Last Name: First Name:

Matriculation Number:

Exam
Artificial Intelligence 2

August 2, 2022

	To be used for grading, do not write here											
prob.	1.1	1.2	2.1	2.2	3.1	3.2	4.1	4.2	4.3	5.1	Sum	grade
total	7	10	8	14	15	10	10	6	6	9	95	
reached												

Organizational Information

Please read the following directions carefully and acknowledge them with your signature.

1. Please place your student ID card and a photo ID on the table for checking.
2. You can reach 95 points if you fully solve all problems. You will only need 90 points for a perfect score, i.e. 5 points are bonus points.
3. No resources or tools are allowed except for a pen.
4. You have 90 min (sharp) for the exam.
5. Write the solutions directly on the sheets, no other paper will be graded.
6. If you have to abort the exam for health reasons, your inability to sit the exam must be certified by an examination at the University Hospital. Please notify the exam proctors and have them give you the respective form.
7. Please make sure that your copy of the exam is complete (18 pages excluding cover sheet and organizational information pages) and has a clear print. **Do not forget to add your personal information on the cover sheet and to sign this declaration.**

Declaration: With my signature I certify having received the full exam document and having read the organizational information above.

Erlangen, August 2, 2022 ....
(signature)

Organisatorisches

Bitte lesen die folgenden Anweisungen genau und bestätigen Sie diese mit Ihrer Unterschrift.

1. Bitte legen Sie Ihren Studierendenausweis und einen Lichtbildausweis zur Personenkontrolle bereit!

2. Sie können 95 Punkte erreichen, wenn Sie alle Aufgaben vollständig lösen. Allerdings zählen 90 Punkte bereits als volle Punktzahl, d.h. 5 Punkte sind Bonuspunkte.
3. Es sind keine Hilfsmittel erlaubt außer einem Stift.
4. Die Bearbeitungszeit beträgt genau 90 min.
5. Schreiben Sie die Lösungen direkt auf die ausgeteilten Aufgabenblätter. Andere Blätter werden nicht bewertet.
6. Wenn Sie die Prüfung aus gesundheitlichen Gründen abbrechen müssen, so muss Ihre Prüfungsunfähigkeit durch eine Untersuchung in der Universitätsklinik nachgewiesen werden. Melden Sie sich in jedem Fall bei der Aufsicht und lassen Sie sich das entsprechende Formular aushändigen.
7. Überprüfen Sie Ihr Exemplar der Klausur auf Vollständigkeit (18 Seiten exklusive Deckblatt und Hinweise) und einwandfreies Druckbild! **Vergessen Sie nicht, auf dem Deckblatt die Angaben zur Person einzutragen und diese Erklärung zu unterschreiben!**

Erklärung: Durch meine Unterschrift bestätige ich den Empfang der vollständigen Klausurunterlagen und die Kenntnisnahme der obigen Informationen.

Erlangen, August 2, 2022 ....

(Unterschrift)

This page was intentionally left blank for extra space

1 Probabilities

Problem 1.1 (Python) 7 pt

Consider the Python program below.

1. Which *mathematical function* does the method `foo` compute?

2. Assume random variables X with domain $\{0, \ldots, m-1\}$ and Y with domain $\{0, \ldots, n-1\}$. Assume the following Python objects
   - C holds the conditional probability distribution $P(X \mid Y)$, i.e., $C[i][j] = P(X = i \mid Y = j)$.
   - D holds the probability distribution $P(Y)$, i.e., $D[j] = P(Y = j)$.

   Complete the definition of E in the program below in such a way that it holds the probability distribution $P(X)$, i.e., $E[i] = P(X = i)$. (Hint: This requires relatively little code.)

```
def foo(a,b):
    l = len(a)
    m = len(a[0])
    n = len(b[0])
    res = []
    for i in range(l):
        row = []
        for j in range(n):
            s = 0
            for k in range(m):
               s += a[i][k] * b[k][j]
            row.append(s)
        res.append(row)
    return res


E =
```

Problem 1.2 (Calculations) 10 pt

Assume three Boolean random variables X, Y, Z, whose joint distribution $P(X, Y, Z)$ is given by

x	y	z	$P(X = x, Y = y, Z = z)$
$true$	$true$	$true$	a
$true$	$true$	$false$	b
$true$	$false$	$true$	c
$true$	$false$	$false$	d
$false$	$true$	$true$	e
$false$	$true$	$false$	f
$false$	$false$	$true$	g
$false$	$false$	$false$	h

1. In terms of a, b, c, d, e, f, g, h, give $P(X = true, Y = false)$.

2. In terms of a, b, c, d, e, f, g, h, give $P(X = true \mid Y = false)$.

3. Which of the following are true if Y and Z are conditionally independent given X?
   - ☐ $a + b + e + f = a + c + e + g$
   - ☐ T

   $a = (a + b) \cdot (a + c)/(a + b + c + d)$
   - ☐ $e = (f \cdot g)/h$
   - ☐ $a = e$

This page was intentionally left blank for extra space

2 Bayesian Reasoning

When working with an upper case Boolean random variable X, you may abbreviate the event $X = true$ by the corresponding lower-case letter x. If you do that, make sure the distinction between upper and lower case letters is clear in your writing.

