Exercise 1 :

a. Let π be the interpretation that assigns the following truth values:

$$\pi(a) = \textit{true}, \pi(b) = \textit{false}, \pi(c) = \textit{false}, \pi(d) = \textit{true}$$

Determine the truth values for the following propositions:

$$\neg a \rightarrow b$$
$$(\neg b \vee c) \wedge (d \rightarrow a)$$
$$(a \rightarrow c) \rightarrow c$$

b. For the following propositions, find an interpretation in which they are true:

$$(a \vee (a \rightarrow c)) \rightarrow b$$
$$(a \wedge (\neg b \vee c)) \wedge (a \rightarrow (c \rightarrow b))$$

Solution:

a. All three propositions are true.

b. The first proposition is true in any interpretation with $\pi(b) = \textit{true}$.

The second proposition is true for $\pi(a) = \textit{true}$, $\pi(b) = \textit{true}$, $\pi(c) = \textit{true}$. It is also true for $\pi(a) = \textit{true}$, $\pi(b) = \textit{false}$, $\pi(c) = \textit{false}$. These are the only two solutions.

Exercise 2 :

Consider the experiment of flipping a coin, and if it lands heads, rolling a four-sided die, and if it lands tails, rolling a six-sided die.

(i) Suppose that the coin and both dice are fair. As we are interested only in the number rolled by the die, and the possible worlds $\mathcal{S}_A$ for the experiment could thus be the numbers from 1 to 6. Another set of possible worlds could be $\mathcal{S}_B = \{t1, \dots, t6, h1, \dots, h4\}$, with for example $t2$ meaning "tails and a roll of 2" and $h4$ meaning "heads and a roll of 4." Choose either $\mathcal{S}_A$ or $\mathcal{S}_B$ and associate probabilities with it. According to your chosen set of possible worlds and probability distribution, what is the probability of rolling either 3 or 5.

(ii) Let $\mathcal{S}_B$ be defined as above, but with a loaded coin and loaded dice. A probability distribution is given in Table 1. What is the probability that the loaded coin lands "tails"? What is the conditional probability of rolling a 4, given that the coin lands tails? Which of the loaded dice has the highest chance of rolling 4 or more?

Solution:

$t1$	$\frac{5}{18}$	$t6$	$\frac{1}{18}$
$t2$	$\frac{1}{9}$	$h1$	$\frac{1}{24}$
$t3$	$\frac{1}{9}$	$h2$	$\frac{1}{24}$
$t4$	$\frac{1}{18}$	$h3$	$\frac{1}{8}$
$t5$	$\frac{1}{18}$	$h4$	$\frac{1}{8}$

Table 1: Probabilities for $\mathcal{S}_B$.

(i) Probabilities for $\mathcal{S}_\mathcal{A}$: $P_A(1) = \cdots = P_A(4) = \frac{5}{24}$ and $P_A(5) = P_A(6) = \frac{1}{12}$.
Probabilities for $\mathcal{S}_\mathcal{B}$: $P_B(t1) = \cdots = P_B(t6) = \frac{1}{12}$ and $P_B(h1) = \cdots = P_B(h4) = \frac{1}{8}$.
$P_A(3) + P_A(5) = \frac{7}{24}$.
$P_B(t3) + P_B(t5) + P_B(h3) = \frac{7}{24}$.

(ii) $P(t) = \frac{5}{18} + \frac{1}{9} + \frac{1}{9} + \frac{1}{18} + \frac{1}{18} + \frac{1}{18} = \frac{2}{3}$.
$P(4\,|\,t) = P(t4)/P(t) = \frac{1}{12}$.
$P(4\,|\,t) + P(5\,|\,t) + P(6\,|\,t) = \frac{1}{4}$.
$P(4\,|\,h) = P(h4)/P(h) = 1/8/(\frac{1}{24} + \frac{1}{24} + \frac{1}{8} + \frac{1}{8}) = \frac{3}{8}$.
The four-sided die thus has a higher probability of rolling 4 or more, than the six-sided die.

Exercise 3 :

Calculate $P(A)$, $P(B)$, $P(A|B)$, and $P(B|A)$ from the joint probability distribution $P(A, B)$ given in Table 2.

	b_1	b_2	b_3
a_1	0.05	0.10	0.05
a_2	0.15	0.00	0.25
a_3	0.10	0.20	0.10

Table 2: The joint probability distribution for $P(A, B)$.

