Aalto University
School of Science

Decision making and problem solving – Lecture 1

- Decision trees
- Elicitation of probabilities

Liesiö, Punkka, Salo, Vilkkumaa

Motivation

- You have just revised some key concepts of probability calculus
  - Conditional probability
  - Law of total probability
  - Bayes' rule

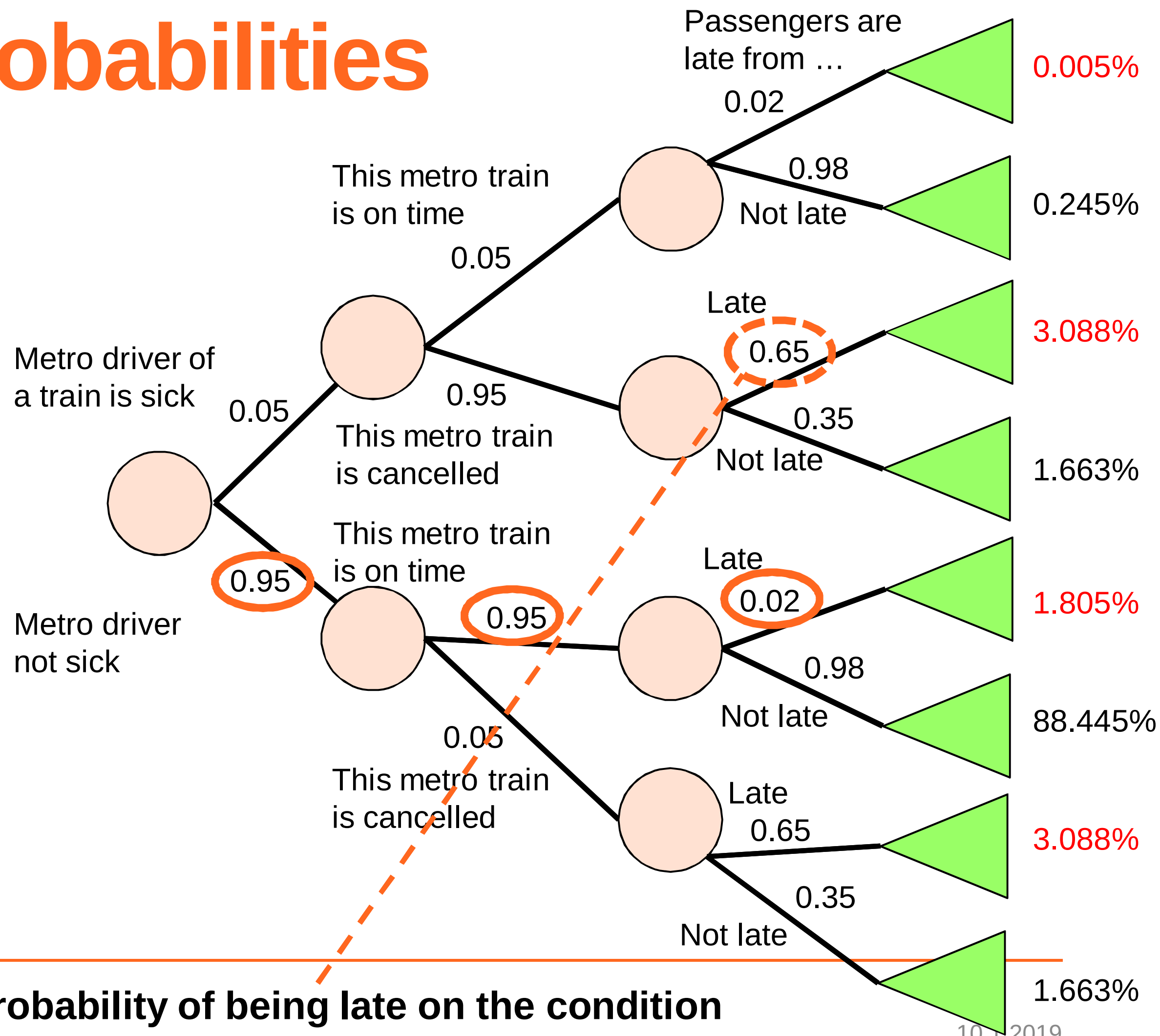
- This time:
  - How to build a probability-based model to support decision-making under uncertainty?
  - How to elicitate the probabilities needed for these models?

Why probabilities for modeling uncertainty?

- Decisions are often made under uncertainty
  - "How many train drivers should be trained, when future traffic is uncertain?"
  - "Should I buy an old or a new car, given that I only need an operational one and want to minimize costs = purchase price, maintenance & repair costs, selling price, etc.?"
  - "Should I buy my first my apartment now or postpone the decision, given that future interest rates, mortgage costs, personal income and apartment prices are uncertain?"
- Probability theory dominates the modeling of uncertainty in decision analysis
  - Well established rules for computations, understandable
  - Other models (e.g., evidence theory, fuzzy sets) exist, too

Conditional probabilities

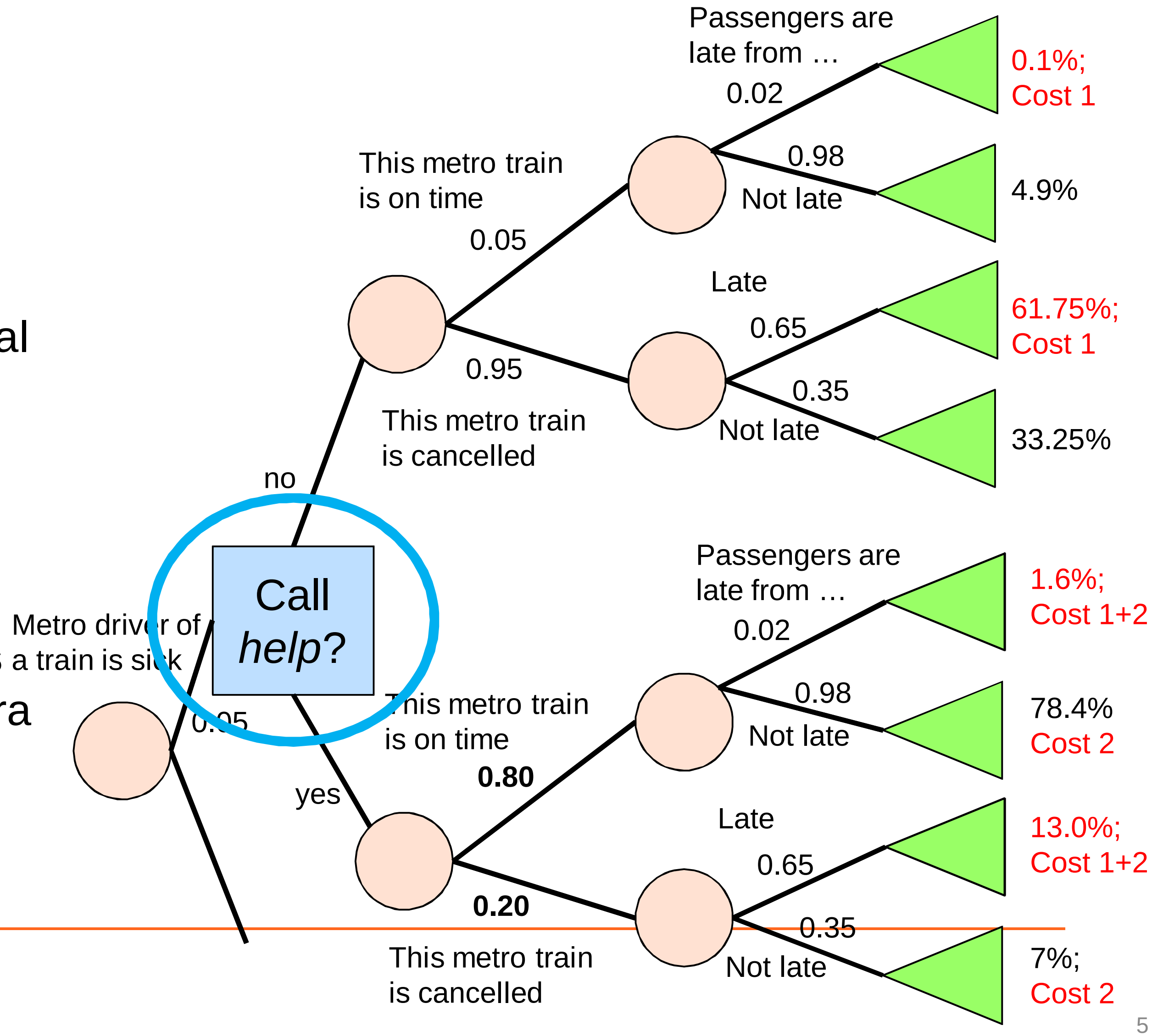
- The probabilities of sequential, mutually exclusive and collectively exhaustive events can be represented if form of a tree
- The probability of a sequence of events is obtained my multiplying the probabilities on the path
  - 0.95 x 0.95 x 0.02 = 1.805 %
- The total probability of being late is 7.985 %