Problem 2.1 (Bayes' Rule) 8 pt

Assume you are trying to predict whether a particular topic comes up in an exam. You have collected the following data:

- 30% of all topics come up in the exam.
- 50% of all topics come up in the assignments.
- If a topic comes up in an exam, it was covered by an assignment 60% of the time.
- If a topic comes up in an exam, it also came up in a recent exam 80% of the time.

You model this situation using 3 Boolean random variables E (comes up in exam), A (covered by assignments), and R (came up in a recent exam).

1. By filling in the gaps below, state for each number in the text above, which probability it describes.

   1. $P($ [ ] $) = 0.3$
   2. $P($ [ ] $) = 0.5$
   3. $P($ [ ] $E = true\ |{=})0.6$
   4. $P($ [ ] $E = true\ |{=})0.8$

2. When modeling this situation, is it reasonable to assume that A and R are stochastically independent? Why (not)?

3. The topic you are interested in was covered by an assignment. Using Bayes' rule, calculate the probability that it will come up in the exam.

Problem 2.2 (Bayesian Networks) 14 pt

Consider the following situation:

- Covid can cause a sickness and/or fever.
- Fever itself is dangerous and can cause sickness.
- Tests can detect Covid. But a false-positive Covid test may cause sickness via a kind of Placebo effect.
- There are no other causal relationships.

You want to model this situation using Boolean random variables C (Covid infection), F (fever), S (sickness), and T (positive Covid test).

1. Give a good variable ordering for forming a Bayesian network for this situation.

2. Give the resulting network.

3. You have a fever, feel sick, and have tested positive for Covid. Now you want to determine if you have Covid. What are the query, evidence, and hidden variables?

4. Assume your network is $C \to F \to T \leftarrow S$ (which *may or may not* be a correct solution to the above question). Which probabilities are stored in the conditional probability table of node T?

5. Again using the network $C \to F \to T \leftarrow S$, give the formula for

$$P(C = true, T = true, S = true)$$

in terms of the entries of the conditional probability table of that network.

This page was intentionally left blank for extra space

3 Markovian Reasoning

Problem 3.1 (Hidden Markov Models) 15 pt

Consider the following situation:

- You make daily observations about your business B. Each day business is either good (b_1) or bad (b_2).
- You know this is caused by the weather W, which can be rainy (w_1), cloudy (w_2), or sunny (w_3).
- You have previously obtained the following information:
  - when the weather is rainy, your business is good 36% of the time,
  - when the weather is cloudy, your business is good 84% of the time,
  - when the weather is sunny, your business is good 90% of the time,
  - half the time, the weather is the same as on the previous day,
  - when the weather changes from one day to the next, each change is equally likely.

You want to model this situation as a hidden Markov model with two families of random variables indexed by day number d.

1. Give the state and evidence variables and their domains.

2. How can you tell that the sensor model is stationary here?

3. What order does the model have?
4. Complete the following sentences:
   5. The transition model T is given by the matrix

$$T = \begin{pmatrix} \quad \end{pmatrix} \qquad \text{where} \qquad T_{ij} = P(W_{d+1} = w_j \mid W_d = w_i).$$

   6. The sensor model S is given by the matrix

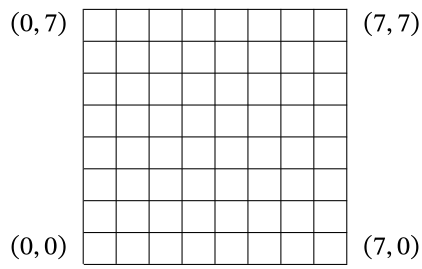
$$S = \begin{pmatrix} \quad \end{pmatrix} \qquad \text{where} \qquad S_{ij} = P(B_d = b_j \mid W_d = w_i).$$

7. Let T be as above and let $\mathbf{v}$ be a 3-dimensional vector whose coefficients sum to 1 What is the intuitive meaning of the property $T \cdot \mathbf{v} = \mathbf{v}$?

