2019 Database System Principles
Test One

Class_____No_____Name_____

1. (22 points) Fill in blanks

（1）Among the following statements, the correct one/ones is/are B,C/BC E .

A. As a type of open-source database systems, SQL Server is developed and distributed by IBM.

B. MySQL and PostgreSQL are two typical open-source database systems.

C. The relational model is applicable to managing structured data such as the table data, while XML provides a way to represent semi-structured data, e.g. the data with nested structures.

D. A on-line shopping site **has a three-tier Browser-Server(B/S)-like architecture. Its application programs are programmed** in Java, and access Oracle database server via the ODBC interface.

E. *Big Data* is now a fashionable term, and the relational databases are able to efficiently manage various types of big data in the forms of tables, texts, web pages, voices, images and videos.

（2）The data model defines the specification of managing data items in database. It is a collection of conceptual tools for describing data structure, data relationships, data semantics, data operations and consistency constraints.

(3) Database design involves the following phases: requirements analysis, conceptual schema design, logical design and physical design.

(4) Here is an example of logical data Independence. A transaction T accesses a relational table *Student*(*SID*, S*Name*, *Age*, *Department*) by the SQL query:

update ***Student***
set *age* = *age* +1
where department='CS'

The table ***Student*** is then renamed as ***StudentNew*** and its schema is changed to ***StudentNew*** (*s#, name, age, college, sex*), to enable T to remain unchanged and need not be rewritten, a view *Student* is created as,

create view (*SID, SName, Age, Department*) as
select *s#, sname, age, college*
from ***StudentNew***

, and T access this view to accomplish its query.

(5)The collection of information stored in the database at a particular moment is called an _instance_ of the database.

(6) As human-machine interfaces, the database language consists of two parts, i.e the data definition language (DDL) and DML (data manipulation language) .

(7) DBMS can be divided into two main parts, that is query processor
and transaction manager .

(8) (2 points) For the entity set Student(#student, sname, department, course, grade), the primary key is C , the primary attributes are D .

A. #student B.{#student} C. {#student, course} D. #student, course E. #student F. course

(9) The six fundamental operations in the relational algebra are select, project , union , set difference, Cartesian-Product, and rename.

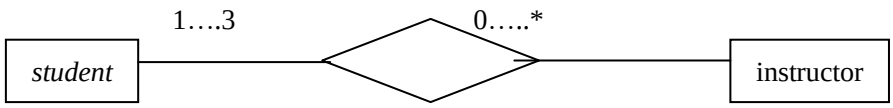
(10) A key/ superkey/primary key/ candidate key is a set of one or more attributes that, taken collectively, can be used to identify uniquely a tuple in the relation.

(11) An entity set that does not have a primary key is referred to as a weak entity set .

(12) If X is one or more attributes in relation R1, and X is also the primary-key of another relation schema R2, X is called a foreign key from R1 referencing R2.

(13) (2 points)For the entity sets *student and instructor* and the relationship set *advisor* among them in the following figure, the mapping cardinality from *student* to *instructor* is many-to-many , and the participation constraints of *instructor* in *advisor* is **partial** .

,



(14) (2 points) Convert the entity set **"学生"**, in which the attribute "老乡" is a multivalued attribute, into two relational **tables**

student-id	籍贯	老乡	性别	年龄
07494	北京	07596, 07611	男	20

答案：

student-id	籍贯	性别	年龄
07494	北京	男	20

注意：主键下的的下划线，如缺少下划线，扣 0.5 分

student-id	老乡
07494	07596
07494	07611

(15) There are three types of pure query languages related to the relational model, that is, relational algebra , tuple relational calculus, and domain relational calculus.

(16) (2 points) The relational algebra expression corresponding to the following SQL statement is：

$loan \rightarrow loan - \sigma_{(amount>=0 \text{ and } amount<=50)}(loan)$.

```
delete
from student
where studentID between 2000200 and 2000500
```

(17) (2 points) The SQL statements corresponding to the following algebra expression is

```
insert into loan
        select * from loan where amount between 0 and 50
```

.

