Name ______________________________ Email ___________________________________

. `Write answers in space provided. Show your work. Each Question is worth 5 points.`

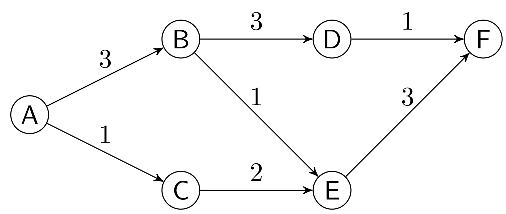


Figure 1: Use for Questions 1, 2, 3, 4. Break any ties and add items to the FRINGE using alphabetical order, i.e. the first expansion always leaves the FRINGE as A←[B,C].

1. (BFS)

   Using the search tree in Figure 1, where A is the start state and F is the goal state, give an ordered list of the nodes as they are visited using BFS. Also give the shortest path found from A to F.

   visited: [A,B,C,D,E,E,F] shortest-path: A→B→D→F

2. (DFS)

   Using the search tree in Figure 1, where A is the start state and F is the goal state, give an ordered list of the nodes as they are visited using DFS. Also give the shortest path found from A to F.

   visited: [A,C,E,F] shortest-path: A→C→E→F

3. (Uniform Cost Search)

   Using the search tree in Figure 1, where A is the start state and F is the goal state, give an ordered list of the nodes (as well as their associated costs) as they are visited using UCS. Also give the lowest cost path found from A to F.

   visited: $[A_0, C_1, B_3, E_3, E_4, D_6, F_6]$ shortest-path:A→C→E→F

4. (Iterative Deepening)

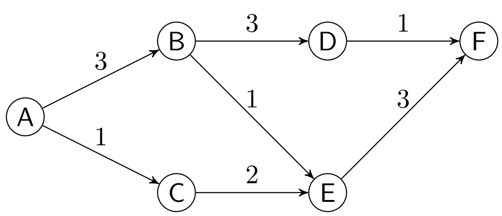
   Using the search tree in Figure 1, where A is the start state and F is the goal state, how many times will state D be visited if we use Iterative Deepening search?

   just once: A,C,B; A,C,E,B,E,D; A,C,E,F stop

5. (Search Optimality)

   If we are asked to implement an uninformed search algorithm which will use the smallest amount of space (polynomial space complexity) but will also be sure to find the goal (optimal), which of the uninformed search methods that we've studied should we choose?

   Iterative Deepening



	$h(x)$
A	0
B	1
C	2
D	1
E	1
F	0

Figure 2: Use for Questions 6, 7. Break any ties using alphabetical order.

6. (A*)

   Use the search tree in Figure 2, and let A be the start state and F the goal. Show two steps of A* search, (dequeuing two nodes and enqueueing their children). Give the nodes visited and the state of the queue (including associated priority values) at each step. The first step has been started for you.

$$\begin{aligned} \mathrm{A} &\leftarrow [\mathrm{B}_4, \mathrm{C}_3] \\ \mathrm{C}_4 &\leftarrow [\mathrm{B}_4, \mathrm{E}_4] \\ \mathrm{B}_4 &\leftarrow [\mathrm{E}_4, \mathrm{D}_7, \mathrm{E}_5] \end{aligned}$$

7. (Admissible Heuristics)

   If, in the table in Figure 2, we change the value of $h(\mathrm{E})$ to 3, is h still an admissible heuristic? Why or why not?

   Yes, since $h(\mathrm{E}) = h^*(\mathrm{E}) = 3$ so $\forall s, h(s) \leq h^*(s)$.

8. (Hill Climbing)

   In the course of performing a hill-climbing gradient search, a state is found for which a step in any direction in feature space does not produce an increase in the score value, i.e.,

$$f(x') \leq f(x_{current}) \ : \ \{x' \in \text{Neighbors}(x_{current})\}$$

   a) [3pt] Have we found a global maximum? If not, what two other things we've discussed could it be?

   No, we may be at local min or on a plateau.

   b) [2pt] What relatively simple additional method should we add to the basic hill-climbing algorithm in order to attempt to verify or disprove this?

   random restarts

9. (Simulated Annealing)

   a) [2pt] If we start the search with $T_0 = 1$ and cooling rate 1.1, will we be more or less likely to accept suboptimal s' over time as we iterate?

   We will be more likely, as temperature T, and so $e^{-\frac{|f(s)-f(s')|}{T}}$, will go up over time.

   b) [3pt] Let $succ(x) = x'$, score $f(x) = f(x') = 1$ and $T = .9$. What is the probability we accept x'?

   Probability(of accepting x') $= e^{-\frac{|f(x)-f(x')|}{T}} = e^{-\frac{0}{T}} = e^0 = 1$.

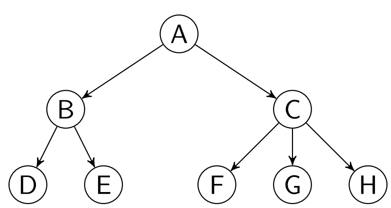


Figure 3: Use for Questions 10, 11. Break any ties using alphabetical order.

