浙江大学 2020－2021 学年 春夏 学期

《数据库系统》课程期末考试试卷参考答案和评分标准

课程号： ___21121350___ ， 开课学院：_计算机学院_____

考试试卷：√A 卷、B 卷（请在选定项上打√）

考试形式：√闭、开卷（请在选定项上打√），允许带一张 A4 纸笔记入场

考试日期： _2021_ 年 7 月 2 日,考试时间： ___120___ 分钟

诚信考试，沉着应考，杜绝违纪。

考生姓名：_____ 学号：_____ 所属院系：_____

题序	一	二	三	四	五	六	七	八	总 分
得分									
评卷人									

Problem 1: Relational Model and SQL (18 points)

Following are the relational schemas of a SRTP (Student Research Training Program) project database.

    student (sId, sName, dId)
    teacher (tId, tName, dId)
    department (dId, dName)
    project (pId, pName, tId, startTime, endTime)
    participate (pId, sId, role)

The underlined attributes are primary keys, and foreign keys are listed as follows:
    "dId" in "student" references "department";
    "dId" in "teacher" references "department";
    "tId" in "project" references "teacher";
    "pId" and "sId" in "participate" reference "project" and "student", respectively.
In "participate", only two different roles are permitted: "leader" and "member". Based on the above relational schemas, please answer the following questions:

(1) Write a relational algebra expression to find the names of the projects that are instructed by a teacher from the department "Computer Science". (4 points)

(2) Write SQL statements to create tables project and participate with all the necessary

constraints (Note: Tables student, teacher, and department have already been created and can be referenced). (6 points)

(3) Write a SQL statement to find the names of the teachers that instruct at least one project started in the year 2020. (4 points)

(4) Write a SQL statement to find the names of the students participating more than 2 projects. (4 points)

Answers of Problem 1:

(1)
$\prod_{pName}(project \bowtie teacher \bowtie (\sigma_{dName=\text{"Computer Science"}}(department)))$
评分细则：每个操作错扣 **1** 分，全对给 **4** 分

(2)
CREATE TABLE project
　　(pId　char(10),
　　pName　varchar(20),
　　tId　char(10),
　　startTime　date,
　　endTime　date,
　　primary key (pId),
　　foreign key (tId) references teacher);
评分细则：写出 **schema**、**primary**、**foreign key** 给 **2** 分，类型全对给 **3** 分

CREATE TABLE participate
　　(pId　char(10),
　　sId　char(10),
　　role　varchar(20),
　　primary key (pId, sId),
　　foreign key (pId) references project,
　　foreign key (sId) references student,
　　check (role="leader" or role="member"));
评分细则：写出 **primary**、**foreign key**、**check** 给 **2** 分，全对给 **3** 分

(3)
select distinct tName
from project, teacher
where project.tId=teacher.tId and startTime between '2020-01-01' and '2020-12-31'
评分细则：每个 **where** 条件错扣 **1** 分，全对给 **4** 分

(4)
select sName
from student

where sId in
    (select sId
    from participate
    group by sId
    having count(pId) > 2)

评分细则：写出 having…且对给 1 分，内表写对给 2 分，全对给 4 分

Problem 2: E-R Model (11 points)

Based on the SRTP project management scenario in Problem 1, some new requirements are added as follows:

(1) There are two kinds of SRTP projects, i.e., school-level projects and national-level projects, and a project is either school-level or national-level.

(2) National-level projects have budget information, and school-level projects have mid-term check information.

(3) A school-level project is associated with exactly a department that is in charge of the management of the project.

