



Indian Institute of Technology Kharagpur

Class Test I 2025-26

Date of Examination: Feb, 2026

Duration: 45 Minutes

Subject No.: CS60092

Subject: Information Retrieval

Department/Center/School: Computer Science

Credits: 3

Full marks: 30

Name: _____

Roll Number: _____

Question:	1	2	3	Total
Points:	10	10	10	30
Score:				

Write the answers in the space provided after the questions only.

1. (a) (10 points) Given the following positional index (format term:: docid1: pos1, ...; docid2: pos2, ...; ...), compute the document-ids for retrieved documents corresponding to the phrase query “world wide web”:

```
world :: {1: 7, 18, 33, 70, 85, 131;
          2: 3, 149; 4: 17, 190, 291, 430, 434 }
wide  :: {1: 12, 19, 40, 72, 86, 231; 2: 2, 17, 74, 150, 551;
          3: 8, 16, 191, 429, 435}
web   :: {1: 20, 22, 41, 75, 87, 200; 2: 18, 32, 45, 56, 77, 151;
          4: 25, 192, 300, 332, 440}
```

2. (10 points) Consider the following algorithm for Logarithmic merging. Each token (termID, docID) is initially added to in-memory index Z0 by LMergeAddToken. Fill in the blanks. Assume that Merge routine can merge two indexes.

```
LMergeAddToken(indexes, Z0, token)
Z0 ← Merge(Z0, {token})
if |Z0| = n
  then for i ← 0 to infty
    do if Ii in Indexes
      then _____
        Indexes ← _____
      else _____
        Indexes ← _____
        Break
Z0 ← Null
```

3. (10 points) Consider the following document collection. Calculate the average number of bytes per token and average number of bytes per term for this collection.

```
D1: data science machine learning
D2: machine learning data
D3: science data
```



Rough space

