University of Zagreb
Faculty of Electrical Engineering and Computing

21 June, 2016

Artificial Intelligence Final Exam (AY 2015/2016)

The exam has ***20 questions*** *for a total of* ***35 points****. All multi-choice questions carry 1 point each (1/2 point subtracted for incorrect answer), while problem questions carry 4 points each. The exam duration is* ***120 minutes****. You must turn in the exam questions with your solutions.*

Part I: Multi-choice questions ***(15 pts)***

1. (*1 pt*) The Rosenblatt's perceptron update rule is as follows (t – target value, o – output value):
   (a) $w(i+1) = w(i) + \eta(t+o)x(i)$ (d) $w(i+1) = w(i) - \eta(t+o)x(i)$
   (b) $w(i+1) = w(i) + \eta(t-o)w(i)$ (e) $w(i+1) = w(i) + \eta(t-o)x(i)$
   (c) $w(i+1) = w(i) - \eta(t-o)w(i)$
2. (*1 pt*) What is the linguistic expression corresponding to the linguistic modifier of *concentration*?
   (a) "more or less" (b) "not" (c) "less" (d) "very" (e) "or"
3. (*1 pt*) The backpropagation algorithm is used for training:
   (a) a single TLU perceptron (d) a feedforward perceptron network
   (b) a feedforward neural network with sigmoid units (e) a naive Bayes classifier
   (c) an arbitrary artificial neuron
4. (*1 pt*) Which of the following R&D areas is the main consumer of machine learning today?
   (a) computer graphics (c) computer security (e) software engineering
   (b) digital education (d) robotics
5. (*1 pt*) We use the genetic algorithm to minimize $h(x)$. Which of the following we can use as a *fitness function*?
   (a) $1 + h(x)$ (b) $h(x)h(x)$ (c) $-h(x)$ (d) $h(x)/2$ (e) $h(x)$
6. (*1 pt*) In an ACO alogrithm, what does τ denote?
   (a) strength of the pheromone trail (c) time step (e) probability of selection
   (b) probability of mutation (d) value of the heuristic function
7. (*1 pt*) In temporal difference learning, the goal is to find:
   (a) an optimal path (c) optimal values (e) optimal Q-values
   (b) optimal policy (d) optimal learning rate
8. (*1 pt*) You are given a set of examples (x_1, x_2), of which examples $(0,0)$, $(2,0)$, $(1,2)$, $(3,1)$ are labeled as positive. The inclusion of which of the following examples, labeled as negative, will render the problem unlearnable for a single TLU-perceptron?
   (a) $(1,3),(5,2)$ (b) $(3,3),(5,5)$ (c) $(0,1),(1,3)$ (d) $(4,1),(5,2)$ (e) $(1,3),(2,1)$
9. (*1 pt*) For hypotesis H and evidence E, we are given the following probabilities: $P(E|H) = 0.7$, $P(E|\neg H) = 0.1$, $P(H) = 0.4$. Using the Bayes rule, compute the probability of hypothesis H given evidence E. (Round the result to two decimals.)
   (a) 0.12 (b) 0.921 (c) 0.46 (d) 0.82 (e) 0.75

10. (*1 pt*) What form of training data is used for supervised learning?

    (a) only inputs (c) only ouptus (e) $(input, output)$ pairs

    (b) $(output, weight)$ pairs (d) $(input, weight)$ pairs

11. (*1 pt*) The update rule of value iteraton is as follows:

    (a) $V_{k+1}(s) \leftarrow \max_a \sum_{s'} T(s,a,s') \cdot [R(s,a,s') - \gamma V_k(s')]$

    (b) $V_{k+1}(s) \leftarrow \max_a \sum_{s'} T(s,a,s') \cdot [R(s,a,s') + \gamma V_k(s')]$

    (c) $V_{k+1}(s) \leftarrow \max_a \sum_{s'} T(s,a,s') \cdot R(s,a,s')$

    (d) $V_{k+1}(s) \leftarrow \max_a \sum_{s'} T(s,a,s') \cdot [R(s,a,s') - \gamma V_k(s)]$

    (e) $V_{k+1}(s) \leftarrow \max_a \sum_{s'} T(s,a,s') \cdot [R(s,a,s') - \gamma$

12. (*1 pt*) In an *Ant System* algorithm, the evaporation of the pheromone trail between nodes i and j is modeled as:

    (a) $\tau_{ij} \leftarrow \tau_{ij}(1-\eta)$ (c) $\tau_{ij} \leftarrow \tau_{ij}(1+\rho)$ (e) $\tau_{ij} \leftarrow \eta\tau_{ij}(1-\rho)$

    (b) $\tau_{ij} \leftarrow \tau_{ij}(1-\rho)$ (d) $\tau_{ij} \leftarrow \tau_{ij} - \rho$

13. (*1 pt*) The membership function of a fuzzy set A is defined as:

    (a) $\mu_A : X \to [-\infty, +\infty]$ (c) $\mu_A : X \to [-1, 1]$ (e) $\mu_A : X \to \{-1, 0, 1\}$

    (b) $\mu_A : X \to [0, 1]$ (d) $\mu_A : X \to [0, 1)$

14. (*1 pt*) One company developed a system for estimating the number of calories from a still photo of a dish. What machine learning task is that?