8. It was sunny yesterday, and your business is good today. You want to use filtering to obtain the probability distribution of today's weather.
   You proceed as follows:
   9. Give the recursive filtering equation for $f_{1:d+1}$.
   10. Give the initial value $f_{1:0}$ to use in this case.
   11. Give the diagonal sensor matrix O_1 to use in this case.
   12. Compute the resulting distribution.
       Fully compute all values including the normalization. (This does not require approximations or a calculator.)

Problem 3.2 (Utility and Decision Processes) 10 pt

Consider an agent moving on an 8×8 grid as indicated in the picture below. The agent can move up, down, right, and left except where restricted by the edges of the grid. Every action results in moving one step in that direction with probability 75% and no move otherwise. The agent's goal is to get to the location $(7, 7)$.



1. Choose an appropriate reward function and model this situation as a Markov Decision Process.

2. Give an optimal policy π^*.

3. Now ignore the rewards, and assume we use a fixed utility $i + j$ for the field (i, j). Compute the expected utility of moving up once in state $(1, 1)$.

4. Now assume the agent is unable to tell whether an action resulted in a move or not. Explain informally how that would change the modeling.

4 Learning

Problem 4.1 (Decision Trees and Lists) 10 pt

You observe the set of values below for 6 games of a sports team. You want to predict the result based on weather, location, and opponent.

#	Weather	Location	Opponent	Result
1	Rainy	Home	Weak	Win
2	Sunny	Away	Weak	Win
3	Sunny	Home	Strong	Loss
4	Sunny	Away	Weak	Win
5	Cloudy	Away	Strong	Loss
6	Sunny	Home	Strong	Loss

1. Draw the decision tree that arises if attributes are chosen according to the priority $Location, Weather, Opponent$.

2. Give all minimal sets A of attributes such that $A \succ Result$.

3. How can such a minimal set A be exploited to find a decision tree?

4. You want to build a decision list for the result using tests with literals of the form $attribute = number$. Give the shortest possible decision list.

5. Now stop using the above observations. Instead, assume some set of observations for which a decision list exists, and assume that the determination $Weather, Location \succ Result$ holds. In the worst case, what is the number of tests in the shortest decision list?

Problem 4.2 (Statistical Learning) 6 pt

You observe the values below for 20 games of a sports team. You want to predict the result based on weather and opponent.

Weather	Opponent	Number of wins	Number of losses
Rainy	Weak	3	1
Cloudy	Weak	0	1
Sunny	Weak	4	2
Rainy	Strong	0	2
Cloudy	Strong	2	3
Sunny	Strong	0	2

1. What is the hypothesis space for this situation, seen as an *inductive learning problem*?

2. Explain whether we can learn the function by building a decision tree.

3. To apply Bayesian learning, we model this situation as a Bayesian network $W \rightarrow R \leftarrow O$ using random variables W (weather), O (opponent), and R (game result). What are the resulting entries of the conditional probability table for the cases
   1. $P(W = rainy) =$ ☐
   2. $P(R = win \mid O = weak) =$ ☐

Problem 4.3 (Support Vector Machines) 6 pt

Consider a set of points $\mathbf{x} = \langle \mathbf{x}_1, \mathbf{x}_2 \rangle$ in $\mathbb{R}^2$ that are classified as either $y = +1$ or $y = -1$.

1. Give the hypothesis space for finding a linear separator.

2. Given a linear separator $h(\mathbf{x})$, which formula computes the classification y of a vector $\mathbf{x}$?

3. What is the point of transforming a dataset into a higher-dimensional space?

4. In this context, briefly discuss the usefulness of the transformation $F(\langle \mathbf{x}_1, \mathbf{x}_2 \rangle) = \langle \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_1 + \mathbf{x}_2 \rangle$.

5 Natural Language Processing

Problem 5.1 (Information Retrieval) 9 pt

Consider the following three texts

- d_1: "The man is tall."
- d_2: "The tall man sees the woman."
- d_3: "The woman shouts."

Let $D = \{d_1, d_2, d_3\}$.

Below we use alphabetical order for the vector components:

is, man, sees, shouts, tall, the, woman

Simplify all results as much as possible without introducing approximate values.

1. What is the idea of cosine similarity for comparing a query against the documents in D?

2. Give the vector $tf(\_, d_2)$

3. Give the vector $idf(\_, D)$.

4. For $d \in D$ and a word t, give the definition of $tfidf(t, d, D)$.

5. What is the benefit of using $tfidf$ instead of tf for using cosine similarity?

This page was intentionally left blank for extra space