

School of Computing and Data Science
The University of Hong Kong
COMP7106 Assignment 1 (10% of total marks)
Due date: Mar 7, 2025 (11:59am HKT)

Q1. [4%]

An online library manages a database with the following schemas:

- **Customer (cID, name, age, gender)**

stores the details of customers, where cID is the ID of a customer; gender is "Mr." or "Ms."

- **Book (bID, title, language, author, genre)**

stores the information of books, where bID is the book ID; language is the language used in the book (e.g., Chinese); author is the name of the author of that book; genre is the category of the book (e.g., Novel)

- **Borrow (cID, bID, dueDate)**

stores which customer (cID) borrowed which book (bID), and the return date (dueDate) of that book.

Answer questions (a) and (b) below.

(a) Express the following queries in **Relational Algebra**.

- 1) Find the titles of the Novel books which use Chinese as the language.
- 2) Retrieve the names of female customers who have borrowed Novel books and are due for return on 01-01-2025.

(b) Express the following queries in **SQL**.

- 1) Display the distinct genres of books borrowed by Mr. customers whose ages are between 40 and 60.
- 2) For each genre of books, display the genre and the average age of customers.

Q2. [6%]

Consider a set of 10 buildings on HKU (e.g., {Library, Gym, Cafeteria, ..., Auditorium}) that are spatial points $\{a, b, c, \dots, i, j\}$. These building points are indexed by an R-tree as illustrated in the two sub-figures of Figure 1. Table 2 shows the student satisfaction ratings of these buildings. Rating values are integers in $[1, 10]$.

Suppose that you are location at position q in Figure 1 (left) and want to find a building that **must** have rating > 6 , **and** as close to q as possible based on Euclidean distance.

(1) Fill in the steps of finding your desired building using incremental Best First Nearest Neighbor Search with the help of a priority queue Q into Table 1. Elements in Q are ordered by the distance to the query point q in ascending order. (In each step, you need to specify which node to access, the content of Q after access action at current step.)

(2) In the end, show the final result of the query. Show how many R-tree nodes are accessed during this search.

(The first step of accessing *root* node of R-tree is provided for your reference. After accessing *root*, MBR M_1 (distance to q is 1) is at the front of Q , while M_2 (distance to q is 2) and M_3 (distance to q is 4) are behind M_1 sequentially.)

When some elements in Q have the same distance, the following rules apply:

1. If the nodes are at different levels, the nodes at lower levels are prioritized first. For example, m_2 would come before M_1 .

2. If the nodes are at the same level, the nodes with smaller indices are prioritized first. For example, m_1 would come before m_2 .
3. In the case of objects(buildings), alphabetical order is followed. For example, "a" would come before "c".

Access action	Content of Q after access action
visit root	$M_1(1), M_2(2), M_3(4)$
	(add more rows if necessary)

Table 1: please fill your steps of finding your desired building into this table

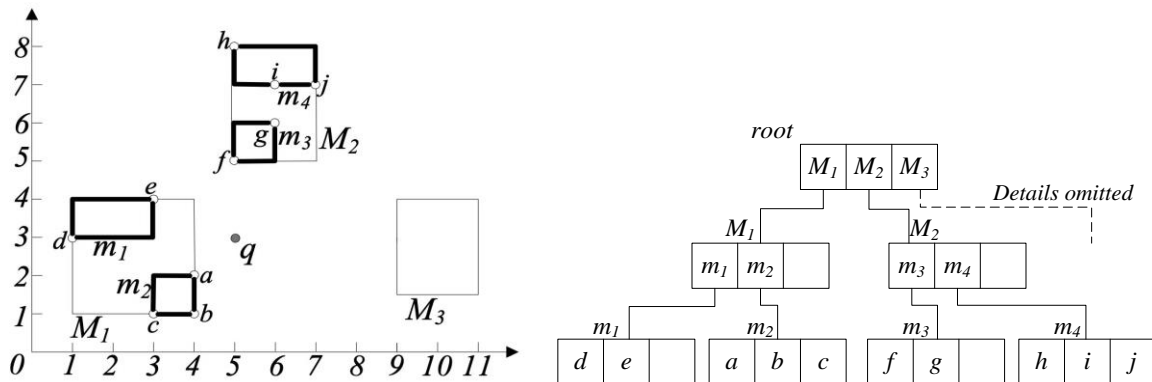


Figure 1: (left) spatial building points and enclosing rectangles; (right) corresponding R-tree

Building	a	b	c	d	e	f	g	h	i	j
Rating	6	7	7	5	5	7	8	6	4	7

Table 2: ratings of the ten buildings

Notes:

Please submit one file for this assignment to Moodle:

1) A PDF file, containing your answers for Q1 and Q2.

Please DO NOT submit any other materials.

Late submission will **NOT** be accepted!