Exercise: 3

Meeting on January 19th / 21st

Problem 1: Expectation Values

You receive an offer from a big software company for a project as an independent developer. Instead of a fixed sum for your work the offer you a 50% share on the expected profit of 1,000,000 €. However, the project is risky since you know from statistics with similar projects that the probability for success lies only around 20%. You measure your labor investment with 150,000 €.

(a) Is the project worth it for you?

Lösung: Without consultation the expectation value for rejecting the project is

$$E(\text{Reject project}) = 0€$$

and for accepting the project

$$E(\text{Accept project}) = 20\% \cdot (50\% \cdot 1.000.000€) - 150.000€ = -50.000€.$$

Without the help of the agency (prior) it is better to *reject* the project since losses are more likely than profits. If you reject the project, the expectation value is at 0 €.

(b) An IT consulting agency with experience in such project divides software companies into 4 groups. The worst group, which 40% companies belong to, have a success probability of 10%. The average group, to which 40% of firms belong to, has a success probability of 20%. The good group, to which 10% of firms belong to, has a success probability of 30%, and the best group with the remeining firms has a success probability of 50%. You trust that the consulting agency is realistic in its assessment of groups.

The agency now offers you to determine the company's group. How much money should you spend at most for this assessment?

Lösung: The following table gives and overview of groups and success probabilities:

	G_1	G_2	G_3	G_4
Group	40%	40%	10%	10%
Probability of success	10%	20%	30%	50%

Lösung: Computing the expectation values under the assumption that we got the group assessment about the company from the agency:

$$
\begin{aligned}
E(\text{Accept project}|G_1) &= 10\% \cdot 500.000€ - 150.000€ = -100.000€ \\
E(\text{Accept project}|G_2) &= 20\% \cdot 500.000€ - 150.000€ = -50.000€ \\
E(\text{Accept project}|G_3) &= 30\% \cdot 500.000€ - 150.000€ = 0€ \\
E(\text{Accept project}|G_4) &= 50\% \cdot 500.000€ - 150.000€ = 100.000€ \\
E(\text{Reject project}|G_x) &= 0€
\end{aligned}
$$

A rational agent would reject the project for the prediction of G_1 or G_2 and otherwise accept it.

The expected profit compared to the best prior action averaged over all groups is computed as

$$
\begin{aligned}
E &= \sum_{i=1}^{4} P(G_i) \cdot E(\text{Best action}|G_i) - E(\text{Beste prior action}) \\
&= 40\% \cdot 0€ + 40\% \cdot 0€ + 10\% \cdot 0€ + 10\% \cdot 100.000€ - 0€ \\
&= 10.000€
\end{aligned}
$$

You should only accept the agency's offer for *less than 10,000 €*.

Problem 2: Decision network for an exam

You are writing a test tomorrow but have neglected studying for it. You are now thinking about your best options for passing the test well. The exam can be easy or hard. Since your tutor has a good reputation, you are estimating that the test will be easy with 60% and hard with 40% probability. The result of the test can be either good, average or bad. You rank their utilities with +10, +5 and -5 respectively. For your preparation you have the options of getting enough sleep or studying through the night, where you estimate the cost with 0 and -2 respectively.

The conditional probabilities mentioned above are listed below:

(a) What does the decision network look like for this problem? Consider, which nodes flow into the utility.

(b) Compute the expected utility for the options "sleep" and "study".

Lösung:

(a) The decision network includes to nodes for random variables: One for the difficulty of the exam and one for the result. The decision node is whether you will study through the night or sleep (titled simply as study in the network) and directly influences the result probabilities. Both the result and the decision affect the utility. The decision netork then looks like:

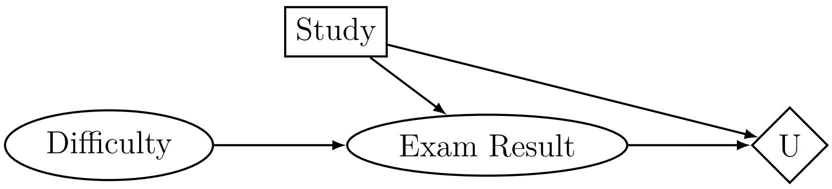
Difficulty	Probability
easy	60%
hard	40%

	Sleep		Study	
Difficulty/Result	easy	hard	easy	hard
good	70%	40%	60%	60%
average	20%	40%	30%	30%
bad	10%	20%	10%	10%

