



**Tuesday 07 February 2023
14:00-16:00 GMT
Duration: 1 hour 30 minutes
Additional time: 30 minutes
Timed exam – fixed start time**

DEGREE of BSc Software Engineering (Graduate Apprenticeship)

**Data Storage and Retrieval
COMPSCI2029**

Answer all questions

This examination paper is an open book, online assessment and is worth a total of 60 marks.

1. This question is about Entity Relationship Diagrams. [Total: 20]

Consider the following scenarios:

- (a) **Scenario 1:** Every student has a tutor that is a course coordinator. Students and coordinators have unique IDs. A student must be enrolled on a course by a course coordinator, and a student can have many course coordinators that enrol them in different courses. All coordinators have students to enrol but not all coordinators are tutors. A course has a unique ID, a start time, end time and a derived duration. Each course has a single coordinator but coordinator run many courses.

- (i) Draw an entity relationship diagram to model the scenario showing the cardinalities and the participation for each of the relationships. [7]
- (ii) List the resulting headings for the schema relations. For each relation in the schema, make reasonable choices for attribute domains, underline the primary key and clearly name any foreign keys to show the constraints. [3]

- (b) **Scenario 2:** A theatre wants to develop a database about its plays. Each play includes some props but a prop can only be used in 1 play. A prop has a unique ID. There are several plays, each with a unique name, for example “Macbeth”. The theatre’s actors have unique actor’s union IDs, as well as their name. A production of a play is led by a director who produces a screenplay (identified by a unique ID) for the play. A director can lead several plays. Directors have unique IDs and a name. An actor can use one or more props and can act in any number of plays. The theatre has rehearsals, and therefore there is a relationship between directors, actors and props. In this relationship, each entity can participate multiple times.

- (i) Draw an entity relationship diagram to model the scenario showing the cardinalities and the participation for each of the relationships. [7]
- (ii) List the resulting headings for the schema relations. For each relation in the schema, make reasonable choices for attribute domains, underline the primary key and clearly name any foreign keys to show the constraints. [3]

2. This question is about SQL and relational Algebra. [Total: 20]

Assume the following database schema:

Employee(ID, name, age)

Business(ID, name, type)

WorksFor(employeeID, businessID, location)

- (a) Write the SQL statement that will return a list of the IDs of businesses that are located in both Glasgow and Edinburgh. What SQL design pattern should be used for this query?

[4]

- (b) Write the SQL statement that will return a list of locations and the number of businesses that are in each location. What SQL design pattern should be used for this query?

[4]

- (c) Write the SQL statement that will return the names and IDs of people that work in restaurants in Glasgow. What SQL design pattern should be used for this query?

[4]

- (d) Write the SQL statement that will return a list containing each location and the number of employees that work in that location, ordered from the location with the largest number of employees to the location with the smallest number of employees.

[4]

- (e) Adhering to best practices for readability, write the relational algebra to return the IDs and names of employees that work in Glasgow and are younger than 21 years old.

[4]

3. This question is about sets and relations.

[Total: 10]

Consider the following relations:

R1=

First	Last	Age
Elisha	Prince	96
Neil	Riggs	90

R2=

First	Last	Age
Cesar	Odom	84
Kian	Arm	65
Elisha	Prince	96
Ali	Gai	84

R3=

First	Last	Age
Neil	Riggs	90
Kian	Arm	65
Ali	Day	72

- (a) What is the relation that results from $R2 \cap R3$? [2]
- (b) What is the relation that results from $R1 \cup R2$? [2]
- (c) What is the relation that results from $R2 \oplus R3$? [2]
- (d) What is the relation that results from $R3 \times R1$? [2]
- (e) Consider the sets

$A = \{\text{Amy, Pete, Kai, Frank}\}$

$B = \{\text{North, South, East, West}\}$

and the relation

$R = \{ \langle \text{Amy, South} \rangle, \langle \text{Pete, North} \rangle, \langle \text{Frank, South} \rangle, \langle \text{Pete, East} \rangle, \langle \text{Kai, South} \rangle, \langle \text{Frank, East} \rangle \}$

What is the set

$\{x \mid \langle x, \text{South} \rangle \notin R \text{ and } \langle x, \text{East} \rangle \in R\}$? [2]

4. This question is about processing in a basic Web search engine. [Total: 10]

(a) Consider the following document:

‘The University’s students are enrolled in many courses. To enroll a student in a course, an administrator checks the student’s grades and enrolls the student to the relevant course.’

Show the result of applying **each** of the three text representation transformation operations that are usually applied to a document before it is indexed by a (basic) Web search engine.

In your answer, you should show the result of applying each of the transformation operations in the correct order, with the output of one transformation as the input to the following transformation. For each transformation, clearly state the transformation operation that is being applied, provide a short description (~1-2 sentences) of what the transformation operation does and **show the resulting transformed text**.

Three marks are available for each transformation and an additional mark is available for showing the correct order of transformations.

[10]