2018 EXAMINATIONS

Lancaster University

Part II

COMPUTING AND COMMUNICATIONS [2.5 Hours]

SCC.201 Databases

Candidates are asked to answer ALL THREE questions on this paper; each question is worth a total of 25 marks.

Use a separate answer book for each question.

Question 1

The following questions refer to the railway database in Appendix A at the end of this paper.

1.a Being careful to declare any keys, give an SQL statement that would create the Locomotives table.

[7 marks]

1.b After some historical research you discover the locomotive commonly known as *The Pilling Pig* was actually named *The Farmers Friend*. Produce an SQL statement to correct this error in the database.

[3 marks]

1.c Write two SQL statements to produce:

i. A list of railway companies and for each, the count of how many of their locomotives are in the database

ii. A list of trains named after royal children along with their original number and company

The expected output of these two queries is shown below.

Company	Locomotives
G&KER	1
LMS	8
LNER	6
LNWR	1

Company	OrigNum	Name
LMS	6201	Princess Elizabeth
LMS	6203	Princess Margaret Rose
LNER	2553	Prince of Wales

[7 marks]

Question 1 continues on next page…

Question 1 continued…

1.d Provide a SQL statement showing how an *inner join* could be used in to produce the following output. Note that the locomotives are given alphabetically and that only six rows are output.

Locomotive	Company
City of Lancaster	London, Midland and Scottish Railway
Coronation	London, Midland and Scottish Railway
Dick Turpin	London and North Eastern Railway
Duchess of Buccleuch	London, Midland and Scottish Railway
Duchess of Hamilton	London, Midland and Scottish Railway
Duchess of Kent	London, Midland and Scottish Railway

[5 marks]

1.e Illustrate with sufficient example rows taken from the dataset provided how the output of a *right outer join* would differ from the output of the *inner join* given above.

[3 marks]

Total 25 marks

Question 2

2.a With reference to the following dataset

ID	Film	Rating	Length	Language	Date	Genre
187	12 Angry Men	U	1:32	English	1957	Courtroom Drama
190	A Few Good Men	15	2:17	English	1992	Drama
191	A Fistful of Dollars	15	1:35	English	1964	Classic Spaghetti Western
192	A View to a Kill	PG	2:05	English	1985	Spy Adventure
196	All the Presidents Men	15	2:12	English	1976	Political Thriller
197	Amadeus	PG	2:33	English	1984	Biographical Drama
201	Amelie	15	1:56	French	2001	Romantic Comedy Drama
251	Diamonds are Forever	PG	1:54	English	1971	Spy Adventure

i. Assuming a relational model, give examples for each of:

- Attribute
- Tuple
- Domain

[3 marks]

ii. Identify the keys(s) and candidate key(s) that would hold for this sample data. You must justify your answers and any decisions you make.

[3 marks]

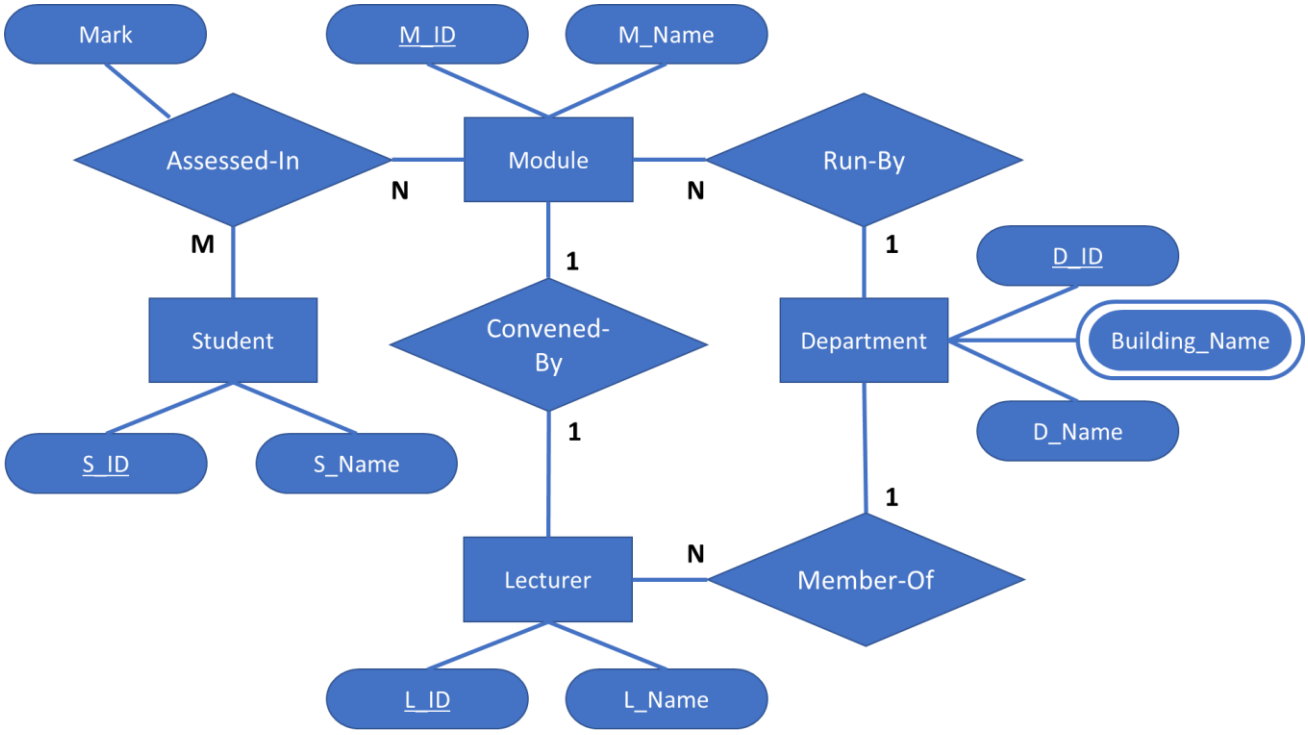
iii. Again justifying your answers, given more data describing a larger set of films, which key(s) might no longer function? Which of the key(s) you identified in part (ii) above would you select as your primary key?