The probability of being late on the condition that the train is cancelled (and the original driver is sick)

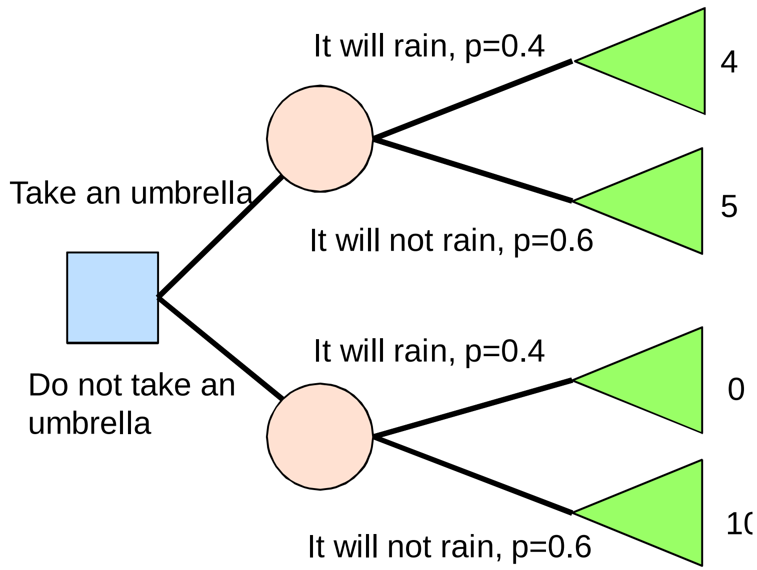
What if…

- We are interested in financial aspects and assume that being late results in unwanted financial consequences (Cost 1)?
  - numerical outcomes for states
- You had a possibility to influence the probability **p(this metro train is on time | metro driver of this train is sick)** by use of extra personnel (*help*) at a cost (Cost 2)?
  - Now the event probabilities depend on your decision



Decision trees

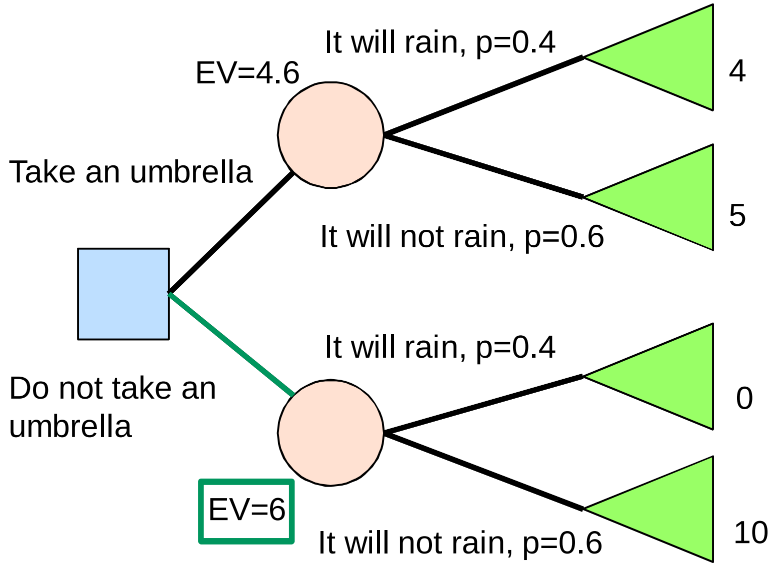
- Decision-making under uncertainty can be modeled by a decision tree
- Decision trees consist of
  - Decision nodes (squares) – DM can choose which arc to follow
  - Chance nodes (circles; cf. states of nature) – chance represented by probabilities dictates which arc will be followed (states of nature). The probabilities following a chance node must sum up to 1
  - Consequence nodes (triangles; resulting consequences) – at the end of the tree; describe the consequence (e.g., profit, cost, revenue, utility) of following the path leading to this node
- Decisions and chance events are displayed **in a logical temporal sequence from left to right**
  - Only chance nodes whose results are known can precede a decision node
- Each chain of decisions and chance events represents a possible outcome



Solving a decision tree

- A decision tree is solved by starting from the leaves (consequence nodes) and going backward toward the root:
  - At each chance node: compute the expected value at the node
  - At each decision node: select the arc with the highest expected value

- The optimal strategy consists of the arcs selected at decision nodes

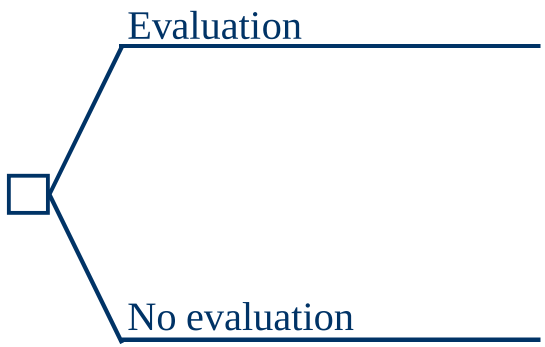


Example: Decision tree (1/12)

- Your uncle is going to buy a tractor. He has two alternatives:
  1. A new tractor (17 000 €)
  2. A used tractor (14 000 €)
- The engine of the old tractor may be defect, which is hard to ascertain. Your uncle estimates a 15 % probability for the defect.
- If the engine is defect, he has to buy a new tractor and gets 2000 € for the old one.
- Before buying the tractor, your uncle can take the old tractor to a garage for an evaluation, which costs 1 500 €.
  - If the engine is OK, the garage can confirm it without exception.
  - If the engine is defect, there is a 20 % chance that the garage does not notice it.
- Your uncle maximizes expected monetary value

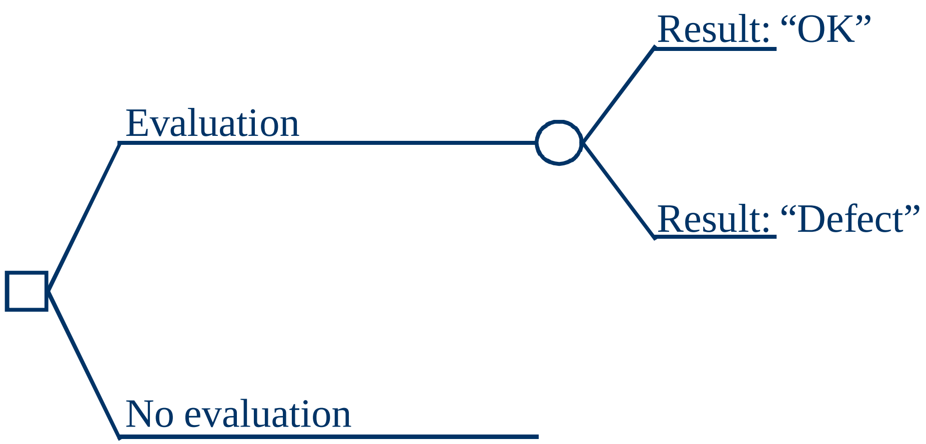
Example: Decision tree (2/12)

