Aalto University
School of Science

MS-E2135
Decision Analysis
Lecture 1

- *Decision trees*
- *Elicitation of probabilities*

Salo, Liesiö, Punkka, Vilkkumaa

Motivation

- On Monday, you revisited key concepts of probability calculus
  - Conditional probability
  - Law of total probability
  - Bayes' rule

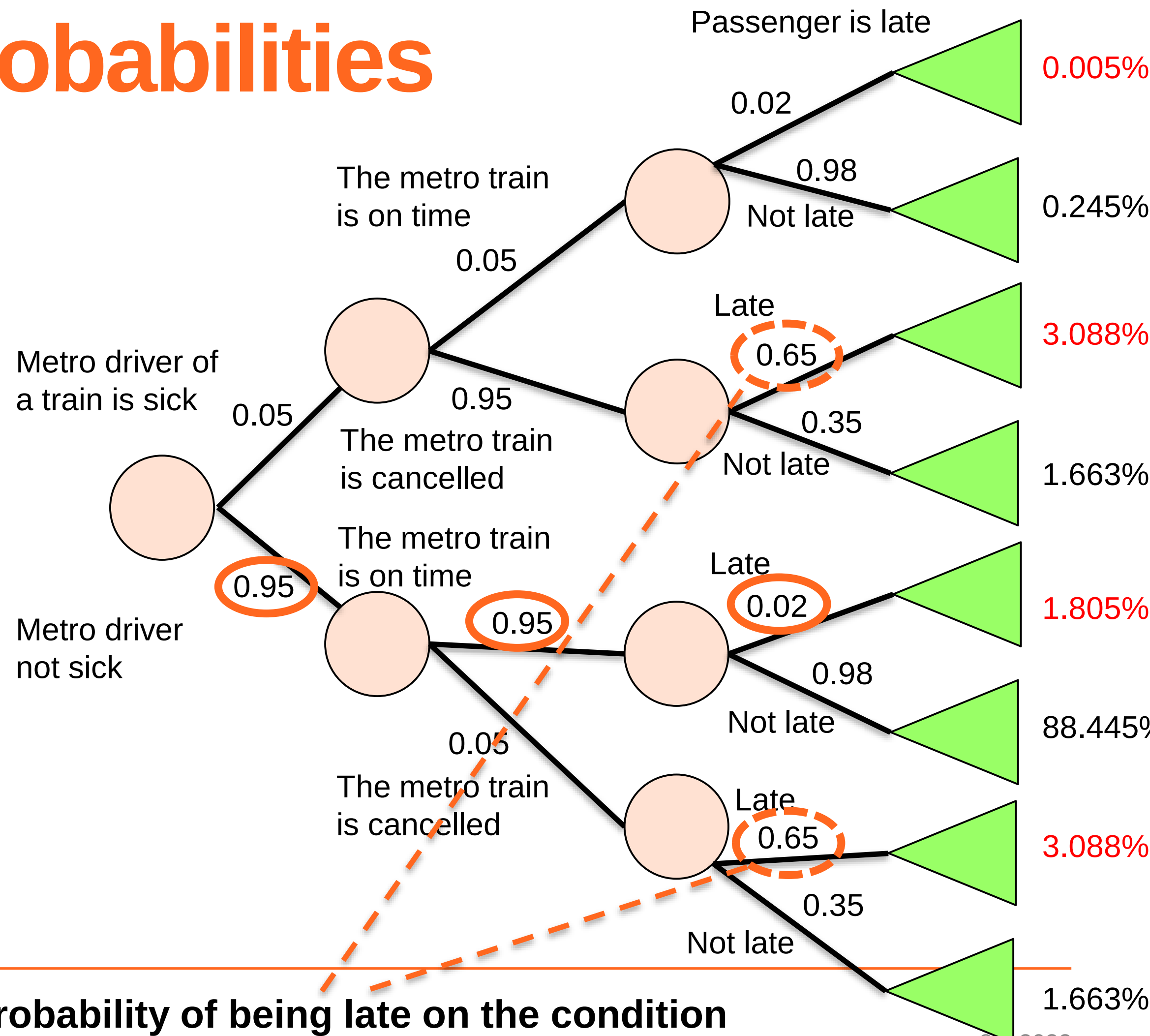
- We next address following questions
  1. How to build a probability-based model to support decisions under uncertainty?
  2. How to elicit <u>subjective</u> probabilities that are needed for these models?

Why probabilities for modeling uncertainty?

- ❑ Most decisions involve uncertainties
  - ❑ *“How many employees should be hired, when future demand is uncertain?”*
  - ❑ *“Should I buy an old or a new car, assuming that I need a functional one and want to minimize total costs (incl. purchase price, maintenance & repair, fuel, insurance, selling price)?”*
  - ❑ *“Should I buy my first my apartment now or postpone this decision, given the uncertainties about interest rates, mortgage costs, income level and housing market?”*
- ❑ In decision analysis, uncertainties are modelled through probabilities
  - – Theoretically rigorous ➔ Sound rules for inference
  - – Understandable, explainable, compatible with statistical analyses
  - – Yet there are other models as well (e.g., evidence theory, fuzzy sets) well

Conditional probabilities

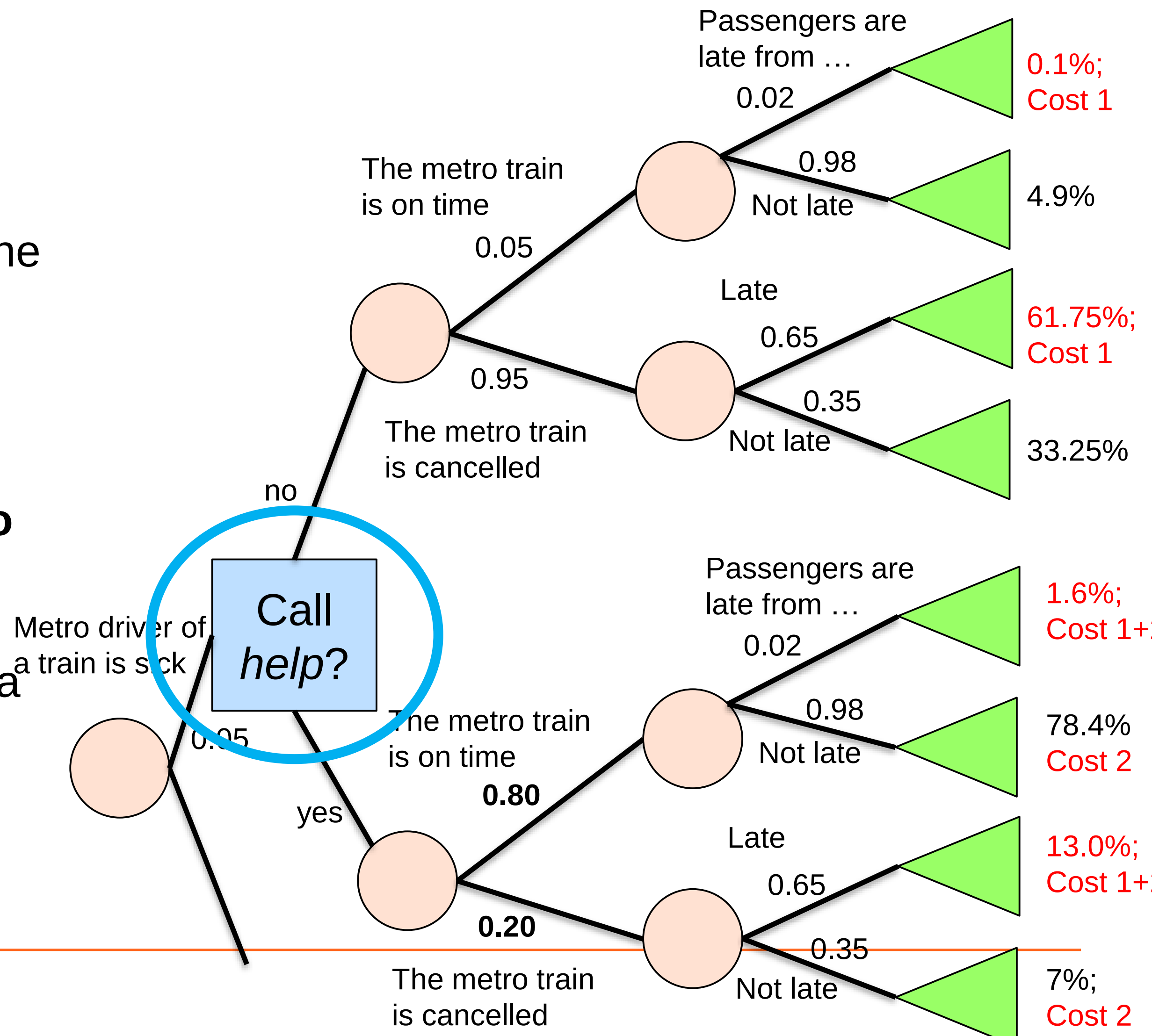
- Probabilities of sequential, mutually exclusive and collectively exhaustive events can be represented as a tree
- The probability of a sequence of events is obtained my multiplying the probabilities on the path
  - *0.95 x 0.95 x 0.02 = 1.805 %*
- The total probability of being late is 7.985 %