Result	Utility
good	10
average	5
bad	-5

Decision	Utility
Sleep	0
Study	-2

Table 1: Probabilities and Utilities for studying for an exam



Decision Network

(b) Let $S = \{\text{easy, hard}\}$ be the difficulty, $E = \{\text{good, average, bad}\}$ be the result and $D = \{\text{sleep, study}\}$ the decision of studying or not. The expected utility is computed as follows:

$$EU(D) = \sum_S \sum_E U(E) \cdot P(E|S,D) \cdot P(S) + U(D) \tag{1}$$

Damit ergibt sich für die einzelnen Entscheidungen:

$$
\begin{aligned}
EU(\text{sleep}) = \quad & [U(\text{good}) \cdot P(\text{good}|\text{easy, sleep}) \\
& + U(\text{average}) \cdot P(\text{average}|\text{easy, sleep}) \\
& + U(\text{bad}) \cdot P(\text{bad}|\text{easy, sleep})] \cdot P(\text{easy}) \\
& + [U(\text{good}) \cdot P(\text{good}|\text{hard, sleep}) \\
& + U(\text{average}) \cdot P(\text{average}|\text{hard, sleep}) \\
& + U(\text{bad}) \cdot P(\text{bad}|\text{hard, sleep})] \cdot P(\text{hard}) \\
& + U(\text{sleep}) \\
= \quad & (10 \cdot 0{,}7 + 5 \cdot 0{,}2 - 5 \cdot 0{,}1) \cdot 0{,}6 \\
& + (10 \cdot 0{,}4 + 5 \cdot 0{,}4 - 5 \cdot 0{,}2) \cdot 0{,}4 + 0 \\
= \quad & 6{,}5
\end{aligned}
$$

$$
\begin{aligned}
EU(\text{study}) = \quad & [U(\text{good}) \cdot P(\text{good}|\text{easy, study}) \\
& + U(\text{average}) \cdot P(\text{average}|\text{easy, study}) \\
& + U(\text{bad}) \cdot P(\text{bad}|\text{easy, study})] \cdot P(\text{easy}) \\
& + [U(\text{good}) \cdot P(\text{good}|\text{hard, study}) \\
& + U(\text{average}) \cdot P(\text{average}|\text{hard, study}) \\
& + U(\text{bad}) \cdot P(\text{bad}|\text{hard, study})] \cdot P(\text{hard}) \\
& + U(\text{study}) \\
= \quad & (10 \cdot 0{,}6 + 5 \cdot 0{,}3 - 5 \cdot 0{,}1) \cdot 0{,}6 \\
& + (10 \cdot 0{,}6 + 5 \cdot 0{,}3 - 5 \cdot 0{,}1) \cdot 0{,}4 - 2 \\
= \quad & 5
\end{aligned}
$$

Hence, the option "sleep" has a better expected utility and should be preferred.

Problem 3: Markov-Decision-Problem

Consider the following *Markov Decision Problem* with three states $Z \in 1, 2, 3$ and their rewards B with $B_1 = -1$, $B_2 = -2$ and $B_3 = 0$. There is no discount, so $\gamma = 1$. The state $Z = 3$ is a ending state. There are two possible actions a and b in the two states $Z = 1$ and $Z = 2$. The indeterministic transition model is given in Fig. 1.

