2023 Final Exam Questions for isys2120. The exam duration was 2 hours; there were 17 questions, worth a total of 100 points. The questions were grouped in 5 sections.

The questions in section A and B are all based on the following relational logical schema about inventory.

```
CREATE TABLE Part (PId INTEGER,
                   PDesc VARCHAR(30),
                   PPrice INTEGER NOT NULL,
              PRIMARY KEY (PId));
CREATE TABLE Customer (CId INTEGER,
                       CName VARCHAR(20),
                       CAddr VARCHAR(40),
                       CLoyaltyDiscount INTEGER,
               PRIMARY KEY (CId));
CREATE TABLE Shipment (SId INTEGER,
                       SDate DATE,
                       CId INTEGER NOT NULL,
              PRIMARY KEY (SId),
              FOREIGN KEY (CId) REFERENCES Customer(Cid));
CREATE TABLE Contains (PId INTEGER,
                       SId INTEGER,
                       SPQty INTEGER NOT NULL,
              PRIMARY KEY (PId, SId),
              FOREIGN KEY (PId) REFERENCES Part(PId),
              FOREIGN KEY (SId) REFERENCES Shipment(SId));
```

In **Part** a row (31, ‘Large Widget’, 3050) represents that the part whose PId is 31, is described as “Large Widget” and the price is $30.50 (that is, PPrice is measured in cents). In **Customer**, a row (27, 'Jay Lui', '20 Pine St, Sydney', 25) represents that the customer whose CId is 27 is named “Jay Lui”, with address “20 Pine St, Sydney”, and gets 25% loyalty discount on every purchase. In **Shipment**, a row (110, '2022-01-20', 27) represents that the shipment whose SId is 110 was shipped on 20 January 2022 for customer with CId 27. In **Contains**, a row (31, 110, 3) represents that the shipment with SId 110 contains 3 of the part whose PId is 31.

Section A [SQL queries, 25 points in total]

Q1. [4 points] Write a SQL query to run against the inventory schema on page 1, that will output a list of the SId of every shipment which contains 2 or more copies of a part whose price is greater than 1000 cents.

Q2. [4 points] Write a SQL query to run against the inventory schema on page 1, that will produce a report showing, for each customer, the number of shipments for that customer which were shipped on '2022-01-20'.

Q3. [5 points] Write a SQL query to run against the inventory schema on page 1, that will find the amount charged (in cents) for shipment with SId which is 110. The total charge is calculated by taking the price of each part in the shipment, multiplied by the number of copies of that part in the shipment, adding all these multiplications, and then deducting from this the loyalty bonus appropriate for the customer. For example, if the shipment contains 3 copies of part 31 (each with price 3050 cents), and 2 copies of part 48 whose price is 115 cents, and the loyalty discount for this customer is 25%, then the cost of the shipment is (3*3050+2*115)*(100-25)/100, that is 7035 cents.

Q4. [6 points] Write a SQL query to run against the inventory schema on page 1, that will consider the customers whose CAddress contains the substring “Sydney”, and among those it should find which has the greatest loyalty discount. You should output the customer’s Cid and CName. If more than one of the customers with appropriate CAddress have the same greatest loyalty discount, your query should list all of the considered customers who have that discount, in alphabetical order by CName.

Q5. [6 points] Write a SQL query to run against the inventory schema on page 1, that will find the PId and PName of any part which costs more than 1000 cents and is present in every one of the shipments of customer whose Cid is 27.

Section B [Relational concepts, 25 points in total]

Q6. [5 points] In the inventory schema shown on page 1, the table **Contains** has a primary key which is (PId, SId). Describe in English, what restrictions this puts on the contents of the table.

Q7. [5 points] Write a relational algebra expression involving the tables in the inventory schema shown on page 1, whose result will answer the request "Find the SId of every shipment which contains 2 or more copies of a part whose price is greater than 1000 cents".

Q8. [7 points] Consider the SQL query below, that refers to the inventory schema shown on page 1.

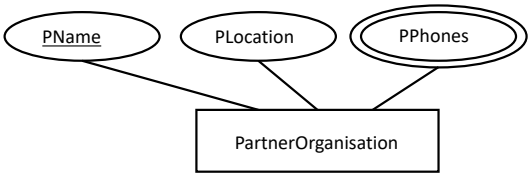
```
SELECT SId

FROM Shipment

WHERE CId < 50 AND SDate = '2022-01-20'
```

Write SQL to create an index that will allow much faster execution for the query above, when the database has a large amount of data in the table. Also, **explain how** the index would allow the query to be calculated.