- Before making the buying decision and before you get to know the result of any uncertain event, you must **decide** upon taking the old tractor to a garage for an **evaluation**.
- The decision node ‘evaluation’ is placed leftmost in the tree



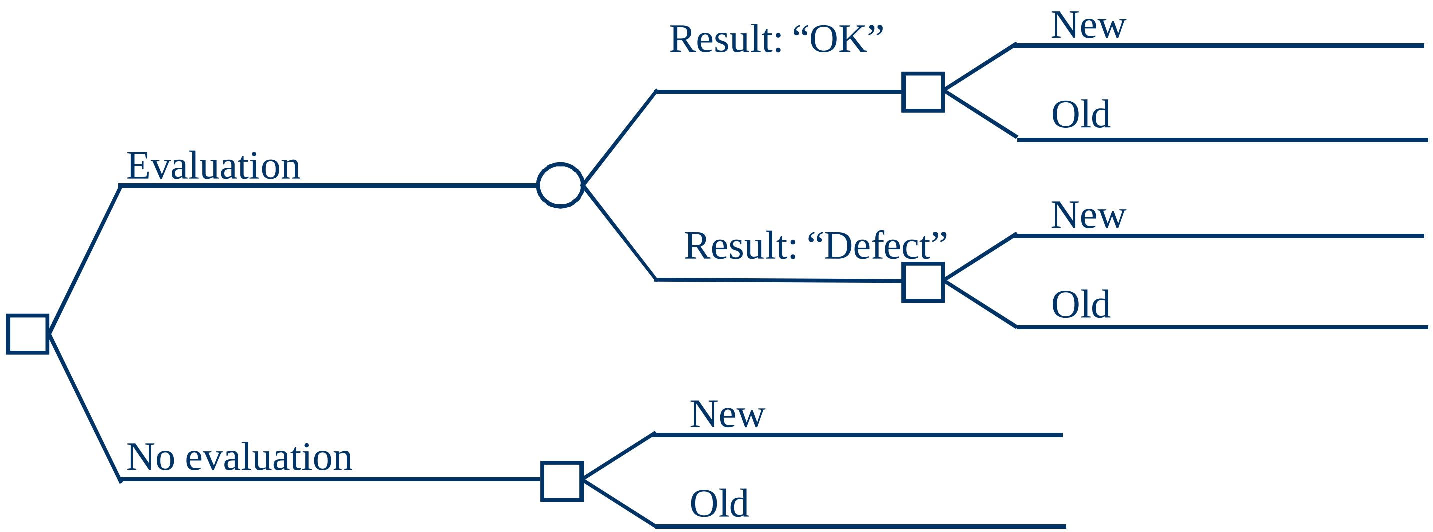
Example: Decision tree (3/12)

- If the old tractor is evaluated, your uncle receives the **results of the evaluation**



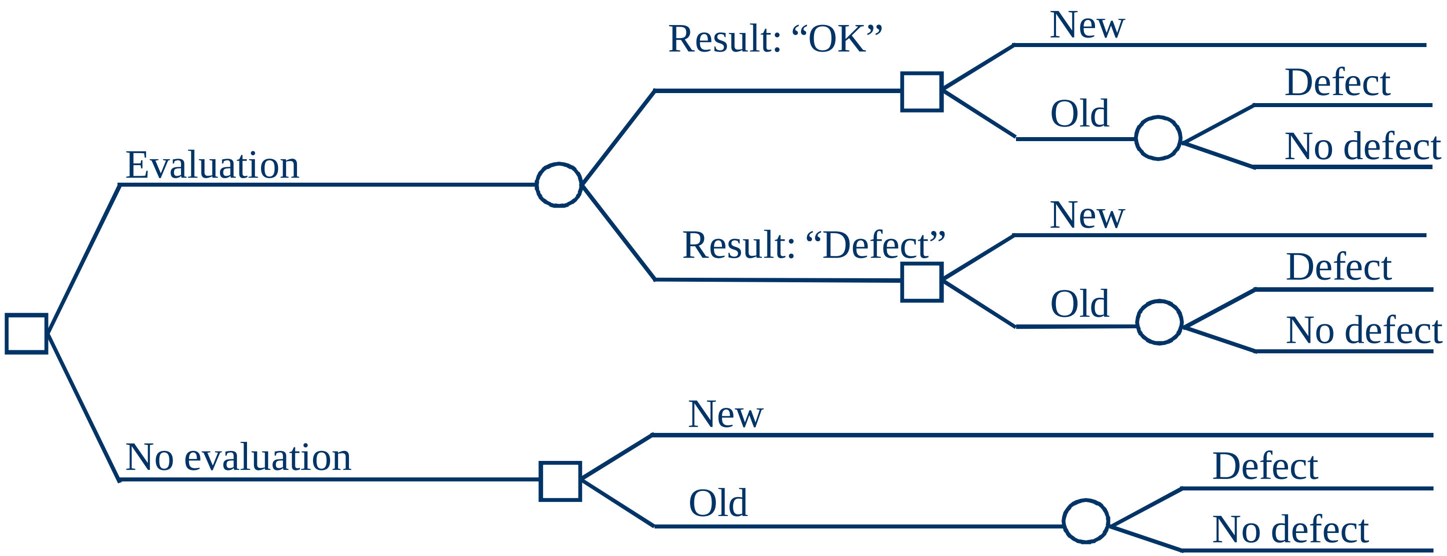
Example: Decision tree (4/12)

- The next step is to **decide which tractor** to buy



Example: Decision tree (5/12)

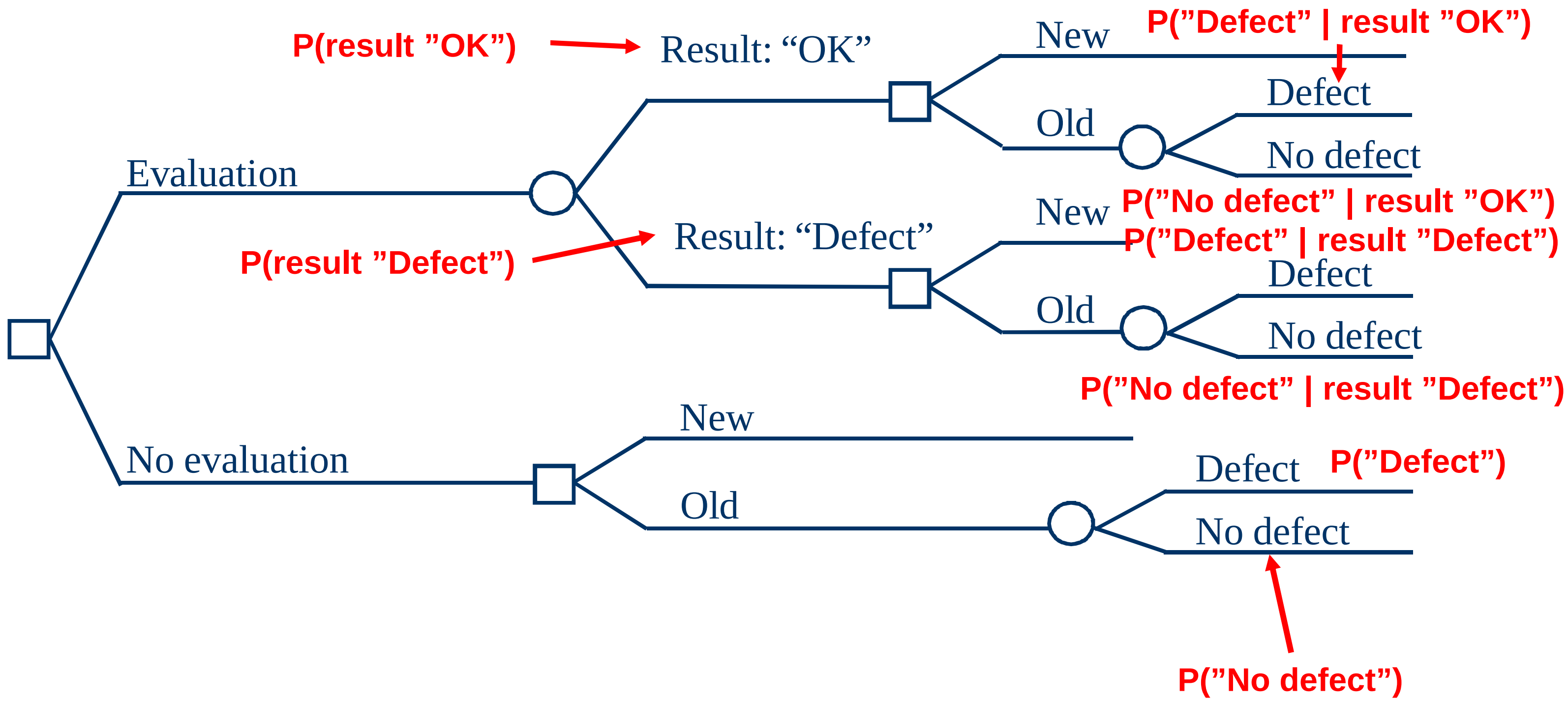
- …But the engine of the old tractor can be defect