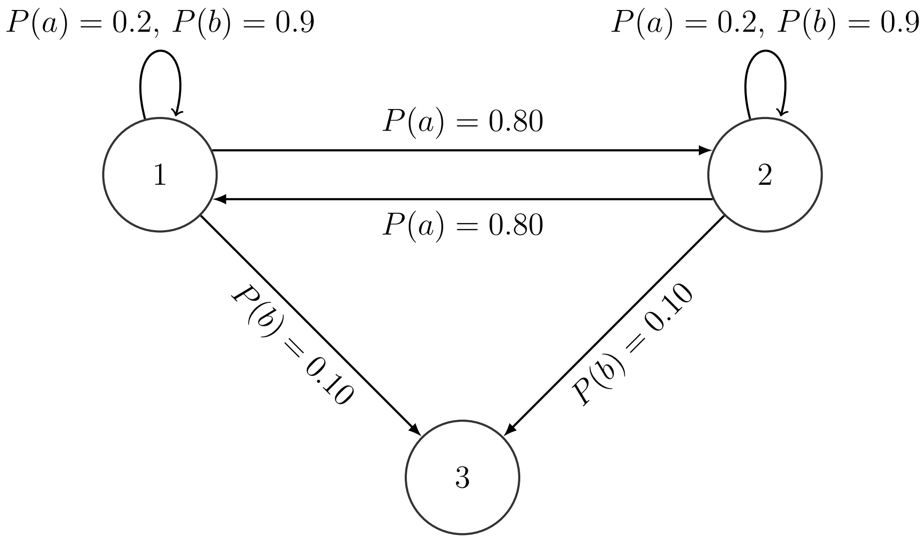


Figure 1: The given indeterministic transition model

(a) What qualitative remarks can be made about the optimal policy in the states $Z = 1$ und $Z = 2$?

> **Lösung:** The agent should strive to reach the sate $Z = 3$ as fast as possible since staying in one of the other two states only gives negative rewards. In state $Z = 1$ it tries action b to reach state $Z = 3$. However, in state $Z = 2$ it tries action a in order to reach state $Z = 1$ and reach $Z = 3$ from there, since state $Z = 1$ gives less "punishment" for staying in it.

(b) Use policy iteration to find the optimal policy and with corresponding utility scores in states $Z = 1$ and $Z = 2$. Show the full derivation. Use an initial policy of action b in both states.

> **Lösung:** Policy iteration is performed in alternating between value determination and policy update.
> Utility U in state Z at time i is given by:
>
> $$U_i(Z) = R(Z) + \gamma \sum_{Z'} T(Z, \pi_i(Z), Z') U_i(Z') \tag{2}$$
>
> The iteration is performed by the following steps:
>
> - *Initialization*:
>
> $$U_0 = \langle -1; -2; 0 \rangle, \pi_0 = \langle b; b \rangle \tag{3}$$

- *Value determination* (Policy $\langle b; b\rangle$):

$$
\begin{aligned}
U_1(1) &= -1 + 0{,}1 \cdot U_1(3) + 0{,}9 \cdot U_1(1)\\
U_1(2) &= -2 + 0{,}1 \cdot U_1(3) + 0{,}9 \cdot U_1(2)\\
U_1(3) &= 0 \qquad (4)
\end{aligned}
$$

Hence:

$$
\begin{aligned}
U_1(1) &= -10\\
U_1(2) &= -20 \qquad (5)
\end{aligned}
$$

- *Policy Update* For $Z = 1$ it is

$$\sum_{j=1}^{3} T(1,a,j)U_1(j) = 0{,}8 \cdot (-20) + 0{,}2 \cdot (-10) = -18 \qquad (6)$$

$$\sum_{j=1}^{3} T(1,b,j)U_1(j) = 0{,}1 \cdot 0 + 0{,}9 \cdot (-10) = -9, \qquad (7)$$

so action b is still preferred. For $Z = 2$ it is

$$\sum_{j=1}^{3} T(2,a,j)U_1(j) = 0{,}8 \cdot (-10) + 0{,}2 \cdot (-20) = -12 \qquad (8)$$

$$\sum_{j=1}^{3} T(2,b,j)U_1(j) = 0{,}1 \cdot 0 + 0{,}9 \cdot (-20) = -18, \qquad (9)$$

so action a is chosen now, since the "punishment" is smaller.

- *Value determination* (Policy $\langle b; a\rangle$):

$$
\begin{aligned}
U_2(1) &= -1 + 0{,}1 \cdot U_2(3) + 0{,}9 \cdot U_2(1)\\
U_2(2) &= -2 + 0{,}8 \cdot U_2(1) + 0{,}2 \cdot U_2(2)\\
U_2(3) &= 0 \qquad (10)
\end{aligned}
$$

Hence:

$$U_2(1) = -10 \qquad (11)$$

$$U_2(2) = -12{,}5 \qquad (12)$$

- *Policy update*: For $Z = 1$ it is

$$\sum_{j=1}^{3} T(1,a,j)U_2(j) = 0{,}8 \cdot (-12{,}5) + 0{,}2 \cdot (-10) = -12 \qquad (13)$$

$$\sum_{j=1}^{3} T(1,b,j)U_2(j) = 0{,}1 \cdot 0 + 0{,}9 \cdot (-10) = -9, \qquad (14)$$

so action b is still preferred. For $Z = 2$ it is

$$\sum_{j=1}^{3} T(2, a, j)U_2(j) = 0{,}8 \cdot (-10) + 0{,}2 \cdot (-12{,}5) = -10{,}5 \tag{15}$$

$$\sum_{j=1}^{3} T(2, b, j)U_2(j) = 0{,}1 \cdot 0 + 0{,}9 \cdot (-12{,}5) = -11{,}25, \tag{16}$$

so action a is chosen again.