Solution: In order to find $P(A)$ we need to marginalize (sum) out the variable B. This is done for each state of A. Thus, for $A = a_1$ we get:

$$P(A = a_1) = \sum_B P(A = a_1, B) = 0.05 + 0.10 + 0.05 = 0.20$$

In total we end up with $P(A) = (0.2, 0.4, 0.4)$. Using a similar procedure to find $P(B)$ (now summing out A) we get $P(B) = (0.3, 0.3, 0.4)$.

In order to find $P(A|B)$ we use

$$P(A|B) = \frac{P(A, B)}{P(B)},$$

hence for $B = b_1$ we get

$$P(A|b_1) = \frac{P(A, B = b_1)}{P(B = b_1)} = \frac{(0.05, 0.15, 0.10)}{0.3} = (0.167, 0.5, 0.333).$$

The operations are similar for the other states of B, which gives $P(A|b_2) = (0.333, 0, 0.667)$ and $P(A|b_3) = 0.125, 0.625, 0.25)$.

Last we need $P(B|A)$. We can calculate these probabilities following the same steps as for $P(A|B)$, resulting in

$$\begin{aligned} P(B|a_1) &= (0.25, 0.5, 0.25) \\ P(B|a_2) &= (0.375, 0, 0.625) \\ P(B|a_3) &= (0.25, 0.5, 0.25). \end{aligned}$$

Exercise 4 :

Consider the binary variable A and the ternary variable B. Assume that B has the probability distribution $P(B) = (0.1, 0.5, 0.4)$ and that A has the conditional probability distribution given in Table 3.

	b_1	b_2	b_3
a_1	0.1	0.7	0.6
a_2	0.9	0.3	0.4

Table 3: The conditional probability distribution $P(A|B)$.

Questions:

(i) Verify that Table 3 specifies a valid conditional probability distribution.

(ii) Calculate $P(B|A)$. *Hint:* Consider Bayes rule illustrated by the temperature-sensor example that we discussed in the lecture

---

Solution:

1. We need to check that $\sum_A P(A|B = b_i) = 1$ for each state b_i of B. This is the case in Table 3, hence the we have a valid conditional probability distribution.

2. According to Bayes rule we have

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{\sum_B P(A, B)} = \frac{P(A|B)P(B)}{\sum_B P(A|B)P(B)}$$

Based on the probabilities given in the exercise, we have for the numerator that

$$P(A|B)P(B) = \begin{array}{|c|c|c|c|} \hline & b_1 & b_2 & b_3 \\ \hline a_1 & 0.1 & 0.7 & 0.6 \\ a_2 & 0.9 & 0.3 & 0.4 \\ \hline \end{array} \cdot (0.1, 0.5, 0.4) = \begin{array}{|c|c|c|c|} \hline & b_1 & b_2 & b_3 \\ \hline a_1 & 0.01 & 0.35 & 0.24 \\ a_2 & 0.09 & 0.15 & 0.16 \\ \hline \end{array}$$

	$A = yes$	$A = no$
$T = yes$	0.99	0.001
$T = no$	0.01	0.999

Table 4: Table for Exercise 5. Conditional probabilities $P(T \mid A)$ characterizing test T for A.

Note that

- $P(A|B)P(B)$ is also equal to $P(A, B)$ (an instance of the *chain rule*)!
- the multiplications are done component-wise so that we only multiply entries that match wrt. the state labels.

For the denominator, we can use the intermediate result ($P(A, B)$) that we just calculated above:

$$P(A) = \sum_B P(A, B) = \sum_B \left(\begin{array}{c|ccc} & b_1 & b_2 & b_3 \\ \hline a_1 & 0.01 & 0.35 & 0.24 \\ a_2 & 0.09 & 0.15 & 0.16 \end{array} \right) = (0.6, 0.4).$$

Plugging these results into Bayes rule we get:

$$\begin{aligned} P(B|A) &= \begin{array}{c|ccc} & b_1 & b_2 & b_3 \\ \hline a_1 & 0.01 & 0.35 & 0.24 \\ a_2 & 0.09 & 0.15 & 0.16 \end{array} \Big/ (0.6, 0.4) \\ &= \begin{array}{c|ccc} & b_1 & b_2 & b_3 \\ \hline a_1 & 0.01/0.6 & 0.35/0.6 & 0.24/0.6 \\ a_2 & 0.09/0.4 & 0.15/0.4 & 0.16/0.4 \end{array} \\ &= \begin{array}{c|ccc} & b_1 & b_2 & b_3 \\ \hline a_1 & 0.0167 & 0.5833 & 0.4000 \\ a_2 & 0.2250 & 0.3750 & 0.4000 \end{array} \end{aligned}$$

Exercise 5 :

Table 4 describes a test T for an event A. The number 0.01 is the frequency of *false negatives*, and the number 0.001 is the frequency of *false positives*.

