UNIVERSITY OF LONDON

BSc EXAMINATION 2024

For Internal Students of
Royal Holloway

DO NOT TURN OVER UNTIL TOLD TO BEGIN

CS2855: Databases
CS2855R: Databases – for FIRSTSIT/RESIT CANDIDATES

Time Allowed: **TWO hours**

Please answer **ALL** questions

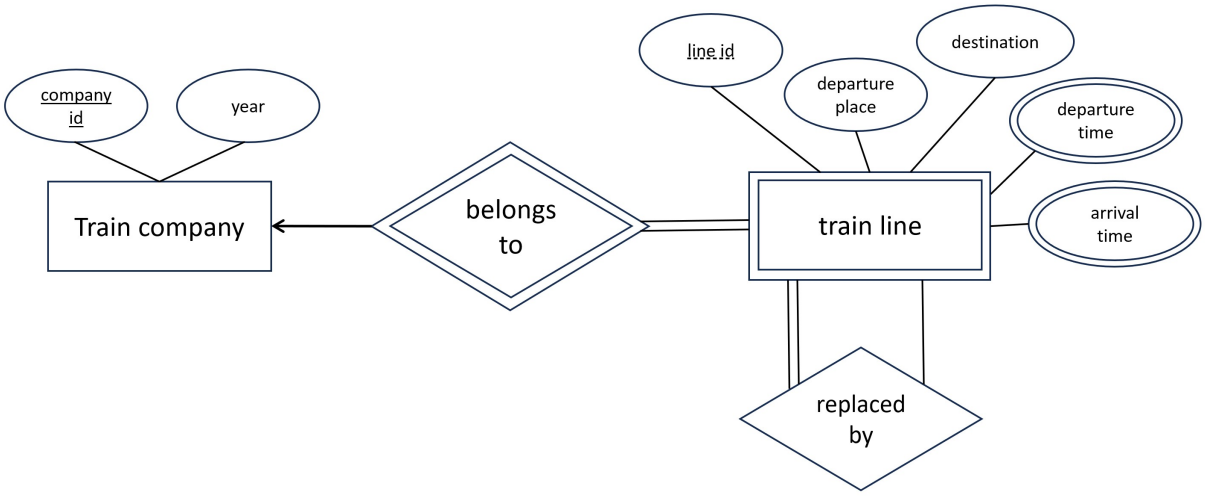
Calculators are not permitted

1. (a) *Innovex* is an IT company that needs a database to store its employee' information and their positions. The desired design is described below.

   - For every employee, the system stores their full name (which is unique), a list of emails, the employee ID, and the date they were hired. Every employee can be associated with several positions.
   - For each employee position, the system stores the position ID, position level, and the salary associated with that position. Note that the salary does NOT depend on the level.
   - The system must store the date each employee takes up their position.
   - Depending on their performance, the employees could be entitled to receive bonuses. For each position, there is a set of bonuses associated with the specific position; each bonus has an ID, the date that it was granted, and an amount. A bonus cannot exist without being associated with a position.

   Draw an E-R diagram according to the above design. Remember to include all constraints. [15 marks]

   (b) Convert the following E-R diagram into a relational model. Write down the relational model, including primary key and foreign key constraints. Minimize as far as possible



[15 marks]

NEXT PAGE

2. The following shows the relations that store information for a transportation company. Attribute DriverID is the primary key of table Driver. Attribute Registration is the primary key of table Vehicle. Every attribute that shares its name with the primary key of a different relation is a foreign key.

Driver		
DriverID	Name	Salary
2334	Alex	27,000
3299	Mary	32,100
1926	Rohit	35,200

Vehicle		
Registration	Vehicle_Mileage	Year
BD32 XYZ	3110	2023
LV63 3BX	22100	2017
FG12 AKD	40088	2020

Transport		
DriverID	Registration	Driver_Mileage
2334	LV63 3BX	505
2334	FG12 AKD	130
3299	LV63 3BX	1007

(a) Translate the following into an equivalent expression in Relational Algebra. Find the DriverID of all drivers that have driven a vehicle with mileage at least 20000. [5 marks]

(b) Write SQL statements for the following tasks over the above relations. Your statements should be correct for general instances of the above schema, and may not depend on the specific contents of the example tables above.

i. The company has decided to store in its database a list of phones associated with each driver. Modify the database in order to be able to store this information. [5 marks]

ii. For each driver, find the number of different vehicles they have driven. [5 marks]

iii. Increase the salary of every driver as follows. Every driver with a salary less than or equal to 20,000 should get an increase of 5%, every driver with a salary greater than 20,000 but less than 30,000 should get an increase of 3%, and every other driver should get an increase of 1%. [5 marks]

iv. For each vehicle, find their Registration and the names of the drivers that have driven it more than 3 times. [5 marks]

v. Find the names of the drivers with salaries less than 15,000 that have driven cars with Vehicle_Mileage at most 50000 and Year after 2019. [5 marks]

vi. Find all vehicles with Vehicle_Mileage higher than the average Vehicle_Mileage that have been driven by at least 3 drivers. [5 marks]

3. (a) Consider the following schedule of transactions T1 and T2. Is the schedule serialisable? If yes, provide an equivalent serial schedule. If no, briefly explain why. [5 marks]

```
T1       | T2
-----------------
         | Read(B)
Read(A)  |
         | Read(A)
         | Write(A)
Write(A)|
         | Write(B)
Read(B)  |
Write(B)|
```

(b) Consider the following transactions T1 and T2 working concurrently on items A, B, and C, without any transaction control. At time step 0, the value of A is 10, B is 20, and C is 40. What are the values of A, B, and C after time step 18? What value does sum have? [5 marks]

Time	T_1	T_2
1	`read item(A);`	
2	`A=A-10;`	
3		`sum=0;`
4		`read item(A);`
5	`write item(A);`	
6		`sum=sum+A;`
7		`A=A+20;`
8	`read item(B);`	
9		`write item(A);`
10		`read item(B);`
11		`sum=sum+B;`
12	`B=B+20;`	
13	`write item(B);`	
14	`read item(C);`	
15	`C=C-10;`	
16	`write item(C);`	
17		`read item(C);`
18		`sum=sum+C;`

(c) In the abbreviation ACID of properties of correctly performed transactions, explain the "D" (durability) briefly. [5 marks]

(d) Write the definition of BCNF (Boyce-Codd Normal Form). [5 marks]

4. Given below is the set F of functional dependencies for the relation schema $R = (A, B, C, D, E)$:

$$F = \{AB \to C,\ BD \to A,\ DC \to E\}\,.$$

(a) Find a candidate key for R and explain your answer. [10 marks]

(b) Is the following decomposition of R a lossless join? Explain your answer by showing the criterion for lossless-join decompositions and why the criterion is or is not met.

$$R_1 = (ABC) \qquad R_2 = (ABDE).$$

[5 marks]

END