Engineering, Built Environment and IT
Department of Computer Science

Database Systems
COS 221

Examination Opportunity 2 (EO2)

16 May 2022 (17:30 to 19:30)

Examiners

Internal: Dr Linda Marshall and Mr Andrew Broekman

Instructions

1. Read the question paper carefully and answer all the questions.

2. The assessment opportunity comprises of **6** questions on **7** pages. Refer to the table below for the marks per question.

3. **2** hours have been allocated for you to complete this paper. You will download the paper and can write it offline. You need to photograph (make sure your photographs are clear) or scan all your pages and create a **single PDF** of your answers.

4. An upload slot will be open on the module ClickUP page under the **Examination Opportunities** menu option for the duration of the examination opportunity (17:30 to 19:30) and then for an additional 30 min to give enough time to download this paper, create the PDF containing your answers and then upload your PDF. **No late submissions will be accepted.**

5. Make sure you include your student number, name and a picture of your student card on the first page of your submission. Leave some space for your marks to be written on the first page. Each subsequent page of your submission must include at least your student number and name at the top of the page.

6. This paper is **take home** and is subject to the University of Pretoria Integrity statement provided below.

   - You are allowed to consult any literature.
   - You are not allowed to discuss the questions with anyone.
   - You may not copy from online resources. All answers must be in your own words.

7. If you have any queries when writing the paper, post them in good time to the COS 221 WhatsApp group or use the chat functionality on the Blackboard Collaborate session for EO2. Make sure your post in general enough as not to give away any answers. These two channels will be used to deliver any general information during the Examination Opportunity.

Integrity statement:
The University of Pretoria commits itself to produce academic work of integrity. I affirm that I am aware of and have read the Rules and Policies of the University, more specifically the Disciplinary Procedure and the Tests and Examinations Rules, which prohibit any unethical, dishonest or improper conduct during tests, assignments, examinations and/or any other forms of assessment. I am aware that no student or any other person may assist or attempt to assist another student, or obtain help, or attempt to obtain help from another student or any other person during tests, assessments, assignments, examinations and/or any other forms of assessment.

Question:	1	2	3	4	5	6	Total
Marks:	14	17	23	10	20	12	96

Full marks is: 85

Scenario

The University of Pretoria's Faculty of Health Sciences has requested the assistance of the Department of Computer Science to redesign the database systems used by the School of Medicine at the Steve Biko Academic HOSPITAL. After consultation with the appropriate stakeholders the following requirements where gathered and submitted to you as the database designer:

The hospital admits various PATIENTs. Each PATIENT can either be a South African CITIZEN or a FOREIGN patient, but not both. All patients must have a full name and date of birth captured. South Africa patients are uniquely identified by their South African identification number while foreign patients are uniquely identified by a passport number. Keep in mind that patients may belong to a medical aid, thus requiring to capture the scheme name (e.g. Medihelp, Bonitas, Discovery), a plan name (e.g. MedSave, Classic Saver), a medical aid number (e.g. 3004577) an a dependent code (e.g. 00 for the principal member, 01 for the first dependent etc.)

A patient is intially traiged and assigned a fixed category. In the context of hospital admission, *triage* is the process of identifying patients requiring immediate assistance from those requiring non-critical assistance. The hospital makes use of four categorisations: *RED* — Patients requiring resuscitation (0min); *ORANGE* — Patients requiring urgent assistance (15min); *Yellow* — Patients requiring less urgent assistance (60min); and *GREEN* — Patients requiring non-urgent assistance (180min). Additional to assigning a categorisation, the initial assessment of the patient is recorded as well as the possibility of contamination and the contaminant. A NULL value for contamination signifies that there is no risk of contamination.

Each patient is assigned an overseeing or RESPONSIBLE doctor with a start date and end date. Each doctor can be categorised based on their qualification and employment type.

- A REGISTERED_DOCTOR is an individual that is registered at the Health Professions Council of South Africa (HPCSA) and has been allocated an HPCSA number.

- A RESIDENT_DOCTOR is an individual who has completed their studies and is currently completing their 24 months residency in a public or private hospital. A resident doctor is registered with the HPCSA and is issued an intern HPCSA number. A resident doctor always has at least one responsible doctor assigned to them, which may change over time as they rotate in the hospital.

- A STUDENT_DOCTOR does not register with the HPCSA and thus have not been allocated a HPCSA number. Further more a student doctor is under the guidance of a REGISTERED_DOCTOR. The institution, student number allocated by the institution and degree the individual is studying towards needs to be captured for compliance reasons.