- Now all chance nodes and decisions are in chronological order such that in each node, we can follow the path to the left to find out what we know

Example: Decision tree (6/12)

- We next need the probabilities for all outcomes of the chance nodes



Remember: Law of total probability

- If $E_1,\ldots,E_n$ are mutually exclusive and A = $\cup_i E_i$, then

$$P(A)=P(A|E_1)P(E_1)+\ldots+P(A|E_n)P(E_n)$$

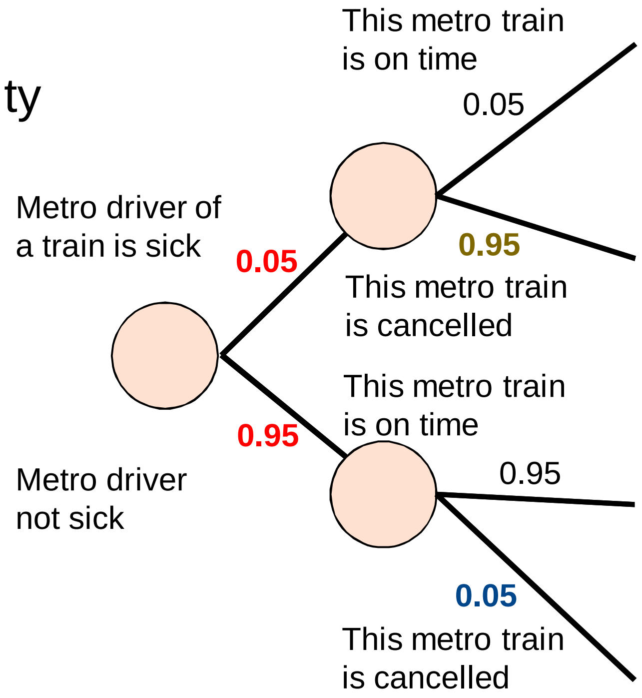
- Most frequent use of this law:
  - Probabilities P(A|B), P(A|B^c), and P(B) are known
  - These can be used to compute P(A)=P(A|B)P(B)+P(A|B^c)P(B^c)

Remember: Bayes' rule

- **Bayes' rule:** $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$
- Follows from
  1. The definition of conditional probability: $P(A|B) = \frac{P(A \cap B)}{P(B)}$, $P(B|A) = \frac{P(B \cap A)}{P(A)}$,
  2. Commutative laws: $P(B \cap A) = P(A \cap B)$.

Example: Bayes' rule

A metro train is cancelled (event *C*) and we have not had the opportunity to *call help*. What is the probability that the driver originally allocated to drive the train is sick (event *S*)? = What is $P(S|C)$?



Solution:

- P(S)=0.05, P(S^c)=0.95, P(C|S)=0.95, P(C|S^c)=0.05

Law of total probability: **P(C)**=P(C|S)P(S)+P(C|S^c) P(S^c)=
0.95 x 0.05 + 0.05 x 0.95 = 0.095

Bayes' rule: $P(S|C) = \frac{P(C|S)P(S)}{P(C)} = \frac{0.95 \cdot 0.05}{0.095}$ = 50%

Example: Decision tree (7/12)

- Solve all probabilities. You know that
  - ”Your uncle estimates a 15 % probability for the defect.” => P(Defect)=0.15
  - “If the engine is OK, the garage can confirm it without exception.” => P(result “OK” | No defect)=1
  - “If the engine is defect, there is a 20 % chance that the garage does not notice it.” => P(result “OK” | Defect)=0.20

$$P(\text{result "OK"}) = P(\text{result "OK"} \mid \text{No defect}) \cdot P(\text{No defect}) + P(\text{result "OK"} \mid \text{Defect}) \cdot P(\text{Defect})$$

$$= 1.0 \cdot 0.85 + 0.20 \cdot 0.15 = 0.88$$

$$P(\text{result "defect"}) = 1 - P(\text{result "OK"}) = 0.12$$

$$P(\text{Defect} \mid \text{result "OK"}) = \frac{P(\text{result "OK"} \mid \text{Defect}) \cdot P(\text{Defect})}{P(\text{result "OK"})} = \frac{0.20 \cdot 0.15}{0.88} \approx 0.034$$

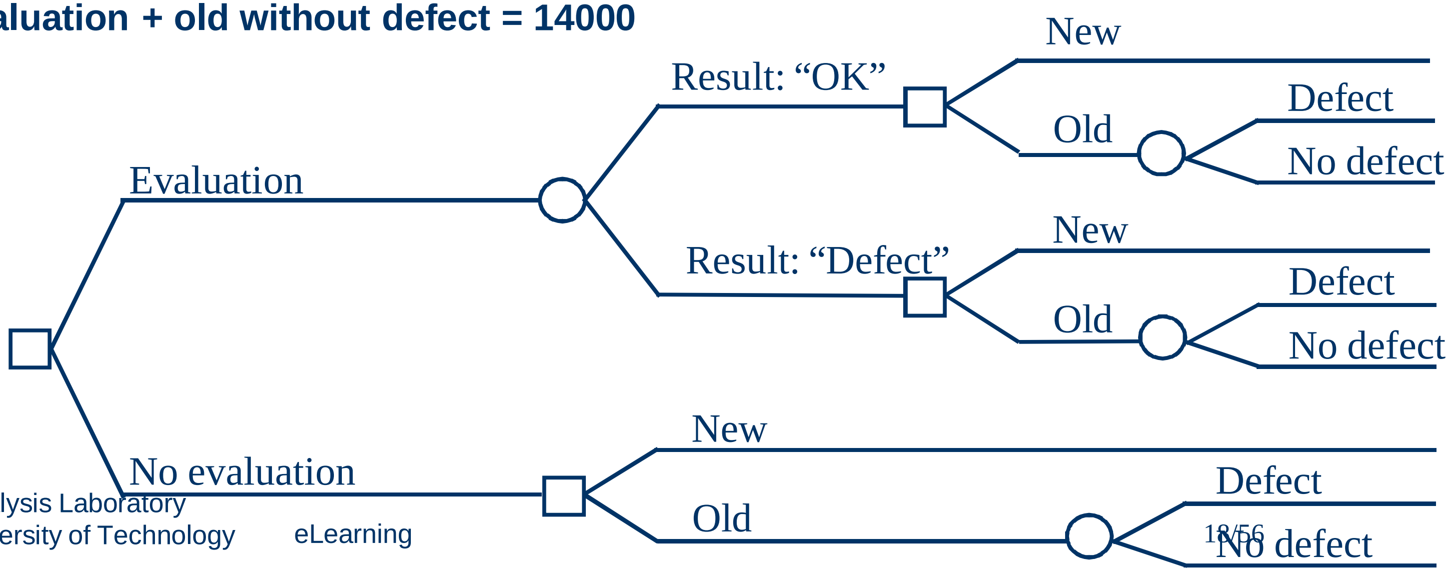
$$P(\text{No defect} \mid \text{result "OK"}) = 1 - 0.034 = 0.966$$