$loan \leftarrow loan \cup \sigma_{amount \geq 0 \text{ and } amount \leq 50}(loan)$

假设 *loan(loan-number, branch, amount)*

2. (30 points) Consider the following relations containing airline flight information:

Flights(flightno, from, to, **distance**, departs_time, arrives_time)
Aircraft(aircraft_id, aname, **cruisingrange**)
Pilots(Pilot_id, name, salary)
Certified(Pilot_id, aircraft_id)

/* distance:指飞行距离，cruisingrange:指飞行最大距离

Every pilot is certified for some aircrafts.

For the following queries, give **relational algebra expressions** for (1), **SQL statements** for (2)~(4):

(1) Find the highest salary of pilots. (5 points)

$$\prod_{salary}(\text{Pilots}) - \prod_{salary}(\sigma_{Pilots.salary < d.salary}(Pilots \times \rho_{d}(Pilots)))$$

差操作 1 分，其余操作及条件各 0.5 分

或者：用扩展关系代数中的聚集函数 max

$$G_{max}(salary)$$

(2) Find the Pilot_ids and names of the pilots who **are certified for more than three aircrafts** but are **not certified on any Boeing aircraft**. (8 points)

```
Select Pilot_id, name
From pilots                                    (2 分)
Where pilot_id in                              （下面两个查询条件的判断各占 3 分）
    (select pilot_id
    from (
          select Pilot_id, count(aircraft_id)   /*或：  其它聚集函数
           from certified
          group by Pilot_id
          having count(aircraft_id)>=3)
          )
And pilot_id not in
    (select Pilot_id
      From certified, aircraft
      Where certified.aircraft_id= aircraft.aircraft_id
          And aname='boeing')
```

(3) Find the flightno of the flight that can be piloted by every pilot whose salary is more than $400,000. (6 points)

```
Select flightno
From flights, aircraft , pilots, certified        (2 分)
Where salary>=400000 and certified.aircraft_id= aircraft.aircraft_id
        And certified. Pilot_id= pilots. Pilot_id
        And cruisingrange>distance  （4 个条件每个 1 分）
```

(4) Create the table *Flights*, in which {*flightno*} is the primary key, and {*from*, *to}* are not permitted to be null; It is also required that the distance is not below 0. (5 points)

create table *Flights*

(flightno char(20) primary key,
From varch(20) not null,
to varch(20) not null,
distance integer,
departs_time time,
arrives_time time,
check (distance>0)
)

数据类型取合适的就行。
表定义语法 1 分，属性及定义 1 分，约束定义 1 分。

(5) 针对 Pilots(pilot_id, name, salary)，使用 SQL 语句，判断 name 是否为表 Pilots 的 super_key.
要求：如果 name 不是 super_key，找出表中导致 name 不是 super_key 的元组
(6 points)

3. (25 points) Convert the following E-R diagram to the relation schemas and identify the primary key of each relation by underlining the primary attributes.

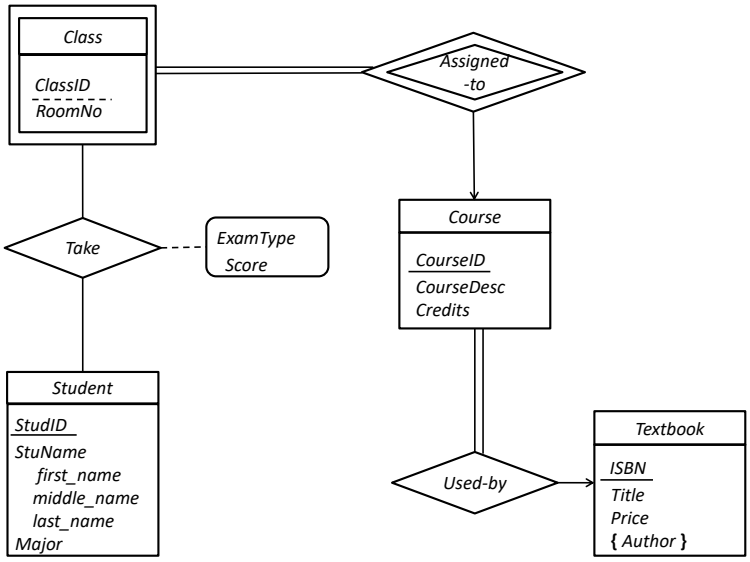


Figure 1 E-R diagram

Student(<u>*StudID*</u>, *first_name*, *middle_name*, *last_name*, *Major*)

