Database Management Systems (COP 5725)

Spring 2017

Instructor:
Dr. Markus Schneider

TAs: Keke Zhai, Man Mohan Devineni

Exam 1 Solutions

Name:	
UFID:	
Email Address:	

Pledge (Must be signed according to UF Honor Code)

On my honor, I have neither given nor received unauthorized aid in doing this assignment.

_____
Signature

For scoring use only:

	Maximum	Received
Question 1	20	
Question 2	25	
Question 3	30	
Question 4	25	
Total	100	

Question 1 (Knowledge Questions) [20 points]

Explain the following notions as precisely as possible in your own words.

a) What is the difference between *physical data independence* and *logical data independence*? [4 points]

Physical data independence: changes of the physical schema do not have impact on the conceptual schema (and thus also not to external schemas);
Logical data independence: changes of the conceptual schema do not have impact on external schemas.

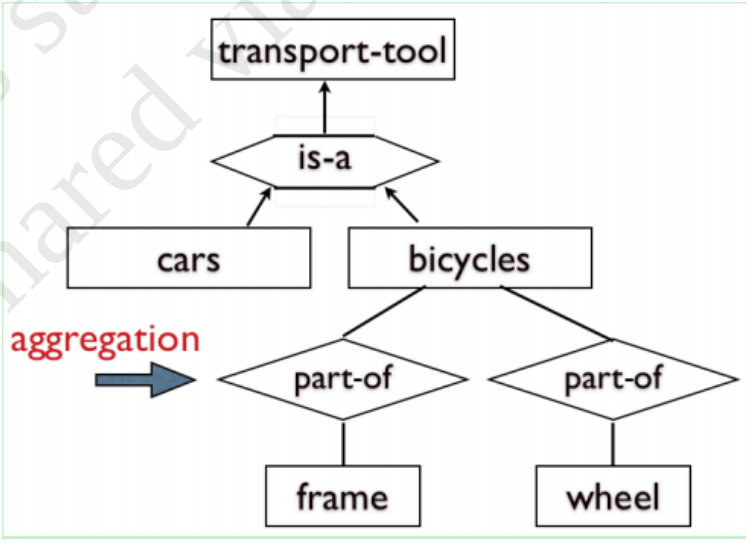
b) Define formally what a *key* is. What is a *candidate key*? What is a *primary key*? [4 points]

Key – Given $R(A_1, A_2, ..., A_n)$ and let $X \subseteq \{A_1, A_2, ..., A_n\}$, X is called key, if the following conditions are fulfilled:

- uniqueness: for all relation instances r_R of R holds:
$\forall t_1, t_2 \in r_R : t_1[X] = t_2[X] \Rightarrow t_1 = t_2$ . . . . .
- minimality: there is no $Y \subset X$, so that the uniqueness is fulfilled.

Candidate key – Same as "key".
Primary key – One of possible keys is selected as the primary key

c) Present a simple example of an ER diagram that includes both an aggregation abstraction and a generalization abstraction. Please point out which part utilizes the aggregation abstraction? [4 points]



d) List all basic operations of the Relational Algebra with their name and their correct symbolic notation? [4 points]

The basic operations are: $\cup$ (union), - (difference), $\times$ (Cartesian product), π (projection), and σ (selection). Further, we have the ρ (rename) operation.

e) What is the difference between an *inner join* and an *outer join*? [4 points]

Inner join: the result does not contain those tuples that did not find a partner;
Outer joins: the result contains those tuples that did not find a partner. The result tuples are filled with null values

Question 2 (ER Model) [25 points]

Based on the following requirements, design an ER diagram for a university database used to keep track of students' transcripts. For each entity set, mark clearly the primary key (If the primary key is not specified by requirement, use your best knowledge to add your own key or use existing attributes). For each relationship you have identified, state its cardinalities in (min, max)-notation.

- The university keeps track of each student's name (including first name, middle name, last name), student number, social security number (ssn), current address (including city, state, and zip) and phone, birthday, sex, class (freshman, sophomore, ..., graduate), major department, minor department (if any), and degree program (B.A., B.S., ..., Ph.D.). Both social security number and student number have unique values for each student.

