Exam Artificial Intelligence for Module 6 (201700269) Bachelor TCS and BIT.

January 10 2019, 8:45-11:45

Adapted version: Thursday 19th December, 2019 at 17:02

Name and student number

Name: ____________________

Student number: ____________________

Introduction

Important message: When you have finished your exam do not leave your table to hand in your exam material but raise your hand and wait until one of the supervisors picks up your exam material. After that you are allowed to leave the room.

This exam is closed book, you may only use a simple calculator (addition, substraction, multiplication, division and exponentiation).

This examination consists multiple-choice questions, for which you have 3 hours.

Tips:

- Read each question carefully keeping the possible answers covered.
- Try to answer the question yourself, before you look at the answers you are given to choose from. Make a note of your first thoughts and calculations on a scribbling-paper (kladpapier).
- Beware of double negations (negatives) as these can be confusing.
- Do not stay on any one question too long. If you do not know the answer and have spent more than 10 minutes on the question, move on to the next question and come back to this one later.
- If you have any time over at the end, check your answers.
- Fill in your answers on this question form first and transfer them to the answer form at the end.
- At the last page of this exam you can find a table with values for $-p\log_2(p)$ and the `Logistic` or `sigmoid` function.

Good luck!

Multiple-choice questions

Questions about AI in general and Intelligent Agents

1. Can a simple reflex agent be perfectly rational for a certain performance measure and task environment?

   (a) This is dependent **only** on the performance measure.

   (b) This is dependent on **both** the task environment and the performance measure.

   (c) This is dependent **only** on the task environment.

   (d) A simple reflex agent is **never** perfectly rational.

2. Consider the following statements about search:

   (i) A^*-Search is **not** applicable (the found solution is not necessary optimal) in a deterministic, static and fully observable environment.

   (ii) A^*-Search is applicable (the found solution is optimal) in a non-deterministic, static and fully observable environment.

   Which of the above statements are true?

   (a) Both statements (i) and (ii) are false.

   (b) Only statement (i) is true.

   (c) Only statement (ii) is true.

   (d) Both statements (i) and (ii) are true.

3. Consider an A* search algorithm for which the heuristic function h is equal to 0, i.e. $h(n) = 0$ for all nodes n. To which of the following search algorithms is this kind of A^* search algorithm equivalent (meaning an equal search strategy)?

   (a) Greedy best-first search.

   (b) Depth-First Search.

   (c) Breadth-First Search.

   (d) None of the above.

4. Underlying each search problem is a *search graph* in which the states are the vertices of the search graph and the edges (connections) are determined by the possible actions; there is an edge from s to s' with label a in the search graph, if and only if there is an action a which leads from state s to state s'. These two conditions together completely define the *search graph*. A cycle in the search graph is a vertex (state) s and a non-empty sequence of actions as such that if we start in s and execute the sequence of actions as then we will end up in s again. Consider the following statements about search problems and graphs:

   (i) If the **Breadth-First Search Graph Algorithm** does terminate on a given search problem then the corresponding search graph contains no cycles.

   (ii) If the corresponding search graph is finite and contains no cycles then the **Depth First Search Tree Algorithm** always terminates on the corresponding search problem.

   Which of the following claims is true?

   (a) Both statements (i) and (ii) are false.

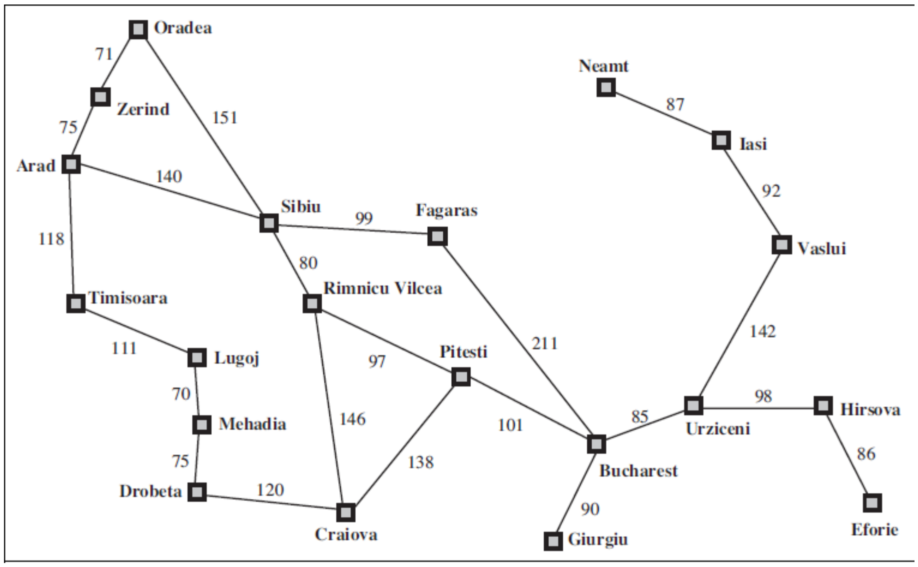
   (b) Only statement (i) is true.

