## CS 1571 Introduction to AI
## Lecture 21

# Decision making in the presence of uncertainty

**Milos Hauskrecht**
milos@cs.pitt.edu
5329 Sennott Square



## Administration

- **Final exam**
  - **December 11, 2002 at 2:00-3:50pm**
  - **25 % of the grade**

# Decision-making in the presence of uncertainty

- Many real-world problems come down to the problem of **making decisions about our future actions in the presence of uncertainty**
- **Examples:** patient management, investments

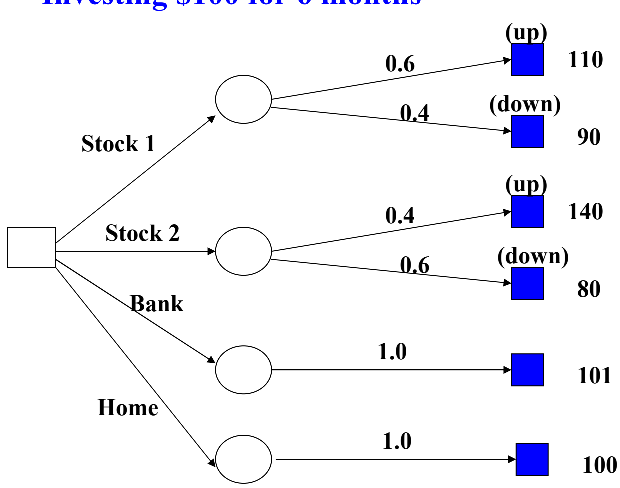
**Main issues:**

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



# Decision tree representation

**Investing $100 for 6 months**



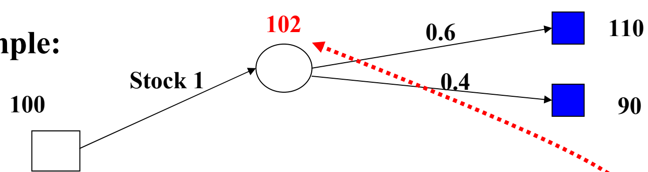
# Expected value

- Let X be a random variable representing the monetary outcome with a discrete set of values $\Omega_X$ .
- **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:**



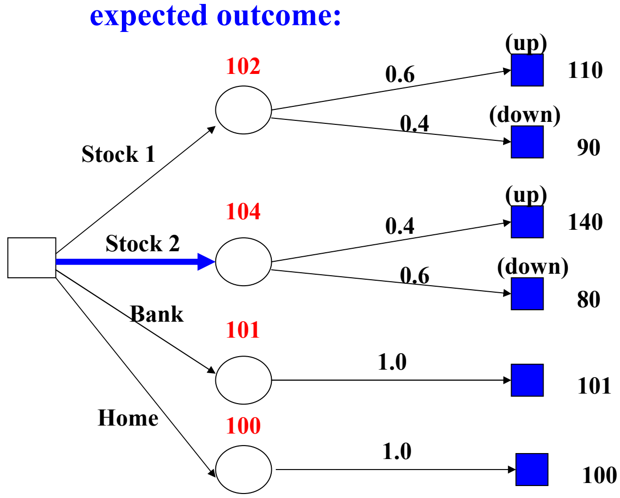

Expected value for the outcome of the Stock 1 option is:

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



# 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:**

1. Choose an action
2. Observe the stochastic outcome
3. And repeat step 1

**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

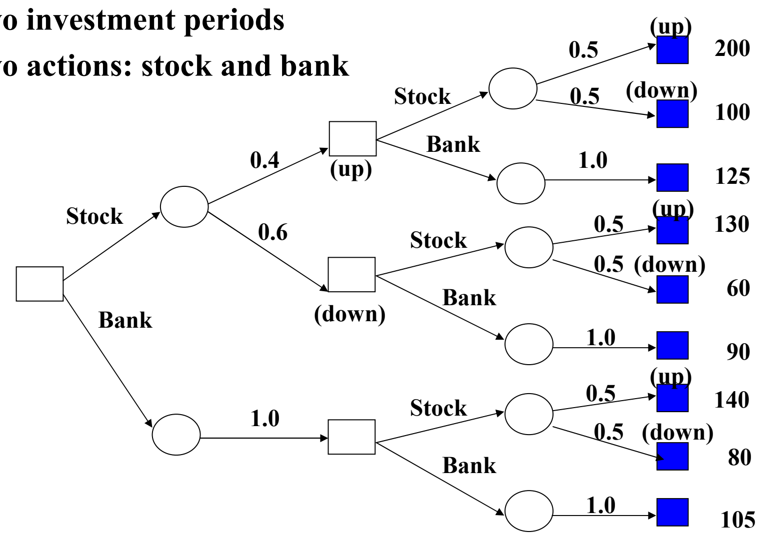
Algorithm is sometimes called **expectimax**



