

Last Name:

First Name

Matriculation Number:

Birth Date:

Seat:

## Exam Artificial Intelligence 2

Feb 17, 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	Sum	grade
total	10	14	15	10	15	15	5	10	94	
reached										

Exam Grade:

Bonus Points:

Final Grade:

## 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. The point distributions for the problems are provisional.
3. You can reach 94 points if you fully solve all problems. You will only need 89 points for a perfect score, i.e. 5 points are bonus points.
4. No resources or tools are allowed except for a pen.
5. You have 90 min (sharp) for the test
6. Write the solutions directly on the sheets, no other paper will be graded.
7. 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.
8. Please make sure that your copy of the exam is complete (17 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, Feb 17, 2022

.....  
(signature)

## Organisatorisches

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

1. Bitte legen Sie Ihren Studentenausweis und einen Lichtbildausweis zur Personenkontrolle bereit!
2. Die angegebene Punkteverteilung gilt unter Vorbehalt.
3. Sie können 94 Punkte erreichen, wenn Sie alle Aufgaben vollständig lösen. Allerdings zählen 89 Punkte bereits als volle Punktzahl, d.h. 5 Punkte sind Bonuspunkte.
4. Es sind keine Hilfsmittel erlaubt außer eines Stifts.
5. Die Bearbeitungszeit beträgt genau 90 min.
6. Schreiben Sie die Lösungen direkt auf die ausgeteilten Aufgabenblätter. Andere Blätter werden nicht bewertet.
7. 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.
8. Überprüfen Sie Ihr Exemplar der Klausur auf Vollständigkeit (17 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, Feb 17, 2022

.....

(Unterschrift)

# 1 Bayesian Reasoning

---

**Note:** 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 1.1 (Bayesian Rules)

10 pt

5 pt

1. Assume a prevalence of SARS-CoV-2 infections in your area of  $1/100$  (share of people who are infected) and a positive-test rate of 19% (share of tests that return positive). Moreover, assume you are using a test with a false-positive rate of 1% (share of uninfected people who test positive) and a false-negative rate of 5% (share of infected people who test negative). You have tested positive. Apply Bayes' rule to determine the probability that you are infected.
2. Assume three random variables  $A, B, C$  such that  $A$  and  $B$  are conditionally independent given  $C$ . You know
  - the probability distribution of  $C$ ,
  - the conditional probability distribution of  $A$  given  $C$ ,
  - the conditional probability distribution of  $B$  given  $C$ .

5 pt

In terms of the above, give the formula for the probability distribution of  $C$  given the event  $A = a, B = b$ .

This page was intentionally left blank for extra space

**Problem 1.2 (Bayesian Networks)**

14 pt

Consider the following situation:

- To attend university, people need academic qualification and interest.
- Both of those are more likely for people with university-educated fathers.
- Interest is also generated by special schooling programs.
- There are no other causal relationships.

You want to model this situation using Boolean random variables  $U$  (attend university),  $Q$  (qualified),  $I$  (interested),  $F$  (father attended university), and  $S$  (covered by a school program).

1. You do not know if a person attends university and want to determine that by asking about their father's education. Which variables are the evidence, query, and hidden variables? 2 pt
2. Give a good variable ordering for forming a Bayesian network for this situation. 3 pt
3. Give the resulting network. 3 pt
4. Now assume your network is  $Q \rightarrow F \rightarrow I \leftarrow U \rightarrow S$  (which *may or may not* be a correct solution to 3. above) and assume that  $I$  is deterministic.
  - (a) What does being deterministic mean for the probability distribution of  $I$ ? 2 pt
  - (b) Which entries in the conditional probability table do we save by exploiting that  $I$  is deterministic? 2 pt
  - (c) Give the formula for
$$P(I|F = true, U = true, S = false)$$
in terms of the entries of the conditional probability table of the network. 2 pt

This page was intentionally left blank for extra space

## 2 Markovian Reasoning

### Problem 2.1 (Hidden Markov Models)

15 pt

Consider the following situation:

- You record daily how your mood was. Each day you feel good or bad.
- You know this is caused by your daily workload, which is either high, medium, or low.
- You have previously assessed the following information:
  - high workload makes you feel bad 80% of the time,
  - medium workload makes you feel bad 60% of the time,
  - low workload makes you feel bad 10% of the time,
  - from day to day, your workload never jumps from high to low or vice versa,
  - 40% of the time, your workload stays the same from one day to the next,
  - if your workload is medium, an increased and a decreased workload on the next day are equally likely,
  - 2 days ago your workload was high.

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

2 pt

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

1 pt

2. How can you tell whether this model is stationary?

1 pt

3. How can you tell whether this model has Markov order 1?

4. Complete the following sentences:

3 pt

(a) The transition model is given by the matrix

$$T = \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \quad \text{where} \quad T_{ij} = P(\text{state} = j \mid \text{state} = i).$$

3 pt

(b) The sensor model is given by the matrix

$$M = \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \quad \text{where} \quad M_{ij} = P(\text{evidence} = j \mid \text{state} = i).$$

where we map values to matrix indices via  $good = 1$  and  $bad = 2$  as well as  $low = 1$ ,  $medium = 2$ , and  $high = 3$ .

5 pt

5. Your mood was good yesterday ( $d = 1$ ) and bad today ( $d = 2$ ). Give the matrix form of the recursive filtering equation and state precisely which concrete values to plug in to obtain the probability distribution of your workload at  $d = 2$ . (You do not have to actually compute the distribution.)