   (c) Both statements (i) and (ii) are true.

   (d) Only statement (ii) is true.

Questions about search and problem solving

5. Consider the simplified roadmap of Romania. The number above a road gives the road distance between the connecting cities.



Assume that a tourist wants to drive from Oradea to Bucharest and he/she applies A^* graph search with the following straight-line heuristic function:

Arad	366	**Mehadia**	241
Bucharest	0	**Neamt**	234
Craiova	160	**Oradea**	380
Drobeta	242	**Pitesti**	100
Eforie	161	**Rimnicu Vilcea**	193
Fagaras	176	**Sibiu**	253
Giurgiu	77	**Timisoara**	329
Hirsova	151	**Urziceni**	80
Iasi	226	**Vaslui**	199
Lugoj	244	**Zerind**	374

What will be third node expanded in the A^* graph search algorithm. The start node corresponding to Oradea will be the first.

(a) The node corresponding to Zerind.

(b) The node corresponding to Sibiu.

(c) The node corresponding to Rimnicu Vilcea.

(d) None of the above.

6. Consider the following statements about the A^* graph search algorithm for which the underlying search space is finite:

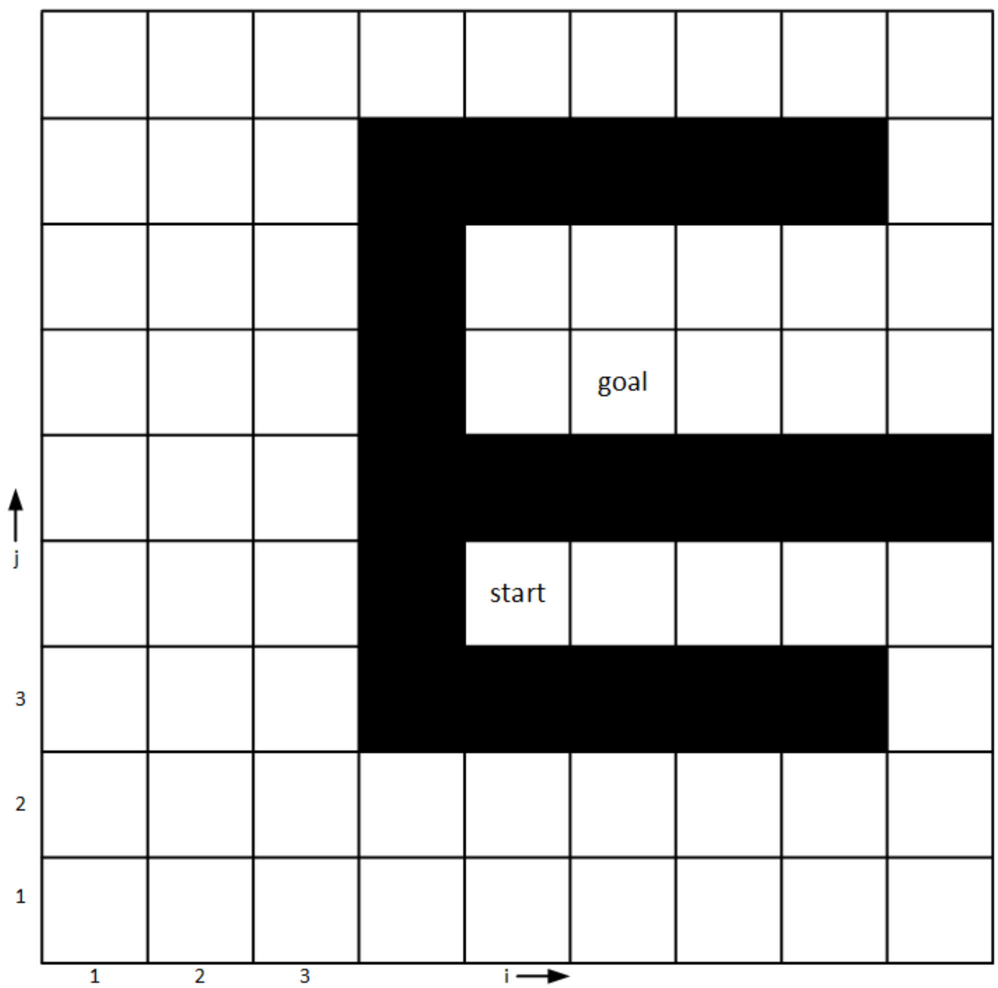
   (i) A^* is complete.