# Multi-step problem example

**Assume:**

- **Two investment periods**
- **Two actions: stock and bank**

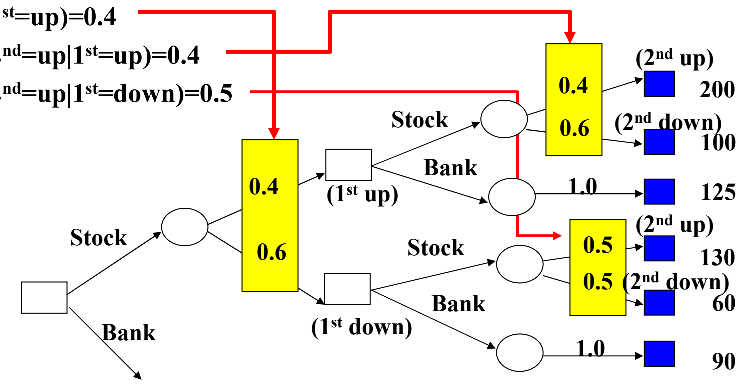


## Constructing the decision tree

- **Probabilities:**
  - **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**





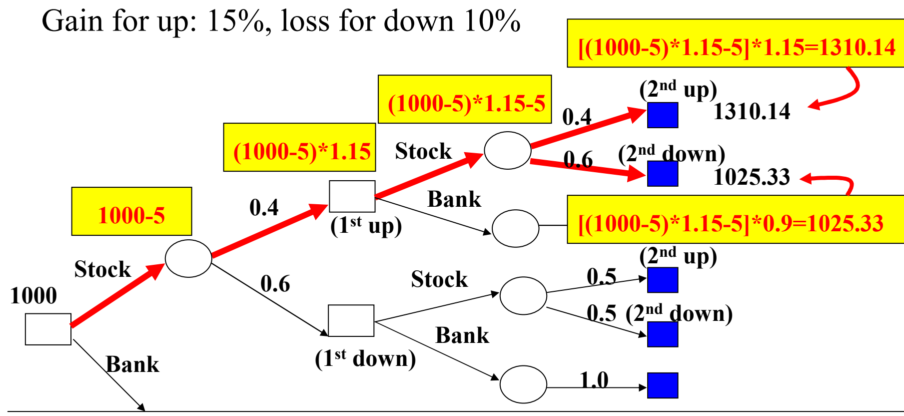

## Constructing the decision tree

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

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

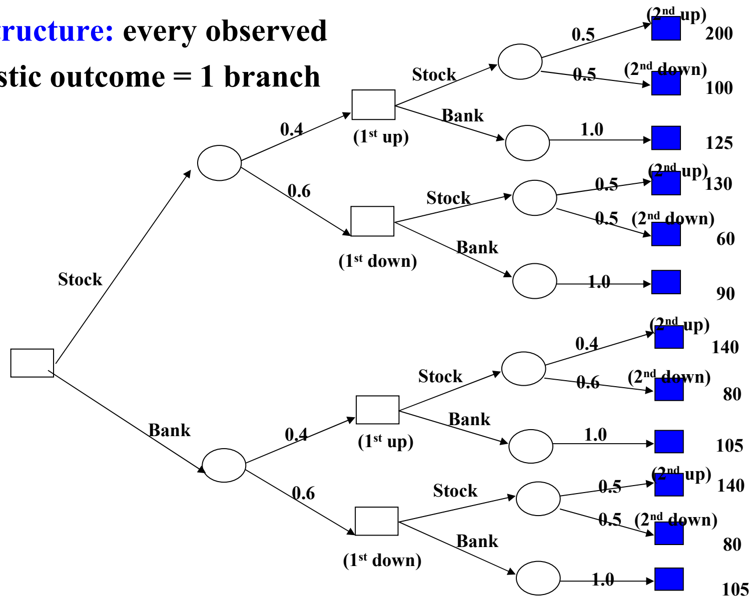
Fee per period: $5 paid at the beginning