- A LOCUM_DOCTOR is another REGISTERED_DOCTOR who fulfils the duties and responsibilities of another doctor. If a given doctor is unavailable, they may nominate a LOCUM_DOCTOR in their position.

The hospital needs to know about all NURSE personnel working in the hospital. Each NURSE can either be a STUDENT or REGISTERED nurse. The classification of nurses is as follows:

- A REGISTERED_NURSE is an individual that is registered at the South African Nurses Council (SANC) and has been allocated a registration number.

- A STUDENT_NURSE does not register with the SANC and thus has not been allocated a number. The institution, student number allocated by the institution and qualification the individual is studying towards needs to be captured for compliance reasons.

The database must track all MEDICINE that is in the possession of the hospital. MEDICINE is identified by the name **and** concentration of the medication, e.g. *Flagyl* is an antibiotic available in concentrations of 250mg, 500mg and 2000mg. In addition the active ingredient, manufacturer and schedule of each medication needs to be captured. Medicines are categorised from schedule 0, available over the counter at a supermarket/pharmacy (e.g. Panado) to schedule 8, which are strictly controlled substances (e.g. anaesthetics),

A HEALTH_CARE_PROFESSIONAL can prescribe MEDICINE to a patient. A HEALTH_CARE_PROFESSIONAL is either a REGISTERED_DOCTOR or REGISTERED_NURSE. For item on the PRESCRIPTION, the medicine with associated dosage information must be specified, Dosage information consists of a quantity and a window

period. Assume that a doctor may PRESCRIBE fractional quantities of the medication; assume window period to be a string. Thus Flagyl at 200mg concentration can be prescribed at a quantity of 1,5 with a window period of "Strictly every 8 hours".

A nurse is responsible to ADMINISTER the MEDICINE to a PATIENT. The date and time of administration, batch number of the medication and administered quantity must be recorded. You may assume only MEDICINE that has a PRESCRIPTION will be ADMINISTERed.

Before you launch into answering the questions, read them carefully and draw the (E)ER-diagram for the HOSPITAL database.

Questions

1. (a) What is the the *Impedance Mismatch Problem*? Give an example to explain your answer. (3)

   (b) Study the Java program below and answer the questions which follow.

```
1  import java.sql.*;
2  class MySQLCon {
3    public static void main (String args[]) {
4      try {
5        Class.forName("org.mariadb.jdbc.Driver");
6        Connection con =
             DriverManager.getConnection("jdbc:mariadb://localhost:3306/SteveBiko",
             "user", "password");
7        Statement stmt = con.createStatement();
8        ResultSet rs = stmt.executeQuery("select * from MEDICINE");
9        while (rs.next()) {
10         String name = rs.getString(1);
11         int concentration = rs.getString(2)
12         System.out.println(name, concentration);
13       }
14       con.close();
15     } catch (Exception e) {
16       e.printStackTrace();
17     }
18   }
19 }
```

   i. What is the main purpose of the program? (3)
   ii. What is the role of the driver manager class in line 5? (2)
   iii. Explain the use of the Connection object, Statement object and ResultSet object. (6)

2. (a) Compare "inheritance" in the EER model to inheritance in the OO model. (4)

   (b) Is the following statement True or False? Provide a reason for your choice. (2)
      Objects in object oriented databases are uniquely identified in the OODB by the "primary key" of the object.

   (c) What is the difference between persistent and transient objects? (2)

   (d) Using examples, briefly explain the differences between structured, semi-structured, and unstructured data? (6)

   (e) Under which category/categories given in Question 2. (d) do XML documents fall? (1)

   (f) What are the differences between the use of tags in XML versus HTML? (2)