   (ii) A^* is optimal given that the heuristic function h is consistent.

   Which of the above statements are true?

   (a) Only statement (i) is true.
   (b) Only statement (ii) is true.
   (c) Both statements (i) and (ii) are false.
   (d) Both statements (i) and (ii) are true.

7. Consider the following path finding problem in which an agent wants to go from the start cell $(5,4)$ (first index is horizontal, second index vertical) to the goal cell $(6,6)$. The agent can only make the following moves: *one cell up, down, left or right* and each move has a cost of 1. The black cells form a barrier which the agent cannot pass. Assume the agent applies **Greedy Graph Search** with heuristic function h the Manhattan distance.



Which of the following nodes will **eventually** be in the list of open nodes, called *frontier* and will **not** be expanded (removed from this list)?

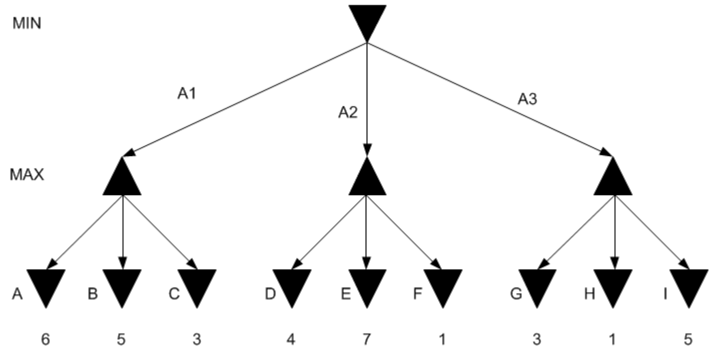
(a) node corresponding to cell (state) $(2,6)$.

(b) node corresponding to cell (state) $(2,2)$.

(c) node corresponding to cell (state) $(4,9)$.

(d) None of the above.

8. Consider the following part of a two-player game tree.



What will be the value of the top MIN node

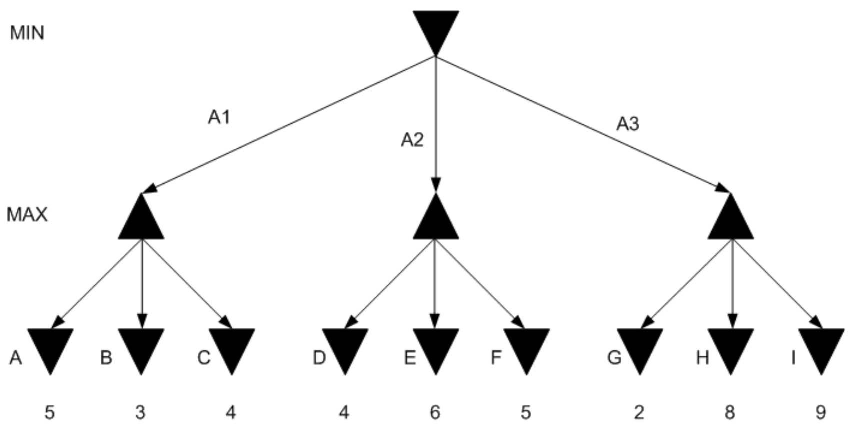
(a) 3

(b) 1

(c) 5

(d) 4

9. Consider the following two-player game tree



If one does not apply alpha-beta pruning then the values for lower level nodes are given. Now assume that one applies alpha-beta pruning, for which of the following collection of nodes will **all** nodes being explored?