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



# Constructing the decision tree

**Tree Structure:** every observed stochastic outcome = 1 branch

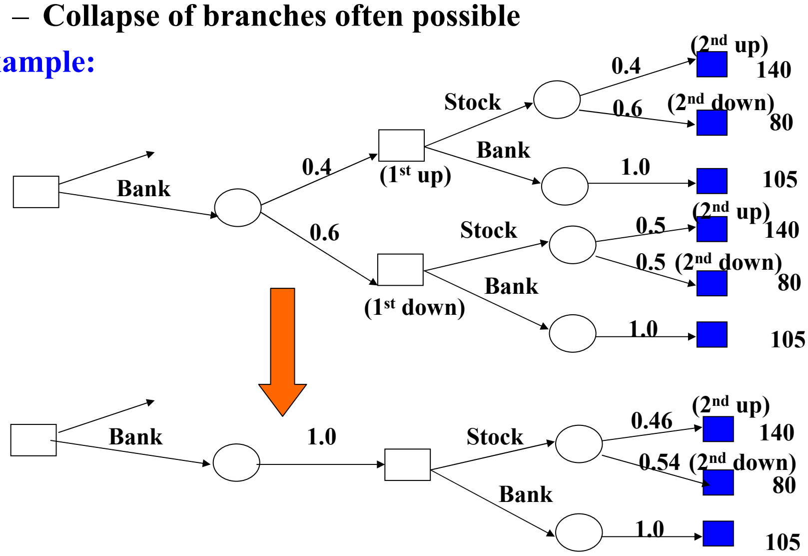





# Constructing the decision tree

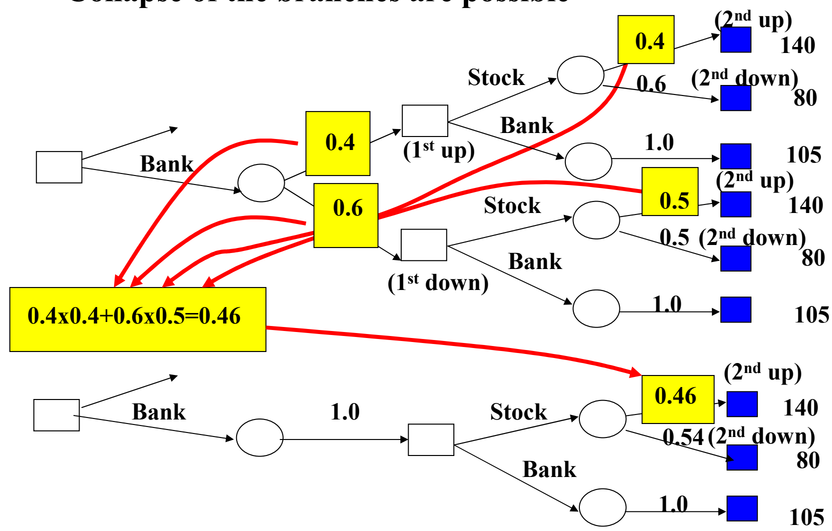
- **Structure**
  - Collapse of branches often possible

**Example:**



# Constructing the decision tree

- **Structure**
  - **Collapse of the branches are possible**






# Information-gathering actions

- **Some actions and their outcomes irreversibly change the world**
- **Information-gathering (exploratory) actions:**
  - **make an inquiry about the world**
  - **Benefit:** reduction in the uncertainty
- **Example: medicine**
  - Assume a patient is admitted to the hospital with some set of initial complaints
  - We 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:** Reduce the uncertainty of outcomes of treatments so that better treatment option can be chosen