(*i*) The police can order a blood test on drivers under the suspicion of having consumed too much alcohol. The test has the above characteristics. Experience says that 20% of the drivers under suspicion do in fact drive with too much alcohol in their blood. A suspicious driver has a positive blood test. What is the probability that the driver is guilty of driving under the influence of alcohol?

(*ii*) The police block a road, take blood samples of all drivers, and use the same test. It is estimated that one out of 1,000 drivers have too much alcohol in their blood. A driver has a positive test result. What is the probability that the driver is guilty of driving under the influence of alcohol?

Hint: Structure-wise this exercise is closely connected to the temperature-sensor example that we discussed in the lecture.

Solution: The solution procedure is the same as for the previous exercise, i.e., use Bayes rule to calculate the probabilities.

$$\begin{aligned} P(A = y \mid T = p) &= \frac{P(A = y, T = p)}{P(T = p)} \\ &= \frac{P(T = p \mid A = y)P(A = y)}{P(T = p)} \end{aligned}$$

From the problem specification, we have the numbers to put into the numerator, but we still need to calculate the denominator:

$$\begin{aligned} P(T = p) &= P(T = p, A = t) + P(T = p, A = f) \\ &= P(T = p \mid A = t)P(A = t) + P(T = p \mid A = f)P(A = f) \end{aligned}$$

By plugging this into the previous expression we get:

$$\begin{aligned} P(A = y \mid T = p) &= \frac{P(T = p \mid A = y)P(A = y)}{P(T = p \mid A = t)P(A = t) + P(T = p \mid A = f)P(A = f)} \\ &= \frac{0.99 \cdot 0.2}{0.99 \cdot 0.2 + 0.001 \cdot 0.8} \\ &= 0.996 \end{aligned}$$

In the second part of the exercise we change the prior probability $P(A)$ for A and update the calculations using the new probabilities. This gives $P(A = t) = 0.498$. Notice how the change in prior distribution affected the posterior distribution for A; the accuracy of a test should always be considered relative to the frequency of the event which you are trying to predict.

Exercise 6 :

A routine DNA test is performed on a person (this exercise is set in the not too distant future!). The test T gives a positive result for a rare genetic mutation M linked to Alzheimer's disease. The mutation is present in only 1 in a million people. The test is 99.99% accurate, i.e. it will give a wrong result in 1 out of 10000 tests performed. Should the person be worried, i.e., what is the probability that the person has the mutation given that the test showed a positive result?

Hint: This is partly a modeling exercises and partly a calculation exercise. First you need to formalize the problem:

- What are the relevant variables and what states do they have?
- Based on the description above, what probability distributions can you infer for the variables?

Based on this formalization, you need to find the rules required to answer the question about the probability of a mutation given a positive test result.

	b_1	b_2
a_1	$(0.006, 0.054)$	$(0.048, 0.432)$
a_2	$(0.014, 0.126)$	$(0.032, 0.288)$

Table 5: $P(A, B, C)$ for Exercise 7.

Solution: We analyze the problem using the two random variables $\mathit{Test} \in \{\mathit{positive}, \mathit{negative}\}$, $\mathit{Mutation} \in \{\mathit{present}, \mathit{absent}\}$. What our patient should be interested in is the probability of having the mutation, given everything he/she knows, i.e. the fact that there was a positive test result. Thus, we need to compute

$$P(\mathit{Mutation} = \mathit{present} \mid \mathit{Test} = \mathit{positive}).$$

According to Bayes rule, this is equal to

$$P(\mathit{Test} = \mathit{positive} \mid \mathit{Mutation} = \mathit{present}) \frac{P(\mathit{Mutation} = \mathit{present})}{P(\mathit{Test} = \mathit{positive})}$$