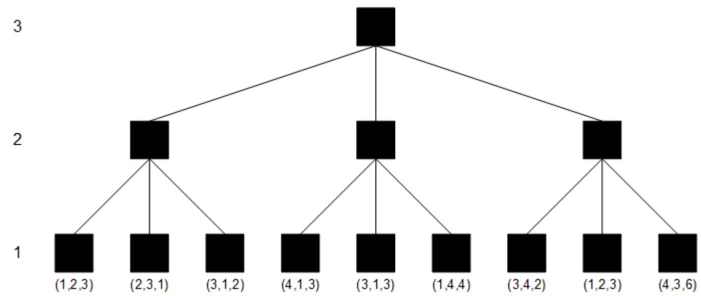
(a) $\{A, F, G\}$.

(b) $\{G, H, I\}$.

(c) $\{D, E, H\}$.

(d) None of the above.

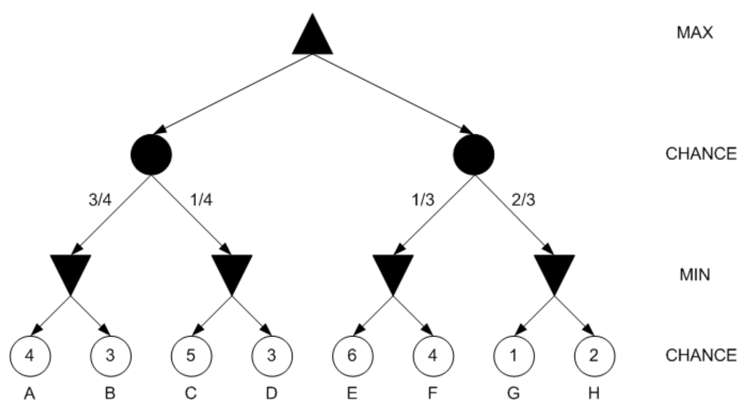
10. Consider the following game tree for three players, player 1, 2 and 3.



At the lowest nodes the evaluation value of the relevant node is given (the value of the evaluation function) for each of the players. In the i position the evaluation value of the node for player i is given, $i = 1, 2, 3$. What is the value of the top node?

(a) $(1, 4, 4)$
(b) $(3, 4, 6)$
(c) $(4, 3, 6)$
(d) None of the above.

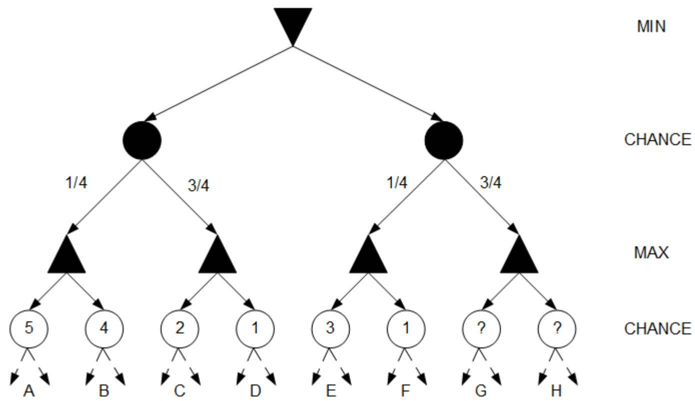
11. Consider the following two-player game tree in which the game has an element of chance, which is shown by the so-called probability nodes in the game tree.



What is the correct value for the top MAX node if one applies the expectiminimax algorithm?

(a) 6
(b) 5
(c) 4
(d) 3

12. Consider the following game tree with an element of chance.



The letters under the bottom row of chance nodes are labels for the nodes just above the letter. One can also apply $\alpha - \beta$ pruning on this game tree. The numbers inside the chance nodes on the bottom row are the computed values of these chance nodes; ? indicates not computed yet. Which of the following statements is true?

(a) For all values of node G which are less than 8/3, $\alpha - \beta$ pruning will expand node H.

(b) For all values of node G which are larger than 8/3, $\alpha - \beta$ pruning will expand node H.

(c) For all values of node G which are less than 11/4, $\alpha - \beta$ pruning will expand node H.

(d) For all values of node G which are larger than 8/4, $\alpha - \beta$ pruning will expand node H.

Questions about Propositional Logic

13. How many models does this propositional formula have:

$$(B \Rightarrow A) \vee A$$

(a) 4
(b) 3
(c) 2
(d) 1

14. Consider the following two statements:

(I) $(A \Rightarrow B) \vee B$ entails $\neg A \vee B$.
(II) $\neg A \vee \neg B$ entails $(A \Rightarrow B) \vee B$.