The probability of being late on the condition that the train is cancelled

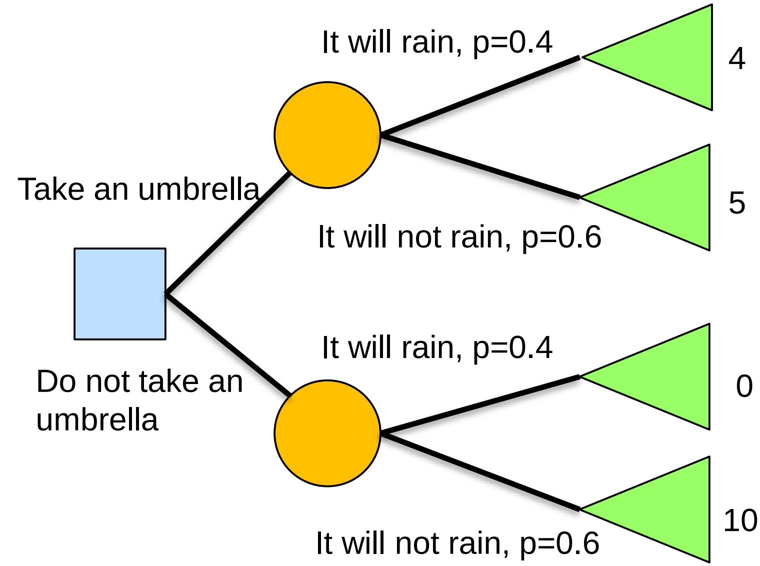
What if…

- The operator is concerned with the unwanted financial consequences caused by the train being late (Cost 1)
  - Numerical *outcomes* for states (consequences)
- The probability **p(the metro train is on time | metro driver is sick)** can be made higher by calling extra personnel (*help*) at a cost (Cost 2)?
  - Now the event probabilities depend on our **decision**



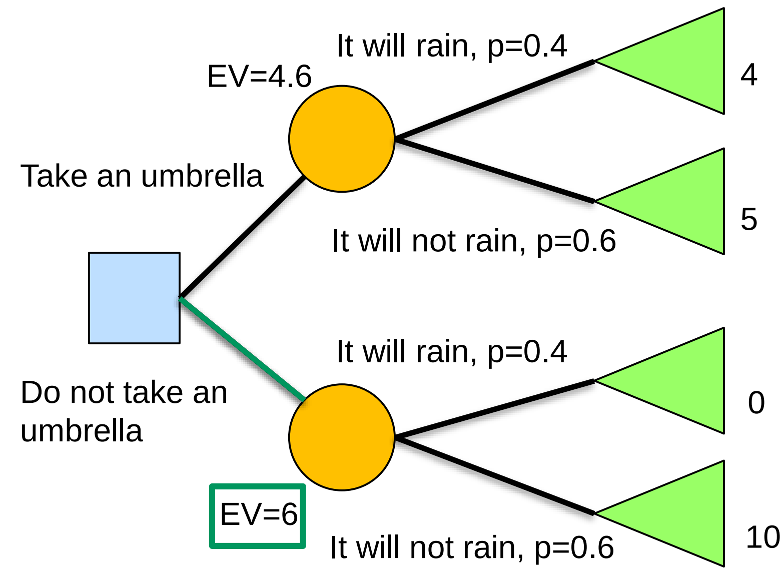
Decision trees

- Decisions under uncertainty can be modeled as decision trees
- A decision tree consists of
  - Decision nodes (squares) represent alternative actions the DM can choose.
  - Chance nodes (circles; cf. states of nature) represent alternative realizations of uncertainties associated with the chance event. The probabilities following a chance node sum up to 1.
  - Consequence nodes (triangles; resulting consequences) at the end of the tree represent decision consequences (e.g., profit, cost, revenue, utility) associated with the path leading to the node.
- Decisions and chance events are displayed logically **in the temporal sequence from left to right**
  - Only chance nodes whose outcomes are known can precede a decision node
- Each path through decisions and chance events represents a possible decision outcome



Solving a decision tree

- A decision tree is solved by starting from the leaves (consequence nodes) and reverting towards the root:
  - At each chance node: **compute** the **expected value** of consequences at the node
  - At each decision node: **select** the arc with the **highest expected value**



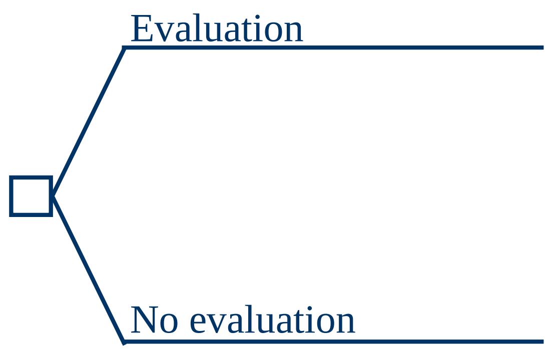
- The optimal strategy is defined by the selected options (arcs) at decision nodes
- A strategy maps available information to choices among alternative actions

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

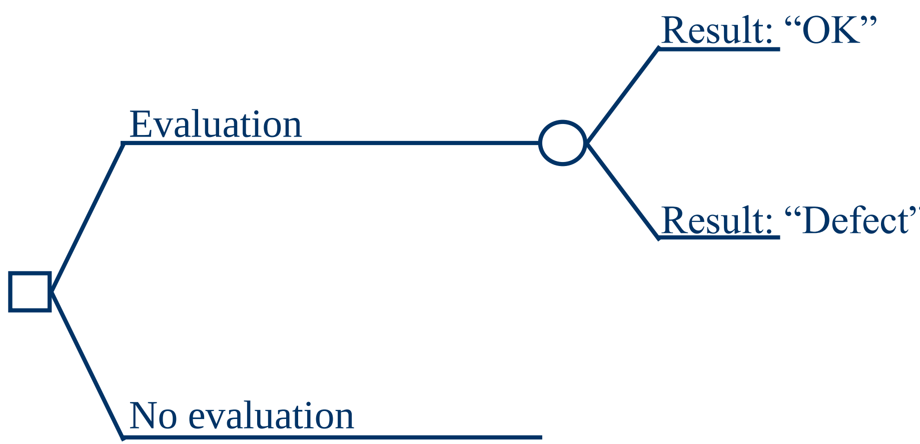
Example: Decision tree (2/12)

