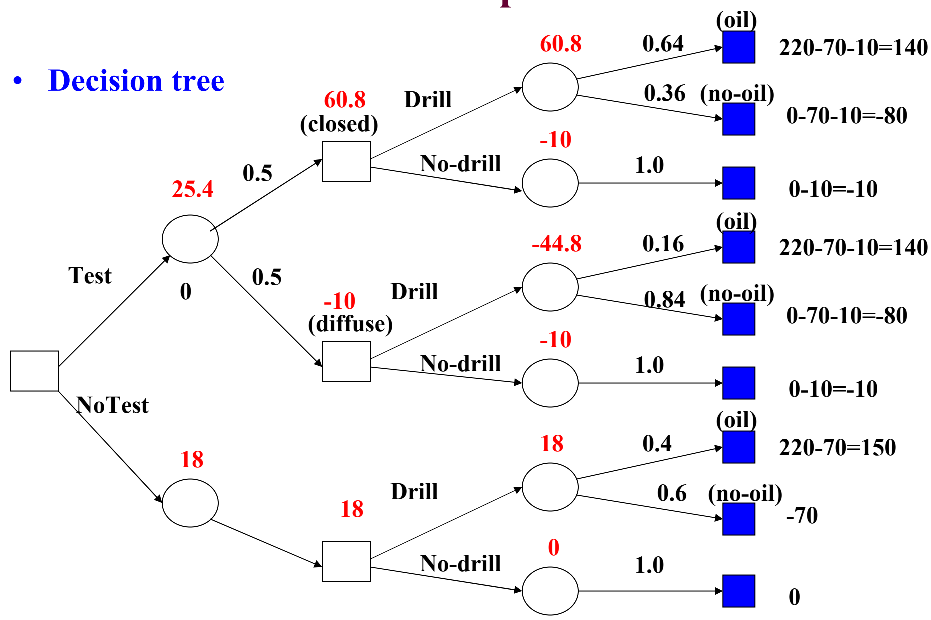
$\mathbf{P}(Test)$

$$P(Test = closed) = P(Test = closed \mid Oil = F)P(Oil = F) + P(Test = closed \mid Oil = T)P(Oil = T)$$

$$P(Test = diff) = P(Test = diff \mid Oil = F)P(Oil = F) + P(Test = diff \mid Oil = T)P(Oil = T)$$

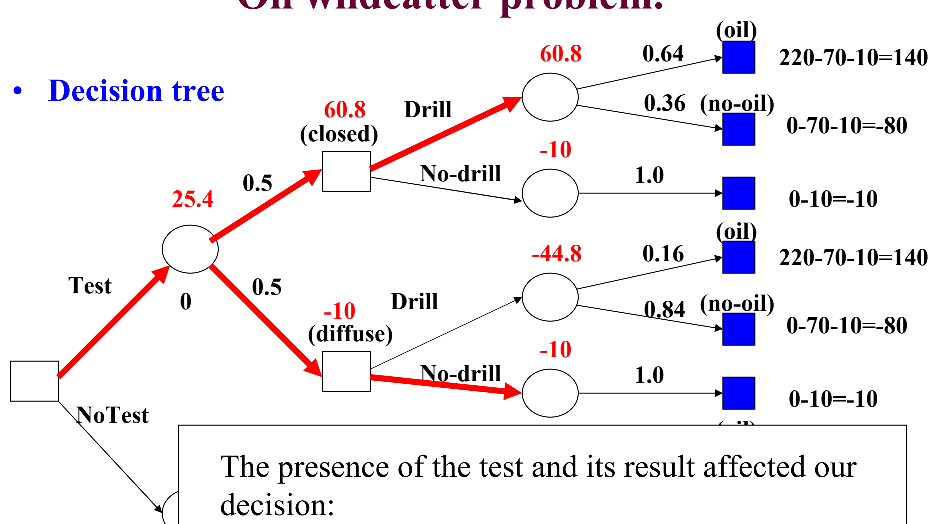
Oil wildcatter problem.

- **Decision tree**



Oil wildcatter problem.

- **Decision tree**



The presence of the test and its result affected our decision:

if test =closed then <u>drill</u>

if test=diffuse then <u>do not drill</u>

Value of information

- **When the test makes sense?**
- Only when its result makes the decision maker to change his mind, that is he decides not to drill.
- **Value of information:**
  - Measure of the goodness of the information from the test
  - Difference between the expected value with and without the test information
- **Oil wildcatter example:**
  - Expected value without the test = 18
  - Expected value with the test =25.4
  - Value of information for the seismic test = 7.4