Which of the two following statements is true?

(a) (I) is true, (II) is false
(b) Both are true
(c) Both are false
(d) (I) is false, (II) is true

Typo in original exam, which results the C is correct.

15. A knowledge base KB contains the following statement (i.e. considered to be true).

- $\neg(W \wedge R) \Rightarrow (Q \vee T)$

The question is whether we can prove $Q \vee T$ from this KB. Which of the following answers is correct?

(a) Yes, we can prove $Q \vee T$.
(b) No, we cannot derive $Q \vee T$, but if we add the premiss W to KB the statement $Q \vee T$ can be derived.
(c) No, we cannot derive $Q \vee T$, but if we add the premiss $\neg W$ to KB the statement $Q \vee T$ can be derived.
(d) None of the above.

16. Which formula do we obtain when we transform the formula

$$(W \Rightarrow R) \vee (\neg Q \wedge R)$$

into conjunctive normal form?

(a) $(\neg W \vee R) \vee (\neg Q \wedge R)$

(b) $(W \vee R \vee \neg Q) \wedge (W \vee R)$

(c) $(\neg W \vee R \vee \neg Q) \wedge (\neg W \vee R)$

(d) None of the above

Questions about Predicate Logic

17. In the context of the Valentine puzzle the predicate $Send(x, c, y)$ means "person x sends card c to person y". What is implied by the following formula in First Order Predicate Logic?

$$\forall x_1, x_2, c_1, c_2, y \; Send(x_1, c_1, y) \land Send(x_2, c_2, y) \Rightarrow x_1 = x_2$$

   (a) Every person gets a card from at most one person.
   (b) Every person sends a card to at most one person.
   (c) Every person sends at most one card.
   (d) Every card is sent by at most one person.

18. We want to formalise the sentence "There exists a student who attends the Artificial Intelligence course but is not intelligent" in first-order logic. Given are the following predicates:

   - $S(x)$: x is a student.
   - $A(x, y)$: x attends course y.
   - $I(x)$: x is intelligent.

   We further introduce a term constant, AI, that stands for the Artificial Intelligence course. We are given two different formalisations:

   I. $\exists x \; S(x) \land A(x, \; AI) \Rightarrow \neg I(x)$
   II. $\exists x \; (S(x) \land A(x, \; AI) \land \neg I(x))$

   Which of the following assertions is correct?

   (a) Formalisation I is correct while formalisation II is incorrect.
   (b) Formalisation I is incorrect while formalisation II is correct.
   (c) Both formalisations are correct.
   (d) Both formalisations are incorrect.

19. We want to Skolemise the following sentence in first-order logic:

$$\forall x, y, z \; \exists w \; P(x, y) \land Q(w, z) \Rightarrow (R(x, z) \land S(z, w, x))$$

   Of the following four substitutions, only one produces a correct Skolemisation. Which one?

   (a) $\{w/S(x)\}$
   (b) $\{w/S(x, y)\}$
   (c) $\{w/S(x, y, z)\}$
   (d) $\{w/S(x, z)\}$

Questions about Reasoning under Uncertainty

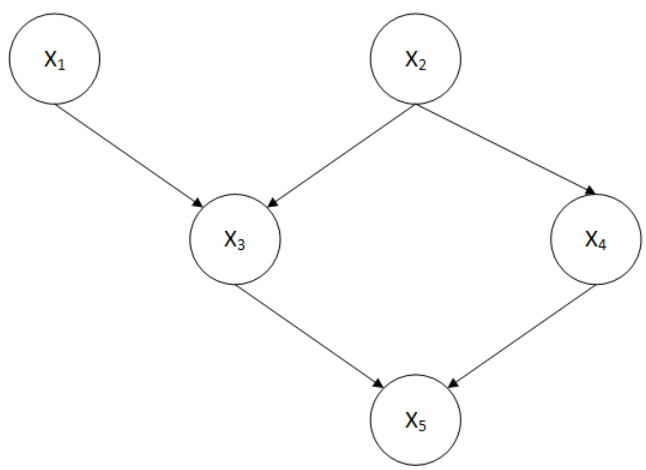
20. A full joint distribution for the *Toothache, Cavity, Catch World* is given by the table below, copied from Figure 13.3 in the book of Russel and Norvig.

	toothache		¬ toothache	
	catch	¬catch	catch	¬catch
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

What is the value of $P(\neg cavity|catch)$? Select the alternative which closest to your answer.