- Before making the decision to buy – and before any uncertainties can be resolved – your uncle will have to **decide** whether or not to take the old tractor to a garage for an **evaluation**.
- The decision node 'evaluation' is thus leftmost in the tree



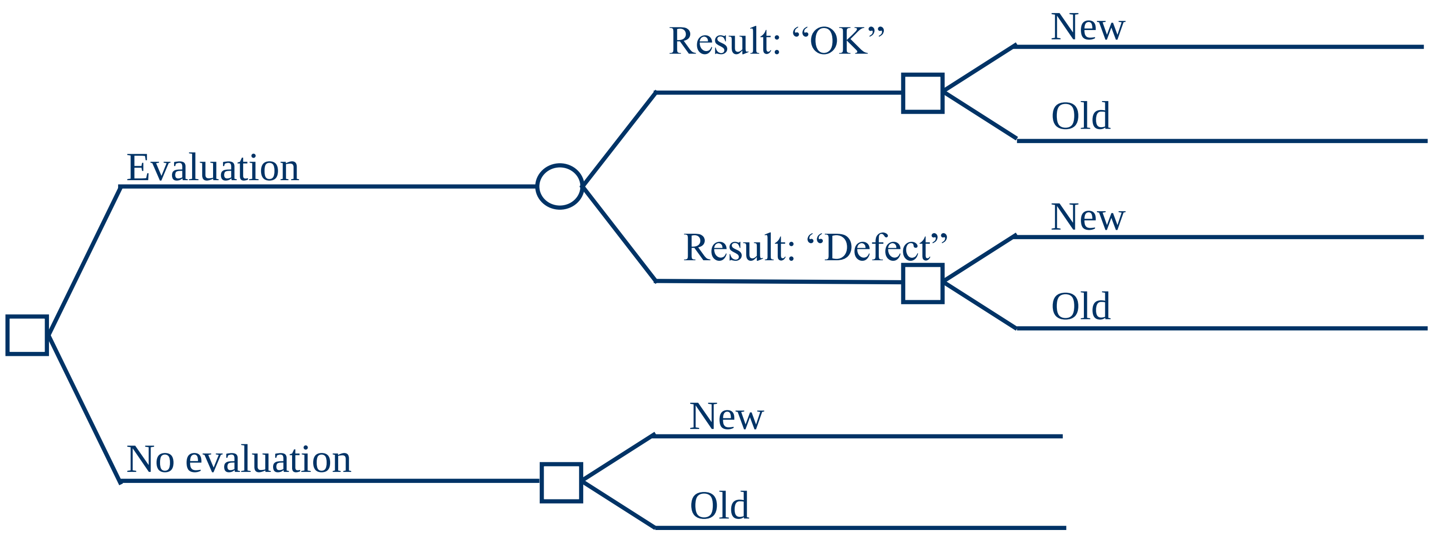
Example: Decision tree (3/12)

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



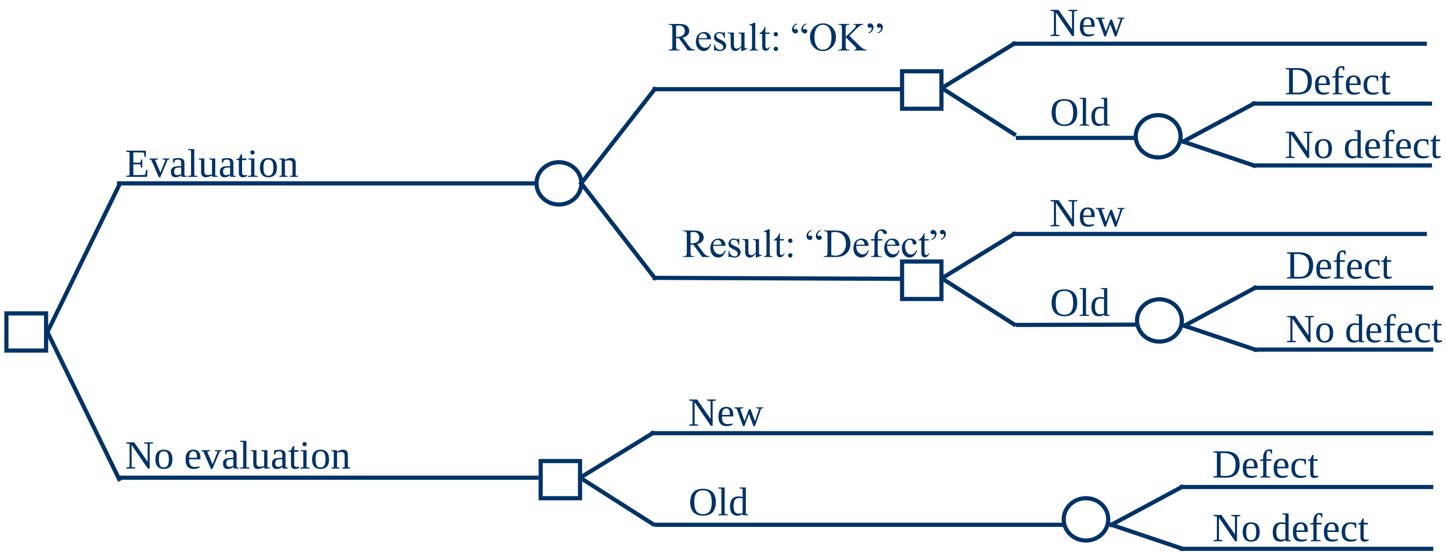
Example: Decision tree (4/12)

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



Example: Decision tree (5/12)