[3 marks]

Question 2 continues on next page…

Question 2 continued…

2.b Given the following Entity-Relationship (ER) diagram



Convert this into a set of relations. You MUST show all working and, each and every transformation step.

[16 marks]

Total 25 marks

Question 3

3.a What is the difference between a primary index and a secondary index?

[2 marks]

3.b

i. In hashing, what is meant by a collision? Briefly describe the two collision resolution techniques known as open addressing and chaining.

[3 marks]

ii. Illustrate the way chaining operates using the following example.

A PARTS file with Part# as the hash key includes records with the following Part# values:

23, 32, 14, 27, 13, 36, 26, 16, 25

The file uses 4 buckets, numbered 0 to 3. Each bucket is one disk block and holds two records. Load these records into the file in the given order using the hash function

$$h(K) = K \bmod 4.$$

To assist you, here are the results of applying the hash function to the nine Part# values shown above.

Part #	23	32	14	27	13	36	26	16	25
h (Part#)	3	0	2	3	1	0	2	0	1

You do not need to show the individual parts being added to the hash file, just its final state.

[3 marks]

iii. Describe three main drawbacks of conventional (static) hashing?

[3 marks]

Question 3 continues on next page…

Question 3 continued…

3.c

i. What does the "B" in B-trees and B+-trees stand for? What is the major difference between the internal nodes of a B-tree and a B+-tree?

[3 marks]

ii. Describe the process of adding the following data to a B+-tree of order 3, with the aid of diagrams. The data is in the following order:

204, 312, 76, 632, 189, 451

The tree begins as a single, empty leaf node. Show clearly the state of the tree after each insertion and describe any transformation that occurs. Full working must be shown for full credit.

[11 marks]

Total 25 marks

Appendix A: *Railway Database for Question 1*

Locomotives

ID	Company	Class	Name	Orig-Num
11302	G&KER	Hudswell Clarke 0-6-0ST	The Pilling Pig	1885
45134	LNWR	Renown	John of Gaunt	1921
46201	LMS	Princess Royal	Princess Elizabeth	6201
46203	LMS	Princess Royal	Princess Margaret Rose	6203
46212	LMS	Princess Royal	Duchess of Kent	6212
46220	LMS	Coronation	Coronation	6220
46229	LMS	Coronation	Duchess of Hamilton	6229
46230	LMS	Coronation	Duchess of Buccleuch	6230
46233	LMS	Coronation	Duchess of Sutherland	6233
46243	LMS	Coronation	City of Lancaster	6243
60017	LNER	A4	Silver Fox	2512
60022	LNER	A4	Mallard	4468
60007	LNER	A4	Sir Nigel Gresley	4498
60103	LNER	A3	Flying Scotsman	4472
60054	LNER	A3	Prince of Wales	2553
60080	LNER	A3	Dick Turpin	2579

Companies

Company	Name	TrackMiles
GWR	Great Western Railway	3797
G&KER	Garstang and Knot-End Railway	11.5
LMS	London, Midland and Scottish Railway	7790
LNER	London and North Eastern Railway	6590
SR	Southern Railway	2186

--- End of Paper ---