(a) 0.471
(b) 0.016
(c) 0.160
(d) 0.144

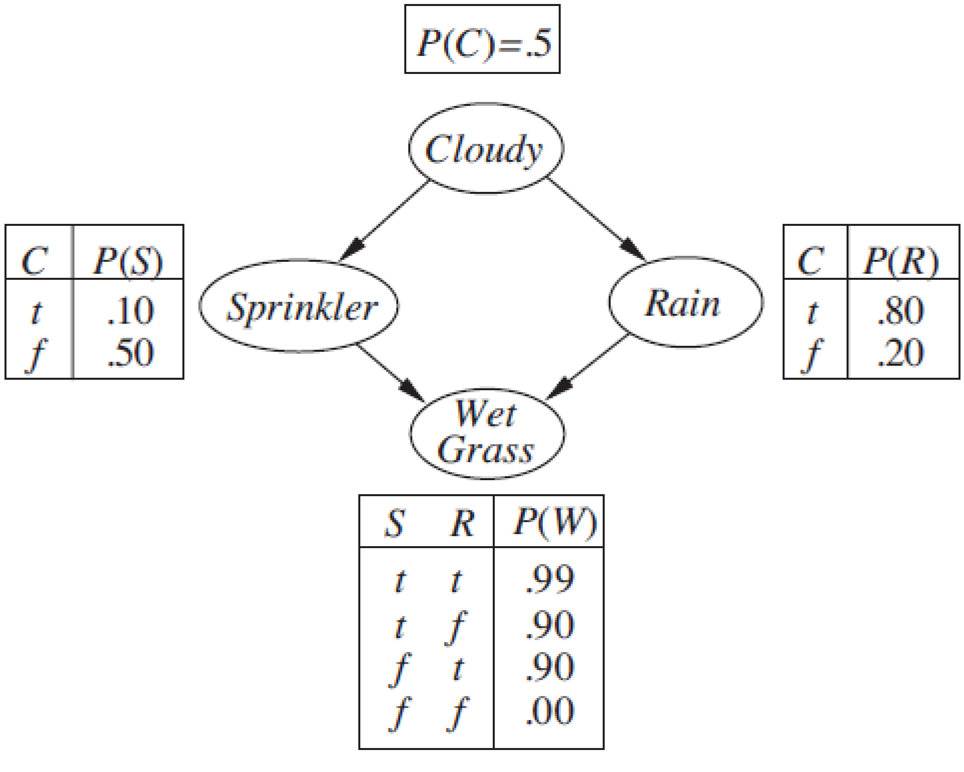
21. Given the Bayesian network shown below.



Only one of the following statements is true, which one?

(a) X_2 and X_5 are independent.
(b) X_2 and X_5 are independent given X_3.
(c) X_2 and X_5 are independent given X_3 and X_4.
(d) None of the above.

22. Given the Sprinkler network shown below.



What is the best approximation of the value of $P(R|W, C)$ (the probability that it is **R**aining given that the grass is **W**et and that it is **C**loudy)?

(a) 0.9758

(b) 0.3636

(c) 0.6364

(d) 0.8920

Questions about Machine Learning

23. A data analyst has collected data (see table below) about customer loans. The goals is to predict, based on the customer profile, if a loan for a customer has a high risk or not.

payment history	debt	guarantee	income	risk
average	low	no	15-35 KEuro	low
average	low	no	0-15 KEuro	high
average	low	no	> 35 KEuro	low
average	low	sufficient	> 35 KEuro	low
bad	low	no	0-15 KEuro	high
bad	low	sufficient	> 35 KEuro	low
good	high	sufficient	> 35 KEuro	low
good	high	no	0-15 KEuro	high
good	high	no	15-35 KEuro	low
good	high	no	> 35 KEuro	low
bad	high	no	15-35 KEuro	high

What is the initial entropy of the above dataset w.r.t. the attribute *risk*?

(a) 1.00
(b) 0.95
(c) 0.50
(d) 0.05

24. The analyst wants to learn the above classification problem (classifying risk) using decision trees. If he uses "information gain" as selection criteria what will be the the attribute at the root of the decision tree (top node)?