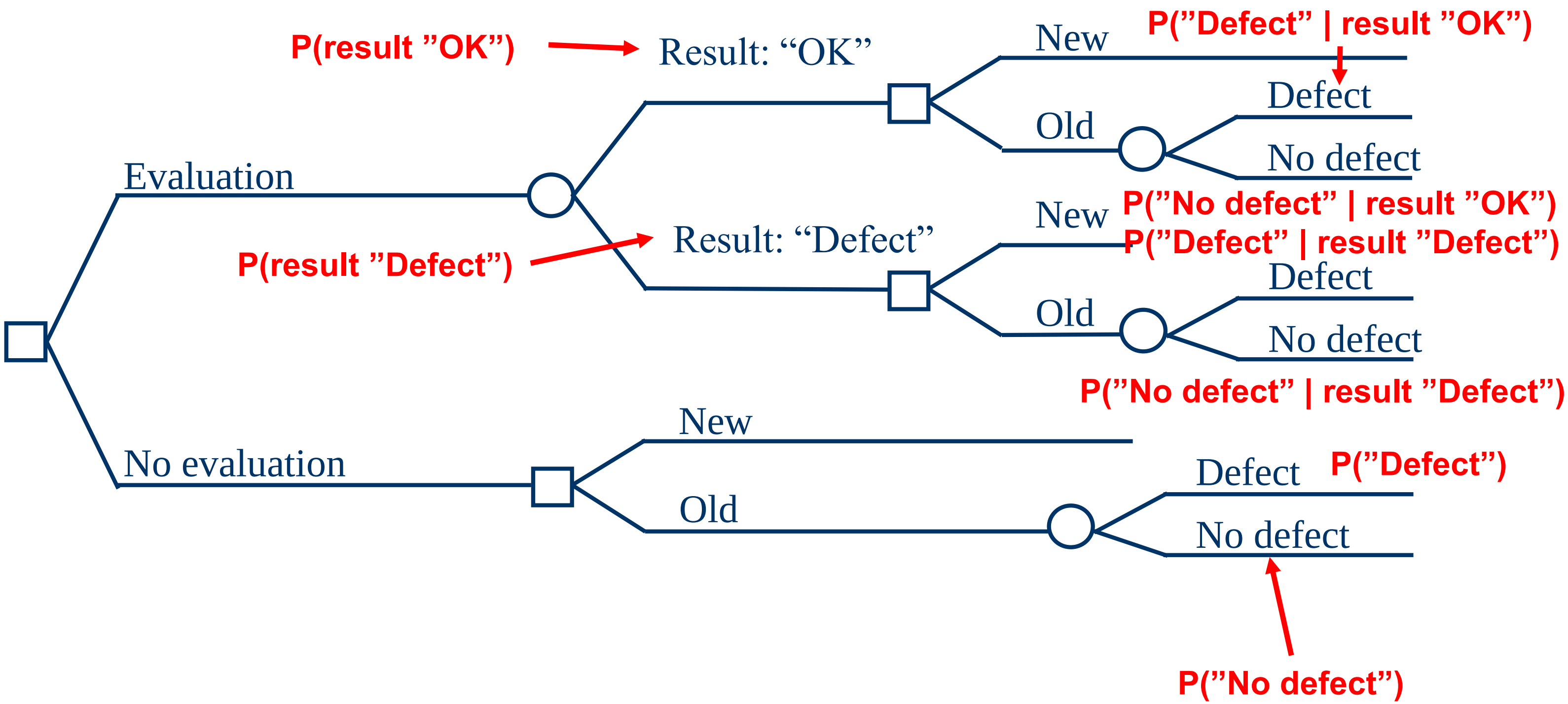
- However, the engine of the old tractor can be defect



- Now all chance nodes and decisions are in chronological order such that at each decision node, the path to the left indicates what is known to inform this decision

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 so that $E_i \cap E_j = \emptyset$, $i \neq j$ and exhaustive so that $\mathrm{A} = \bigcup_i E_i$, then

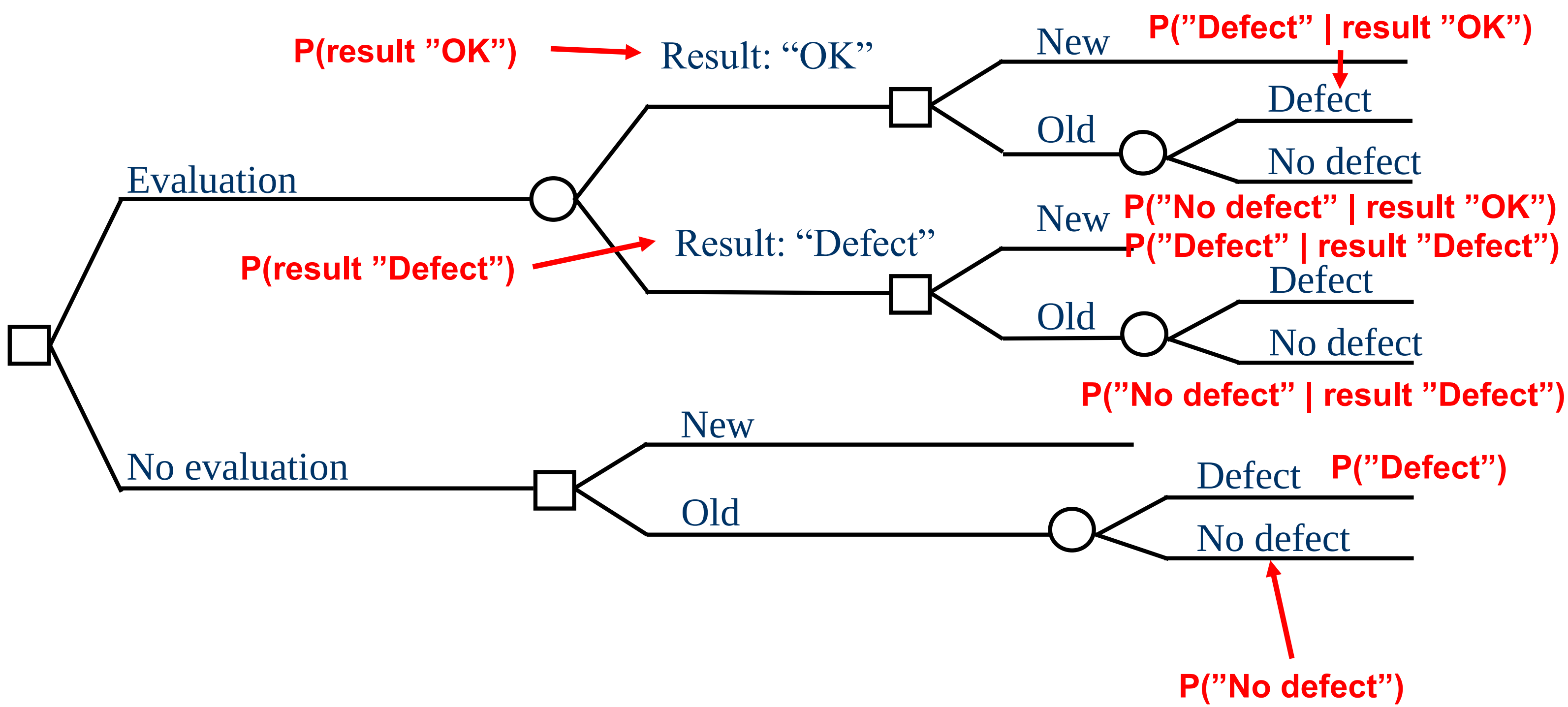
$$P(A)=P(A|E_1)P(E_1) + \ldots + P(A|E_n)P(E_n)$$

- This law is frequently employed through
  - 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:** $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$
- 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: Decision tree (6/12)



Example: Decision tree (7/12)