Of the probabilities on the right, we have $P(\mathit{Test} = \mathit{positive} \mid \mathit{Mutation} = \mathit{present}) = 0.9999$, and $P(\mathit{Mutation} = \mathit{present}) = 0.000001$. We still need $P(\mathit{Test} = \mathit{positive})$. This can be computed as

$$\begin{aligned} P(\mathit{Test} = \mathit{positive}) &= P(\mathit{Test} = \mathit{positive}, \mathit{Mutation} = \mathit{present}) \\ &\quad + P(\mathit{Test} = \mathit{positive}, \mathit{Mutation} = \mathit{absent}) \\ &= P(\mathit{Test} = \mathit{positive} \mid \mathit{Mutation} = \mathit{present}) P(\mathit{Mutation} = \mathit{present}) \\ &\quad + P(\mathit{Test} = \mathit{positive} \mid \mathit{Mutation} = \mathit{absent}) P(\mathit{Mutation} = \mathit{absent}) \\ &= 0.9999 \cdot 0.000001 + 0.0001 \cdot 0.999999 \sim 0.0001 \end{aligned}$$

We now get

$$P(\mathit{Mutation} = \mathit{present} \mid \mathit{Test} = \mathit{positive}) = 0.9999 \cdot \frac{0.000001}{0.0001} = 0.009999.$$

Thus, there is only about a 1% chance that the patient has the mutation.

Exercise 7 :

In Table 5, a joint probability table for the binary variables A, B, and C is given.

- Calculate $P(B, C)$ and $P(B)$.
- Are A and C independent given B?

Solution:

For the first part of the exercise, consider calculating $P(B, c_1)$. To do that we marginalize out A, hence

$$P(B, c_1) = (0.006 + 0.014, 0.048 + 0.032) = (0.02, 0.08).$$

The remaining probabilities are calculated similarly, and we get:

(i) $P(B, c_1) = (0.02, 0.08), P(B, c_2) = (0.18, 0.72), P(B) = (0.2, 0.8)$

For the second part of the exercise we can check that

$$P(A|C = c_1, B) = P(A|C = c_2, B)$$

to verify that A and C are independent given B. These probabilities can be calculated as

$$P(A|B, C) = \frac{P(A, B, C)}{P(B, C)},$$

where the denominator was calculated in the first part of the exercise. When doing the calculations we get

(ii) $P(A|b_1, c_1) = (0.3, 0.7) = P(A|b_1, c_2), P(A|b_2, c_1) = (0.6, 0.4) = P(A|b_2, c_2)$,

hence A and C are independent given B.

Exercise 8 :

In a university far, far away… every year a lot of students take the Artificial Intelligence (AI) lecture. From years and years of experience it is known that there are three (distinct) types of students at the university:

- *H(ard-working)*, who solve all exercises and qualify for the exam,
- *L(azy)*, who get enough points to qualify for the exam, but then stop working on the exercises,
- *U(nfortunate)*, who do not qualify for the exam.

Typically, 90% of the H-students pass the AI exam with a good grade, but only half of the L-students get a good grade. Every student that qualifies for an exam takes it.

a) Last year, 7 out of 10 students belonged to the L category and there were 10% U-students. What was the probability that a student who got a good grade in the AI exam was of $Type\ H$?

b) In another year 400 students took the AI course. We know that 36 H-students got a good grade, and that 80% of those who got a good grade were L-students. What's the probability that a student in that year was a U-student?

c) Assume there exists a second lecture at the university, namely Machine Learning (ML). This year, we expect 35% of the students to be working hard and 27% of the students to be lazy. Overall, 40% of the students get a good grade in ML and 45% of the students get a good grade in AI. Assume the latter two events to be independent. Furthermore, 80% of the students that get a good grade in ML are H-students, and 5% of the students with a good grade in both ML and AI are lazy.

   Are the events of getting a good grade in AI and getting a good grade in ML conditionally independent given the type of student? Justify your answer by checking if the condition from the definition of conditional independence holds.