(a) payment history
(b) debt
(c) guarantee
(d) income

Questions about Neural Networks

25. Assume that we are training a logistic classifier (linear classifier with logistic regression) using the `Logistic` function (called the `sigmoid` function in the slides) and that the current logistic classifier has the weights $(w_0, w_1, w_2) = (1, 2, -2)$. The next feature point in our training set is given by $x = (1, 2)$. What is the output of this logistic classifier? You can find a table values for the `Logistic` function at the last page of the exam.

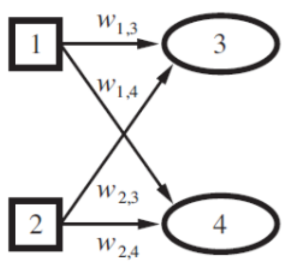
    (a) 1.00

    (b) 0.73

    (c) 0.50

    (d) 0.27

26. Consider the Neural Network depicted below, where the dummy ("bias") inputs and weights are not shown!



We assume the following weights: $w_{0,3} = 1,\ w_{1,3} = -2,\ w_{2,3} = -1,\ w_{0,4} = 2,\ w_{1,4} = -1,\ w_{2,4} = -2$, The activation function of the neurons 3 and 4 is the `Logistic` function (called the `sigmoid` function in the slides). What is the output of this NN on the input $x = (1, 1)$? First component is the output of neuron 3, second component the output of neuron 4.
You can find a table values for the `Logistic` function at the last page of the exam.

(a) $(0.73, 0.88)$

(b) $(0.50, 0.88)$

(c) $(0.12, 0.27)$

(d) $(0.88, 0.73)$

27. Once again consider the above NN and weights. Assume that for a given input x the output is $(0.6, 0.2)$ (first component is the output of neuron 3) and the target output is $(1, 0)$. Moreover we assume a L_2 loss function $1/2[(y_3 - a_3)^2 + (y_4 - a_4)^2]$. What is the value for the delta of neuron 3: $\Delta[3]$?
(formula book: $\Delta[j] \leftarrow g'(in_j) \times (y_j - a_j)$)

   (a) 0.10
   (b) 0.14
   (c) 0.40
   (d) -0.40

28. Once again consider the above NN and weights. Assume that for a given input $x = (2, 1)$ the value for $\Delta[3] = 0.6$. What will be the new value for the weight $w_{2,3}$ (old value was -1) if one uses a learning rate $\alpha = 0.5$?

   (a) -1.3
   (b) -0.7
   (c) -2.3
   (d) -1.4

Questions about Reinforcement Learning

29. Consider the following 4x3 environment with reward and transition function as described in the caption and the figure.

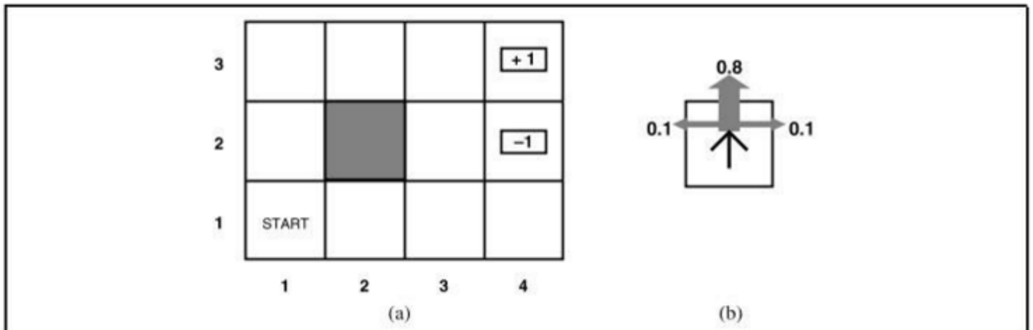


Figure 1 (a) A simple 4×3 environment that presents the agent with a sequential decision problem. (b) Illustration of the transition model of the environment: the "intended" outcome occurs with probability 0.8, but with probability 0.2 the agent moves at right angles to the intended direction. A collision with a wall results in no movement. The two terminal states have reward +1 and –1, respectively, and all other states have a reward of –0.04.

