Problem 1: Change This! [9 marks]

For the following questions in this section, if applicable, you can leave your answer in equation form if you wish. For example, you can leave your answer in the form of $2^2 + 2^2$ instead of calculating that out to become 8. Don't forget you have a **binary conversion table** that may be of use!

1. [4 marks] Write an algorithm to convert a number from hexadecimal to decimal. You can assume that you have already written an algorithm that details how to convert from binary to decimal and can refer to it inside your hexadecimal to decimal algorithm if necessary (using the binary to decimal algorithm in your algorithm is not mandatory).

2. [1 mark] What is the sum of $1111_2 + 1111_2$? This question requires an answer that is not in equation form (i.e., you will not receive marks for having $1111_2 + 1111_2$ as your answer).

3. [1 mark] Convert the binary number 100111011 to hexadecimal.

4. [1 mark] Convert the binary number 1101 1100 to decimal.

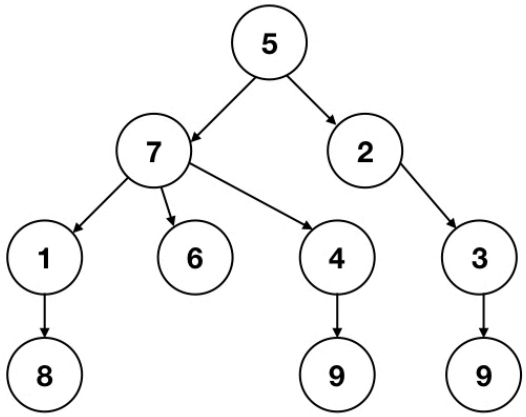
5. [2 marks] If I had an image that was 4 pixels long and 10 pixels wide and I wanted it to fill up a screen that is 20 pixels long and 40 pixels wide, how many screen pixels would I need to display each image pixel? Show your work.

The rest of this page has been intentionally left blank.

Problem 2: True or False? [12 marks]

Clearly indicate the answer to the following questions by circling True or False and then writing a 1-2 sentence justification for why you chose that answer. Ambiguously answered questions will be considered incorrect.

1. [2 marks] The following picture is a valid example of a rooted tree.



True False

Justification:

2. [2 marks] If you use the k-means algorithm to cluster the same dataset with the same number of clusters, you will always obtain the same grouping. For example, if you choose to cluster 100 data points into 3 groups, running the k-means algorithm on these data points multiple times will result in the same data points clustered together.

True False

Justification:

3. [2 marks] In order to load a webpage, the server has to go to the client to fetch the webpage.

   True False

   Justification**:**

4. [2 marks] 0x100 and 100 evaluate to the same number.

   True False

   Justification**:**

5. [2 marks] If I zoom in on a picture and it becomes pixelated, the picture has been saved using a lossless algorithm.

   True False

   Justification**:**

6. [2 marks] A computer stores the value of each pixel by storing the amount of red, gray, and blue that the colour is composed of.

True False

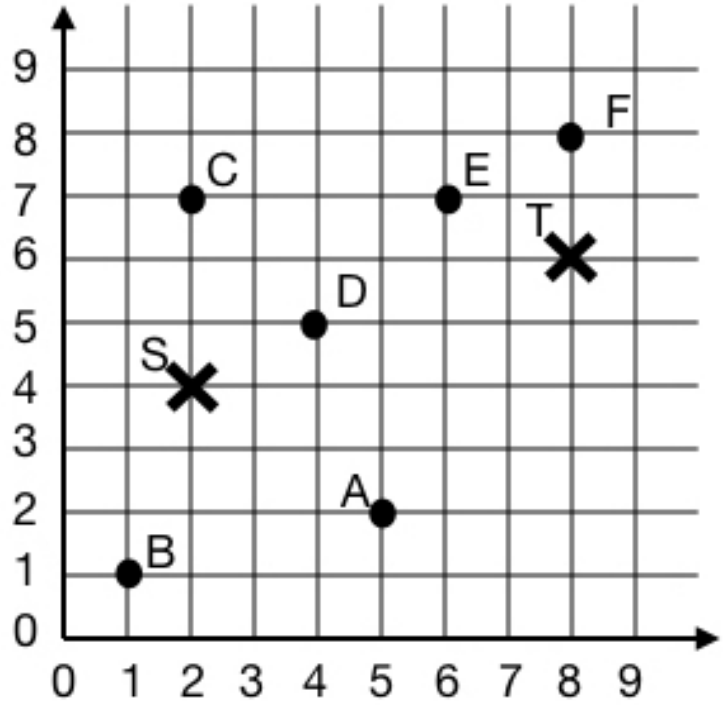
Justification**:**

The rest of this page has been intentionally left blank.