Solution:

We know that:

(1) $P(Grade = good \mid Type = H) = 0.9$

(2) $P(Grade = good \mid Type = L) = 0.5$

a) Further, for a) we know that:

(3a) $P(Type = L) = 0.7$

(4a) $P(Type = U) = 0.1$

We look for: $P(Type = H \mid Grade = good)$
Using the Bayes' rule,

$$P(Type = H \mid Grade = good) = \frac{P(Grade = good \mid Type = H)P(Type = H)}{P(Grade = good)}$$

To compute this we need $P(Type = H)$ and $P(Grade = good)$.

$$P(Type = H) = 1 - P(Type = L) - P(Type = U) = 0.2 \qquad /(3a, 4a)$$

$$\begin{aligned} P(Grade = good) &= P(Grade = good \wedge Type = H) + P(Grade = good \wedge Type = L) = \\ &= P(Grade = good \mid Type = H)P(Type = H) + \\ &\quad + P(Grade = good \mid Type = L)P(Type = L) = \\ &= 0.9 * 0.2 + 0.5 * 0.7 = 0.53 \qquad /(1, 2, 3a, 4a) \end{aligned}$$

Therefore

$$P(Type = H \mid Grade = good) = \frac{0.9 * 0.2}{0.53} = 33.96\%.$$

The probability that a student who got a good grade in the AI exam was of $Type\ H$ is 33.96%.

b) (3b) *# students* $= 400$

(4b) *# H-students with good Grade* $= 36$

(5b) $P(Type = L \mid Grade = good) = 0.8$

We look for: $P(Type = U)$

$$P(Type = U) = 1 - P(Type = H) - P(Type = L)$$

So we need $P(Type = H)$ and $P(Type = L)$.

Computing $P(Type = H)$:

$$P(Type = H) = \frac{P(Type = H \wedge Grade = good)}{P(Grade = good \mid Type = H)} = \frac{0.09}{0.9} = 0.1 \qquad /(1)$$

since

$$P(Type = H \mid Grade = good) = 1 - P(Type = L \mid Grade = good) = 1 - 0.8 = 0.2, \qquad /(5b)$$

and

$$P(Type = H \wedge Grade = good) = \frac{36}{400} = 0.09. \qquad /(3b, 4b)$$

Computing $P(Type = L)$:

Using the rearranged Bayes' rule:

$$P(Type = L) = \frac{P(Type = L \mid Grade = good)P(Grade = good)}{P(Grade = good \mid Type = L)} = \frac{0.8 * 0.45}{0.5} = 0.72 \qquad /(2, 5b)$$

since

$$P(Grade = good) = \frac{P(Type = H \wedge Grade = good)}{P(Type = H \mid Grade = good)} = \frac{0.09}{0.2} = 0.45$$

Therefore

$$P(Type = U) = 1 - 0.1 - 0.72 = 0.18.$$

c) We use "X = g" to abbreviate $GradeX = good$.

$$\begin{aligned}\mathbf{P}(AI = g, ML = g \mid Type) &= \frac{\mathbf{P}(Type \mid AI = g, ML = g)P(AI = g, ML = g)}{\mathbf{P}(Type)} \\ &= \frac{\mathbf{P}(Type \mid AI = g, ML = g)P(AI = g)P(ML = g)}{\mathbf{P}(Type)}\end{aligned}$$

$$\mathbf{P}(AI = g, ML = g \mid Type = H) = \frac{0.95 \cdot 0.45 \cdot 0.4}{0.35} \approx 0.49$$

We know that $P(AI = g \mid Type = H) = 0.9$

$$\mathbf{P}(ML = g \mid Type) = \frac{\mathbf{P}(Type \mid ML = g)P(ML = g)}{\mathbf{P}(Type)}$$

$$P(AI = g \mid Type = H)P(ML = g \mid Type = H) = 0.9 \cdot \frac{0.8 \cdot 0.4}{0.35} \approx 0.82 \neq 0.49$$

Thus, the two events are not conditionally independent given the variable $Type$.

Exercise 9 :

Solve Exercise 1(a-d) in PM.

Solution:

$$P(A = t \,|\, G = m) = \frac{37}{69} = 0.54$$
$$P(A = t \,|\, G = f) = \frac{14}{31} = 0.45$$
$$P(A = t \,|\, G = m, D = 1) = \frac{32}{50} = 0.64$$
$$P(A = t \,|\, G = f, D = 1) = \frac{7}{10} = 0.7$$
$$P(A = t \,|\, G = m, D = 2) = \frac{5}{19} = 0.26$$
$$P(A = t \,|\, G = f, D = 2) = \frac{7}{21} = 0.3$$