2019 EXAMINATIONS

Lancaster University

Part II

COMPUTING AND COMMUNICATIONS [2.5 Hours]

SCC.201 Databases

Candidates are asked to answer THREE questions from FOUR; each question is worth a total of 25 marks.

Use a separate answer book for each question.

Question 1

1.a Illustrating your answers with examples, explain the following relational database terms:

i. Functional Dependency **[2 marks]**
ii. Partial Dependency **[2 marks]**
iii. Transitive Dependency **[3 marks]**

1.b The tables in the appendix at the end of this paper represent scheduled flight departures from London Heathrow between 16:35 and 16:55, and details of actual flights on specified dates.

i. Giving clear justification for your answers, state whether each of the tables given in the appendix is in First Normal Form (1NF) and if not, explain how each table can be converted into 1NF. **[3 marks]**

ii. Identify the dependencies present in the tables, and in each case state the form of dependency. **[4 marks]**

iii. With reference to your answer to part ii above, state whether each of the tables is in Second Normal Form (2NF) and explain why this is the case. If not, show how each table can be converted into 2NF. You **must** show the final set of relations and their primary keys but there is **no need** to include any of the data. **[6 marks]**

iv. State whether the relations created in part iii are now in 3NF, justifying your answer with reference to the expectations for Third Normal Form (3NF). If not, giving the final set of relations, show what would need to be done to convert them to 3NF. **[5 marks]**

Total 25 marks

Question 2

You are asked to help develop a new database for a bus company intending to move into a new area at the start of next year. The database is intended to hold information about the proposed depots and the bus routes they will serve.

After initial discussion it was decided that the company will continue to use its existing systems to create timetables for services, while route numbers must be stored, at this stage, no bus stop or operating times are to be held in the database.

The company plans to establish several depots and operates several different types of buses. Each depot will typically have a few buses of each type. The company doesn't make modifications to buses, but the vehicle registration, and both the service/ maintenance and certification dates must be held for each individual bus.
Each bus route is operated from a single depot.

Not all drivers will be able to operate all the different types of buses and so must be authorised for the vehicles they can handle. The date at which the check was completed and the staff number of the person who carried out the check must be recorded.

Similarly, drivers must be trained for and learn a given route. Again, information about the date of the training and who supervised that training must be recorded.

Each driver will be based at a single depot and the database needs to hold the first and last names for each driver, their date of birth and contracted hours.

2.a Provide an complete Entity-Relationship diagram for the above database, including all entities and their relationships. **[14 marks]**

2.b Provide a final set of relations that would meet the above requirements, showing each and every step of the process. All working must be shown; you must identify each step and explicitly state when any step in the Entity-Relationship modelling process is not relevant to the design. **[11 marks]**

Total 25 marks

Question 3

3.a Below are two simple transactions that can potentially result in a lost update problem :

T1	**T2**
`read_item(X);`	`read_item(X);`
`X := X - N;`	`X := X + M;`
`write_item(X);`	`write_item(X);`
`read_item(Y);`	
`Y := Y + N;`	
`write_item(Y);`	

i. Give a schedule for the two transactions that will result in the lost update problem. Clearly indicate the point at which the problem arises. **[4 marks]**

ii. Apply the two-phase locking protocol to the two transactions so that issues such as the lost update problem cannot arise. You may use simple binary locks. **[2 marks]**

3.b Describe the concept of fragile and broken locks with respect to transactions. Why do these two states exist? **[2 marks]**

3.c What happens at the commit point of a transaction? Why might a transaction that has reached the commit point be aborted instead? **[2 marks]**

3.d If a transaction is aborted, how are the contents of the system log used to undo the effects of the transaction? **[4 marks]**

3.e Convert the following SQL query into a canonical query tree involving the relational algebra operators.

```
SELECT station_name
FROM TOWN T, RAILWAY_STATION R
WHERE (T.town_id = R.town_id) AND (T.town_name = "GLASGOW")
```

Apply simple heuristics to optimise the tree. Describe the heuristics applied and show the tree after each transformation. (Ignore heuristics applicable to the project operator.)