## Decision-making with exploratory actions

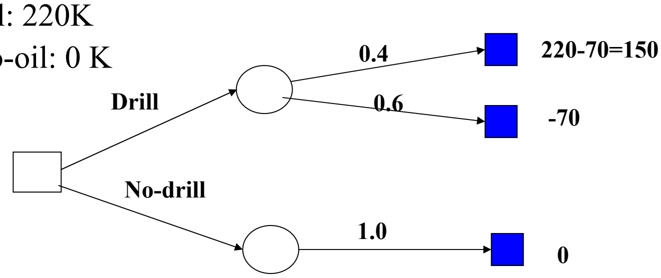
- **Exploratory actions can be represented and reasoned about the same way as other actions**.
- **Effect:** Information obtained through exploratory actions effects the probabilities of outcomes
  - Recall that the probabilities on later outcomes are conditioned on past observed outcomes and past actions



## Oil wildcatter problem.

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

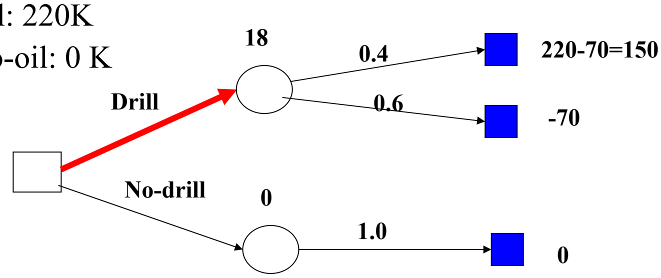
- **Chance of hitting an oil deposit:**
  - Oil: 40% $P(Oil = T) = 0.4$
  - No-oil: 60% $P(Oil = F) = 0.6$
- **Cost of drilling:** 70K
- **Payoffs:**
  - Oil: 220K
  - No-oil: 0 K