10. (Minimax)

    If Figure 3 represents a 2-player Zero-Sum game where node A indicates the max player's turn, should the max player be more interested in knowing max{F,G,H} or min{F,G,H}?

    min{F,G,H}, as that's what min will play at C.

11. (Alpha/Beta Pruning)

    If Figure 3 represents a 2-player Zero-Sum game, node A indicates the max player's turn, and the tree is traversed from left to right at each level, what node(s) have a possibility of never being visited?

    G,H as the first time we can prune is after visiting F.

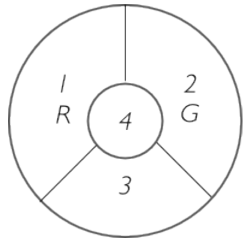


Figure 4: Use for problems 12, 13, 14. We'd like to color this map with three colors:{R,G,B} so that no two regions which share an edge share the same color. So far in our search we've assigned 1=R and 2=G.

12. (Forward Checking) Without assigning any additional values to variables, if we are doing Backtracking Search with Forward Checking, will we backtrack at this point? (show your work)

    No, 3 and 4 both still have available values (B).

	R	G	B
1	o	x	x
2	x	o	x
3	x	x	
4	x	x	

13. (Arc Consistency) Without assigning any additional values to variables, if we are doing Backtracking Search with Arc Consistency, will we backtrack at this point? (show your work)

    Yes. Checking constraints:
    (2,3)→3 $\neq$ G, (3,1)→3 $\neq$ R, (3,4)→4 $\neq$ B, (4,1)→4 $\neq$ R, (4,2)→4 $\neq$ G. No values left for 4 →backtrack.

14. (CSP) Can we find a coloring that will satisfy our constraints? Describe how we can modify the problem, using appropriate CSP terminology, so that we can find such a coloring.

    No. We could expand our variable domain to include 4 colors.

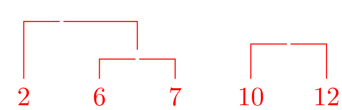
15. (ML)

    Explain in one sentence why, given a that we'd like to train and evaluate our learner on a dataset, we split the data into a Training Set and a held aside Test Set instead of training and evaluating on the full dataset?

    We want to avoid overfitting on the Training Set, which can lead to poor generalization.

16. (HAC)

    Run the first three iterations of hierarchical clustering using complete linkage on the 1D dataset $X_{1:6} = \{2, 6, 7, 10, 12, 17\}$ and draw the resulting (partial) dendrogram.



17. (k-means)

    We are using k-means to cluster the 1D dataset $X_{1:5} = \{1, 2, 4, 9, 11\}$, $k = 2$. On the previous iteration we assigned the points to clusters $\{1, 1, 2, 2, 2\}$ (i.e. 1,2 are in cluster 1 while 4,9,11 are in cluster 2).

    a) [3pt] What are the new cluster centers c_1, c_2?

    $c_1 = \frac{1+2}{2} = 1.5, \quad c_2 = \frac{4+9+11}{3} = 8$

    b) [2pt] Then what are the new cluster assignments for points $X_{1:5}$

    {1,1,1,2,2}: 4 has changed clusters.

18. (kNN)

    What is the predicted class for item $x = (2, 2)$ using k-NN with $k = 3$, using Manhattan distance, and dataset $\{(1, 1, -), (1, 7, +), (4, 3, +), (5, 4, -)\}$.

    Predict $-$: (1,1), (4,3) and (5,4) are closest so majority vote returns $-$.

19. (Perceptron)

    Give a set of weight values for a single Linear Threshold Unit with two inputs (plus bias=1) that implements the NOR function, which has the following truth table:

x_1	x_2	y
1	1	0
1	0	0
0	1	0
0	0	1

Just one example: $w_0 = 1, \; w_1 = -1.1, \; w_2 = -1.1$.

20. (Gradient Descent)

    We are given a single training item $\mathbf{x}' = [x_1, x_2]' = [2, 2], \; y = 1$ and asked to train a linear perceptron initialized with $\mathbf{w}' = [w_0, w_1, w_2]' = [1, 0, 1]$ and with learning rate $\alpha = .5$.

    a) [1pt] Calculate the initial squared Euclidean error.

    $a = 1 + (2)(0) + (2)(1) = 3 \quad \text{so} \quad \text{Error} = \frac{1}{2}(3 - 1)^2 = 2.$

    b) [4pt] Perform a single gradient descent step and give the new values for $\mathbf{w}$.

$$
\begin{aligned}
w_0 &\leftarrow 1 - (.5)(3 - 1)(1) = & 0 \\
w_1 &\leftarrow 0 - (.5)(3 - 1)(2) = & -2 \\
w_2 &\leftarrow 1 - (.5)(3 - 1)(2) = & -1
\end{aligned}
$$