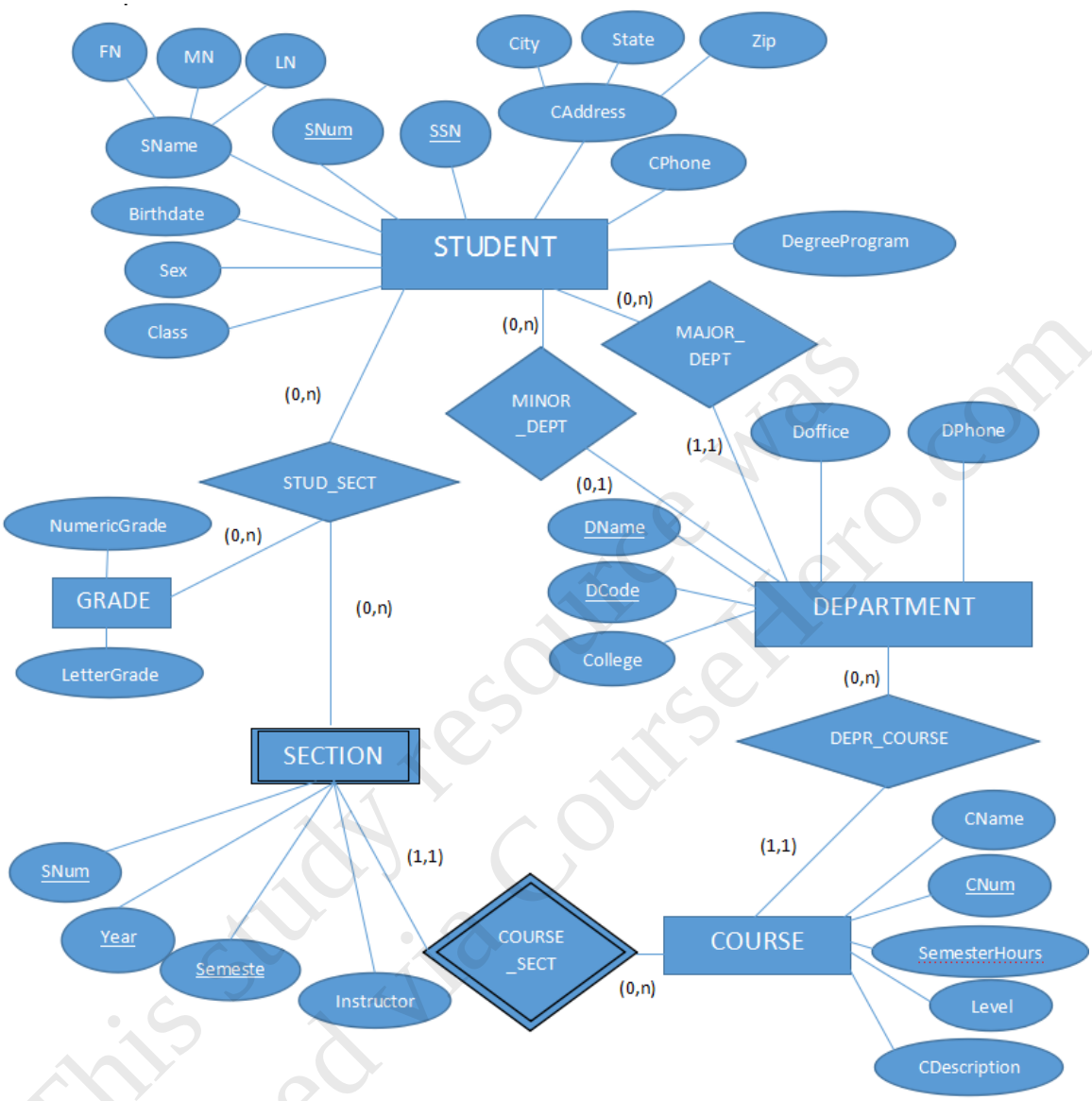
- Each department is described by a name, department code, office number, office phone, and college. Both name and code have unique values for each department.

- Each course has a course name, description, course number, number of semester hours, level, and offering department. The value of course number is unique for each course.

- Each section has an instructor, semester, year, course, and section number. The section number distinguishes different sections of the same course that are taught during the same semester/year; its values are 1, 2, 3, ...; up to the number of sections taught during each semester.