## 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)

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

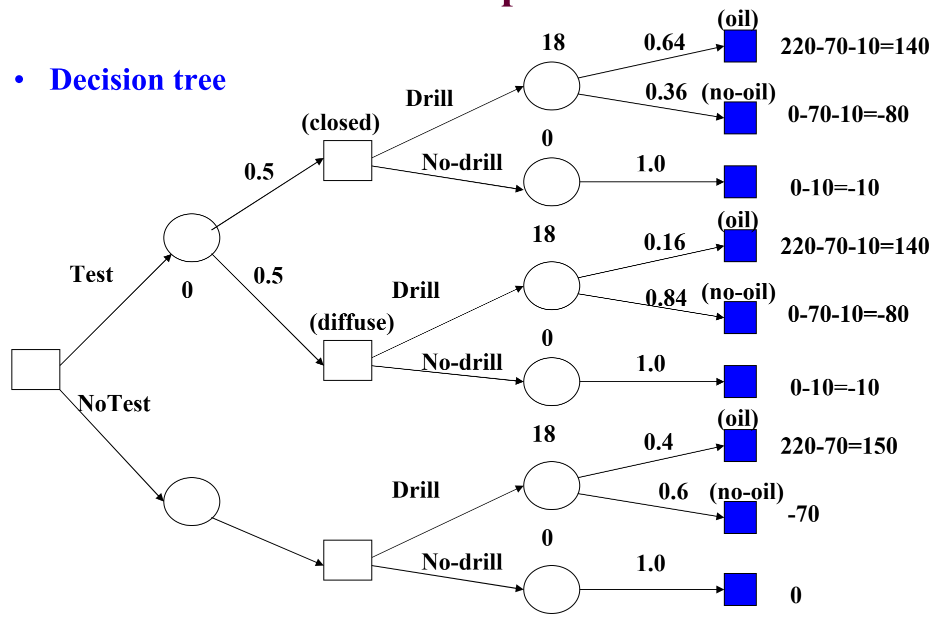
*Seismic resonance test pattern*

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

- **Test cost:** 10K

# Oil wildcatter problem.

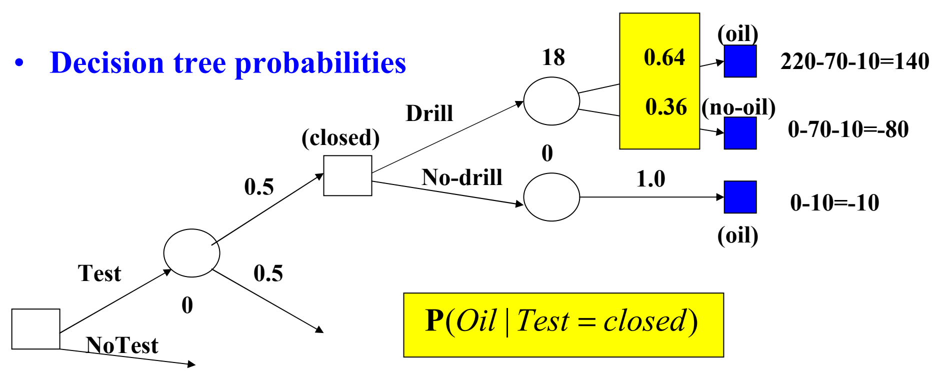
- **Decision tree**






# Oil wildcatter problem.

- **Decision tree probabilities**




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

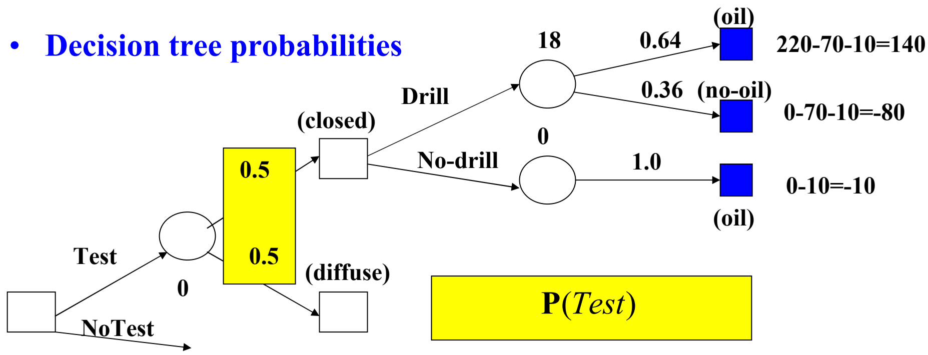
$$P(Oil = T \mid Test = closed) = \frac{P(Test = closed \mid Oil = T) P(Oil = T)}{P(Test = closed)}$$

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

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

# Oil wildcatter problem.

- **Decision tree probabilities**



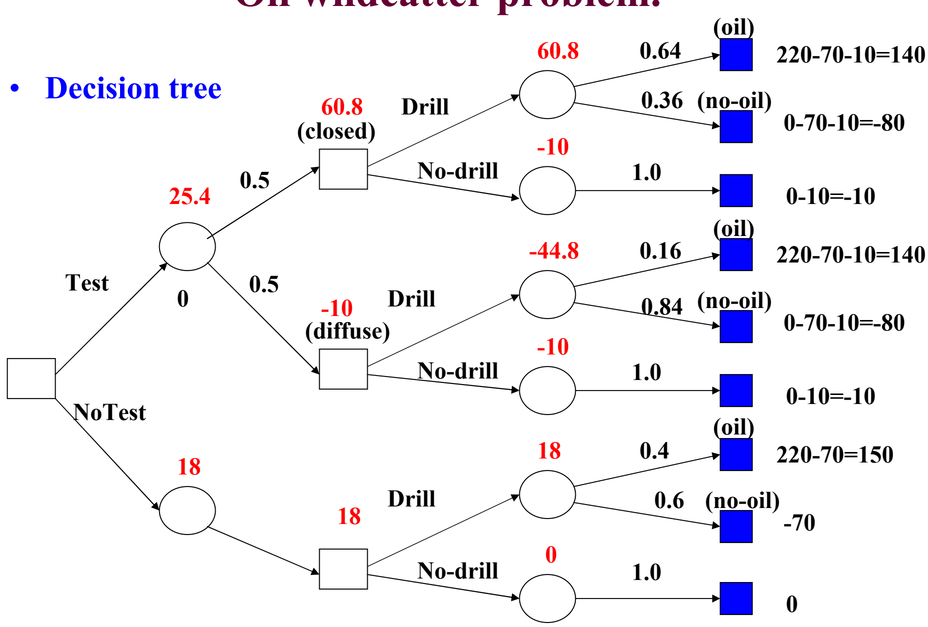

$$P(Test = closed) = P(Test = closed \mid Oil = F)P(Oil = F) + P(Test = closed \mid Oil = T)P(Oil = T)$$

