

TD01 - Information Retrieval and Algorithms

Third Year ISIL - University of M'Sila

Exercise 01 – Information Retrieval (MCQs)

1. Information Retrieval (IR) is mainly concerned with:
 - (a) Storing structured data
 - (b) Retrieving relevant documents from large collections
 - (c) Database transaction management
 - (d) Network security
2. Which type of data is mainly handled by Information Retrieval systems?
 - (a) Structured data
 - (b) Numerical data
 - (c) Unstructured data
 - (d) Binary data
3. Which of the following is an example of an IR system?
 - (a) MySQL
 - (b) Excel
 - (c) Google Search
 - (d) Operating System
4. In Information Retrieval, a document can be:
 - (a) Only a text file
 - (b) Only a PDF
 - (c) Any unit of retrievable information
 - (d) Only a web page
5. What is a query in IR?
 - (a) A document stored in the system
 - (b) A user's information need expressed as keywords
 - (c) A database table
 - (d) A ranked result list
6. Relevance in Information Retrieval is:
 - (a) Always objective
 - (b) Always binary
 - (c) User-dependent and subjective
 - (d) Based only on document length
7. What is the main purpose of indexing?
 - (a) To increase storage size
 - (b) To reduce retrieval speed
 - (c) To improve search efficiency
 - (d) To delete irrelevant documents
8. Ranking in IR systems is important because:
 - (a) Users read all retrieved documents
 - (b) Only the first results matter most
 - (c) Documents have the same relevance
 - (d) Queries are always long
9. Information Retrieval differs from Database systems because IR:
 - (a) Uses exact matching
 - (b) Works only with numbers
 - (c) Uses approximate matching
 - (d) Does not use indexing
10. The main goal of an IR system is to:
 - (a) Retrieve all documents
 - (b) Retrieve documents quickly
 - (c) Retrieve relevant documents
 - (d) Retrieve recent documents only
11. Which component is responsible for ordering results?
 - (a) Indexing module
 - (b) Ranking module
 - (c) Document collection
 - (d) Tokenizer
12. Information need refers to:
 - (a) The query exactly as typed
 - (b) The user's real underlying problem
 - (c) The document title
 - (d) The ranking score

Exercise 02 – Data Structure & Algorithms

01. Sort a binary array in linear time

Write an algorithm to sort a binary array in linear time. The output should print all zeroes (0), followed by all ones (1).

02. Array frequency problem

Given the following code

```
int i = 0, j = 0, m = 0, c = 0;
while (i < array.Count)
{
    if (array[i] == array[j])
        c++;
    j++;
    if (j >= array.Count)
    {
        if (c > m)
            m = c;
        c = 0;
        i++;
        j = i;
    }
}
```

1. Given array = 15, 25, 36, 1, 1, 25, 63, 8, 8, 8, 1, 1, 1, 8, 9, 8, 96, 8, what is the value of the variable m.
2. What does the code do, knowing that the variable m is the result?
3. What is the complexity of the code
4. Describe another algorithm that would be better in time complexity

03. Extract common values in two sorted lists (Linear Time)

Write an algorithm to extract common values in two sorted lists, the algorithm must execute in linear time complexity.

04. Move Zeros to End While Maintaining Order

Write an algorithm to move all zeros to the end of ordered list, the algorithm must execute in linear time complexity.

05. Merge Two Sorted Arrays In Place

Write an algorithm to merge two sorted arrays in place; the algorithm must execute in linear time complexity

06. Display Distinct values in Sorted Array

Write an algorithm to display only distinct elements of a sorted array; the algorithm must execute in linear time complexity

07. Find All Duplicates in Array

Given an array nums of length n where $1 \leq \text{nums}[i] \leq n$, find all elements that appear more than once, under the following constraints: Do not use extra space, The algorithm should run in $O(n)$ time, The input array can be modified