Name:

NetID:

1.) For the graph given below answer the questions. The cost to travel between nodes are given as/on the solid lines and the heuristic distances are on the dashed lines.

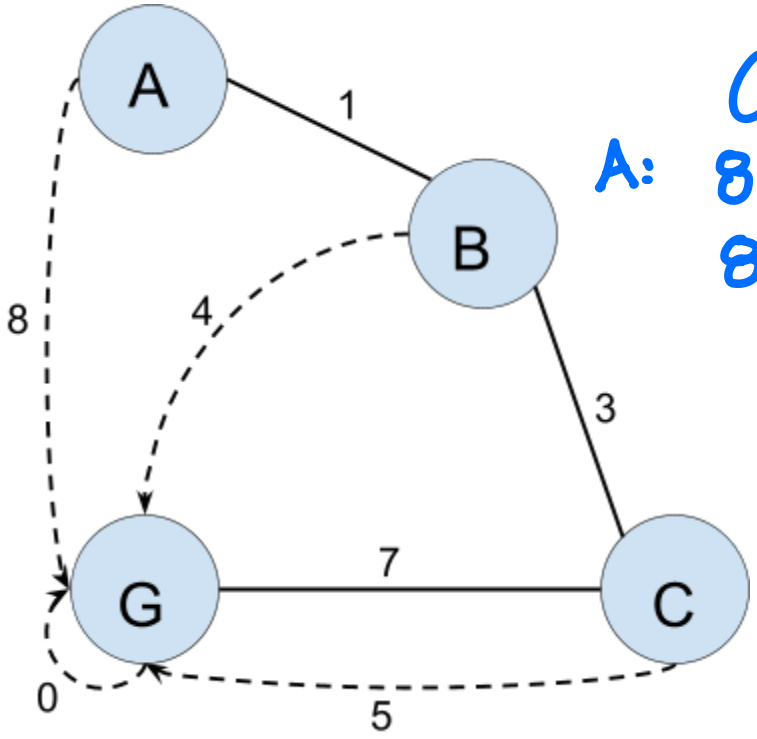
Admissible

A: $8 \leq 11$ ✓

B: $4 \leq 10$ ✓

C: $5 \leq 7$ ✓

G: $0 \leq 0$ ✓



Consistent:

A: $8 \leq 1+4$

$8 \leq 5$ ✗

a.) Is this heuristic admissible? Why or why not?

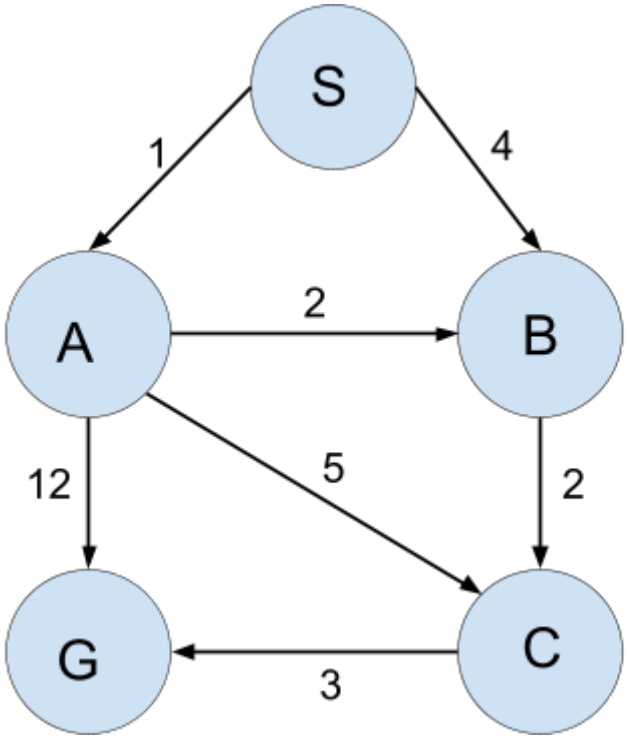
yes, for all nodes, $h(N) \leq d(N, G)$

b.) Is this heuristic consistent? Why or why not?

No! $h(A) \leq d(A,B) + h(B)$ is NOT true!

$h(N) \leq d(N, N') + h(N')$

2.) Given the following graph, starting in Node S and trying to get to Node G, trace the steps the A* Algorithm would take. To make the graph less busy, the heuristic distances are given as a table.



N	h(N)
S	7
A	6
B	2
C	1
G	0

*.) N -> N' = d(N, N') + h(N', G)