$$P(\text{Defect} \mid \text{result "defect"}) = \frac{P(\text{result "defect"} \mid \text{Defect}) \cdot P(\text{Defect})}{P(\text{result "defect"})} = \frac{0.80 \cdot 0.15}{0.12} = 1.00$$

$$P(\text{No Defect} \mid \text{result "defect"}) = 1 - 1 = 0$$

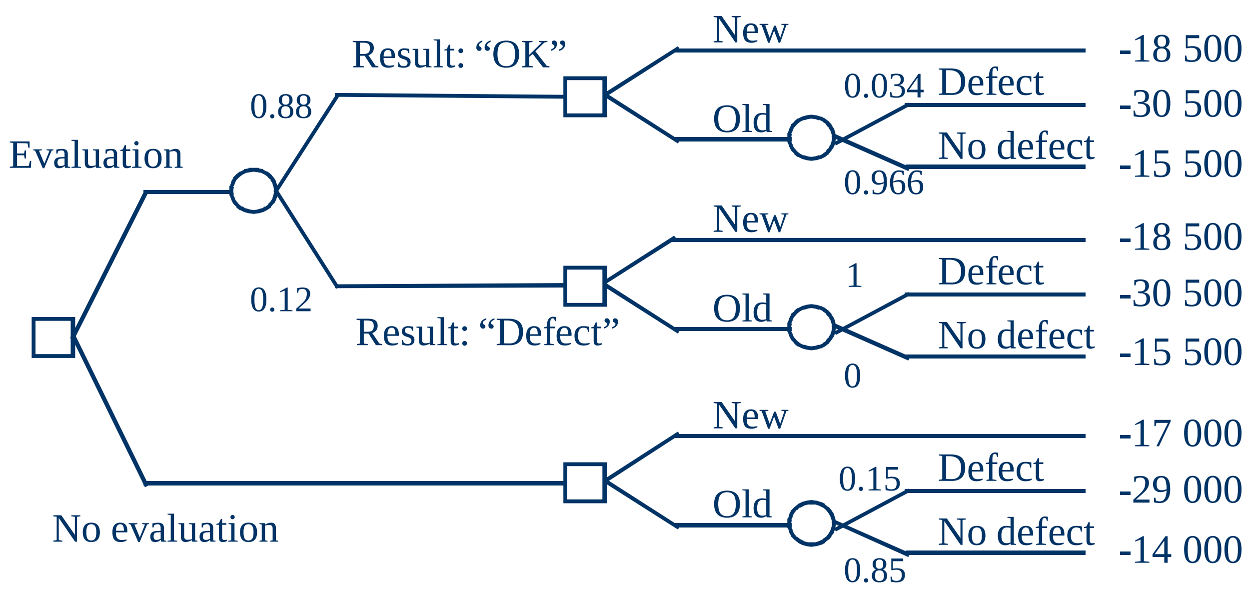
Example: Decision tree (8/12)

- Compute monetary values for each **end node**
  - **Evaluation + new = 1500 + 17000 = 18500**
  - **Evaluation + old with defect = 1500 + 14000 – 2000 + 17000 = 30500**
  - **Evaluation + old without defect = 1500 + 14000 = 15500**
  - **No evaluation + new = 17000**
  - **No evaluation + old with defect = 14000 – 2000 + 17000 = 29000**
  - **No evaluation + old without defect = 14000**



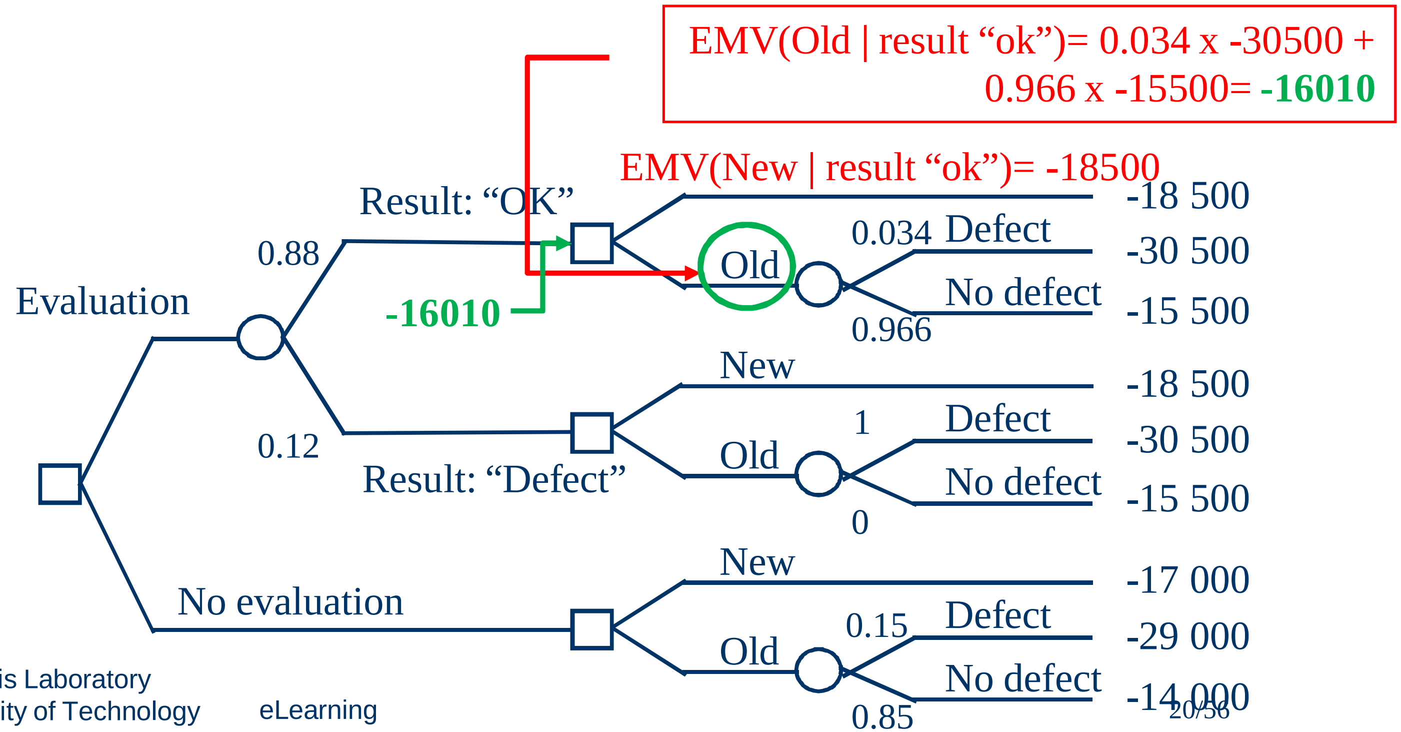
Example: Decision tree (9/12)