Problem 3: Clustering [11 marks]

1. [2 marks] Using one sentence, describe what the k in the k-means algorithm stands for.

2. Consider the following clustering example:



This example has six data points (A, B, C, D, E, and F) that have been represented in the above example by using circles. Daphne wants to use the k-means algorithm to create two clusters. S and T (the two X's in the above example) have been chosen as the initial centroid locations of the two clusters.

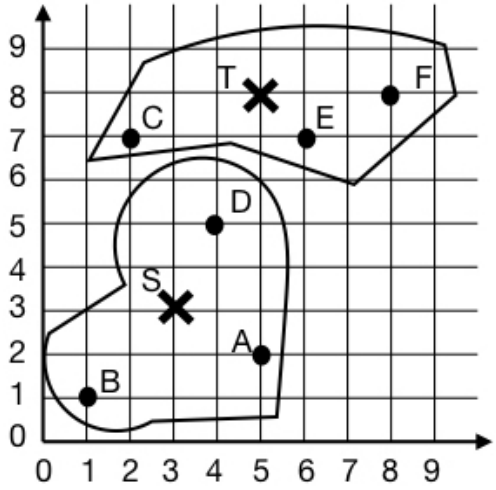
The following chart contains the distances between the various points and the two centroid locations which may or may not be helpful to you as you solve the rest of the questions in this section.

	A	**B**	**C**	**D**	**E**	**F**
S	3.6	3.2	3.0	2.2	5.0	7.2
T	5.0	8.6	6.1	4.1	2.2	2.0

a. [6 marks] Circle which cluster (S or T) points A, B, C, D, E, and F should belong to. Ambiguously circled answers will be considered incorrect.

Point	Cluster S	Cluster T
A	Yes	Yes
B	Yes	Yes
C	Yes	Yes
D	Yes	Yes
E	Yes	Yes
F	Yes	Yes

b. [3 marks] After the k-means algorithm has iterated for some time, the following clustering appears. The circles indicate which data points have been assigned into which cluster.



Can we stop repeating the k-means algorithm now? Why or why not?

Problem 4: Internet [6 marks]

1. [2 marks] The exam schedule has come out and you can finally start to book a vacation for yourself at the end of April! You start shopping around on the Internet for the best deal on hotels at your ideal destination place but being the savvy shopper that you are, you don't want to buy the first package that you see. You want to shop around on different websites to see if you can get a better deal.

   As you sit there browsing, you realize that the price of the hotel seems to go up each time you refresh the page. Assuming that no one else has booked a stay at the hotel you are interested during the time you are browsing for deals online, use the terminology we have learned in class to **briefly** explain (in 2-4 sentences) why the displayed prices are going up.

2.

   a. [1 mark] Think about the times when you have posted an article for the individual portion of In the News. What type of communication is it? Circle your answer from the following four options. Ambiguously circled answers will be considered incorrect.

      i. Synchronous
      ii. Asynchronous
      iii. Both synchronous and asynchronous
      iv. Neither

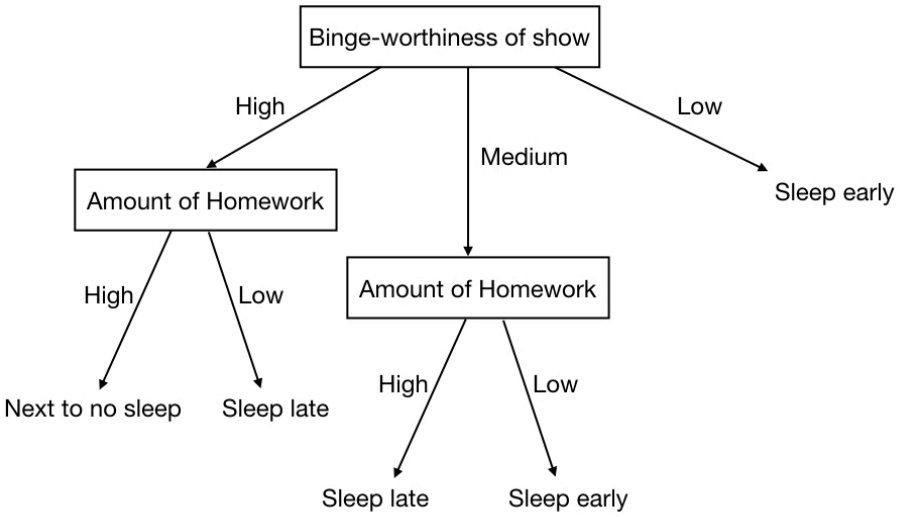
   b. [1 mark] In 2-3 sentences, justify your answer to question 2a.