- "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

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

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

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

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

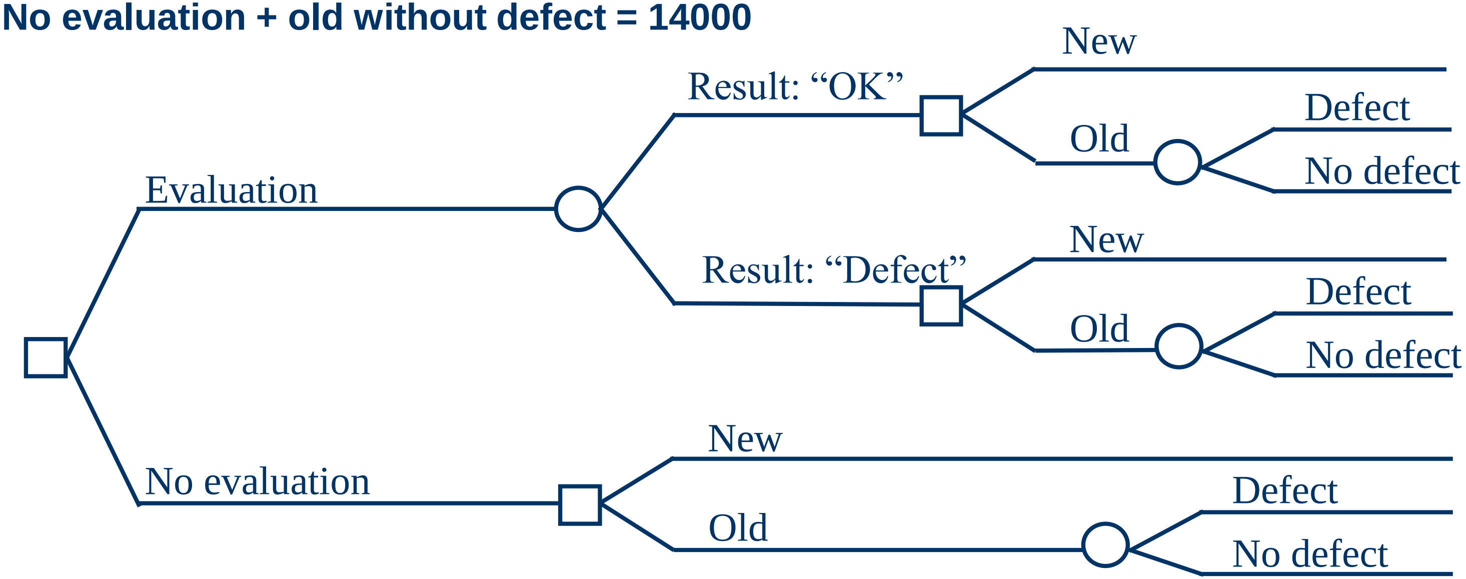
$$P(\text{No defect}|\text{result "OK"}) = 1 - 0.034 = 0.966$$

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

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

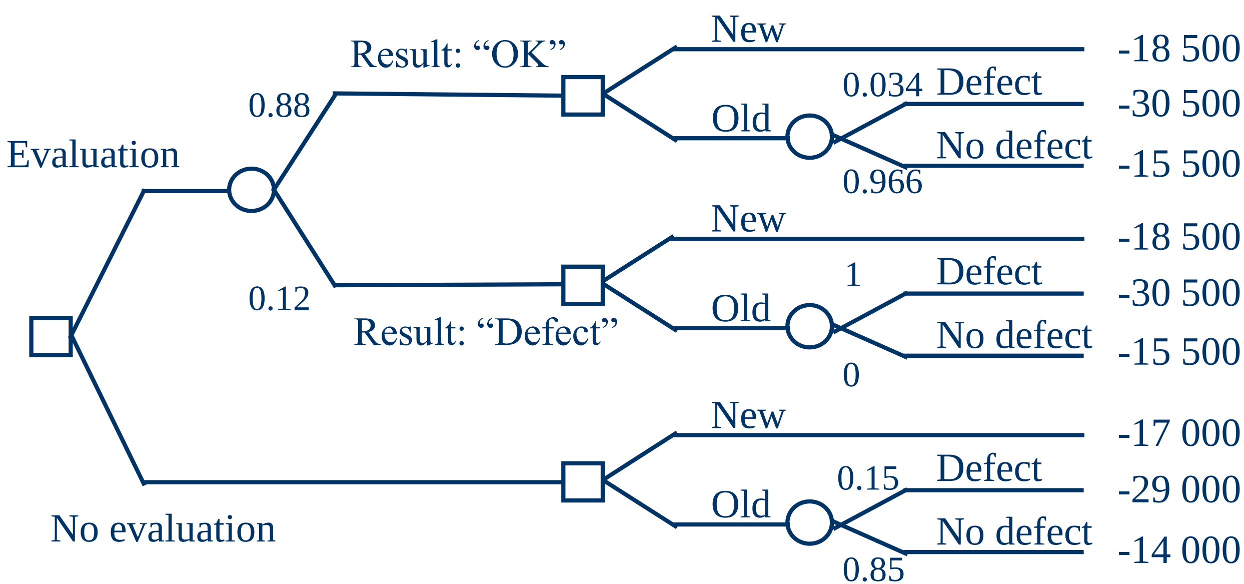
Example: Decision tree (8/12)

- 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**



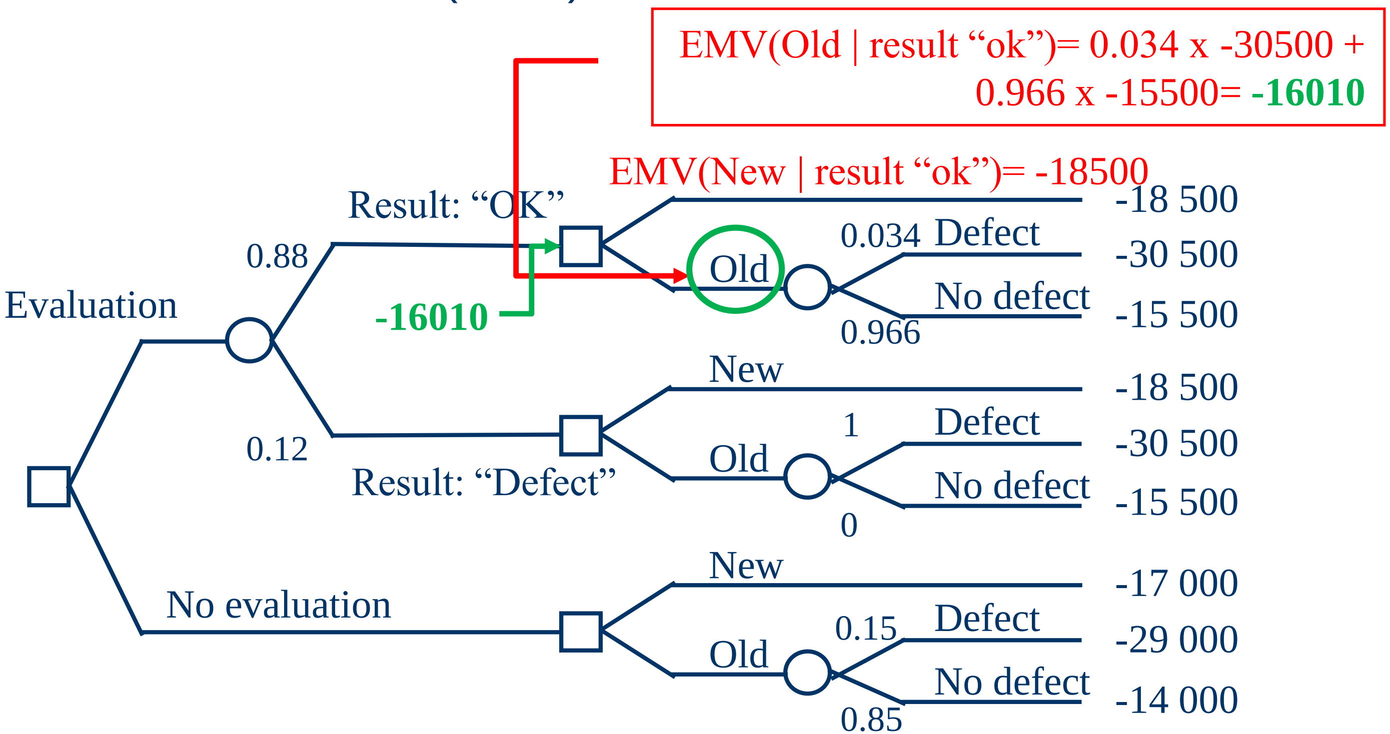
Example: Decision tree (9/12)

