MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
(A constituent unit of MAHE, Manipal)

VI SEMESTER B.TECH. END SEMESTER EXAMINATION

MAY 2022

SUBJECT: DATABASE MANAGEMENT SYSTEMS [MTE 4055]

Date of Exam: **19/05/2022** Time of Exam: **10:00 AM – 1:00 PM** Max. Marks: **50**

Instructions to Candidates:

- Answer ALL the questions & missing data may be suitable assumed

Q.NO.	QUESTION	M	CO	PO	LO	BL
1.	Association rule mining often generates a large number of rules. Discuss one effective method that can be used to reduce the number of rules generated while still preserving most of the interesting rules.	2	4	1, 2, 3	1, 2, 3	Discuss
2.	Consider the market basket of transactions shown below. For a minimum support of 40% perform the Apriori algorithm to find the most frequent item set. Suppose that frequent item sets are saved for a large transaction database, *DB*. Discuss how to efficiently mine the (global) association rules under the same minimum support threshold if a set of new transactions, denoted as Δ*DB*, is (incrementally) added in.	3	4	1, 2, 3	1, 2, 3	Discuss
3.	Construct a decision-tree classifier with binary splits at each node, using tuples in relation *r* (*A, B,C*) shown below as training data; attribute *C* denotes the class. Show the final tree, and with each node show the best split for each attribute along with its information gain value. (1, 2, a), (2, 1, a), (2, 5, b), (3, 3, b), (3, 6, b), (4, 5, b), (5, 5, c), (6, 3, b), (6, 7, c).	5	5	1, 2, 3	1, 2, 3	Construct

<table>
<tr>
<td>4.</td>
<td>
The following Table 2B an example of customer purchase transaction data set.

Calculate the support and confidence of the following association rule. Infer whether the items in the association rule are independent of each other or have negative or positive impacts on each other.

Note: CID = Customer ID and TID = Transactions ID

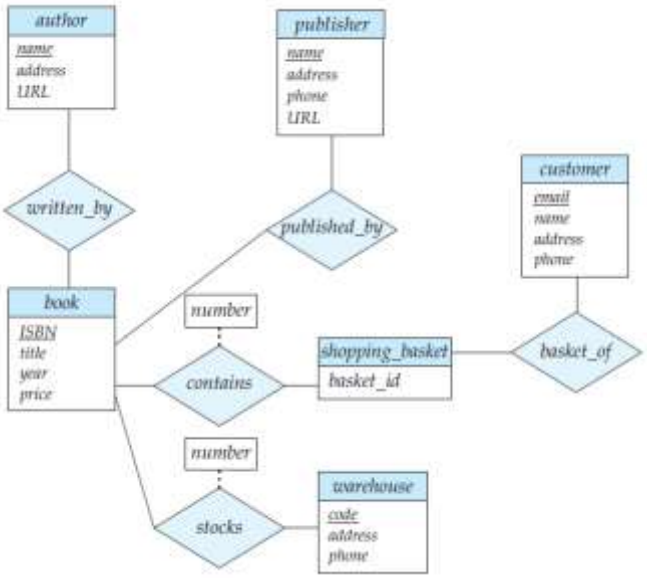
{10} → {50,70}
<table>
<tr><th>CID</th><th>TID</th><th>Date</th><th>Items Purchased</th></tr>
<tr><td>1</td><td>1</td><td>01/01/2001</td><td>10, 20</td></tr>
<tr><td>1</td><td>2</td><td>01/02/2001</td><td>10, 30, 50, 70</td></tr>
<tr><td>1</td><td>3</td><td>01/03/2001</td><td>10, 20, 30, 40</td></tr>
<tr><td>2</td><td>4</td><td>01/03/2001</td><td>20, 30</td></tr>
<tr><td>2</td><td>5</td><td>01/04/2001</td><td>20, 40, 70</td></tr>
<tr><td>3</td><td>6</td><td>01/04/2001</td><td>10, 30, 60, 70</td></tr>
<tr><td>3</td><td>7</td><td>01/05/2001</td><td>10, 50, 70</td></tr>
<tr><td>4</td><td>8</td><td>01/05/2001</td><td>10, 20, 30</td></tr>
<tr><td>4</td><td>9</td><td>01/06/2001</td><td>20, 40, 60</td></tr>
<tr><td>5</td><td>10</td><td>01/11/2001</td><td>10, 20, 30, 60</td></tr>
</table>
</td>
<td>2</td><td>4</td><td>1, 2, 3</td><td>1, 2, 3</td><td>Calculate</td>
</tr>
<tr>
<td>5.</td>
<td>
Let the following relational schemas be given:

$$R = (A, B, C)\ ; S = (D, E, F)$$
Let relations $r(R)$ and $s(S)$ be given. Develop an expression in the tuple relational calculus that is equivalent to each of the following:

