UNIVERSITY OF LONDON

BSc EXAMINATION 2017

For Internal Students of
Royal Holloway

DO NOT TURN OVER UNTIL TOLD TO BEGIN

CS2855: Databases
CS2855R: Databases – PAPER FOR RESIT CANDIDATES

Time Allowed: **TWO hours**

Answer ALL questions
Calculators are not permitted

1. The newly founded UK Train Council is responsible for all trains within the UK. The Council is in the process of setting up its database. The Council first sets out the E-R model for the database.

   In this E-R model they have to record as entities different *train companies* (e.g., South West Trains, Virgin Trains, etc.) and *train lines* for each company. Each train company is identified by its unique alphanumeric code, and each train line is identified by its id number; though the id number for each train line is unique only with respect to the specific train company (i.e., two train lines that belong to different companies can have the same code); also every train line belong to precisely one train company. Moreover, each train line has its own fixed departure place and destination place and departure and arriving times.

   (a) Draw an appropriate E-R diagram for the above setting (use only the information provided in the textual description above). Make sure to identify correctly all the constraints in the diagram. Explain very briefly your design decisions. [13 marks]

   (b) Convert the E-R diagram from section (a) to *tables/relations in the Relational-Data model*. In other words, show all the relevant *table-schemas* (written simply as a sequence of attributes, as we have done in class), specifying their integrity constraints (i.e., primary and foreign keys), while also aiming not to add unnecessary tables and attributes. [10 marks]

   (c) Recurring signaling problems and frequent delays in train services entailed the UK Train Council to assign a set of potential "replacement" train lines for each train line. For example, the fast train line from Waterloo to Egham can be replaced by taking the train line from Waterloo to Staines and then another train line from Staines to Egham.

   Draw the E-R diagram for this part of the database taking into account the constraints mentioned (you do **NOT** need to draw the diagram related to entities and relations that are not effected by the above consideration).

   [7 marks]

2. Consider the following relational tables:

FlightDelays = (flight_code, origin, destination, average_minutes_delay, price)

Foreign Key: origin referencing airport_code in Airport table.

Foreign Key: destination referencing airport_code in Airport table.

Airport = (airport_code, city_name)

Foreign Key: city_name referencing city_name in City table.

City = (city_name, country)

(we assume for simplicity that every city name in the world is unique).

In what follows please **do not** use in your SQL queries any built-in JOIN operation.

(a) Write the SQL definition of the table FlightDelays (including integrity constraints); in other words, create the table in SQL. We assume that flight codes are alphanumeric and airport codes are always three characters. [8 marks]

(b) Express the following query in SQL:

"What are the three flights with the largest total delay time (in minutes) and their corresponding total delay time?"

For example, the output should be:

EZ9342 188
BA2334 140
IB7720 98 [8 marks]

(c) Express the following query in SQL:

"What are the flights whose origin airport and destination airport are in the same city?"

[7 marks]

(d) Express the following query in SQL:

"What are the airports that have flights originating from them whose average delay in minutes is less than 15, and whose price is less than 200 (GBP)?" [8 marks]

(e) Consider the following relational tables:

Airport = (airport_code, destination_city)

Foreign Key: destination_city referencing city_name in City table.

City = (city_name, country)

(we assume for simplicity that every city name in the world is unique, and that all the cities in the world are listed in the City relation).

Express the following query in **Relational Algebra**:

"What are the airports that have outgoing flights to *each and every* city in the world?" [6 marks]

3. (a) Let $F = \{CB \to D, C \to ED, A \to B\}$ be a set of functional dependencies. Is CA a superkey of relation $R = (A, B, C, D, E)$ with respect to F? Explain briefly your answer. [8 marks]

   (b) Let $R = (J, K, L)$ be a relation schema and consider the set of functional dependencies $F = \{JK \to L, L \to K\}$.

      i. Decompose R into relation schemas in BCNF (Boyce-Codd Normal Form). (You should do this by applying the simple procedure for decomposing relations into BCNF we have seen in class.) Briefly justify your decomposition (namely, explain briefly why the decomposition is indeed in BCNF). [13 marks]

      ii. Does your BCNF decomposition in (i) *preserve dependencies*? Explain briefly your answer. [2 marks]

   (c) Determine whether the following relation schema R is in 3NF with respect to the following set F of functional dependencies:

$$R = (airport\_code, city, country, flight\_code, price)$$

$$F = (airport\_code \to city \;\; country, \;\; flight\_code \to price, \\ flight\_code \to airport\_code)$$

[10 marks]

END