

# 北京大学信息科学技术学院考试试卷

考试科目： 数据库概论 姓名： \_\_\_\_\_ 学号： \_\_\_\_\_

考试时间： 2017 年 4 月 25 日 任课教师： 陈立军

题号	一	二	三	四	五	六	七	八	总分
分数									
阅卷									

## 北京大学考场纪律

1、考生进入考场后，按照监考老师安排隔位就座，将学生证放在桌面上。无学生证者不能参加考试；迟到超过 15 分钟不得入场。在考试开始 30 分钟后方可交卷出场。

2、除必要的文具和主考教师允许的工具书、参考书、计算器以外，其它所有物品（包括空白纸张、手机、或有存储、编程、查询功能的电子用品等）不得带入座位，已经带入考场的必须放在监考人员指定的位置。

3、考试使用的试题、答卷、草稿纸由监考人员统一发放，考试结束时收回，一律不准带出考场。若有试题印制问题请向监考教师提出，不得向其他考生询问。提前答完试卷，应举手示意请监考人员收卷后方可离开；交卷后不得在考场内逗留或在附近高声交谈。未交卷擅自离开考场，不得重新进入考场答卷。考试结束时间到，考生立即停止答卷，在座位上等待监考人员收卷清点后，方可离场。

4、考生要严格遵守考场规则，在规定时间内独立完成答卷。不准交头接耳，不准偷看、夹带、抄袭或者有意让他人抄袭答题内容，不准接传答案或者试卷等。凡有违纪作弊者，一经发现，当场取消其考试资格，并根据《北京大学本科考试工作与学术规范条例》及相关规定严肃处理。

5、考生须确认自己填写的个人信息真实、准确，并承担信息填写错误带来的一切责任与后果。

学校倡议所有考生以北京大学学生的荣誉与诚信答卷，共同维护北京大学的学术声誉。

以下为试题和答题纸，共 8 页。

得分

一、简答题（每小题 3 分，共 30 分）

1. DBMS 是通过哪些机制来保证数据独立性的？
2. 已知关系  $R(A, B)$ ，在  $A$  上建有索引  $idxA$ 。现在执行查询 `select B from R where A = 10`，发现基于  $idxA$  比表扫描的执行方式要慢。请解释原因，并给出可以加快这个查询的做法。
3. 请给出一个 ER 模型中采用特化设计的例子，并说明其优点。

4. 有两个表 R(A) 和 S(B), R(A)={1, 2}, S(B)={3, 4, 5, 6}。定义如下触发器:

```
create trigger R_IST before insert on R
reference old table as OT
for each statement
update R set A = A + (select count(*) from OT)
```

当执行完 insert into R (select \* from S)之后, 写出 R 的结果。

5. Codd 提出的全关系系统的十二条准则中, 有一条是关于数据完整性的独立性, 你如何理解这一点? 试列举几个相关的 SQL 关键字。

6. DBMS 提供的数据库控制功能包括哪些?

7. 什么是联系的势? 请列举一个例子。

8. 什么时候会引入弱实体? 请列举一个弱实体的例子。

9. 两个表  $R(A, B)$ ,  $S(A, C)$ , 其中  $A$  是这两个表的主码, 请问下面哪些查询中的 `distinct` 可以去掉?

查询一:

查询二:

查询三:

***select distinct***  $R.A, S.A$   
***from***  $R, S$   
***where***  $R.B = S.C$

***select distinct***  $R.A$   
***from***  $R, S$   
***where***  $R.B = S.C$

***select distinct***  $R.A$   
***from***  $R, S$   
***where***  $R.B = S.A$

10. 如果表的主码过于冗长, 会有哪些害处?

得分

二、关系代数 (20 分)

1. 如何检测包  $R(A, B)$  中是否有重复的行? 。
2. 试用基本的关系代数操作表示出  $R(A, B)$  和  $S(B, C)$  之间的左外连接。
3. 已知关系  $R(A)$ , 请给出计算  $A$  最小值的关系代数表达式。

4. 已知关系 friends (ME, YOU)，表示 ME 和 YOU 两人相互认识，可以假定 ME 和 YOU 都是序列号。试用关系代数给出 friends 去重后的结果，也即如果表中存在两行 (a, b) 和 (b, a)，只保留其中一行。

得分

三、SQL (20 分)

已知有如下关系表（标下划线的列是主码）

员工(员工号，员工姓名，工资，所属部门号)

部门(部门号，部门名，部门经理)

项目(项目号，项目名，所属部门号，资金额，负责人)

参加(员工号，项目号，工时)

试用 SQL 语言完成下列功能：

1. 给出比其部门经理工资高的员工姓名

2. 给出每个部门中参加项目的总工时最多的员工姓名

3. 给出参加了所有属于 **d1** 部门项目的员工姓名

4. 根据每个员工参加某项目的工时占该项目总工时的百分比来分配其资金额，作为员工的劳动报酬。请计算每个员工总的劳动报酬（可以使用临时表或表变量）。

5. 对于一串不连续的数字，请给出其最小的缺失值。比如对于如下一串数字，1, 2, 3, 5, 6, 8, 9，最小的缺失值是 4。使用游标和集合表示方法均可

得分

#### 四、数据库设计（25 分）

1. 在一个数据库中，相关表的建表语句如下：

```
create table T1( a1 int primary key,
                a2 int,
                a3 int foreign key reference T2(a3))

create table T2( a3 int primary key,
                a4 int,
                a5 int foreign key reference T2(a3))

create table T3( a6 int primary key,
                a7 int)

create table T4( a1 int,
                a3 int,
                a6 int,
                a8 int,
                primary key (a1, a3, a6),
                a1 foreign key reference T1(a1),
                a3 foreign key reference T2(a3),
                a6 foreign key reference T2(a6))
```

试画出相应的 E-R 图，使得可以从该 E-R 图推导出上述表定义，其中实体和联系的名称可以自定。



2. 假定我们要建立一个关于篮球职业联盟的数据库，存储如下信息：

- 每个球队有球队名称、所在城市；
- 每位球员有球员姓名、薪酬；
- 每位球员属于一个球队，每个球队拥有多位球员；
- 各个球队之间会进行主客场比赛，需要记录比赛时间、比赛结果（假定两只球队之间只轮流进行一次主客场比赛）；
- 每位球员参加多场比赛，每场比赛有多位球员参加，同时球员参加每场比赛会有相应的数据统计，包括得分、篮板等。

其中带下划线的属性是实体的主码。

(1) 请根据以上描述，试画出相应的 ER 图，不用标注属性。

(2) 将上面的 ER 图转换为相应的关系模式，并用下划线标出其主码，用波浪线标出外码。

(3) 更进一步，如果两只球队之间会进行多轮的主客场比赛，此时需要为比赛定义唯一的比赛编号，请画出这种场景下的 ER 图。