- We now have a fully specified decision tree representation of the problem



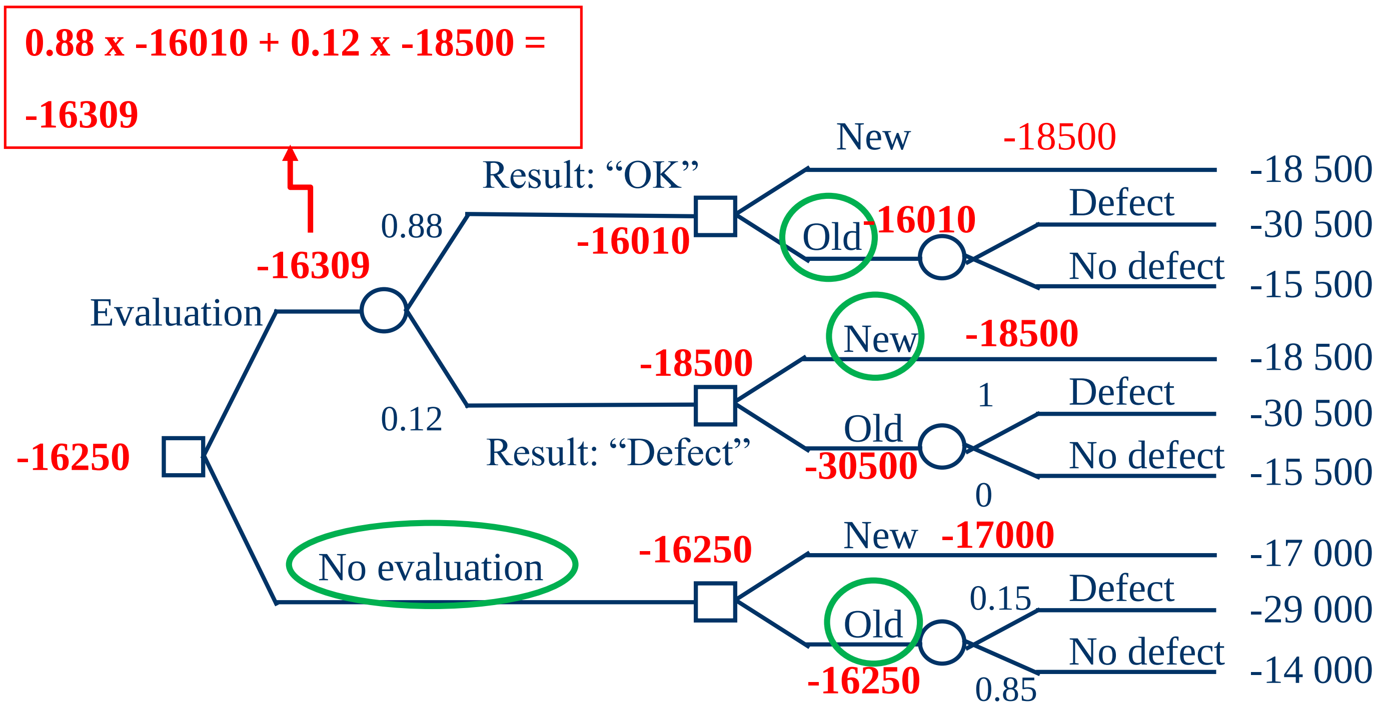
Example: Decision tree (10/12)

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



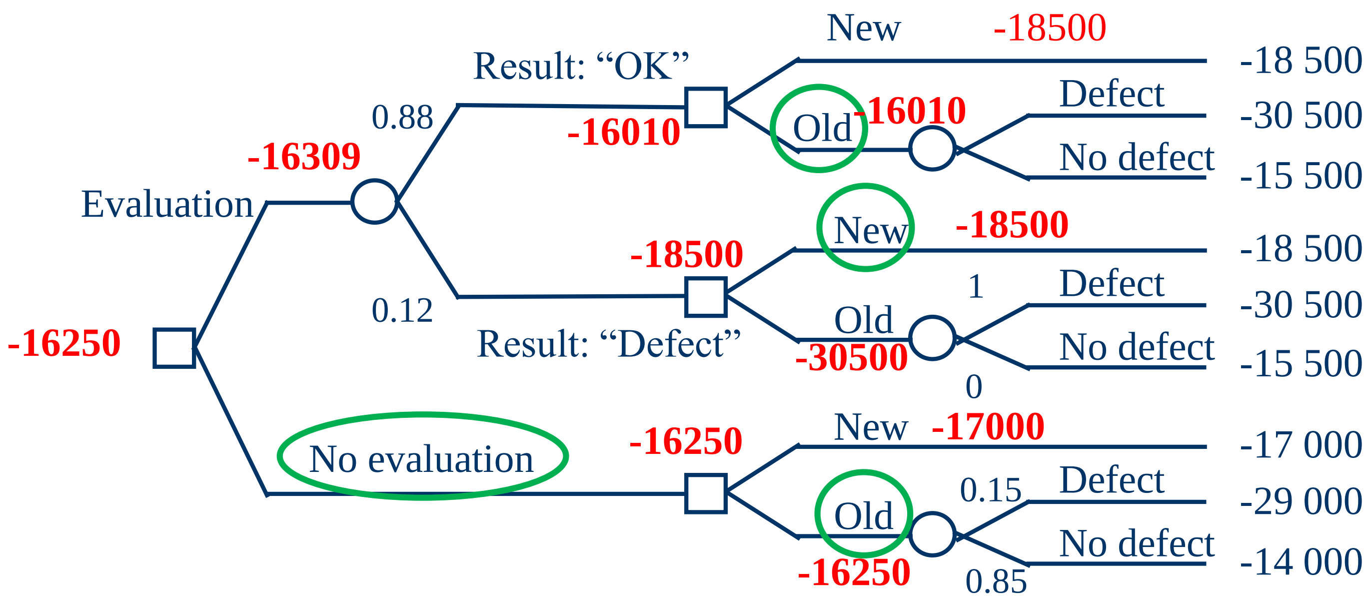
Example: Decision tree (11/12)

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



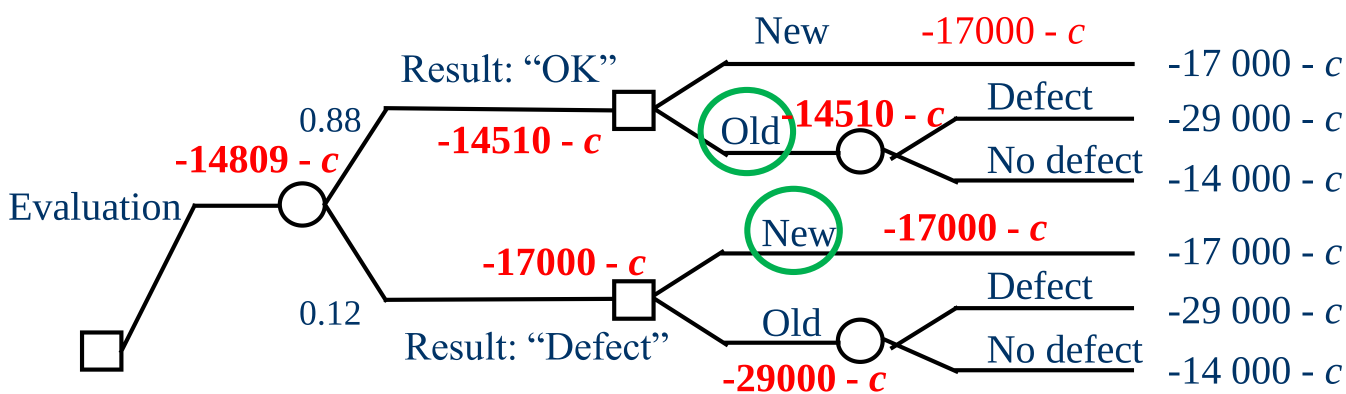
Example: Decision tree (12/12)

- 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 <u>without evaluating the old tractor</u>
- Determine the evaluation cost *c* so that you are **<u>indifferent</u>** 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€
  - ***<u>Expected value of sample information</u>*** *= Expected value with sample information – Expected value without sample information = -14809€ - (-16250€) = 1441€*

Expected value of perfect information

You are participating in a gambling game, where Mia has tossed a coin. Mia knows the result of the toss.