Assume that the agents applies the Bellman update

$$U_{i+1}(s) \;\leftarrow\; R(s) + \gamma \mathbf{max}_a \sum_{s'} P(s' \mid s, a) U_i(s')$$

with $\gamma = 1$ and initial values for U equal to 0 except for (4,2) and (4,3), for which the values are as given: -1 and +1 respectively. What will be the values for $U(3,2)$ after one complete update of U:

(a) 0
(b) -0.04
(c) -0.05
(d) 0.46

30. An agent uses Q-learning, to learn an optimal strategy for a probabilistic game. The current (internal) state of the agent is s. In this state s the agent can do four actions; a, b, c and d. The Q-values, computed by the agent, for these state action pairs are given by:

action x	$Q(s, x)$
a	20.0
b	60.0
c	40.0
d	50.0

Moreover assume that the agent decides to do some exploration and does the action d and receives a reward 10.0. Due to this action d the agent ends up in state s'. In this new state s' the agent can do actions e, f with the following Q-values:

action x	$Q(s', x)$
e	10.0
f	5.0

Assume that the agent applies Temporal Difference Learning with learning parameter $\alpha = 0.6$ and discount factor $\gamma = 0.8$. What will be the new Q-values for state s?

(a) $Q(s, a) = 38.0$, $Q(s, b) = 78.0$, $Q(s, c) = 58.0$ and $Q(s, d) = 68.0$.

(b) $Q(s, a) = 2.0$, $Q(s, b) = 42.0$, $Q(s, c) = 22.0$ and $Q(s, d) = 68.0$.

(c) $Q(s, a) = 20.0$, $Q(s, b) = 60.0$, $Q(s, c) = 50.0$ and $Q(s, d) = 68.0$.

(d) None of the above.

Questions about Cyber Security

31. For this question we use the following terminology:

    - True positive: "normal" input flagged as "normal"
    - True negative: "anomalous" input flagged as "anomalous"
    - False positive: "anomalous" input flagged as "normal" item False negative: "normal" input flagged as "anomalous"

    Which of the following statements is correct?

    (a) A signature-based NIDS produces more true negatives than an anomaly-based NIDS.

    (b) A signature-based NIDS produces fewer true negatives than an anomaly-based NIDS.

    (c) A signature-based NIDS produces more false negatives than an anomaly-based NIDS.

    (d) A signature-based NIDS produces fewer false negatives than an anomaly-based NIDS.

32. Consider an anomaly-based Network Intrusion Detection System (NIDS). Which of the following statements regarding the performance metrics is correct?

    (a) An accuracy (ACC) of 0.0 means that the NIDS produces no false positives nor false negatives

    (b) Having a high False Positive Rate (FPR) will generate more alerts for the security operator to look at.

    (c) The True Positive Rate (TPR) behaves reciprocal to the False Negative Rate (TNR), i.e.: TPR = 1 - FNR

    (d) A system with an accuracy of 0.8 always produces more false positives than a system with an accuracy of 0.9.

Table for $-p\log_2(p)$

p	$-p\log_2(p)$	p	$-p\log_2(p)$	p	$-p\log_2(p)$
0	0	1/8	0.38	1/10	0.33
1	0	2/8	0.50	2/10	0.46
1/2	0.50	3/8	0.53	3/10	0.52
1/3	0.53	4/8	0.50	4/10	0.53
2/3	0.39	5/8	0.42	5/10	0.50
1/4	0.50	6/8	0.31	6/10	0.44
2/4	0.50	7/8	0.17	7/10	0.36
3/4	0.31	1/9	0.35	8/10	0.26
1/5	0.46	2/9	0.48	9/10	0.14
2/5	0.53	3/9	0.53	1/11	0.31
3/5	0.44	4/9	0.52	2/11	0.45
4/5	0.26	5/9	0.47	3/11	0.51
1/6	0.43	6/9	0.39	4/11	0.53
2/6	0.53	7/9	0.28	5/11	0.52
3/6	0.50	8/9	0.15	6/11	0.48
4/6	0.39			7/11	0.42
5/6	0.22			8/11	0.33
1/7	0.40			9/11	0.24
2/7	0.51			10/11	0.13
3/7	0.52				
4/7	0.46				
5/7	0.35				
6/7	0.19				

Table for Logistic or sigmoid function $\sigma(x)$

x	$\sigma(x)$
-4	0.02
-3	0.05
-2	0.12
-1	0.27
0	0.50
1	0.73
2	0.88
3	0.95
4	0.98