

Assessment 3: Short Answer Questions

Q1 a) [2 marks]

Build a B-tree of order **3** with the following elements in the order of {3, 5, 7, 1, 4, 6, 10, 8}.

<https://app.diagrams.net/> (formerly draw.io) is recommended as a good tool for this. You can also use a different app if preferred.

Submission

Export/screenshot your diagram as a PNG or JPG, and then add your file to the workspace using the "Upload File" link. Click "Submit" in the bottom right corner to submit your work.



Keep the drawing tool open after submitting, Q1 c) involves modifying the B-tree

Q1 b) [1 mark]

Question

How many nodes (blocks) are there in the B-tree generated in Question 1 a)?

Q1 c) [1 mark]

Insert **9** into your B-tree that was submitted in Q1 a)

<https://app.diagrams.net/> is recommended as a good tool for this. You can also use a different app if preferred.

Submission

Export/screenshot your diagram as a PNG or JPG, and then add your file to the workspace using the "Upload File" link. Click "Submit" in the bottom right corner to submit your work.

Q2 a) [4 marks]

Build a decision tree with the following data using ID3 algorithm. Please provide the detailed calculation of information gain at each step.



Submit your decision tree in this slide, and your detailed calculations of information gain in the next [Q2 b]

Outlook	Temperature	Decision
Sunny	High	No
Rainy	Mild	No
Rainy	Cool	No
Overcast	Mild	Yes
Overcast	High	No
Sunny	Cool	Yes
Rainy	Mild	No

<https://app.diagrams.net/> is recommended as a good tool for this. You can also use a different app if preferred.

Submission

Export/screenshot your diagram as a PNG or JPG, and then add your file to the workspace using the "Upload File" link. Click "Submit" in the bottom right corner to submit your work.



Q2 a) and Q2 b) make up 4 marks in total

Q2 b)

Question

Please provide the detailed calculation of information gain at each step for the decision tree you made in the previous question [Q2 a)].

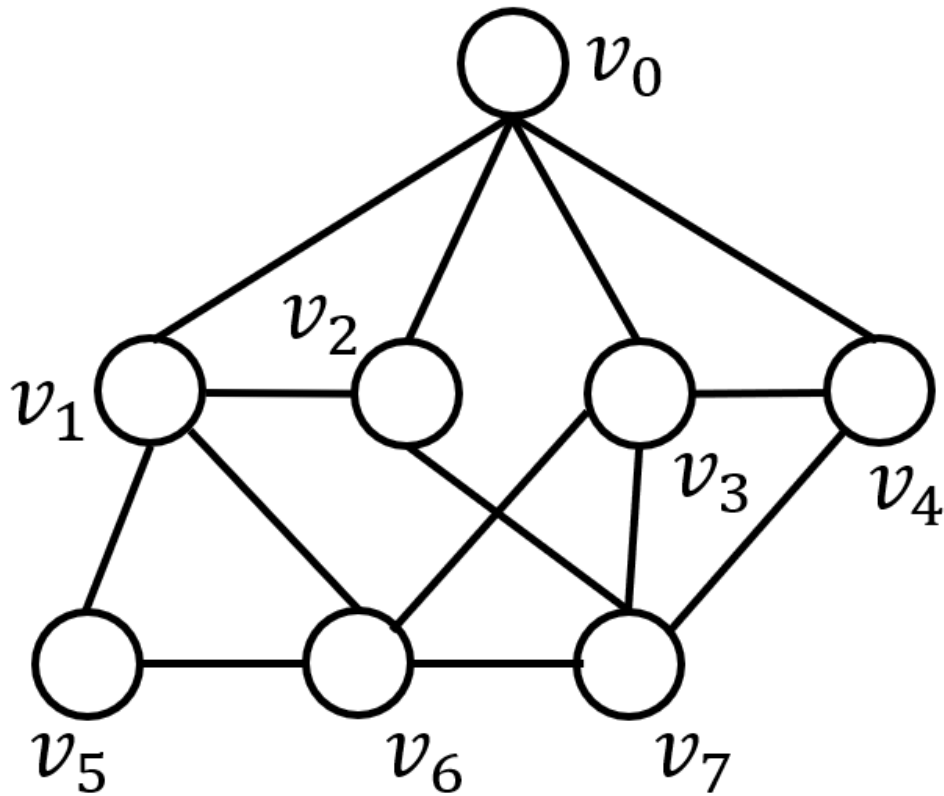
No response

Q3 [2 marks]