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



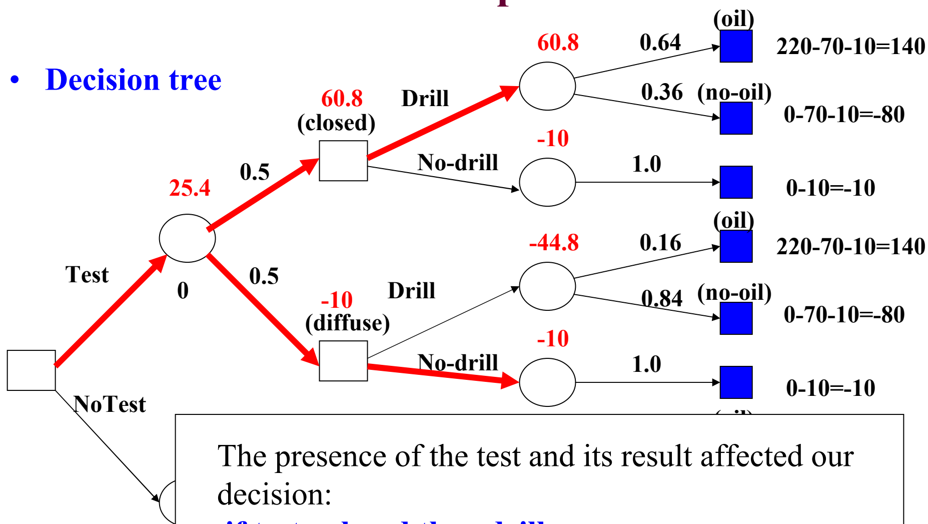
# Oil wildcatter problem.

- **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>**

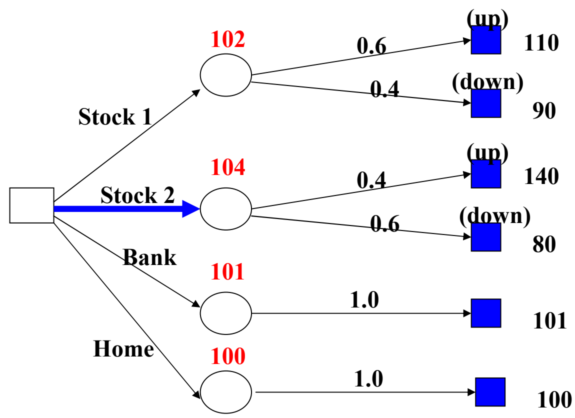


# Value of information

- **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

## Selection based on expected values

- **Until now:** The optimal action choice was the option that maximized the expected monetary value.
- **But is the expected monetary value always the quantity we want to optimize?**






## Selection based on expected values

- **Is the expected monetary value always the quantity we want to optimize?**
- **Answer:** Yes, but only if we are risk-neutral.

- But what if **we do not like the risk (we are risk-averse)?**
- In that case we may want to get the premium for undertaking the risk (of loosing the money)
- **Example:**
  – we may prefer to get $101 for sure against $102 in expectation but with the risk of loosing the money
- **Problem:** How to model decisions and account for risks?
- **Solution:** use **utility function, and utility theory**

# Utility function

- **Utility function (denoted U)**
  - Quantifies how we "value" outcomes, i.e., it reflects our preferences
  - Can be also applied to "value" outcomes other than money and gains (e.g. utility of a patient being healthy, or ill)
- **Decision making:**
  - uses expected utilities (denoted EU)

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

$U(X = x)$ the utility of outcome x

**Important !!!**

- Under some conditions on preferences **we can always design the utility function that fits our preferences**



# Utility theory

- Defines axioms on preferences that involve uncertainty and ways to manipulate them.
- Uncertainty is modeled through lotteries
  - **Lottery:**

    $$[p : A; (1-p) : C]$$

    - Outcome A with probability p
    - Outcome C with probability (1-p)
- The following six constraints are known as the axioms of utility theory. The axioms are the most obvious semantic constraints on preferences with lotteries.
- **Notation:**

  $\succ$ - preferable

  $\sim$ - indifferent (equally preferable)

## Axioms of the utility theory

- **Orderability:** Given any two states, the a 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).