~~1.) S -> A = 1 + 6 = 7~~

~~2.) S -> B = 4 + 2 = 6~~

~~3.) S->B->C = 4+2+1=7~~

~~4.) S->A->B = 1+2+2 = 5~~

~~5.) S->A->C = 1+5+1 = 7~~

~~6.) S->A->G = 1+12+0 = 13~~ new maximum

~~7.) S->A->B->C = 1+2+2+1 = 6~~

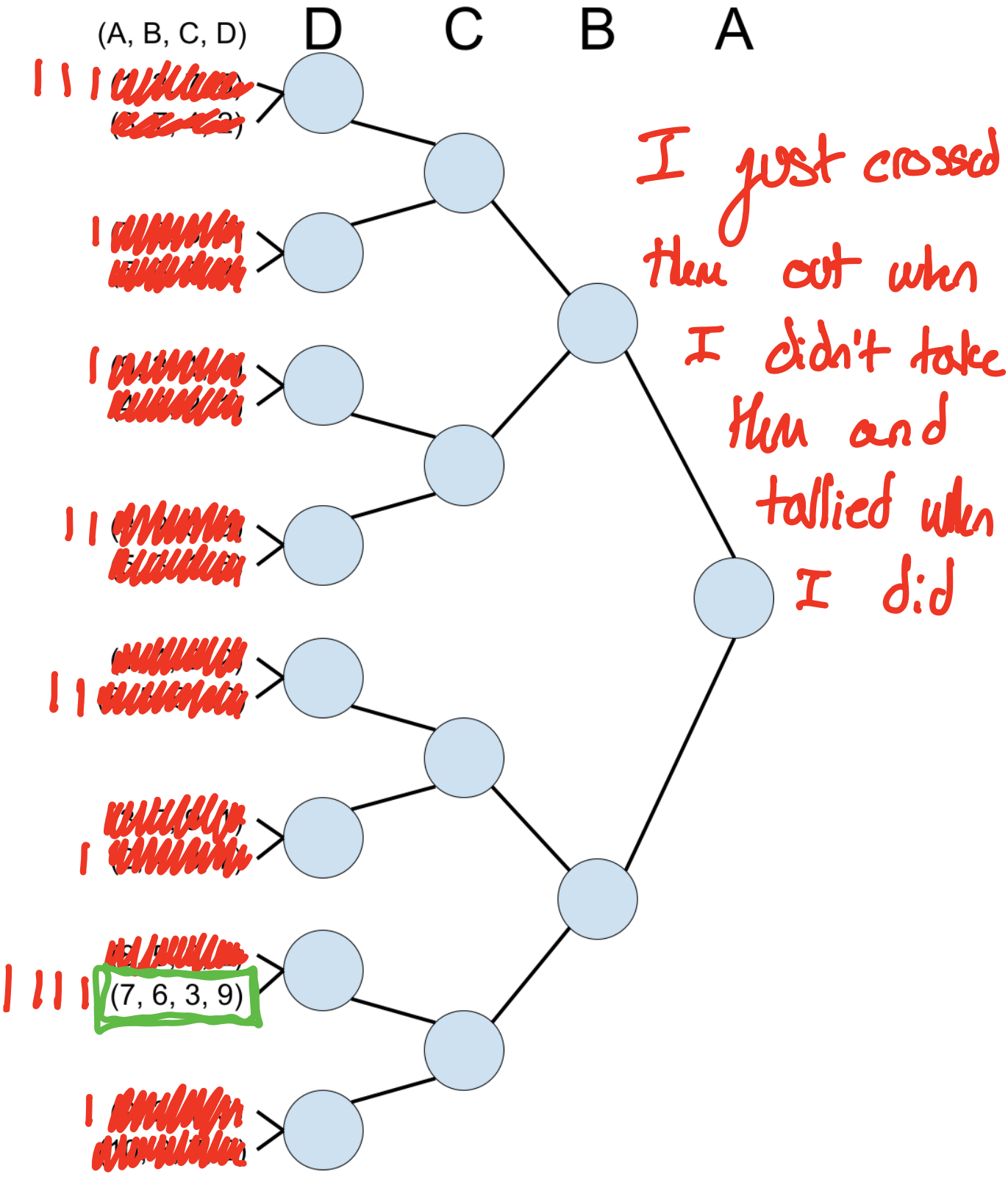
8.) S->A->B->C->G = 1+2+2+3 = 8 new max

~~9.) S->B->C->G = 4+2+3 = 9~~

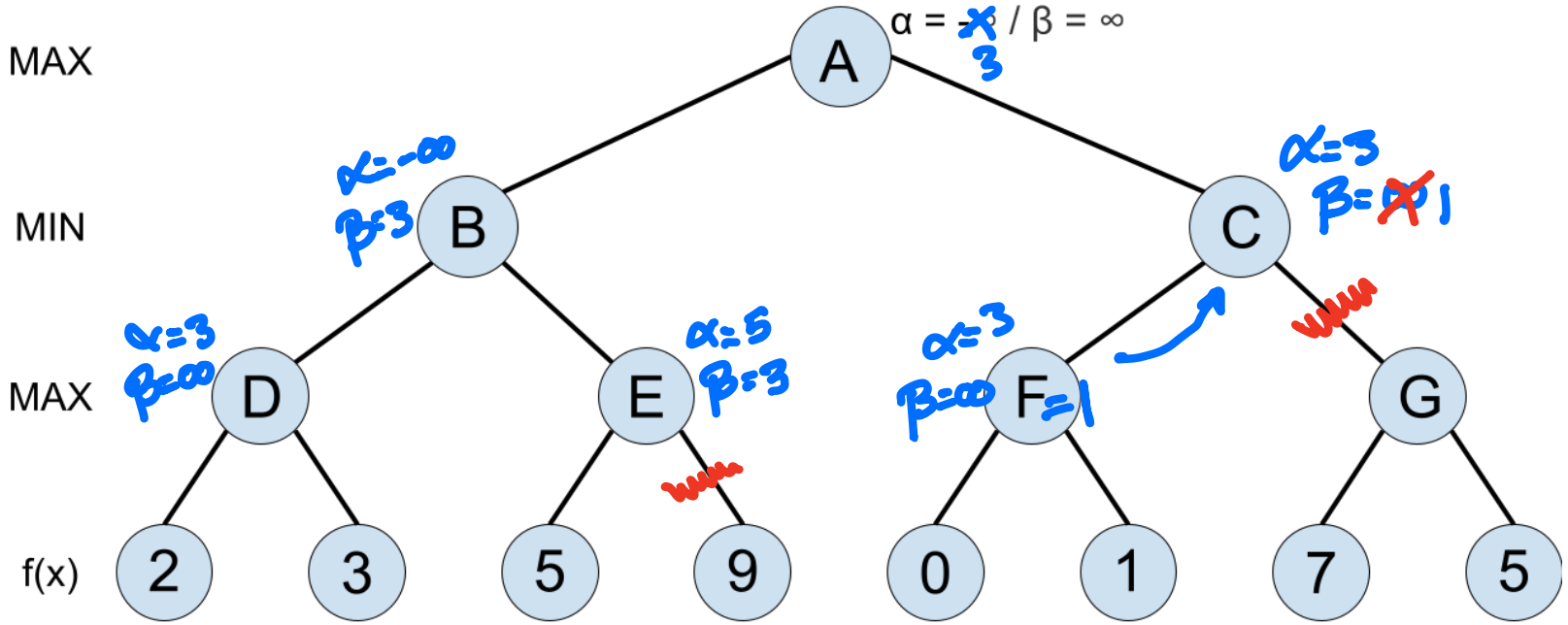
10.) ~~S->A->C->G = 1+5+3 = 9~~

Is the heuristic admissible and/or Consistent?

3.) Show which tuple of numbers gets chosen for each level of a 4 player minimax tree.



4.) Use Alpha/Beta Pruning on the tree below, showing the Alpha and Beta at each node and show which branches don't need to be explored.



I'm gonna be honest I tried the idea we had in class of just ignoring α and β and running minimax and thinking about when to prune and I found it easier and more intuitive so I probably won't make you label α and β on the exam

5.) Answer the following two questions about Constraint Satisfaction Problems (CSPs).

a.) Explain in your own words how the backtracking algorithm for CSPs works and is implemented.

Essentially the backtracking algorithm is just an algorithm to recursively generate all permutations of variable assignments using a DFS. We check every assignment right away and "prune" the branch if it violates a constraint.

b.) How does forward checking or constraint propagation improve on basic backtracking?

With forward checking or constraint propagation immediately after assigning a variable, even if it passes constraints, we reflect the choice we made on the other connected variables to see if our current choice would cause us to fail in the future

6.) Given the following dataset layout, answer the two questions.

	S	X		Y
	S	X_1	X_2	Y
Train	S_1	3.9	4.2	0
	S_2	2.7	8.5	1
	S_3	4.1	4.0	0
	...			
Test	S_{N-1}	3.1	7.9	1
	S_N	4.0	3.8	0

a.) How do we derive the mu and sigma for a gaussian PDF during the training or fitting step of a Naive Bayes classifier?

for each of the two classes, for each feature, we calculate μ and σ