- We now have a decision tree presentation of the problem



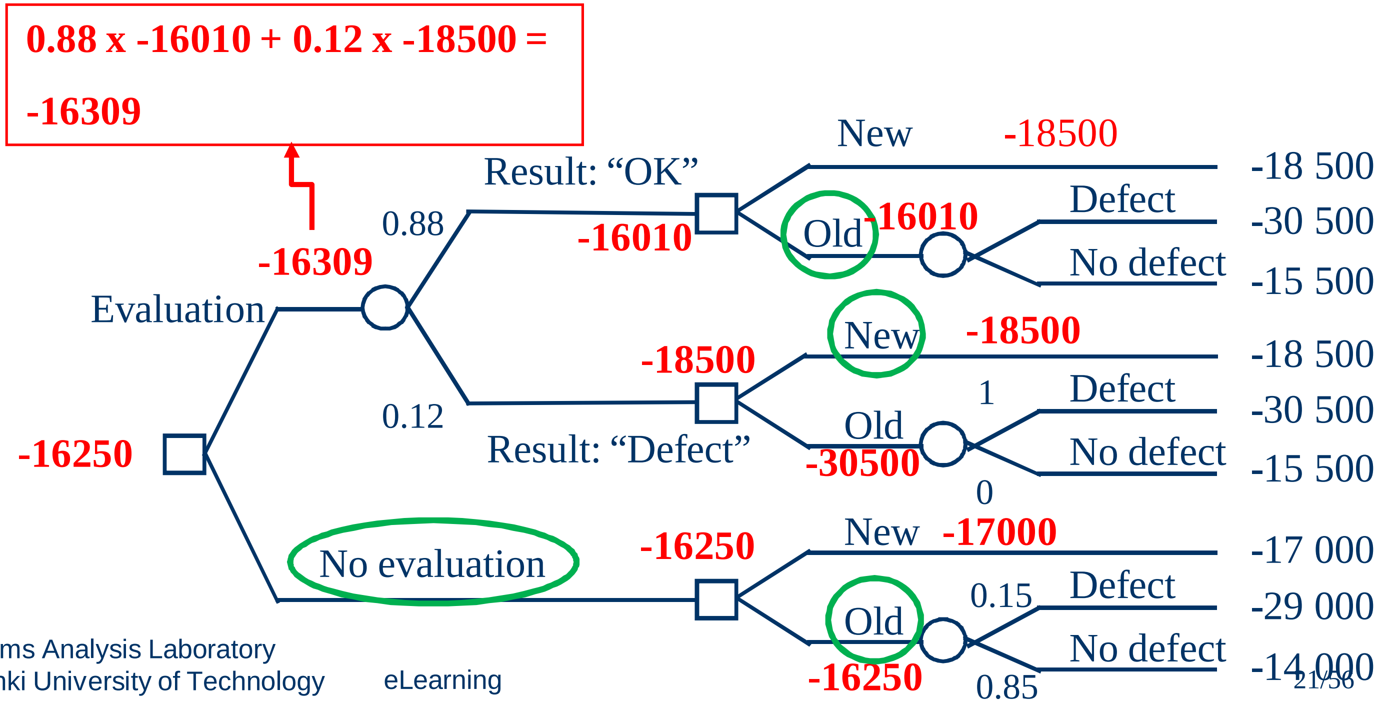
Example: Decision tree (10/12)

- **Starting from the right**, compute expected monetary values for each decision
- Place the value of the better decision to the decision node



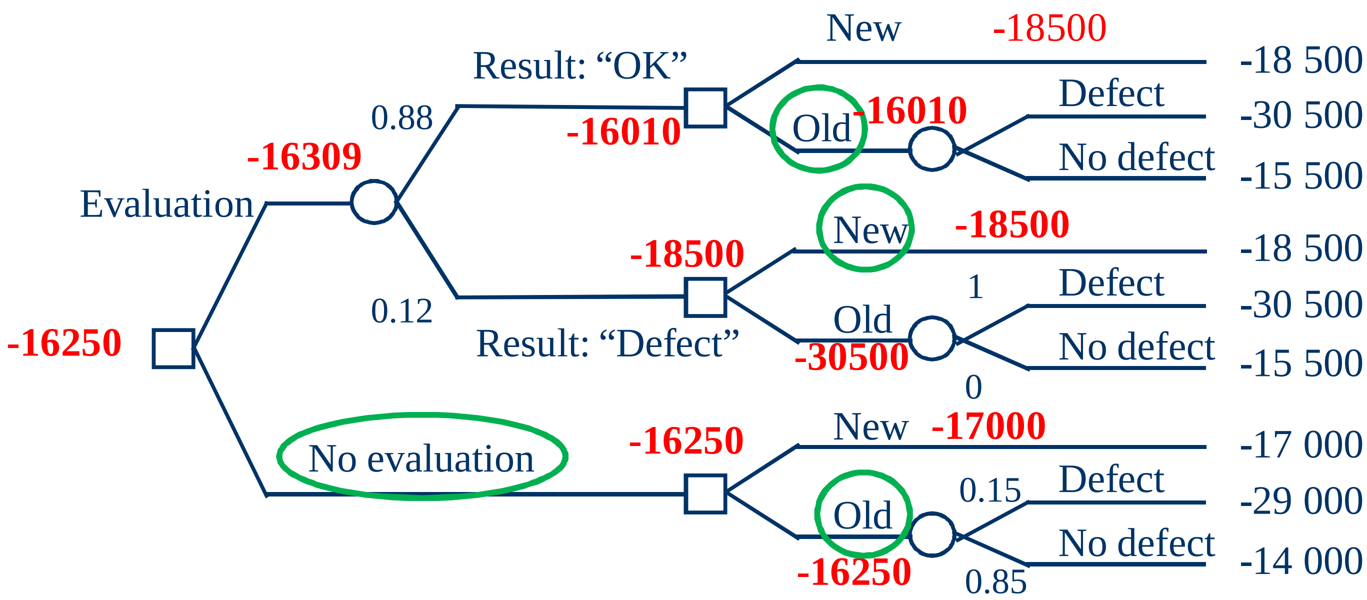
Example: Decision tree (11/12)

- Starting from the right, compute expected monetary values for each decision
- Place the value of the better decision to the decision node



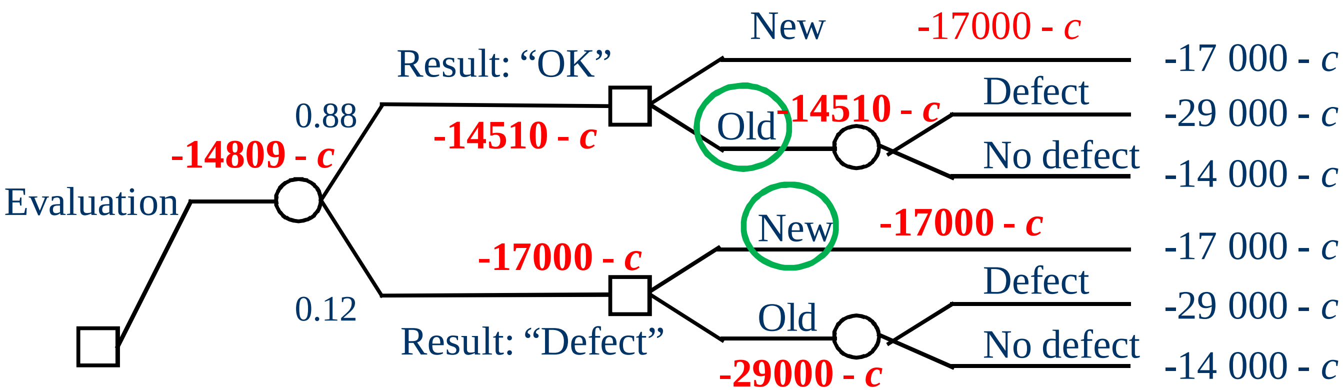
Example: Decision tree (12/12)

- The optimal solution is to buy the old tractor without evaluating it



… How much should we pay for the *sample information* by the garage?