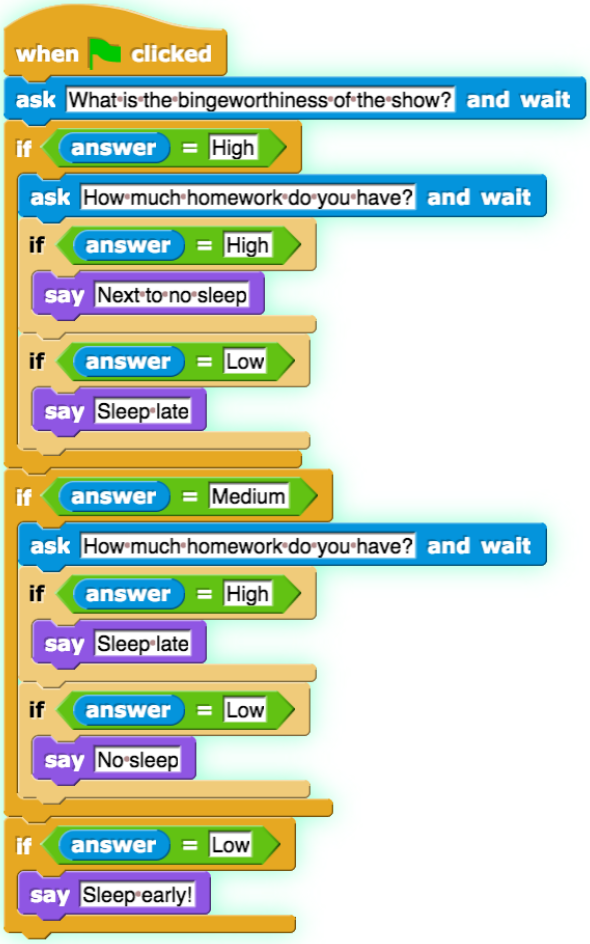
3. [2 marks] If I use the TCP/IP protocol to send a packet from one computer to another, can I predict the packet's path with absolute certainty? Why or why not?

The rest of this page has been intentionally left blank.

Problem 5: Decision Trees [10 marks]

1. [2 marks] Amy has a decision tree below that she coded in Snap. Does the given Snap code snippet match the decision tree? Why or why not?





Write your answer to question 1 here:

2. Yelim would like to create a decision tree for the following data.

Weather	Free Time	Number of People Going	Go Hiking?
Sunny	Very little	Many	Yes
Rainy	Lots	Few	No
Snowy	Very little	Many	No
Sunny	Somewhat	Many	No
Sunny	Lots	Few	Yes

a. [6 marks] What is the entropy for each of the attributes? List the name of each attribute and its entropy.

b. [1 mark] Which attribute should Yelim split on? List all possibilities.

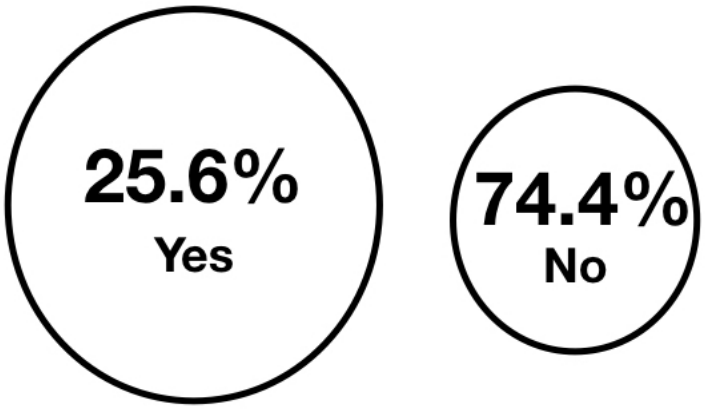
c. [2 marks] Consider what happens when we add the following row to Yelim's original data above.

Rainy	Lots	Few	No
Snowy	Very Little	Many	No

Should she still split her tree based on your answer to question b above? Why or why not?

Problem 6: Data Visualization [8 marks]

Imagine the following infographic was presented to you during an argument for why a pineapple and pickle mix should be added to all pizzas. The pickle association representative has said the following infographic is representative of the survey results from the student population.



1. [4 marks] Should you trust this infographic? List two reasons to support your answer.

2. [4 marks] Are there any issues with the way this infographic is displayed? List two reasons to support your answer.