

1. (10%) To manage and query data, we can either implement a file-based system supported by an ordinary programming language such as C++, or apply specialized database software, such as Oracle. Please list at least two significant differences between these two approaches (or the two types of systems).
2. (10%) Please give two types of integrity constraints defined in the SQL standard. For each type of constraints, you need to give a sample expression, and explain what constraints it imposes on data.

Answer questions 3-4 below based on the following three relational schemas:

employee (ENO, NAME, ADDRESS) *works* (ENO, PNO, HOUR)
project (PNO, LEADER)

The relation **employee** has three attributes: ENO for the employee identifier, NAME for the name, and ADDRESS for the address. The relation **project** has the attribute PNO for the project number and LEADER is the name of the person who leads the project. The relation **works** describes the project (PNO) which an employee (ENO) works for, and the working hours for that project (HOUR).

3. (5%) Give the relational algebra expression for the following question:
Find the names and addresses of those employees whose leader is "John".
4. (20%) Write SQL expressions for the following questions, respectively.
 - (a) Find the ENO of all employees who work for the project with PNO = "P1" or PNO = "P2".
 - (b) Find the employees (ENO) whose average working hours are more than the average working hour of the project with PNO = "P1".
 - (c) Create a view called 'LC', which outputs the names of the leaders and the number of employees led by each leader.
 - (d) Delete those projects which the employee named "Mary" works for. (You only need to delete the data in the *project* relation.)
5. (10%) Please explain what the following trigger tries to do. This trigger is defined in the SQL standard.

```
create trigger test-trigger after update on order
referencing new row as nrow
for each row
when nrow.amount < 0
begin atomic
  insert into warnings values (nrow.id, nrow.amount)
end
```

6. (20%) Draw an ER diagram to design the database for an online store, which needs to satisfy all the following descriptions and constraints:
 - We want to record the information about each customer, such as his email, name, address, etc.
 - This store sells books and CDs. For each item, we need to know its name and price.
 - We want to record the order of each customer. Note that each order belongs to only one customer.
 - Using generalization to represent that each order can contain any combinations of books and CDs.
 (For each entity set, design any appropriate attributes and primary keys for them by yourselves. You can also assign attributes to relationship sets)
7. (15%) Consider the relational schema $R = (A, B, C, D, E, F)$, and the following set of functional dependencies $F = \{A \rightarrow BC; E \rightarrow AF\}$.
 - (a) Show that the schema R is not in BCNF (Boyce-Codd Normal Form).
 - (b) Decompose R into a set of BCNF schemas. (Note that the decomposition process should guarantee lossless-join).
8. (10%) Please explain the differences between 3NF (3rd Normal Form) and BCNF.