- The expected monetary value was higher without evaluating the old tractor
- Determine evaluation cost *c* so that you are **indifferent** between
  1. Not taking the old tractor for an evaluation (EMV = -16250€)
  2. Taking the old tractor for an evaluation



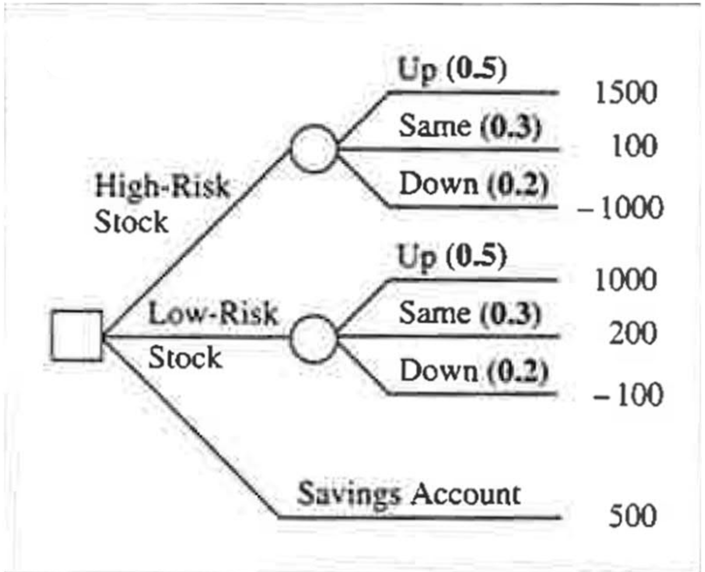
- Indifference, when EMVs equal: -16250 = -14809 – c => c = 1441€
  - Expected value of sample information = Expected value with sample information – Expected value without sample information = -14809€ - (-16250€) = 1441€

Example: expected value of ***perfect information***

- You are considering between three investment alternatives: high-risk stock, low-risk stock, and savings account
- Savings account: certain payoff of 500€
- Stocks:
  - 200€ brokerage fee
  - Payoffs depend on market conditions

	Up	Same	Down
High-risk	1700	300	-800
Low-risk	1200	400	100
Probability	0.5	0.3	0.1

Decision tree



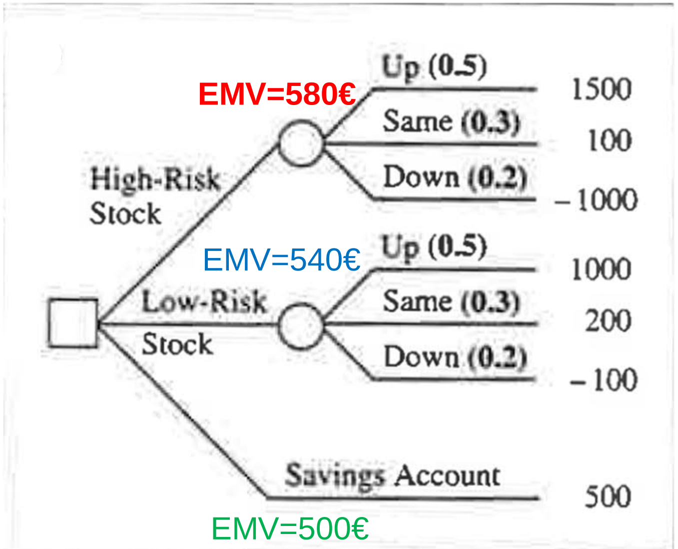
Source: Clemen, R.T. (1996): Making Hard Decisions: An Introduction to Decision Analysis, 2nd edition, Duxbury Press, Belmont.

Example: investing in the stock market

- The expected monetary values (EMVs) for the different alternatives are
  - HRS: 0.5·1500+0.3·100-0.2·1000=580
  - LRS: 0.5·1000+0.3·200-0.2·100=540
  - Savings Account: 500

→ It is optimal* to invest in **high-risk** stock

Decision tree



Source: Clemen, R.T. (1996): Making Hard Decisions: An Introduction to Decision Analysis, 2nd edition, Duxbury Press, Belmont.

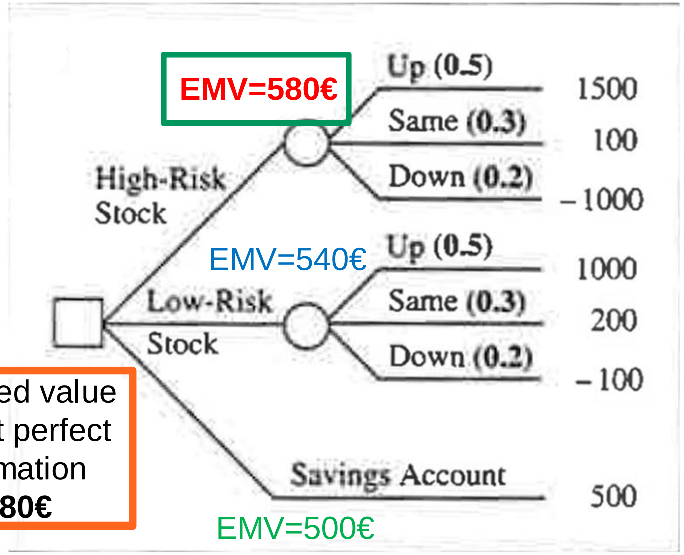
* Assuming you are *risk-neutral* !!! – *risk attitudes* discussed later on this course

Expected value of perfect information

- How much could the expected value be expected to increase, if
  - Additional information about the uncertainties was received before the decision
  - The decision would be made according to this information?
    - Note: this analysis is done before any information is obtained
- Perfect information: certain information about how the uncertainties are resolved – ”if we could choose after we know the state of the world”
  - Expected value of perfect information = Expected value with perfect information – Expected value without perfect information
- **Expected value of perfect information** is computed through a reversed decision tree in which all chance nodes precede all decision nodes

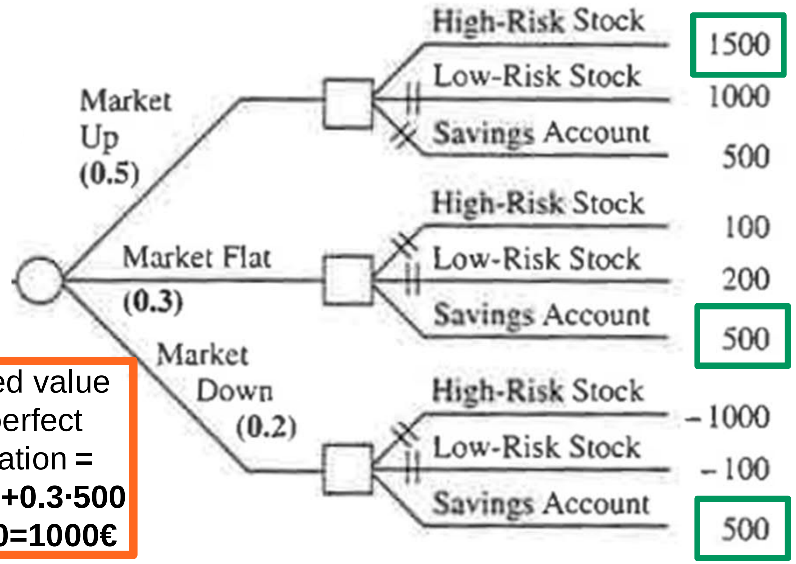
Expected value of perfect information

Decision tree



Expected value without perfect information
=580€

Reversed decision tree: you know the state of the world when making the decision(s)



Expected value with perfect information =
0.5·1500+0.3·500
+0.2·500=1000€

Expected value of perfect information
= 1000€ - 580€ = 420€

Probability assessment

- Use a few minutes to answer ten probability assessment questions
  - You have either questionnaire sheet A or B