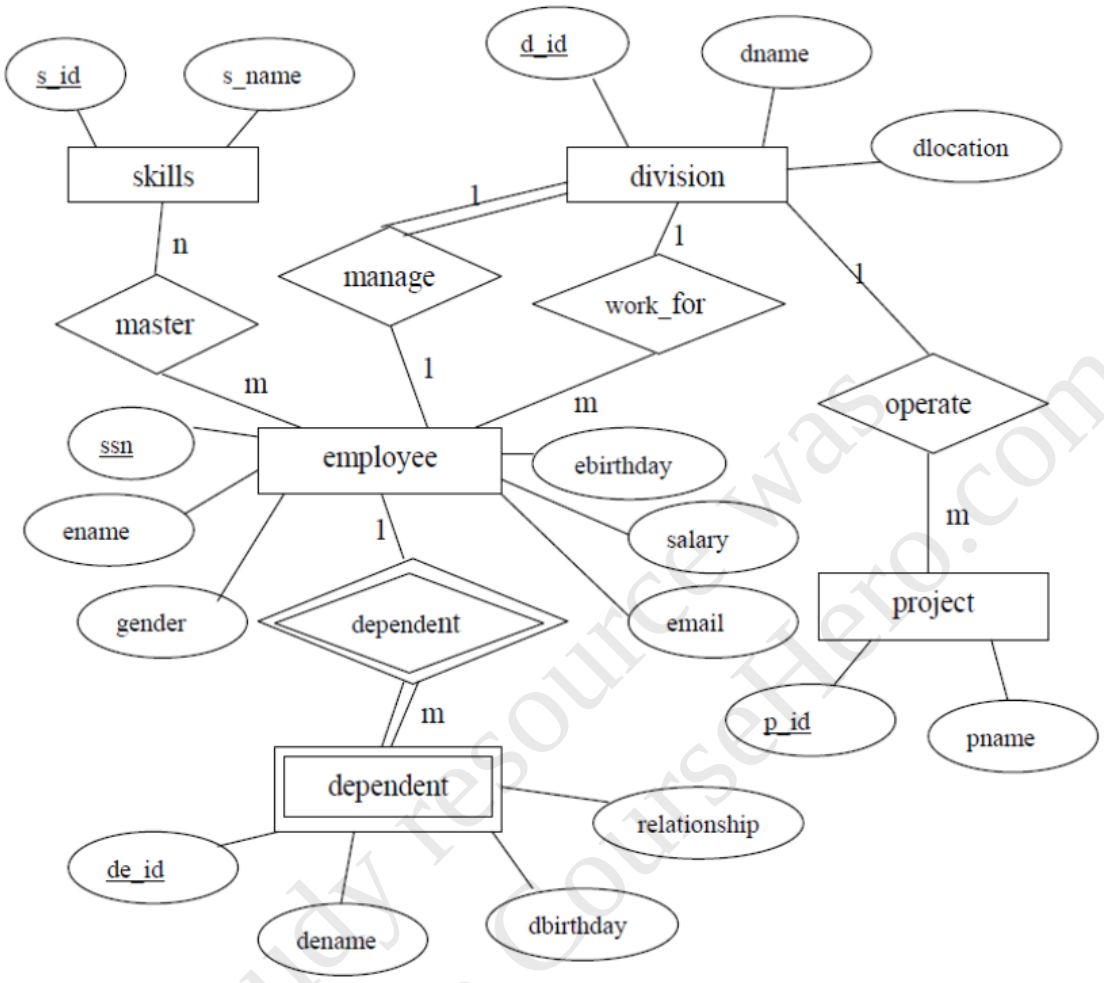
.

- A grade report has a student, section, letter grade, and numeric grade (0, 1, 2, 3, 4 for F, D, C, B, A, respectively).



Question 3 (ER Model) [30 points]

Consider the following ER model:



a) Transform the ER model into a corresponding database schema. Identify the primary keys and foreign keys. You need not include data types in the schema [18 points]

[Grading criteria: 3 points for each schema. incorrect key: -1. Incorrect attribute: -2]

employee (ssn, ename, gender, ebirthday, salary, email, employed_by)
employed_by is the foreign key from table 'division'

project (p_id, pname, made_by)
made_by is the foreign key from table 'division'

skills (s_id, s_name)

division (d_id, dname, dlocation, e_id)
e_id is the foreign key from table 'employee'

dependent(de_id, e_id, dename, dbirthday, relationship)
e_id is the foreign key from table 'employee'

master (skill_id, e_id)
skill_id is the foreign key from table 'skills'
e_id is the foreign key from table 'employee'

