Machine, Data, Learning
End Semester Exam

25th April 2022 — 70 marks, 180 min

1. **Answer any 7 questions. All questions carry equal marks.**
2. **No clarifications during the exam.**
3. **Make *reasonable assumptions* and *clearly state* them to answer *ambiguous* questions.**
4. **Use the common default interpretation of any concept/method as discussed in class.**
5. **Show your steps / reasons. Be concise and organized.**
6. **Floating point answers should be upto 2 significant decimal places.**
7. **Calculators are allowed. Sharing of calculators not allowed.**

Q1. Consider the game of Tic-Tac-Toe (noughts and crosses). Define X_n as the number of rows, columns or diagonals with exactly n X's and no O's. Similarly, O_n is the number of rows, columns or diagonals with just n O's. The utility function thus assigns +1 to any position with $X_3 = 1$ and -1 to any position with $O_3 = 1$. All other terminal positions have utility 0. We will use a linear evaluation function defined as

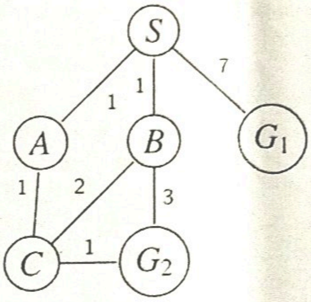
$$\text{Eval} = 3X_2 + 2X_1 - (3O_2 + 2O_1)$$

(a) Show the whole game tree starting from an empty board down to depth 2, (i.e., one X and one O on the board), taking symmetry into account. You should have 3 positions at level 1 and 12 at level 2. [3]

(b) Mark on your tree the evaluations of all the positions at level 2. [3]

(c) Mark on your tree the backed-up values for the positions at levels 1 and 0, using the minimax algorithm, and use them to choose the best starting move. Underline the node to which the first move should be made. [4]

Q2. Which goal is reached and what is the total cost of the solution found for the following state-space graph when using Breadth-First Search and Uniform-Cost Search? (S is the start state, G1 and G2 are the goal states, arcs are bidirectional, no repeated state checking, break any ties alphabetically) [10]



Q3. Prove or disprove: If $h1$ and $h2$ are two admissible heuristics for a given problem, then heuristic $h3(n) = 2\,h1(n) - h2(n)$ for all states, n, must also be admissible. [10]

Q4. Consider a dataset containing six one-dimensional points: {2, 4, 7, 8, 12, 14}. After three iterations of Hierarchical Agglomerative Clustering using Euclidean distance between points, we get the 3 clusters: C1 = {2, 4}, C2 = {7, 8} and C3 = {12, 14}.

(a) What is the distance between clusters C1 and C2 using Single Linkage?
(b) What is the distance between clusters C1 and C2 using Complete Linkage?
(c) What clusters are merged at the next iteration using Single Linkage? [10]

i) $2h_1(n) \le h^*n$
ii) $2h_1(n) \ge h^*(n)$
$h_1(n) \le h^*(n)$

$h_2(n) \le h^*(n)$

~~$-h_2(n) \ge -h^*(n)$~~

$-h_2(n) \le h^*(n)$

Q5. You want to cluster 7 points into 3 clusters using the k-Means Clustering algorithm. Suppose after the first iteration, clusters C1 , C2 and C3 contain the following two-dimensional points:

C1 contains the 2 points: {(0,6), (6,0)}
C2 contains the 3 points: {(2,2), (4,4), (6,6)}
C3 contains the 2 points: {(5,5), (7,7)}

(a) What are the cluster centers computed for these 3 clusters?
(b) What clusters are computed in the next iteration?

[10]

Q6. The table below shows a training set with 10 examples for training a **3-nearest-neighbors classifier** that uses Manhattan distance, i.e., the distance between two points at coordinates p and q is $|p - q|$. The only attribute, X, is real-valued, and the label Y has two possible classes, 0 and 1. What is the 2-fold cross validation accuracy (percentage correct classification)? The first fold contains the first 5 examples, and the second fold contains that last 5 examples. In case of ties in distance, use the example with smallest X value as the neighbor. [10]

X	0	1	2	3	4	5	6	7	8	9
Y	1	0	1	0	1	0	1	0	1	0

Q7. Use the table below that defines a training set containing 4 examples. The two attributes, X1 and X2 , and the class label, Y, are all binary.

[10]

X1	0	0	1	1
X2	0	1	0	1
Y	1	1	1	0

(a) What is the entropy of Y, i.e., H(Y)? Use the convention that $0 \log 0 = 0$.
(b) Using the above training set, a Decision Tree is built that contains only 3 nodes: the root and its 2 children. Each leaf node is assigned the majority class of its associated set of examples; break ties in favor of $Y = 0$. What is the classification accuracy of this Decision Tree on the training set of 4 examples?

Q8. An agent is in state (x,y) in a MDP. The agent can take actions N (North), S (South), E (East) and W (West) leading to states (x-1, y), (x+1, y), (x, y+1) and (x, y-1) respectively. Action is stochastic i.e., whenever action is taken it goes to the intended state with probability of 0.6 and in perpendicular directions with 0.2 probability each. Each step (or action taken) incurs a cost i.e. receives reward of -1. The discount factor is 0.53. The current utility values (i.e., at time step t) for each of the 4 states (in the order mentioned earlier) are 4, 5, 4 and 1 (i.e., U_t(state)). What is the best action for the agent to take from (x,y). Please show (all the) utility computations for each of the 4 actions and then mention the best action as outcome of these computations. [10]

$-\frac{3}{4}(\log 3 - 2) - \frac{1}{4}(0 - 2)$

$\Rightarrow -\frac{3\log 3}{4} + \frac{3}{2} + \frac{1}{2}$

$\Rightarrow 2 - \frac{3\log 3}{4}$

0.311278 + 0.5

$\frac{\frac{1}{2} + \frac{2}{3}}{2}$

0.811278