- Do not communicate with others

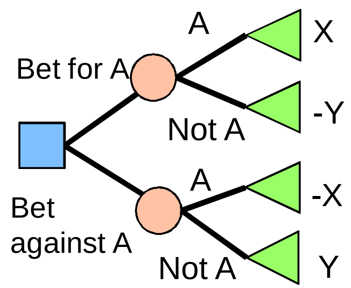
- Do not look up the answers on the internet

Estimation of probabilities

- How to obtain the probabilities needed in decision models?
  1. If possible, use objective data
  2. If objective data is not available, obtain subjective probability estimates from experts through
     - Betting approach
     - Reference lottery
     - Direct judgement

Estimation of probabilities: Betting approach

- **Goal:** to estimate the probability of event A
  - E.g., A="GDP growth is above 3% next year" or A="Sweden will join NATO within the next five years"

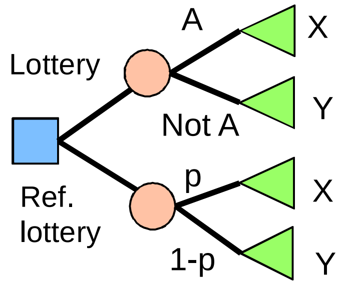


- **Betting approach:**
  - Bet for A: win X € if A happens, lose Y € if not
    - Expected monetary value $X \cdot P(A) - Y \cdot [1 - P(A)]$
  - Bet against A: lose X € if A happens, win Y € if not
    - Expected monetary value $-X \cdot P(A) + Y \cdot [1 - P(A)]$
  - Adjust X and Y until the respondent is indifferent between betting for or against A
  - Assuming risk-neutrality(*, the expected monetary values of betting for or against A must be equal:

$$X \cdot P(A) - Y \cdot [1 - P(A)] = -X \cdot P(A) + Y \cdot [1 - P(A)] \Rightarrow P(A) = \frac{Y}{X + Y}$$

*)A strong assumption

Estimation of probabilities: Reference lottery



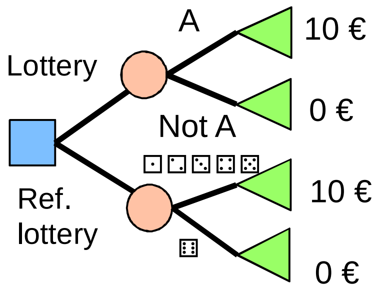
- Lottery:
  - Win X if A happens
  - Win Y if A does not happen
  - X is preferred to Y
- Reference lottery:
  - Win X with (known) probability p
  - Win Y with (known) probability (1-p)
  - Probability p can be visualized with, e.g., a wheel of fortune
- Adjust *p* until the respondent is **indifferent** between the **two lotteries**:

$$X \cdot P(A) + Y \cdot [1 - P(A)] = X \cdot p + Y \cdot [1 - p] \Rightarrow P(A) = p$$

- Here, the respondent's risk attitude does not affect the results (shown later)

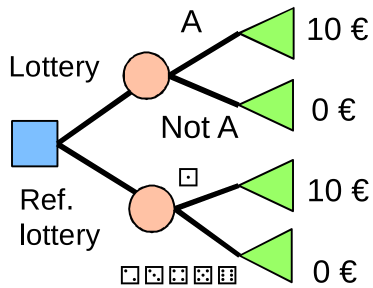
Reference lottery: example

❑ **Event A:** ”HIFK wins Jokerit”



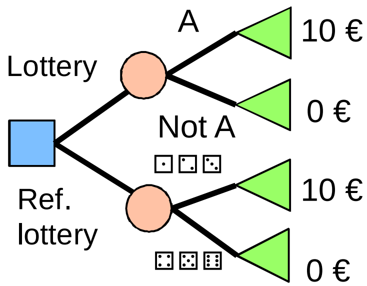
The respondent chooses the reference lottery:

$$10 \cdot P(A) < 10 \cdot \frac{5}{6}$$



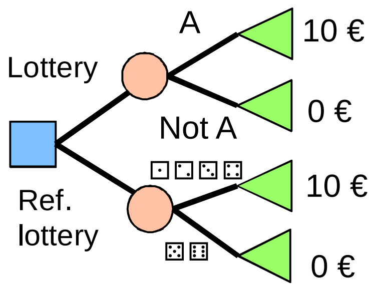
The respondent chooses the lottery:

$$10 \cdot P(A) > 10 \cdot \frac{1}{6}$$



Chooses the lottery:

$$P(A) > \frac{1}{2}$$



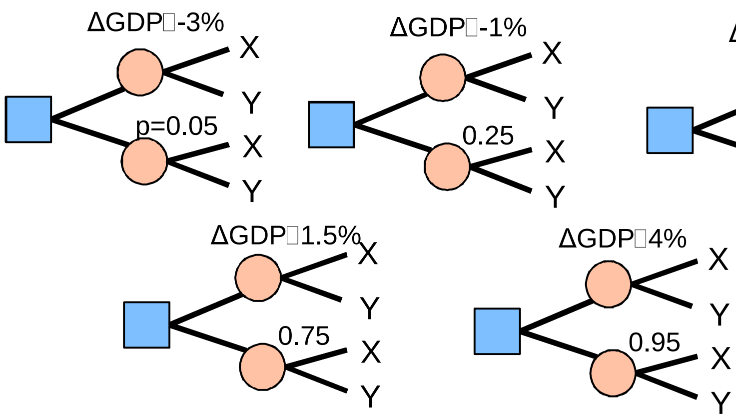
Chooses the reference lottery:

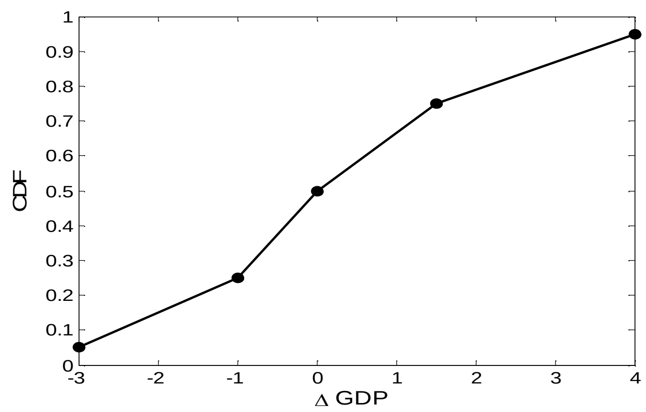
$$P(A) < \frac{2}{3}$$

These four answers revealed to probability estimate of A to be in (0.5, 0.67). Further questions should reveal the respondent’s estimate for P(A)

Estimation of continuous probability distributions

- A continuous distribution can be approximated by estimating several event probabilities (X is preferred to Y)
- Example:
  - Goal: to assess the distribution of the change in GDP (ΔGDP) in Finland next year
  - Means: elicitation of probability p for five different reference lotteries





Estimation of continuous probability distributions

- Often experts assess the descriptive statistics of the distribution directly, e.g.,
  - The feasible range (min, max)
  - Median f_{50} (i.e., $P(X<f_{50})=0.5$)
  - Other quantiles (e.g., 5%, 25%, 75%, 95%)
- In the previous example:
  - ”The 5% and 95% quantiles are f_5 =-3% and f_{95} = 4%”
  - ”The change in GDP is just as likely to be positive as it is to be negative”
  - ”There is a 25% chance that the change in GDP is below -1%”
  - ”There is a 25% chance that the change in GDP is above 1.5%”

Summary

- Decision trees are probability-based models to support decision-making under uncertainty
  - Which decision alternative should I choose?
  - How much would I be willing to pay for perfect information or (imperfect) sample information about how the uncertainties are resolved?

- Subjective probability assessments often required
  - Probability elicitation techniques require some effort