You can choose between *heads* or *tails.* If you choose correctly, Bernard will pay you 50€. Otherwise you pay Bernard 50€. You are maximizing EMV.

Mia suggests that you could offer her money and she'd tell you the correct choice. How much would you <u>at most</u> pay her?

Expected value of perfect information

You can choose between *heads* or *tails.* If you choose correctly, Bernard will pay you 50€. Otherwise you pay Bernard 50€. You are maximizing EMV.

EMV <u>without</u> perfect information = 0.5 x 50€ + 0.5 x (-50€) = 0€

Mia suggests that you could offer her money and she'd tell you the correct choice. How much should you <u>at most</u> pay her?

EMV <u>with</u> perfect information = 1 x 50 € + 0 x (-50€) = 50€

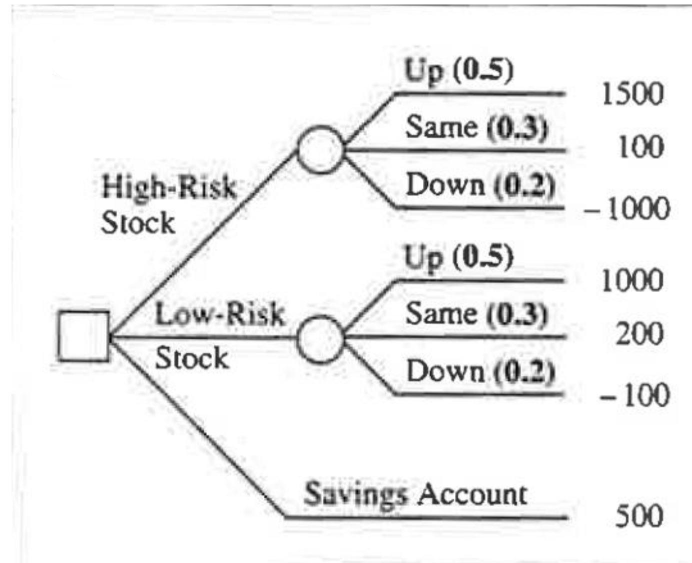
You should pay at most 50€ - 0€ = 50 €

Example: expected value of *perfect information*

- You consider 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	0.2

Decision tree



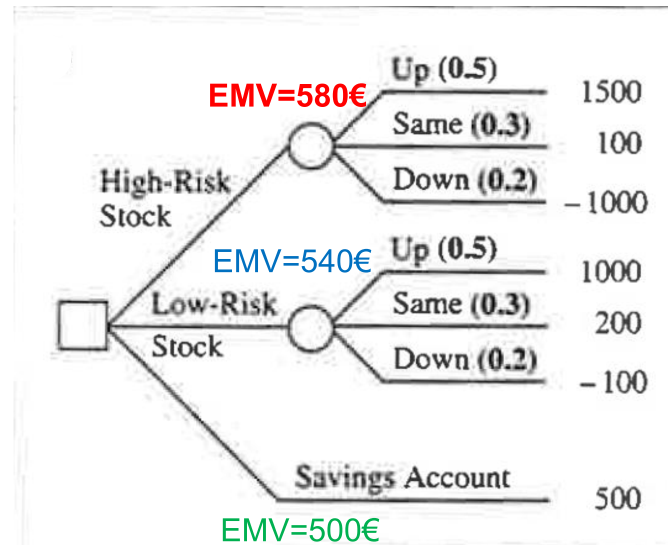
Source: Clemen, R.T. (1996): Making Hard Decisions: An Introduction to Decision Analysis, 2nd edition, Duxbury Press, Belmont.

Example: investing in the stock market

- The expected monetary values (EMVs) for the 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

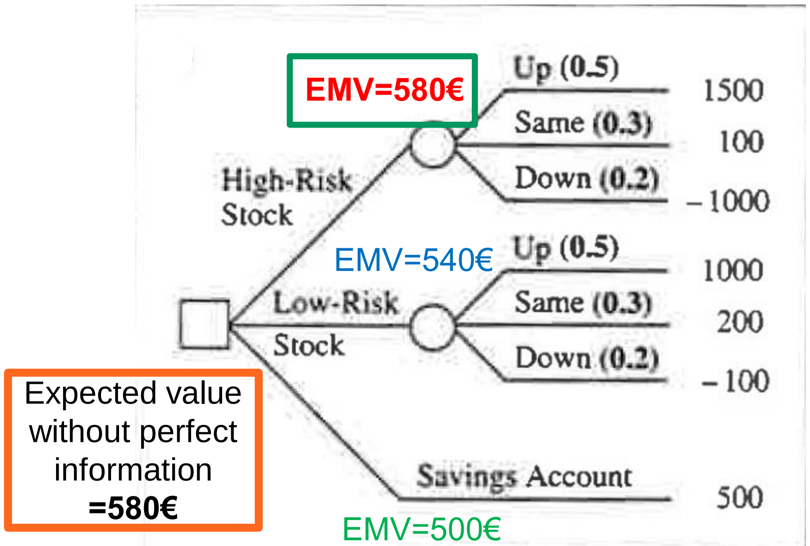


Source: Clemen, R.T. (1996): Making Hard Decisions: An Introduction to Decision Analysis, 2nd edition, Duxbury Press, Belmont.

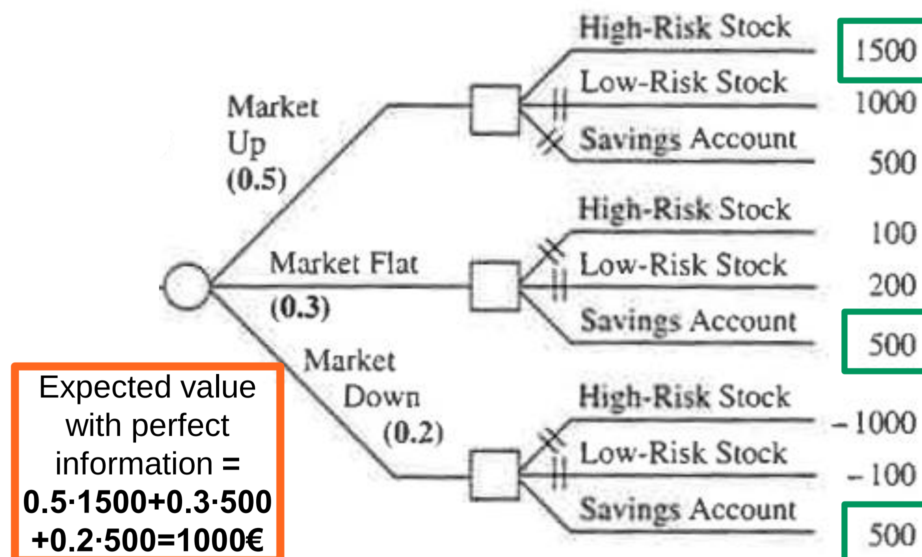
* Assuming you are *risk-neutral* !!! – *risk attitudes* will be discussed later on in this course

Expected value of perfect information

Decision tree



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



Expected value of perfect information

= 1000€ - 580€ = 420€

Expected value of perfect information

- How much would the expected value increase, if:
  - Additional information about the uncertainties were to be received before the decision
  - The decision would be made according to this information?
    - *Note: this analysis can be carried out before any information is obtained*

- Perfect information: complete information about how the uncertainties will play out – "we decide knowing what the state of the world is"
  - **Expected value of perfect information** =
    Expected value with perfect information – Expected value without perfect information

- **Expected value of perfect information** is computed by solving a restructured decision tree in which all chance nodes precede all decision nodes

Probability assessment

- Go to the course's MyCourses site
  - https://mycourses.aalto.fi/mod/feedback/view.php?id=925354

- Use a few minutes to answer
- Do not communicate with others
- Do not look up the answers on the internet

Question 2 should read (the text is too long for the MyCourses module)

Consider two bags X and Y. Bag X contains 30 white balls and 10 black balls, whereas bag Y contains 30 black balls and 10 white balls. Suppose that you select one of these bags at random, and randomly draw five balls one-by-one by replacing them in the bag after each draw. Suppose you get four white balls and one black. What is the probability that you selected bag X with mainly white balls?