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



## Axioms of the 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.

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

- **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.

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

## Utility theory

**If the agent obeys the axioms of the utility theory, then**

1. there exists a real valued function U such that:

$$U(A) > U(B) \Leftrightarrow A \succ B$$

$$U(A) = U(B) \Leftrightarrow A \sim B$$

2. The utility of the lottery is the expected utility, that is the sum of utilities of outcomes weighted by their probability

$$U[p : A; (1-p) : B] = pU(A) + (1-p)U(B)$$

3. Rational agent makes the decisions in the presence of uncertainty by maximizing its expected utility



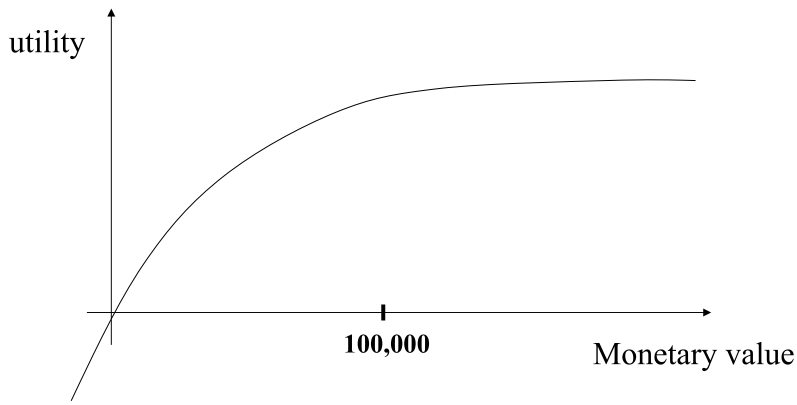
## Utility functions

**We can design a utility function that fits our preferences if they satisfy the axioms of utility theory.**

- **But how to design the utility function for monetary values so that they incorporate the risk?**
- **What is the relations between utility function and monetary values?**
- Assume we loose or gain $1000.
  - Typically this difference is more significant for lower values (around $100 -1000) than for higher values (~ $1,000,000)
- What is the relation between utilities and monetary value for a typical person?

# Utility functions

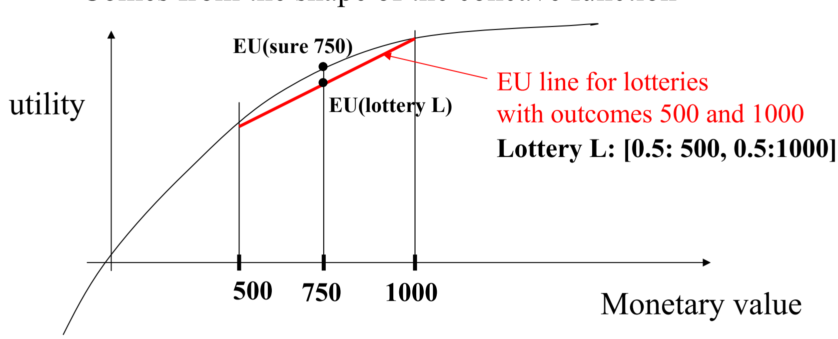
- What is the relation between utilities and monetary value for a typical person?
- Concave function that flattens at higher monetary values






# Utility functions

- How do we obtain the risk aversion (typical for people)?
- Comes from the shape of the concave function




Assume a lottery L [0.5: 500, 0.5:1000]

- Expected value of the lottery = 750
- Expected utility of the lottery EU(lottery L) < EU(sure 750)
- Risk averse – a bonus is required for undertaking the risk