

Problem 1: Number Representation [10 marks]

Please show your work to receive full credit. A conversion table is provided separately. You may use it to help with any calculations. Please note that the additional sheet of paper will NOT be graded.

- a. [2 marks] Convert 200 (in decimal) to binary.

0b11001000

Note: no points removed if did not use 0b notation throughout this question

- b. Convert 0b10010111 to decimal.

151

- c. Convert 0x82 to decimal.

130

- d. Convert decimal 314 to hexadecimal.

0x13A

Note: these are all practice questions

Problem 2: Image Representation: Bitmaps [6 marks]

Consider the following Bitmap Image:



For all colours, please write the representation in hex.

- a. Write the standard (non-compressed) representation that we covered in class of the bitmap above

2x4

000000 000000 FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF 000000 000000

Note: line breaks don't matter.

No points removed for extra punctuation (e.g., parentheses) or adding 0x in front of things, etc.

- b. Write the lossless compression representation that we covered in class of the bitmap above

2x4

000000 2 FFFFFFFF 4 000000 2

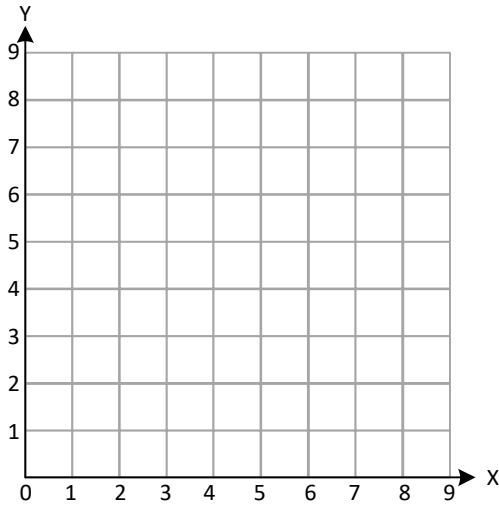
Do not worry about format of numbers here since that will have been graded on the previous sub-part

Problem 3: Vector graphics [4 marks]

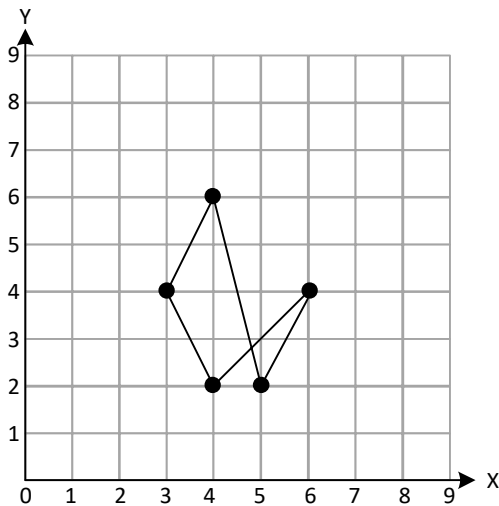
Consider a simple vector representation like the one that we used in the exercise in class – no additional information is provided for thickness, etc.

On the following graph, draw what is represented by the following information:

4, 2, 3, 4, 4, 6, 5, 2, 6, 4, 4, 2



Sample solution:



End of sample solution

Problem 4: Decision tree [9 marks]

Consider that you are trying to decide whether or not to come to class. You have the following data that you are trying to learn from

Tiredness	Nearness to exam	Weather	Come to class?
Low	Soon	Sunny	Yes
Medium	Soon	Sunny	Yes
High	Far	Sunny	No
Medium	Far	Sunny	No

Consider adding each of the following rows to your training data separately (i.e., consider only each row at a time – you are not adding row a then row b, etc.)

Circle the attribute name(s) each optimal attribute to split first on for each of the following additions. For example, (and this is not necessarily the right answer) the example below shows how you would answer if Weather and Tiredness were the optimal attributes to split on for part x, your answer should look like this:

x.