Estimation of probabilities

- How to elicit probabilities for decision trees?

1. If possible, use *objective* data

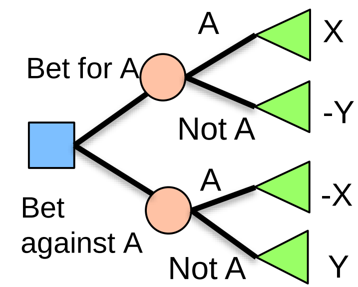
2. If objective data is not available, obtain *subjective* probability estimates from experts through
    - Betting approach
    - Reference lottery
    - Direct judgement

Subjective probabilities reflect the respondent's beliefs

Betting approach

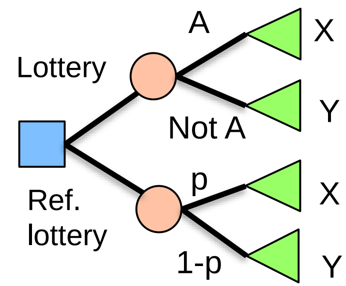
- **Goal:** to estimate the probability of event A
  - E.g., A= “GDP growth is above 2% in 2023” or
    A= “There will be a NATO military base in Finland within the next five years”
- **Betting approach:**
  - Bet for A: win X € if A happens, lose Y € if not
    - Expected monetary value $X \cdot P(A) - Y \cdot [1 - P(A)]$
  - 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

Reference lottery



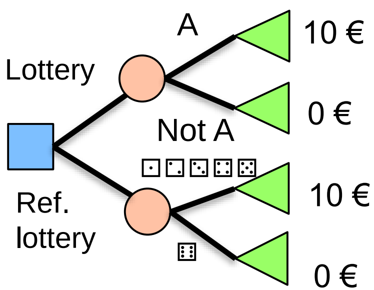
- Lottery:
  - Win X if A happens
  - Win Y if A does not happen
  - X is preferred to Y
- Reference lottery:
  - Win X with (known) probability *p*
  - Win Y with (known) probability (1-*p*)
  - Probability *p* can be visualized with, e.g., a wheel of fortune
- Adjust *p* until the respondent is **indifferent**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

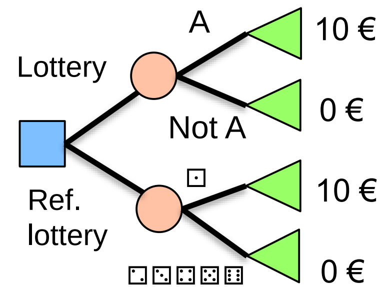
Reference lottery: example

- **Event A:** “Finland beats Sweden in the track-and-field competitions”



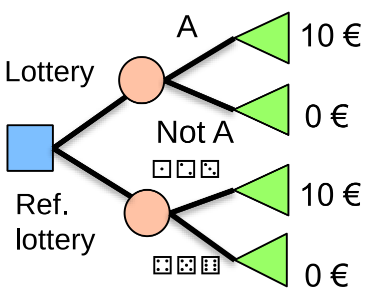
The respondent chooses the reference lottery:

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



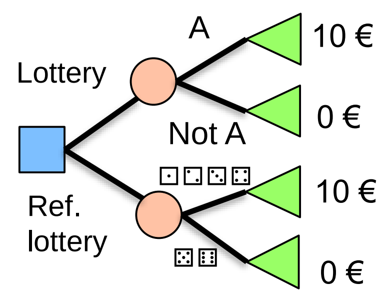
The respondent chooses the lottery:

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



Chooses the lottery:

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



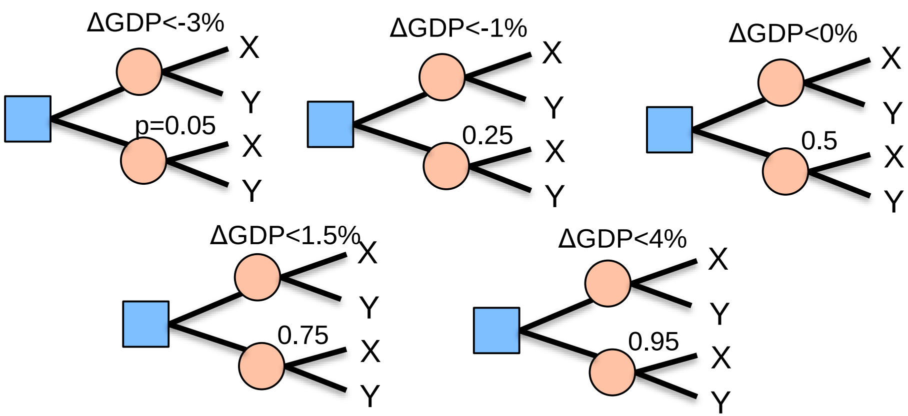
Chooses the reference lottery:

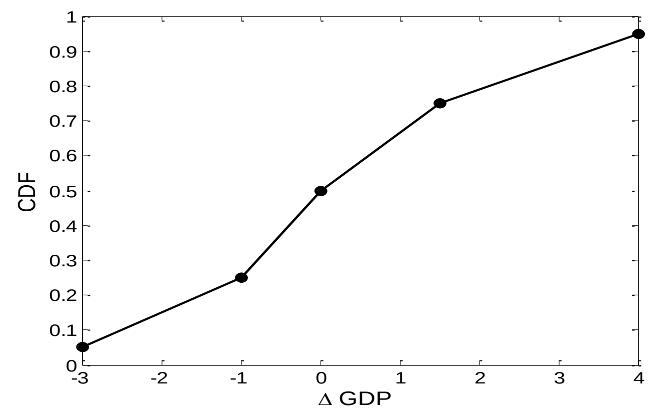
$$P(A) < \frac{2}{3}$$

These four answers put the probability estimate of A in the range (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: Assess the probability distribution for the GDP growth (ΔGDP) in Finland in 2023
  - Means: Elicit the probability *p* for five different reference lotteries





Estimation of continuous probability distributions

- Often experts are asked to 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 provide support for decisions under uncertainty
  - Which decision alternative has the best expected consequences?
  - How much should one be willing to pay for perfect information or (imperfect) sample information about how the uncertainties are resolved?

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

Next time

- **Are ‘experts’ capable of giving good (probability) estimates?**
  - What kind of *elicitation biases* should we be aware of?

- **How to model the DM’s preferences over risky alternatives?**
  - What if the DM does not **maximize expected monetary value**, but has an **attitude towards risk**?
    - *Would you*
      1. choose **5000 €** for sure OR
      2. participate in 50-50 lottery between **0 €** and **12000 €**?
  - What were the **magic numbers** in the umbrella example?