(4 分，没有正确标注主键，扣 1 分)

Class(<u>*ClassID*</u>, <u>*CourseID*</u>, *RoomNo*)

(4 分，没有正确标注主键，扣 1 分)

Take(<u>*StudID*</u>, <u>*ClassID*</u>, <u>*CourseID*</u>, *ExamType*, *Score*)

(4 分，没有正确标注主键，扣 1 分)

Course(<u>*CourseID*</u>, <u>*ISBN*</u>, *CourseDesc*, *Credits*)

(4 分，没有正确归并联系 *Used-by*，扣 1 分)

Textbook(*ISBN*, *Title*, *Price*)　(3 分)

TextbookAuthor(*ISBN*, *Author*) (3 分)

4. (24 points) Consider the following information on the production procedure in facilities.

(1) A *production facility* (制造工厂) is uniquely identified by a *FacilityID* and described by *FcltyName* and *FcltyLocation*.

(2) Every *production line* (生产线) is identified by a *PIID* and described by *LineName* and *LineType*.

(3) Each *line activity* (生产线活动) is identified by *ActivityID*. It also has descriptive attributes *ActvtyDesc* and *ActvtyType*.

(4) A *product* (产品) is distinguished by its *ProdID*, and has attributes *ProdName* and *ModelID*.

(5) An *assembly part* (零部件) in a product is recognized by its *PartID* and has attributes *Partname* and *InspectionDate*. A part may be inspected several times and thus have several inspection dates.

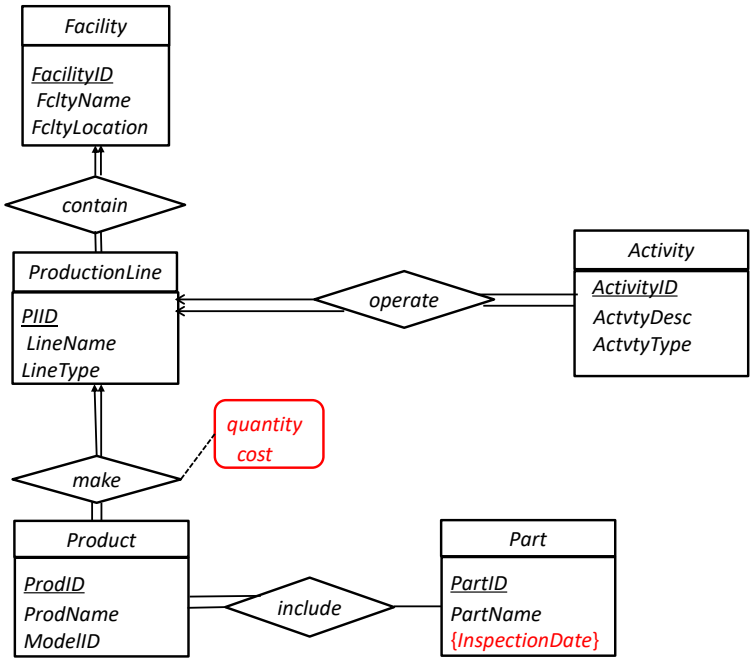
(6) Every production facility *contains* several production lines, and each line must belong to a unique facility.

(7) Each production line *operates* more than one line activity, and an activity has to be attached to a unique line.

(8) A production line can *make* several products, but a product can be produced by only a production line. The quantity of the product made by the line and the production cost must be recorded.

(9) A product *includes* one or more assembly parts, and a part can also be used for several products.

Construct an E-R diagram to depict the above mentioned data items and the associations among them.



Part 实体 4 分，没有标注多值属性{*InspectionDate*}，扣 1 分。
其余 4 个 entity 各 3 分，共 12 分；

4 个 relationship 各 2 分，共 8 分。
make 联系没有标注属性扣 1 分；
联系中映射基、完全/部分参与有误，酌情扣分。