Q9. [8 points] Produce an ER diagram, for a conceptual model from which one could produce the relational inventory schema shown on page 1.

Section C [Conceptual model, 10 points in total]

Q10. [10 points] In an information system that is managing data about an accommodation service for students, there are two kinds of customer. Every customer has an identifier, and a name, and a bank account name. Some of the customers are sponsored by a partner organisation (such as their university); for them there is also information about which partner has sponsored them, what discount they should receive, and an upper limit on the travel time to the partners location. If a customer is not sponsored, the system should keep an upper limit on the rental charge the customer is willing to pay. **Draw an ER diagram** of a conceptual model for the information described here (we already include an Entity type for PartnerOrganisation), and also, **give CREATE TABLE statements** for a relational design that represents the information from your conceptual design as well as possible.



The questions in section D are all based on the following relational design, which collects data about the usage of drugs at several hospitals. There is a single table:

DrugUsage(HospitalName, HospitalID, DiseaseCode, DiseaseName, DrugName, SizeofDose, Cost, ManufacturerName, ManufacturerAddress).

In **DrugUsag**e, a tuple ('RNSH', 25, 319, 'Malaria', 'Quinine', 100, 2500, 'BigPharma', '100 Pharma Place, Geneva, Switzerland') represents that at hospital "RNSH", whose ID is 25, for patients who have the disease "Malaria" whose code is 319, these patients can be treated with drug "Quinine" in 100ml dose, manufactured by "Big Pharma" whose address is 100 Pharma Place, Geneva, Switzerland, and the cost of this treatment will be $25.00 (ie 2500 cents).

The following functional dependencies are valid in this schema:

- **HospitalID → HospitalName**
- **HospitalName → HospitalID**
- **DiseaseCode → DiseaseName**
- **DrugName → DiseaseCode, ManufacturerName**
- **DrugName, SizeofDose → Cost**
- **ManufacturerName → ManufacturerAddress**

Section D [Relational design theory, 25 points in total]

Q11. [4 points] Based on the **DrugUsage** schema and functional dependencies given on page 5, **explain in English the meaning** of the functional dependency DrugName, SizeofDose → Cost. Also, **give an example of data** that would not be allowed in the table **DrugUsage** because of this dependency.

Q12. [4 points] Based on the **DrugUsage** schema and functional dependencies given on page 5, **c**alculate the attribute closure (HospitalName, DrugName, SizeofDose)+ . Show the step-by-step working of the calculation. Write your answer in the box below.

Q13. [5 points] Based on the **DrugUsage** schema and functional dependencies given on page 5, **t**he relation **DrugUsage** is not in BCNF. Justify this statement.

Q14. [12 points] Based on the **DrugUsage** schema and functional dependencies given on page 5,

- give a lossless-join decomposition of **DrugUsage**, into TWO relations;
- for each of the two decomposed relations, state the functional dependencies that hold for that relation, and state a primary key for that relation, and indicate whether or not that relation is in BCNF.
- justify that your decomposition has the lossless-join property,
- indicate whether or not your decomposition has the dependency-preserving property, and justify your decision.

Section E [Database-backed applications, 15 points in total]

Q15. [5 points] In the code you used as a skeleton for asst3 (and which had previously been used in lab for week 8), there are several places where the end-user enters some value (such as a name), and this value is then used in querying the database. Describe the aspects of the application code, which aim to prevent a SQL injection attack from damaging the data.

Q16. [5 points] In the code you used as a skeleton for asst3 (and which had previously been used in lab for week 8), the data about end-users (such as their name and password for accessing the web application) is stored in the PostgreSQL database. **Describe one security mechanism** that is used in this code, to try to protect the confidentiality of the information about end-users. **Discuss one attack** that could be made to breach the security you described.

Q17. [5 points] In the code you used as a skeleton for asst3 (and which had previously been used in lab for week 8), describe the steps by which information returned from a database query is displayed to the end-user. Mention explicitly the files and functions in the code, where each step is done.