Zero	Middling	Rainy	Yes
------	----------	-------	-----

Tiredness

Nearness to exam

Weather

a.

High	Middling	Rainy	No
------	----------	-------	----

Tiredness

Nearness to exam

Weather

b.

High	Soon	Sunny	No
------	------	-------	----

Tiredness

Nearness to exam

Weather

c.

High	Soon	Snowy	No
------	------	-------	----

Tiredness

Nearness to exam

Weather

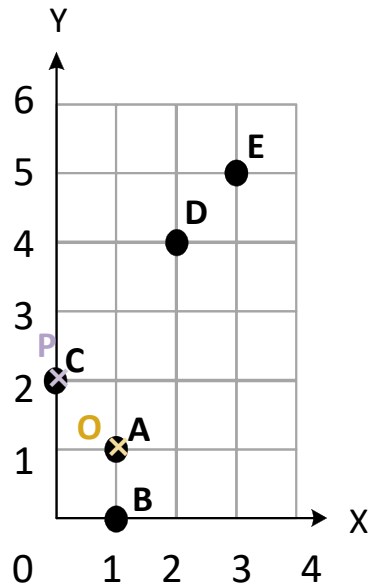
A. The entropy for Nearness to exam is 0, but 4 for the others, so Nearness to exam is the correct answer

B. The entropy for Tiredness in this case = 2. The entropy for Nearness to Exam = 3. The entropy for weather = 5. So the right choice is now Tiredness.

C. The entropy for Tiredness in this case is 2. The entropy for Nearness to exam = 3. Entropy for Weather = 4. So the right answer is tiredness

Problem 5: Clustering [10 marks]

Consider the following clustering example:



There are five data points: A, B, C, D, and E. You are trying to find two clusters, “O” and “P”. Point A has initially been selected as the centroid for cluster O. Point C has initially been selected as the centroid for cluster P.

The following chart contains the distances between various points, which you may find useful in answering the questions in this problem. For any other math that you need to do, you may leave your answer as the expression rather than by doing the calculations, e.g., if your answer is $2^{10} = 1024$, you may leave your answer as 2^{10}

	A	B	C	D	E
A	0	1	1.4	3.2	4.5
B	1	0	2.2	4.1	5.4
C	1.4	2.2	0	2.8	4.2
D	3.2	4.1	2.8	0	1.4
E	4.5	5.4	4.2	1.4	0

- For each point, circle which of the clusters it belongs to. The first row is an example. If non-existent point G belonged to cluster O, you would circle “yes” in the cluster O column as shown.

Point	Cluster O	Cluster P
G	Yes	Yes
A	<u>Yes</u>	Yes
B	<u>Yes</u>	Yes
C	Yes	<u>Yes</u>
D	Yes	<u>Yes</u>
E	Yes	<u>Yes</u>

- b. Explain why you made the answer you did in part A using a single sentence to express your reasoning, e.g., "Points A and E belong to cluster O because they are vowels and points B, C, and D belong to cluster P because they are consonants."

The points should be assigned to the centroid that they are closest to. A and B were closer to the centroid for O, and C, D, and E were closer to the centroid for P, so they were assigned there. Note: sentence 2 is additional explanation for those reading the exam key. It is not required for full credit.

- c. Choose the new centroids for Clusters O and P:

Cluster O's X value:

$(1+1)/2 = 1$ (note: throughout, I show the results of the mathematical equations so that students looking at the solution can know what they are, but it is not required for credit)

Cluster O's Y value:

$$\underline{(1+0)/2=.5}$$

Cluster P's X value:

$$\underline{(0+2+3)/3 = 1.7}$$

Cluster P's Y value:

$$\underline{(2+4+5)/3 = 3.7}$$

This space intentionally left blank

Note that any work that you do on this page will NOT be graded. If you want to do work here you MUST transfer it to the appropriate spot in the rest of the exam. The same holds for the page on the back of the cover sheet.