This page was intentionally left blank for extra space

**Problem 2.2 (Markov Decision Processes)**

10 pt

1. Give an optimal policy  $\pi^*$  for the following MDP:

4 pt

- set  $S$  of states: integers  $\mathbb{Z}$  with initial state 0
- set of actions for  $s \in S$ :  $A(s) = \{-1, 0, 1\}$
- transition model for  $s, s' \in S$  and  $a \in A(s)$ :  $P(s'|s, a)$  is such that
  - \*  $s' = s + a$  with probability  $2/3$ ,
  - \*  $s' = s - a$  with probability  $1/3$ .
- reward function:  $R(10) = 1$  and  $R(s) = -0.1$  for all other states  $s$

2. Assume that all utilities are initialized as  $U(s) = 0$ . Perform one round of value iteration using  $\gamma = 1$ .

2 pt

3. Now assume that utilities are initialized as  $U(s) = R(s)$ . Give the value of  $U(0)$  after one round of value iteration using  $\gamma = 1$ .

4 pt

This page was intentionally left blank for extra space

### 3 Learning

**Problem 3.1 (Decision Trees)**

15 pt

You observe the values below for 6 different football games of your favorite team. You want to construct a decision tree that predicts the result.

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

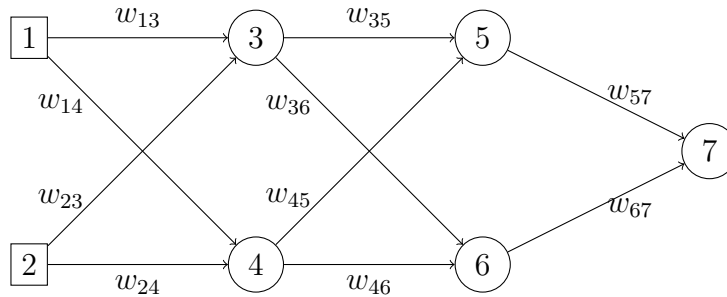
- Assume you choose attributes in the order *Opponent, Location, Weather, Day*. Give the resulting decision tree. 4 pt
- How does the information-theoretic algorithm choose an attribute? 1 pt
- Without using the above observations, give the formula for the information gain of the attribute *Weather*. 3 pt
- Using the above observations, give the results of
  - $I(P(Result)) =$
  - $P(Result = Loss | Weather = Rainy) =$
 You do not have to compute irrational logarithms. 2 pt
- Give all minimal sets  $A$  of attributes such that  $A \succ Result$  hold for the above observations. 3 pt
- Explain *why or why not* the determination  $Result, Location \succ Opponent$  holds for the above observations. 2 pt

This page was intentionally left blank for extra space

**Problem 3.2 (Neural Networks)**

15 pt

Consider the neural network without bias given below where units 1, 2 are inputs, unit 7 is output, weights are given by the labels on the edges, and units 3, 4, 5, 6, 7 are perceptron units with activation function  $T(x) = 1$  for  $x > 0.5$  and  $T(x) = 0$  otherwise.



1. Which nodes are part of a hidden layer? 1 pt
  
2. Which important property of the network changes if we add an edge from 5 to 1? 1 pt
  
3. Give the formula for the activation  $a_5$  of unit 5 in terms of the activations  $a_3$  and  $a_4$  and the weights  $w_{ij}$ . 3 pt
  
4. Assume  $w_{ij} = 1$  for all weights  $w_{ij}$  and  $a_1 = a_2 = 1$ . What is the resulting output  $a_7$ ? 2 pt
  
5. Assume  $a_1, a_2 \in \{0, 1\}$  and  $w_{13} = w_{23} = 1$  and  $w_{14} = w_{24} = 0.3$  and  $w_{35} = w_{46} = 1$  and  $w_{57} = 1$ . Choose appropriate values for the other weights such that the network implements the XOR function, i.e.,  $a_7 = a_1 XOR a_2$ . 4 pt
  
6. Complete the high-level description of the back-propagation algorithm *on the next page*. 4 pt

To learn a target function  $a_7 = f(a_1, a_2)$ , do the following for each input to  $f$ :

- compute the
  
- determine the error between
  
- and propagate
  
- use the propagation results to update

## 4 Natural Language Processing

### Problem 4.1 (Language Models)

5 pt

1. How can we obtain a trigram model for a language? Explain the probability distribution involved. 2 pt

2. Explain informally how we can use trigram models to identify the language of a document  $D$ . 2 pt

3. Explain briefly what named entity recognition is. 1 pt



**Problem 4.2 (Information Retrieval)**

10 pt

Consider the following two texts

- $d_1$ : “The air is cold.”
- $d_2$ : “The sun warms the air.”
- $d_3$ : “The day is over.”

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

Below we use alphabetical order for the vector components:

air, cold, day, is, over, sun, the, warms

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

1. Give the vector  $tf(\_, d_2)$ . 3 pt
2. Give the vector  $idf(\_, D)$ . 3 pt
3. Let  $w$  be the word “is”. Give the value  $tfidf(w, d_1, D)$ . 2 pt
4. Now assume we have computed  $tfidf(x, d, D)$  for every word  $x$  and every  $d \in D$ . 2 pt  
How do we use those values to rank the texts in  $D$  for a query  $q$ ?

This page was intentionally left blank for extra space

This page was intentionally left blank for extra space

This page was intentionally left blank for extra space