Question 1

Find the DFS traversal order starting from v_0 .

i When ordering your answer below, start from the top where v_0 has been placed



v0

v7

v6

v5

v4

v3

v2

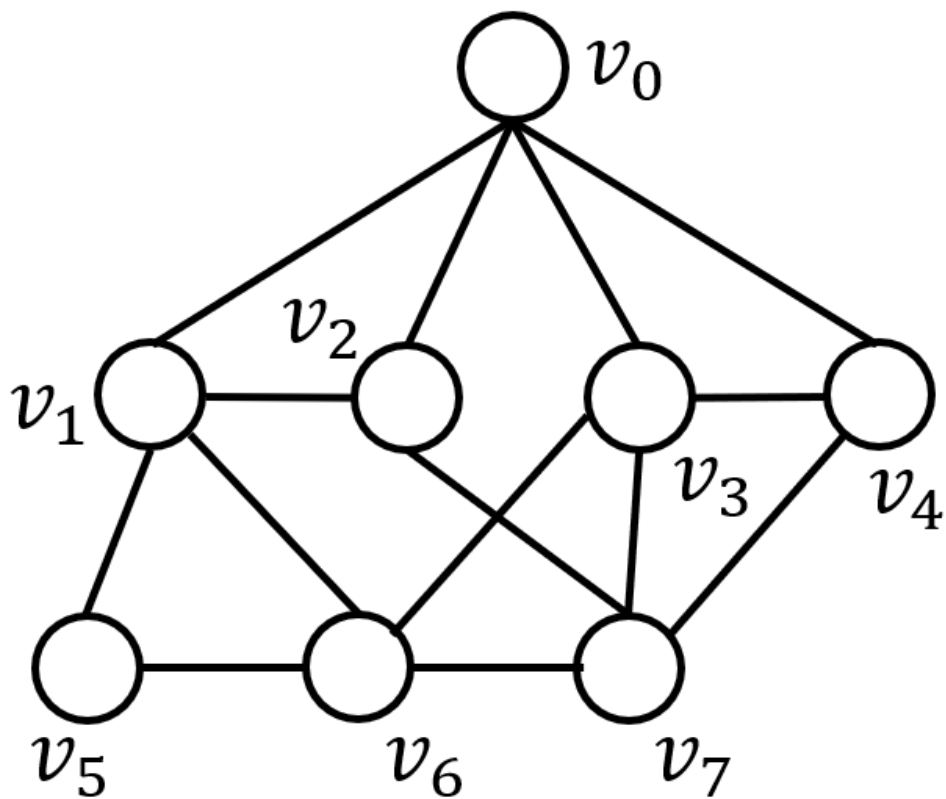
v1

Question 2

Find the BFS traversal order starting from v0



When ordering your answer below, start from the top where v0 has been placed



v0

v7

v6

v5

v4

v3

v2

v1

Q4 [2 marks] and Q5 [3 marks]

Question 1

Consider a Bloom filter of size $m = 12$ (i.e., 12 bits) and 3 hash functions that take a string (lowercase) as input:

$$h_1(\text{str}) = \sum_{c \text{ in } \text{str}} (c - 'a') \bmod 12$$

$$h_2(\text{str}) = \sum_{c \text{ in } \text{str}} ('z' - c + 1) \bmod 12$$

$$h_3(\text{str}) = \text{str.length} \bmod 12$$

For example:

$$h_1(\text{"bd"}) = (1 + 3) \bmod 12 = 4;$$

$$h_2(\text{"bd"}) = (25 + 23) \bmod 12 = 0;$$

$$h_3(\text{"bd"}) = 2 \bmod 12 = 2$$

Answer the following questions:

Q4 (a) [1 mark] Given a set of string $S = \{\text{"unsw"}, \text{"cse"}, \text{"school"}\}$, show the updated result of the Bloom filter.

No response

Question 2

Q4 (b) [1 mark] Given a string "math", use the Bloom filter to check whether it is contained in S , and give the reason.

No response

Question 3

Q5 [3 marks] Given the following temperature data, please compute the maximum temperature for every year-month pair by using MapReduce.

Year	Month	Day	Temperature
2018	4	1	26
2018	4	3	23
2018	5	1	27
2018	5	2	24
2018	5	4	25
2019	1	2	20
2019	1	3	18
2019	2	5	22
2019	2	2	21
2019	2	4	17
2019	6	6	26
2020	7	5	30
2020	7	6	32
2020	7	8	34

No response