[11 marks]

Total 25 marks

Question 4

4.a

i. For a primary index structure to be applied, how should the records in the main file be ordered? **[2 marks]**

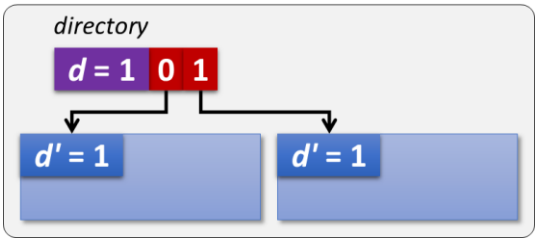
ii. For a clustering index structure to be applied, how should the records in the main file be ordered? **[2 marks]**

iii. For a secondary index structure to be applied, how should the records in the main file be ordered? **[2 marks]**

4.b Explain the drawbacks of conventional hashing. **[3 marks]**

4.c Describe the technique known as *extendible hashing* and detail how it overcomes these drawbacks. **[5 marks]**

4.d Given an extendible hashing scheme, and assuming the following starting point for the structure



add the following values to the structure :

value	pseudo-key (least significant 5 bits)
26	11010
10	01010
19	10011
24	11000

For full credit, you must show the state of the structure after each insertion, and describe clearly what changes each insertion cause to the structure. **[11 marks]**

Total 25 marks

Appendix: ***Flights from Heathrow Airport for question 1***

The following tables represent scheduled and actual flight departure times from London Heathrow.

You should assume that the content represents a representative sample of data.

The primary keys are underlined.

Schedule				*This table represents scheduled departure details*		
Airline	Flight	Destination	IATA_Code	Flight_Times		Terminal
				Depart	Arrive	
VS	011	Boston	BOS	16:35	19:15	3
A3	215	Athens	ATH	16:35	22:20	2
BA	680	Istanbul	IST	16:35	23:30	5
BA	215	Boston	BOS	16:40	19:30	5
BA	912	Frankfurt	FRA	16:40	19:20	5
KU	102	Kuwait City	KWI	16:40	01:55	4
BA	580	Milan	MXP	16:40	19:35	5
BA	820	Copenhagen	CPH	16:45	19:45	5
PK	786	Istanbul	ISB	16:45	05:20	3
BA	789	Montreal	YUL	16:45	19:10	5
VS	461	Johannesburg	JNB	16:45	05:45	3
CX	238	Hong Kong	HKG	16:50	12:55	3
LH	2477	Munich	MUC	16:50	19:35	2
OU	491	Zagreb	ZAG	16:50	20:05	2
TU	791	Tunis	TUN	16:50	20:50	4

Flights			*This table represents actual departure details*		
Airline	Flight	Date	Operator	Plane	Departed
VS	011	10-Jan-19	Virgin Atlantic	G-VLUV	16:56
VS	011	11-Jan-19	Virgin Atlantic	G-VGBR	16:52
A3	215	11-Jan-19	Aegean Airlines	SX-DVP	16:54
BA	680	11-Jan-19	British Airways	G-EUYH	17:07
BA	215	11-Jan-19	British Airways	G-ZZZA	17:23
BA	912	11-Jan-19	British Airways	G-EUOB	17:04
KU	102	11-Jan-19	Kuwait Airlines	9K-AOI	17:54
BA	580	11-Jan-19	British Airways	G-EUPK	17:28
BA	820	11-Jan-19	British Airways	G-EUUT	17:03
PK	786	11-Jan-19	Pakistan International Air	AP-BHX	17:01
BA	789	11-Jan-19	British Airways	G-ZBKB	18:00
VS	461	11-Jan-19	Virgin Atlantic	G-VAHH	17:36
CX	238	11-Jan-19	Cathay Pacific	B-KQZ	17:08
LH	2477	11-Jan-19	Lufthansa	D-AIWB	17:10
OU	491	11-Jan-19	Croatia Airlines	9A-CTH	17:13
TU	791	11-Jan-19	Tunisair	TS-IMF	17:38

--- End of Paper ---