Hence, the policy is now stable at $\pi_i = \langle b; a \rangle$.

Problem 4: Learning Decision Trees

Learn a Decision Tree from the following dataset and determine the assessment for the two test cases.

Fall-Nr.	Clouds	Temperature	Humidity	Wind	Assessment
1	sunny	high	high	no	+
2	sunny	high	high	yes	+
3	cloudy	medium	normal	no	+
4	rainy	high	high	no	-
5	sunny	low	normal	yes	-
6	rainy	low	high	yes	-
7	cloudy	high	high	yes	-
8	sunny	medium	normal	no	+
9	cloudy	medium	high	yes	-
Testfall 1	rainy	medium	high	no	?
Testfall 2	sunny	medium	high	yes	?

Table 2: Dataset and test cases for learning a decision tree

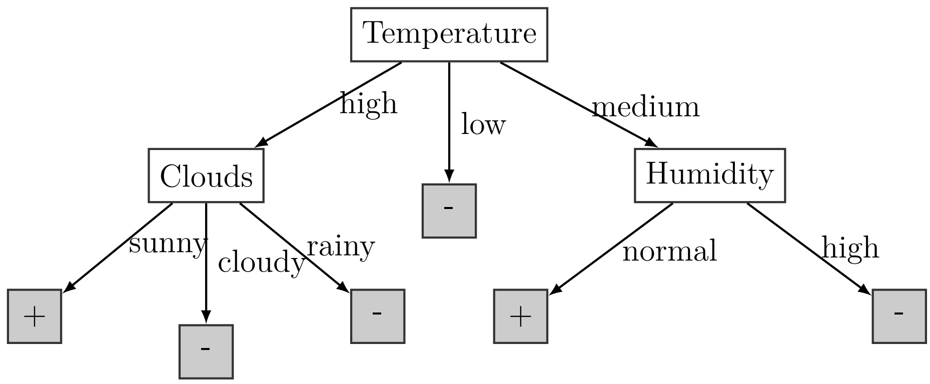
(a) Determine the decision tree intuitively first, i.e. choose approximately only those attributes which allow for a clear classification on most data points.

(b) Now learn the decition tree exactly.

Lösung:

(a) The intuitively best attribute is one for which most values already allow an unambiguous classification. For instance, the value *low* of *Temperature* direclty classifies to "-". In the case of *Temp. = high*, *Clouds* classifies the remaining data points and for *Temp. = medium*, *Humidity* classifies the rest.

Hence, the decision tree looks like this:



The assessments for the test cases are then

Testfall 1: Test case 1: negative, frist via "medium" to Humidity and then via "high"

Testfall 2: Test case 2: negative, like case 1

(b) The decision tree is determined by the most important attributes, i.e. the ones with highest information gain. The *gain* is computed as

$$\begin{aligned}\text{Gain}(A) &= B\left(\frac{p}{p+n}\right) - \text{Remainder}(A)\\ &= B\left(\frac{p}{p+n}\right) - \sum_{k=1}^{d} \frac{p_k+n_k}{p+n} \cdot B\left(\frac{p_k}{p_k+n_k}\right)\end{aligned}$$

with

$$B(q) = -q\log_2 q - (1-q)\log_2(1-q)\ . \tag{17}$$

The attribute values are:

$$\begin{aligned}\text{Clouds} &\in \{\text{sunny, cloudy, rainy}\}\\ \text{Temperature} &\in \{\text{high, medium, low}\}\\ \text{Wind} &\in \{\text{yes, no}\}\\ \text{Humidity} &\in \{\text{high, normal}\}\end{aligned}$$

Step 1: Computing the *Remainder*
Initial is $B\left(\frac{p}{p+n}\right) = B\left(\frac{4}{9}\right) = 0.99$ for all attributes.
For two attributes A and C the Gain(A) to Gain(C) is bigger, if the corresponding Remainder($\cdot$) is smaller:
The Remainder($\cdot$) is determined by the size of the subset, that is created by splitting the examples with that attribute, and the remaining entropy.