3. Consider the XML schema document with the patient as the root and answer the questions that follow.

```
1  <xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
       xmlns:xs="http://www.w3.org/2001/XMLSchema">
2    <xs:element name="patient" type="patientType"/>
3    <xs:complexType name="medicalAidType">
4      <xs:sequence>
5        <xs:element type="xs:string" name="provider"/>
6        <xs:element type="xs:string" name="membershipNumber"/>
7        <xs:element type="xs:int" name="plan"/>
8        <xs:element type="xs:byte" name="dependentCode"/>
```

```
 9        </xs:sequence>
10      </xs:complexType>
11      <xs:complexType name="nameType">
12        <xs:sequence>
13          <xs:element type="xs:string" name="title"/>
14          <xs:element type="xs:string" name="firstName"/>
15          <xs:element type="xs:string" name="lastName"/>
16        </xs:sequence>
17      </xs:complexType>
18      <xs:complexType name="identifierType">
19        <xs:simpleContent>
20          <xs:extension base="xs:string">
21            <xs:attribute type="xs:string" name="type"/>
22          </xs:extension>
23        </xs:simpleContent>
24      </xs:complexType>
25      <xs:complexType name="assessmentType">
26        <xs:simpleContent>
27          <xs:extension base="xs:string">
28            <xs:attribute type="xs:string" name="type" use="optional"/>
29          </xs:extension>
30        </xs:simpleContent>
31      </xs:complexType>
32      <xs:complexType name="assessmentsType">
33        <xs:sequence>
34          <xs:element type="assessmentType" name="assessment" maxOccurs="unbounded"
                minOccurs="0"/>
35        </xs:sequence>
36      </xs:complexType>
37      <xs:complexType name="triageType">
38        <xs:sequence>
39          <xs:element type="xs:string" name="contamination"/>
40          <xs:element type="assessmentsType" name="assessments"/>
41        </xs:sequence>
42        <xs:attribute type="xs:string" name="color"/>
43      </xs:complexType>
44      <xs:complexType name="periodType">
45        <xs:sequence>
46          <xs:element type="xs:date" name="startDate"/>
47          <xs:element type="xs:date" name="endDate"/>
48        </xs:sequence>
49      </xs:complexType>
50      <xs:complexType name="doctorType">
51        <xs:sequence>
52          <xs:element type="nameType" name="name"/>
53          <xs:element type="periodType" name="period"/>
54        </xs:sequence>
55      </xs:complexType>
56      <xs:complexType name="doctorsType">
57        <xs:sequence>
58          <xs:element type="doctorType"/>
59        </xs:sequence>
60      </xs:complexType>
61      <xs:complexType name="medicationType">
62        <xs:sequence>
63          <xs:element type="xs:string" name="name"/>
64          <xs:element type="xs:string" name="concentration"/>
65          <xs:element type="xs:string" name="dosage"/>
66          <xs:element type="xs:string" name="period"/>
67        </xs:sequence>
68      </xs:complexType>
69      <xs:complexType name="medicationsType">
70        <xs:sequence>
```

```
71          <xs:element type="medicationType" name="medication"/>
72        </xs:sequence>
73      </xs:complexType>
74      <xs:complexType name="scriptType">
75        <xs:sequence>
76          <xs:element type="xs:date" name="date"/>
77          <xs:element type="medicationsType" name="medications"/>
78        </xs:sequence>
79      </xs:complexType>
80      <xs:complexType name="scriptsType">
81        <xs:sequence>
82          <xs:element type="scriptType" name="script"/>
83        </xs:sequence>
84      </xs:complexType>
85      <xs:complexType name="patientType">
86        <xs:sequence>
87          <xs:element type="medicalAidType" name="medicalAid"/>
88          <xs:element type="nameType" name="name"/>
89          <xs:element type="xs:string" name="dateOfBirth"/>
90          <xs:element type="identifierType" name="identifier"/>
91          <xs:element type="triageType" name="triage"/>
92          <xs:element type="doctorsType" name="doctors"/>
93          <xs:element type="scriptsType" name="scripts"/>
94        </xs:sequence>
95      </xs:complexType>
96    </xs:schema>
```

(a) Use a diagram to model the information given in the XML schema document, in a hierarchical tree structure. It is not necessary to add all the elements associated with each of the complex types. Adding details to the level of the complex types (not beyond) and relationships between the types will suffice. (8)

(b) If a patient is admitted to the hospital multiple times, how will the information be represented in a hierarchical tree structure? (2)

(c) The University of Pretoria wants to integrate their system with the National Health System (NHS) system, however there is a conflict with the `<doctor>` tag. **List** and provide an **example** of the XML concept that can be used to ensure that the two systems can communicate and exchange information without conflict. (2)

(d) Study the extract given below of a patient's records. Provide an improvement for the bloodPressure assessment tag in terms of a partial XSD schema definition. (3)

```
<triage color="green">
    .....
    <assessments>
        <assessment type="bloodPressure">120mmHg/80mmHg</assessment>
        <assessment type="heartRate">66bpm</assessment>
        <assessment type="tempreture">32.5 deg. Celsius</assessment>
        <assessment type="note">This is a note</assessment>
    </assessments>
  </triage>
```

(e) In the provided schema, you note that the types for the elements inside the `medicationType` element type definition are incorrect. **Provide** and **explain** the **most appropriate** types to be used for the elements whose datatypes are incorrect. (4)

(f) Explain whether the current schema allows for multiple scripts to be listed in the `<patient>` tag. (2)

(g) Provide the missing constraints for the `<doctors>` tag in the above XML schema. (2)

4. (a) Consider the universal relation R given below.