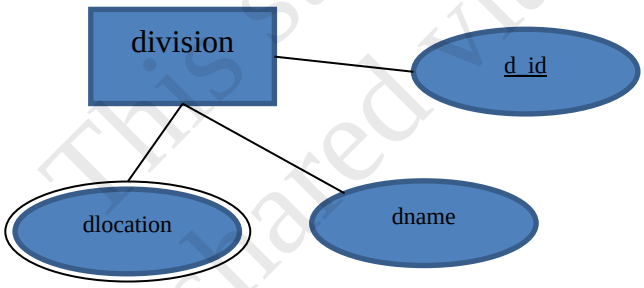
b) Using SQL, create the table Skills. Your answer should be recognized by oracle [4 points].

```
CREATE TABLE SKILLS(
    Sid INTEGER PRIMARY KEY,
    Sname VARCHAR(100)
)
```

c) Is there any weak entity set? If so, give its name(s), partial key and identifying relationship. [4 points]

Weak entity set: dependent
Partial key: de_id
Identifying relationship: employee

d) Suppose the entity set "division" has multiple locations, transform the corresponding ER model below to a relational schema. (you only need to consider "division") [4 points]



division (d_id, dname)
div-location(location, d_id)

Question 4 (Relational Algebra) [25 points]

Consider the following database schema:

Car_Owner(<u>name</u>, <u>date_of_birth</u>, <u>phone_number</u>, address, city, country, registration_date)
Car(<u>Chassis_identifier</u>, model, manufacture_year, color, MSRP, current_value, condition)
Insurance(<u>name, date_of_birth, phone_number, chassis_identifier, insurance_type</u>, date, expiration_date, cost, limit, description)

Write relational algebra expressions that will answer the following questions:

(1) List the car owners′ names and addresses who have had insurance with a cost of less than 500 dollars and expiration date after 2017-06-21. [4 points]

$$\pi_{name,address}(\sigma_{expiration\_date>'2017-06-21' \wedge cost<500}(Insurance) \rhd\lhd Car\_Owner)$$

(2) List the names of the car owners who have a blue car in good condition and whose age is greater than 50 (A person who was born in 2007 is 10 years now). [4 points]

$$\pi_{name}((\sigma_{color='blue' \wedge condition='good'}(Car) \rhd\lhd Insurance) \rhd\lhd \sigma_{date\_of\_birth<1967}(Car\_Owner))$$

(3) List the names of car owners who have more than one car and where the color of one of their cars is red. [4 points]

$$\rho_R(\sigma_{I1.name=I2.name \wedge I1.date\_of\_birth=I2.date\_of\_birth \wedge I1.p\_number=I2.p\_number \wedge I1.c\_identifier \neq I2.c\_identifier}(\rho_{I1}(Insurance) \times \rho_{I2}(Insurance)))$$

$$\pi_{name}(\sigma_{color='red'}(R \rhd\lhd Car))$$

(4) List the chassis identifiers of those cars whose current value is greater than 6000 dollars, the insurance type is ‘full’, and the cost of insurance is between 700 and 1000 Dollars? [4 points]

$$\pi_{chassis\_identifier}(\sigma_{current\_value>6000 \wedge insurance\_type='full' \wedge cost \leq 1000 \wedge cost \geq 700}(Car \rhd\lhd Insurance))$$

(5) List the chassis identifiers of those cars whose model is ‘Ford’ and has the most expensive insurance cost. [5 points]

$$\rho_{c1}(\sigma_{\text{model}='Ford'}(Car \triangleright\triangleleft Insurance))$$
$$\rho_{c2}(c1)$$
$$\pi_{chassis\_identifier}(c1) - \pi_{c2.chassis\_identifier}(\sigma_{c1.\text{cost}>c2.\text{cost}}(c1 \times c2))$$

(6) List the car owners who registered before 2010 and don’t have a car manufactured after 2000. [4 points]

$$\rho_{c1}(\sigma_{registration\_date<2010}(Car\_Owner))$$
$$\pi_{name}(c1) - \pi_{name}(\sigma_{manufacture\_year>2000}(c1 \triangleright\triangleleft Insurance \triangleright\triangleleft Car))$$