**Tony Siu 916209430**

**1. (25 pts)**

The following is the requirement descriptions for a Coronavirus Testing Database for a state.

Create the ER diagram for this database.

* The database will keep track of the residents, testing centers, zip codes, cities and counties in the state.
* For each resident, we will keep track of a unique resident identifier, a unique SSN, his/her name, street address, multiple phones, birthdate, and age. The name is composed of first name and last name. Each resident may have a various number of phones. The age is computed based on the birthdate and the current date.
* For each testing center, we will keep track of a unique testing center identifier, name, phone, and street address.
* For each zip code, we keep track of a unique zip code identifier (a simple number), and the unique zip code. For example, if there are only 60 zip codes stored in the database, then the zip code identifier goes from 1 to 60. The actual zip code could be like 19122, 19123, etc.
* For each city, we keep track of a unique city identifier, the city name, and the population in the city.
* For each county, we keep track of a unique county identifier, the unique county name, and the population in the county.
* Each resident in the database must get tested in at least one testing center and could get tested in multiple testing centers. Each testing center in the database must offer testing to at least one resident, and could offer testing to many residents. For each testing that a resident gets in a testing center, we will record the date when the testing occurs, and the testing result.
* Each resident in the database could get tested in a specific testing center multiple times on different dates. But on the same day, no resident gets tested in the same testing center more than once.
  + - **HINTS: the slides in the following file may be helpful to you.** 
      * [**Ch2-MultipleInstancesBetweenSameEntities-ForFinalExam.pdf**](https://templeu.instructure.com/courses/138900/files/28725659?wrap=1)
* Each resident in the database resides in one and only one zip code. For each zip code in the database, there can be many residents living there, or no resident at all.
* Each testing center in the database is located in exactly one zip code. For each zip code in the database, there can be at most one testing center, or no testing center at all.
* Each city in the database must have at least one zip code, and could have more than one zip code. Each zip code belongs to exactly one city.
* Each city in the database is incorporated into exactly one county. Each county in the database must include one city, and could include many cities.

2.

A diagram of a company

Description automatically generated

3.   A diagram of a computer program

Description automatically generated

1. **(20 pts)**

A white table with black text

Description automatically generated

Observe the following table AUTHERBOOKMIX with the sample data.

This table captures the data about authors, books, and publishers.

In addition, it shows the per-book royalty amount of an individual author.

Each author has a unique author id, last name, first name, and an institution to which he/she is affiliated.

Each book has a unique book id, a title, and a price.

Each publisher has a unique name and a city where it is headquartered. Each publisher has exactly one headquarter.

Each author can write at least one book and could write multiple books.

Each book is written by at least one author and could be written by multiple authors.

Each book is published by exactly one publisher.

Each publisher can publish at least one book, and could publish many books.

**a)**

        Using the table AUTHERBOOKMIX, describe an example of an insertion anomaly, a deletion anomaly, and a modification anomaly.

**Insertion Anomaly**

An **insertion anomaly** occurs when certain information cannot be recorded until other information is provided. In the AUTHERBOOKMIX, you can’t add new book without creating new publisher and additional information pertaining to the book.

**Deletion Anomaly**

A **deletion anomaly** occurs when deleting data about one entity inadvertently removes data about another entity. An example from the AUTHERBOOKMIX would be deleting the book *Introduction to Algorithms.* If we delete *Introduction to Algorithms*, you also have to delete information of 2 of its authors and additional information about the book.

**Modification Anomaly**

A **modification anomaly** happens when the same data is duplicated across multiple rows and needs to be updated in more than one place. In the given table, if the price of "Introduction to Algorithms" changes, it must be updated in every row where this book appears (once for each author associated with it, such as Thomas Cormen and Charles E. Leiserson). Failing to update all instances can lead to data inconsistencies.

**b)**

What are the types of the following functional dependencies (full, partial, and transitive) ?

AuthorID ---> AuthorLastName, AuthorFirstName, AuthorInstitution (Partial)  
BookID ---> BookTitle, BookPrice, BookPublisher, PublisherCity (Partial)  
BookPublisher ---> PublisherCity (Transitive)  
AuthorID, BookID ---> AuthorBookRoyalty (Full)

**c)**

Is the table in 1NF? If not, normalize it to 1NF.

It is in 1NF

**d)**

Show the result of normalizing the table to 2NF.

A black and white diagram

Description automatically generated

**e)**

Show the result of normalizing the table to 3NF.

A computer screen shot of a black background

Description automatically generated

**5. (15 pts)**

Given the following transaction schedule, draw its precedence graph, decide if it is conflict serializable.

If it is conflict serializable, list a serializability order of the transactions based on the precedence graph.

If it’s not conflict serializable, explain why it’s not.

T1: W(D), T5: R(F), T4: W(E), T5: R(D), T3: W(F), T2: R(H), T4: R(D), T6: R(F), T3: R(E), T1: R(D), T6: W(H), T2: R(E)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| T1 | T2 | T3 | T4 | T5 | T6 |
| W(D) |  |  |  |  |  |
|  |  |  |  | R(F) |  |
|  |  |  | W(E) |  |  |
|  |  |  |  | R(D) |  |
|  |  | W(F) |  |  |  |
|  | R(H) |  |  |  |  |
|  |  |  | R(D) |  |  |
|  |  |  |  |  | R(F) |
|  |  | R(E) |  |  |  |
| R(D) |  |  |  |  |  |
|  |  |  |  |  | W(H) |
|  | R(E) |  |  |  |  |

A diagram of a number

Description automatically generated with medium confidence

It is not conflict serializable because the precedence graph is not a acyclic graph