DASC 5333 Database Systems for Data Science
CSCI 4333 Design of Database Systems
Spring 2023
Section 1 Mid-Term Examination

Last Name: _____ **First Name:** _____ **Student Id:** _____

Number: _____ **Circle One: CSCI 4333 or DASC 5333**

Time allowed: *1 hour 20 minutes*. Total score: 100 points. *Closed* book examination. An information sheet prepared by yourself is allowed.

Answer all questions. Turn in both question and answer sheets (if needed).

(1) [25 points] The goal is to build a highly simplified database to store information about employees working for departments and working on projects. Provide an UML class diagram to capture and model the partial requirements below. You should list class names, attributes with multiplicities, and associations with multiplicities. Roles of associations should also be provided when appropriate. Multiplicities should be as specific as possible. Show the stereotypes <<pk>> and <<unique>> (indicating that the value of the attribute must be unique for each object) when applicable. Since this is only a simplified part of the application, model your design in a flexible way.

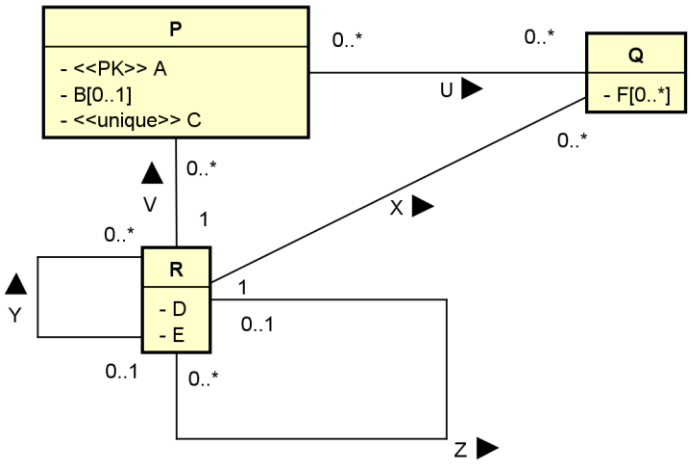
In the company, there is a unique id for an employee. The last names, first names, emails and an optional phone for an employee should be stored. There are departments. A department has a unique id, and it must have a name, an office, an email, and a phone. A department has a head, and possibly many vice-heads and many staff members. They are employees. On the other hand, an employee works for only one department.

There are projects. Each project has a unique id. It has a name and a description. A project has a lead department and many collaborating departments. A project also has a project leader and many team members. A project can be a sub-project of a parent project. A project can also be a top-level project and thus has no parent.

Please answer your question in the next page.

(1) Your answer here:

(2) [15 points] Consider the following data model in the UML class diagram. Attribute multiplicity is included. Construct a reasonable set of relation schema to implement it. For each relation, list its candidate keys, foreign keys, and all attributes you know for sure that are nullable and non-nullable. Ignore data types.



Answer: fill in the table below.

Relation	**Relation**
[CK] [FK] [Nullable] [Non-nullable] [Note]	[CK] [FK] [Nullable] [Non-nullable] [Note]
Relation	**Relation**
[CK] [FK] [Nullable] [Non-nullable] [Note]	[CK] [FK] [Nullable] [Non-nullable] [Note]
Relation	**Relation**
[CK] [FK] [Nullable] [Non-nullable] [Note]	[CK] [FK] [Nullable] [Non-nullable] [Note]

(3) [24 points] True or False. Circle the choice or write 'T' or 'F' clearly.

(a) [T or F] The Excel spreadsheet uses a graph data model.

(b) [T or F] The query design wizard in Microsoft's Access allows users to specify conditions to be satisfied by the result rows in a query.

(c) [T or F] A Relation R(A,B,C,D,E) may have all five attributes as prime attributes.

(d) [T or F] A SQL SELECT statement may have no WHERE clause.

(e) [T or F] Relational Calculus is a procedural language.

(f) [T or F] A foreign key of a relation must be declared as "NOT NULL" in SQL.

(g) [T or F] One advantage of DBMS as compared to a file system is its better concurrent access.

(h) [T or F] In the relational model, a relation schema is a set of tuples.

(i) [T or F] If A is a candidate key of R(A,B,C,D), {A,B} is not a candidate key.

(j) [T or F] 'CREATE TABLE' is an example of a Data Manipulation Language (DML) command in SQL

(k) [T or F] An operation in Relational Algebra returns a relation.

(l) [T or F] Stereotypes can be added to a class in UML.

(4) Short question [4 points + 4 bonus points]

(a) [4 points] The relation R(A,B,C,D,E) has three prime attributes. How many candidate keys can there be? Why? (A prime attribute is an attribute that appears in one or more candidate keys.)

(b) [4 points: bonus] It is known that the relation R has n attributes and two candidate keys. What are the minimum and maximum numbers of superkeys R may have? Show your reasoning.

Question 5 uses the following the toyu database, which is provided separately.

(5) [32 points] Write the *SQL* queries for the following data problems. Result orders are unimportant unless explicitly stated otherwise.

(a) Show all CSCI classes, their students, and grade in the follow manner.

```
+-----------+---------+------------+-------+
| course    | classId | student    | grade |
+-----------+---------+------------+-------+
| CSCI 2315 |   10000 | Tony Hawk  | A     |
| CSCI 2315 |   10000 | Mary Hawk  | NULL  |
| CSCI 2315 |   10000 | David Hawk | B-    |
| CSCI 4333 |   10001 | Tony Hawk  | A     |
| CSCI 4333 |   10001 | Mary Hawk  | A-    |
| CSCI 4333 |   11001 | Tony Hawk  | D     |
| CSCI 5333 |   10002 | Tony Hawk  | B+    |
| CSCI 5333 |   10002 | David Hawk | B+    |
+-----------+---------+------------+-------+
8 rows in set (0.001 sec)
```

(b) Show the classId of all classes with some CSCI students and some CINF students enrolled.

```
+---------+
| classId |
+---------+
|   10003 |
|   10004 |
+---------+
2 rows in set (0.001 sec)
```

(c) Show the ids and names of students who have not enrolled in a CSCI course in the following manner.

stuId	student
100003	Catherine Lim
100004	Larry Johnson
100005	Linda Johnson
100006	Lillian Johnson
100007	Ben Zico
100008	Bill Ching
100009	Linda King
100111	Cathy Johanson

8 rows in set (0.001 sec)

(d) Show the accumulative numbers of students enrolled in all courses in the following manners.

course	title	accumulative number of students
ACCT 3333	Managerial Accounting	0
ARTS 3311	Hindu Arts	1
CINF 3321	Introduction to Information Systems	4
CINF 4320	Web Application Development	4
CSCI 2315	Data Structures	3
CSCI 3352	Advanced Data Structures	0
CSCI 4333	Design of Database Systems	3
CSCI 5333	DBMS	2
ENGL 1311	English II	0
ENGL 1410	English I	2
ITEC 3312	Introduction to Scripting	1
ITEC 3335	Database Development	2

12 rows in set (0.001 sec)