    (a) pattern recognition (b) regression (c) classification (d) clustering (e) reinforcement learning

15. (*1 pt*) We're searching for the optimum of $f(x, y)$ using a genetic algorithm. We use 5-bit chromosome representations, where we use first two for the variable x and the latter three for the variable y. Values of both variables are within hthe $[0, 3]$ interval, and we use binary coding. What is the value corresponding to chromosome 01010?

    (a) x = 1, y = 6/7 (b) x = 2, y = 1/2 (c) x = 1, y = 5/2 (d) x = 2, y = 2/5 (e) x = 2, y = 2

Part II: Problem questions *(20 pts)*

16. (*4 pts*) Fuzzy logic.

    (a) *(2 pts)* According to their income, the citizens of some country are are divided into seven categories: A, B, C, D, E, F, and G. On these categories, the following fuzzy sets are defined:

    $rich = \{0.25/D, 0.5/E, 0.75/F, 1/G\}$
    $poor = \{1/A, 0.75/B, 0.5/C, 0.25/D\}$
    $average = \{0.1/B, 0.7/C, 1/D, 0.7/E, 0.1/F\}$

    Using standard Zadeh's operators, derive the fuzzy set corresponding to the following expression: $(rich \lor average) \land (\neg rich \lor poor)$,

    (b) *(1 pt)* What rules from the classic logic do not hold in fuzzy logic? Illustrate this on the fuzzy set $\{1/a + 0.5/b + 0/c\}$.

    (c) *(1 pt)* Briefly explain how generalized modus ponens differs from the same inference rule in the classic logic. Showcase your explanation with an example.

17. *(4 pts)* Naïve Bayes classifier.

    You are given the following set of training examples:

	Place setting	Main character	Time setting	Time travel	Good SF movie
1.	space	woman scientist	present	yes	no
2.	Earth	outlaw	future	no	no
3.	elsewhere	child	past	yes	no
4.	space	woman scientist	present	no	yes
5.	space	outlaw	past	no	yes
6.	Earth	child	past	yes	yes

For each example, the first four columns contain the features, while the last column is the target classification.

(a) *(1 pt)* Derive the Bayes rule for the case of multiple hypotheses and multiple evidences. Explicitly state all the conditions (assumptions) under which this rule holds.

(b) *(1 pt)* Write down the expression for MAP hypothesis of a naïve Bayes classifier.

(c) *(2 pts)* Using the naive Bayes classifier trained on the above data, determine the classification of a new, previously unseen example: $(\text{space}, \text{child}, \text{past}, \text{yes})$.

18. *(4 pts)* Reinforcement learning.

(a) *(3 pts)*
You are given a Markov decision process with six states $S = (a, b, c, d, e, T)$ and five transitions $A = (n, s, e, w, x)$. The transition function $T(s, a, s')$ is as follows: $T(a, w, b) = 1.0$, $T(a, e, c) = 1.0$, $T(b, x, T) = 1.0$, $T(c, n, d) = 0.8$, $T(c, n, a) = 0.1$, $T(c, n, e) = 0.1$, $T(c, e, e) = 0.9$, $T(c, e, d) = 0.1$, $T(d, x, T) = 1.0$, $T(e, x, T) = 1.0$. The reward function $R(s, a, s')$ is non-zero only in the following cases: $R(b, x, T) = 1.0$, $R(d, x, T) = -50.0$, $R(e, x, T) = 10.0$.
Run three iterations of the value iteration algorithm for the given Markov decision process, with the parameter $\gamma = 1.0$.

(b) *(1 pts)* In the last iteration of the algorithm, also calculate the values for all Q-states and the optimal policy.

19. *(4 pts)* The perceptron.

You are given the following set of training examples:

x_2	x_1	c
-1	-1	-1
-1	1	-1
1	-1	-1
1	1	1

For each example, the values x_1 and x_2 are the feature values, while c is the target classification.

(a) *(1 pt)* Sketch the artificial neuron model, as proposed by McCulloh and Pitts, and label all its parts.

(b) *(1 pt)* Explain what is the stopping criterion for Rosenblatt's perceptron learning algorithm. Explain if there is any condition that the training examples must meet in order for this criterion to be satisfiable. Sketch an illustrative example.

(c) *(2 pts)* Train the TLU perceptron with outputs on the above training set. The initial values of the weights are $w_2 = 0.2$, $w_1 = -0.1$, and $w_0 = -0.1$. The learning rate is 0.5. Assume that $step(0) = 1$.

20. *(4 pts)* Genetic algorithm.

Steady state genetci algorithm is used to find a maximum of the function $f(x) = -(x-3)^2 + 4$. Soultions are represented as 7-bit binary chromosomes. The search space is $[-4, +6]$. Initial population consists of the following units `J1=0101010`, `J2=1111000`, `J3=1010101`, `J4=0000000`, `J5=1111111`.

(a) *(2 pts)* Determine the precision with which the algorithm will search across the search space.

(b) *(2 pts)* Let us use the simplified tournament selection for the algorithm. We randomly select units `J1`, `J2` i `J4` for the tournament. We use the one-point crossover after the 4th bit (the first bit is the leftmost bit in the chromosome). Determine which of these three units will become parents. Perform crossover and mutation, and determine which of the two children chromosomes will replace which of the three randomly selected units. (NB: Only one child survives.) Let mutation in this example affect only the first and the last bit of the chromosome.