b.) What is the role of the gaussian PDF during the prediction step of using a Naive Bayes classifier and what assumption do we make by using a Gaussian distribution?

It gives us the likelihood of the feature we're testing belonging to the class the PDF corresponds to.

NOTE: I don't expect any math here, just use words and pictures if you like.

We assume our features are normally distributed

7.) What is the difference between a regression task and a classification task?

Regression gives a real number,
Classification gives a discrete label

8.) Why is the Viterbi algorithm considered a "greedy" algorithm

Because it always just picks the
Max at each timestep

9.) What would the equation look like for a multiple linear regression with three features?

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

OR $y = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3$

10.) What are the two regularization techniques we learned and how do they affect the weights of a linear regression?

L_1 - Lasso - Some weights end up as 0

L_2 - ridge - weights are small and evenly distributed

11.) What is the purpose of the AC-3 Algorithm?

A "pre-pruning" step for CSPs that can potentially solve the problem directly

12.) What is the "learning" a Markov Babbler does when given an input text?

It calculates all the transition probabilities for words following each other in the text

13.) What are MLE and MAP?

MLE- Maximum Likelihood estimation

Model assumes a certain "shape"

MAP - Maximum a Posteriori

Weights fit a distribution

14.) What is the difference between uninformed and informed searches?

Informed searches have extra heuristic information like in A*, uninformed searches do not.

15.) How does IDS simulate BFS behaviour with a DFS?

It tells our DFS it can only go to a certain depth level on the current search and we repeat that for increasing depth levels

16.) What is a loss function?

A function that gives us a singular number telling us how well our model fits the data