a) $\Pi_A(r)$
</td>
<td>3</td><td>2</td><td>1, 2, 3</td><td>1, 2, 3</td><td>Develop</td>
</tr>
</table>

	b) $\sigma_{B=17}(r)$ c) $\Pi_{A,F}(\sigma_{C=D}(r \times s)$					
6.	Design a database for an airline company. The database must keep track of customers and their reservations, flights and their status, seat assignments on individual flights, and the schedule and routing of future flights. Your design should include an E-R diagram, a set of relational schemas, and a list of constraints, including primary-key and foreign-key constraints.	**5**	**2**	**1, 2, 3**	**1, 2, 3**	**Design**
7.	Consider the relational schema of Figure. *employee* (*person_name*, *street*, *city*) *works* (*person_name*, *company_name*, *salary*) *company* (*company_name*, *city*) *manages* (*person_name*, *manager_name*) Develop an expression in the relational algebra to express the following query; a) List the names of all employees who do not work for "First Bank Corporation". b) List the names of all employees managed by Mark Zuckerberg.	**2**	**3**	**1, 2, 3**	**1, 2, 3**	**Develop**
8.	Consider following Table. *ID* | *name* | *salary* | *dept_name* | *building* 22222 | Einstein | 95000 | Physics | Watson 12121 | Wu | 90000 | Finance | Painter 32343 | El Said | 60000 | History | Painter 45565 | Katz | 75000 | Comp. Sci. | Taylor 98345 | Kim | 80000 | Elec. Eng. | Taylor 76766 | Crick | 72000 | Biology | Watson 10101 | Srinivasan | 65000 | Comp. Sci. | Taylor 58583 | Califieri | 62000 | History | Painter 83821 | Brandt | 92000 | Comp. Sci. | Taylor 15151 | Mozart | 40000 | Music | Packard 33456 | Gold | 87000 | Physics | Watson 76543 | Singh | 80000 | Finance | Painter Explain two problems caused by the design given in the table. Suggest a normalization technique to resolve the anomaly in the table.	**3**	**4**	**1, 2, 3**	**1, 2, 3**	**Suggest**
9.	Consider the bank database of the figure where the primary keys are underlined and the following query holds.	**5**	**3**	**1, 2, 3**	**1, 2, 3**	**Develop**

	branch(branch_name, branch_city, assets) *customer (customer_name, customer_street, customer_city* *loan (loan_number, branch_name, amount)* *borrower (customer_name, loan_number)* *account (account_number, branch_name, balance)* *depositor (customer_name, account_number)* **select** *T.branch_name* **from** *branch T, branch S* **where** *T.assest > S. assests* and *S. branch_city = 'Brooklyn'* Develop an efficient relational algebra expression that is equivalent to this query. Justify your choice					
10.	Explain the distinctions among the terms primary key, candidate key and super key.	**2**	**3**	**1, 2, 3**	**1, 2, 3**	**Explain**
11.	For the set of transactions shown in Table, find the most frequent item set with a minimum support of 3 using the FP Growth algorithm. **Transactions** / **Item sets** 1 — Fries, Apples, Coconuts, Dates, Guavas, Mangoes, Pineapples 2 — Apples, Bananas, Coconuts, Fries, Lychees, Mangoes, Oranges 3 — Bananas, Fries, Hotdogs, Oranges 4 — Bananas, Kitkats, Coconuts, Pineapples 5 — Apples, Fries, Coconuts, Lychees, Pineapples, Mangoes, Nuts	**3**	**4**	**1, 2, 3**	**1, 2, 3**	**Find**
12.	Elaborate on the ACID properties of transactions with suitable examples.	**5**	**4**	**1, 2, 3**	**1, 2, 3**	**Elaborate**
13.	Break down the table shown in the table into Second Normal form. Differentiate between Third Normal form and Boyce Codd Normal Form. **Customer ID** / **Store ID** / **Purchase Location** 1 / 1 / Los Angeles	**2**	**4**	**1, 2, 3**	**1, 2, 3**	**Differentiate**

		1	3	San Francisco				
		2	1	Los Angeles				
		3	2	New York				
		4	3	San Francisco				
14.	Define Weak Entity set. Describe by using suitable illustrations.				**3**	**4**	**1, 2, 3**	**1, 2, 3**	**Define**
15.	Observe the database structure shown in the figure and write SQL queries for the following: i. List all the suppliers in alphabetical order. ii. List the number of customers in each country. Only include countries with 10 or more customers. iii. List the customers in Sweden. iv. Get details of the customer named 'Thomas Cook'. v. List all the products with names that start with 'Ca'. 				**5**	**3**	**1, 2, 3**	**1, 2, 3**	**List**