    R (A, B, C, D, E, F, G, H, I, J)

    Suppose that the following functional dependencies hold on R:
    FD1: {A, B} → {C}
    FD2: {A} → {D, E}
    FD3: {B} → {F}
    FD4: {F} → {G, H}

FD5: $\{D\} \rightarrow \{I, J\}$

Answer the questions which follow.

    i. Determine which attributes form the primary key of R. (2)

    ii. Is the following statement True or False? Provide a reason for your answer. (2)
      R is in 1NF

    iii. Consider the following resultant relations after some normalisation has taken place. (1)

```
R2 (B, F, G, H)
R3 (A, D, E, I, J)
```

      What is the highest normal form that these relations are in?

    iv. Take the following relation to the next normal form. (3)

```
R3 (A, D, E, I, J)
```

(b) Given the following relation R and an extension for the relation. (2)

R

A	**B**	**C**
ID_1	a	m
ID_1	b	n
ID_2	e	m
ID_2	f	n

What is the next normal form to which the relation needs to be taken? Explain your answer.

5. Consider the following relation and functional dependencies. (20)

Patient(ID, Name(Title, FirstName, LastName), MedicalAid(Number, Scheme, Plan, DependentCode), DateOfBirth, {Doctor(HPCSA, DTitle, DFirstName, DLastName, StartDate, EndDate)})

```
ID → {Title, FirstName, LastName}
ID → Number
ID → DependentCode
ID → DateOfBirth
Number → {Scheme, Plan, DependentCode}
HPCSA → {DTitle, DFirstName, DLastName}
{ID, HPCSA} → {StartDate, EndDate}
```

Normalise the relation to third normal form (3NF). Make sure you show all the steps and your reasoning for each of the steps. Make sure that all your final relations are clearly marked.

6. (a) Consider each of the figures below. (6)

Figure 1   Figure 2   Figure 3

Figure 4   Figure 5   Figure 6
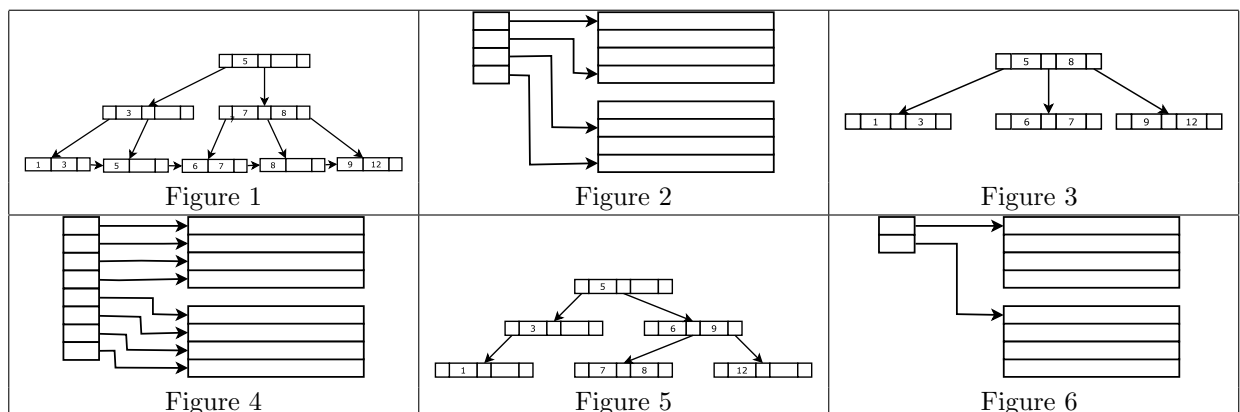
Which description best matches each of the figures. Answer the question by writing the Figure number followed by a right arrow ($\rightarrow$) and the description letter. For example, **Figure 1 $\rightarrow$ (a)**. Note, this may not be the correct answer.

Descriptions:

(a) Dense index

(b) Sparse index

(c) Cluster index

(d) Search tree

(e) B-Tree

(f) B^+-Tree

(b) A disk has been formatted with a block size of 8192 bytes and records are saved as spanned on the disk. (3) If the block is packed full of records, 239 bytes of space remain. For the given records, the disk has a blocking factor of 33. Calculate the size of each record R.

(c) In terms of RAID, specifically when one considers the definition of Redundant Array of Independent Disks. (3) Argue whether the disks can be considered independent or not and what this means for data integrity for a RAID 0 array.