$$\begin{aligned}\text{Remainder(Clouds)} &= \frac{4}{9}B\left(\frac{3}{4}\right) + \frac{3}{9}B\left(\frac{1}{3}\right) + \frac{2}{9}B\left(\frac{0}{2}\right) = 0.67\\ \text{Remainder(Temperature)} &= \frac{4}{9}B\left(\frac{2}{4}\right) + \frac{3}{9}B\left(\frac{2}{3}\right) + \frac{2}{9}B\left(\frac{0}{2}\right) = 0.75\\ \text{Remainder(Wind)} &= \frac{5}{9}B\left(\frac{1}{5}\right) + \frac{4}{9}B\left(\frac{3}{4}\right) = 0.76\\ \text{Remainder(Humidity)} &= \frac{6}{9}B\left(\frac{2}{6}\right) + \frac{3}{9}B\left(\frac{2}{3}\right) = 0.92\end{aligned}$$

Hence Clouds is the best attribute since it got the lowest *Remainder*. It can determine some cases immediately and offers good "pre-decisions" for the rest .

Step 2: Computing the *Remainder* the condition Clouds = sunny
Initial is $B\left(\frac{p}{p+n}\right) = B\left(\frac{3}{4}\right) = 0.81$ for all attributes.

$$\begin{aligned}
\text{Remainder(Temperature)} &= \frac{2}{4}B\left(\frac{2}{2}\right) + \frac{1}{4}B\left(\frac{1}{1}\right) + \frac{1}{4}\left(\frac{0}{1}\right) = 0 \\
\text{Remainder(Wind)} &= \frac{2}{4}B\left(\frac{2}{2}\right) + \frac{2}{4}B\left(\frac{1}{2}\right) = \frac{1}{2} \\
\text{Remainder(Humidity)} &= \frac{2}{4}B\left(\frac{2}{2}\right) + \frac{2}{4}B\left(\frac{1}{2}\right) = \frac{1}{2}
\end{aligned}$$

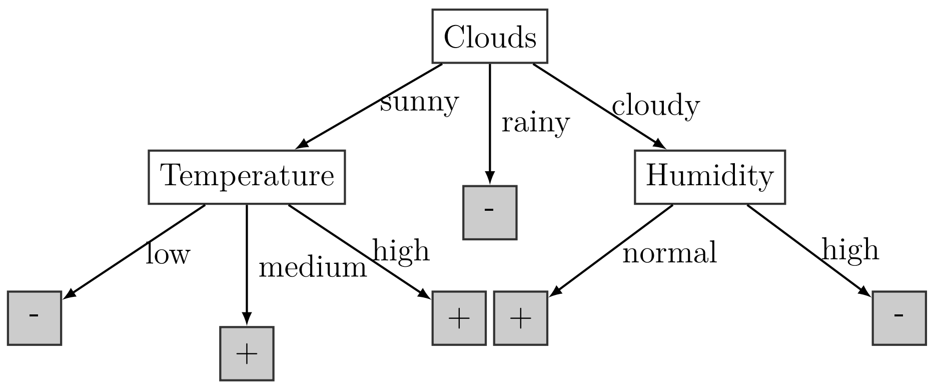
Hence, Temperature is the best attribute under the given conditions. It can now determine the remaining cases.

Step 3: Computing the *Remainder* under the condition Clouds = cloudy
Initial is $B\left(\frac{p}{p+n}\right) = B\left(\frac{1}{3}\right) = 0.92$ for all attributes.

$$\begin{aligned}
\text{Remainder(Temperature)} &= \frac{1}{3}B\left(\frac{0}{1}\right) + \frac{2}{3}B\left(\frac{1}{2}\right) + \frac{0}{3}\left(\frac{0}{0}\right) = \frac{2}{3} \\
\text{Remainder(Wind)} &= \frac{2}{3}B\left(\frac{0}{2}\right) + \frac{1}{3}B\left(\frac{1}{1}\right) = 0 \\
\text{Remainder(Humidity)} &= \frac{2}{3}B\left(\frac{0}{2}\right) + \frac{1}{3}B\left(\frac{1}{1}\right) = 0
\end{aligned}$$

Hence, both Wind and Humidity can be used as the next branch if it is cloudy.

Der gelernte Entscheidungsbaum ist hier zu sehen:



The assessments for the test cases are then

Test case 1: negative, due to ”rainy” Clouds.

Test case 2: positive, first via ”sunny” , then via ”medium” Temperature.