Please draw an E-R diagram for the scenario.
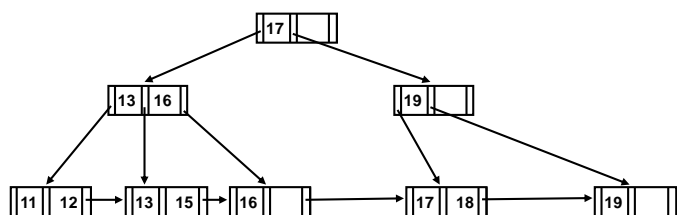
<u>Answers of Problem 2:</u>
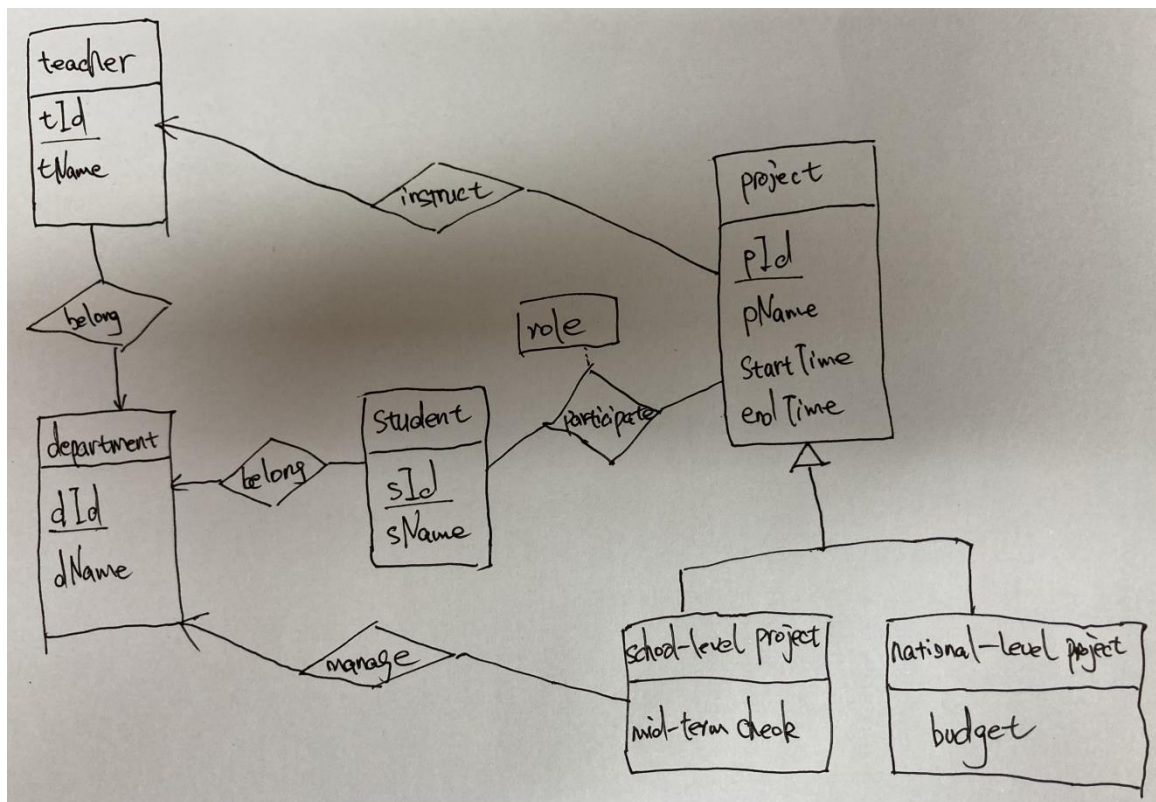
评分细则：联系写成实体扣 1 分，没写联系扣 1 分，属性写错漏写扣 1 分，少写实体或联系扣 1 分，全对给 11 分

Problem 5: B+ -Tree and Query Processing (10 points)

Table student in Problem 1 is stored sequentially on sId. The following B+-tree is built for the table on attribute dId. Please answer the following questions:
(1) Is the built index a primary index? Why? (2 points)
(2) Draw the B+-tree after inserting entry 14. (4 points)
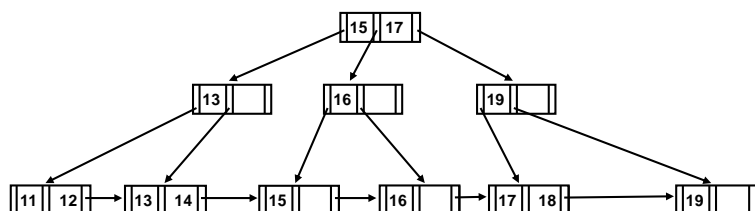(3) Draw the B+-tree after deleting entry 19 from the original B+-tree. (4 points)

Answers of Problem 5:

(1)
The built index is not a primary index, as the search key of the index is not the search key of the sequentially ordered data file.
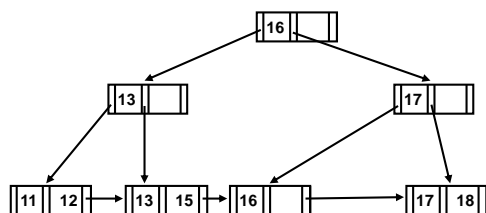评分细则：结论正确给 **1** 分，解释正确给 **1** 分，结论错误不给分

(2)

评分细则：每个 **entry** 插入错扣 **1** 分，图中缺箭头或线扣 **1** 分，全对给 **4** 分

(3)

评分细则：每个**entry**删除错扣**1**分，图中缺箭头或线扣**1**分，全对给**4**分

Problem 6: Query Processing (14 points)

There are two relations r (100 blocks) and s (20 blocks), and hash-join algorithm is used to perform natural join between these two relations (memory size M=6 blocks). Please answer the following questions:

(1) How many partitions can be constructed? Why? (3 points)

(2) Which relation is best to choose as the build relation? Why? (3 points)

(3) Is recursive partition needed? Why? (3 points)

(4) Please compute the cost (numbers of seeks and block transfers) of the hash-join. (5 points)

Answers of Problem 6:

(1)
5 partitions, as the number of partitions is M-1.
评分细则：结论正确给 **1** 分，解释正确给 **2** 分，结论错误酌情给分

(2)
Relation s, as relation s is smaller than relation r.
评分细则：结论正确给 **1** 分，解释正确给 **2** 分，结论错误酌情给分

(3)
Recursive partition is not needed, as the size of the partitions of relation s (i.e., 4) is less than or equal to M-2 (i.e., 4).
评分细则：结论正确给 **1** 分，解释正确给 **2** 分，结论错误酌情给分

(4)
Number of block transfers: $3 \times (100+20) + 4 \times 5$
Note: 4×5 is not necessary, which considers partially filled blocks.
Number of seeks: $2 \times (100+20) + 2 \times 5$
评分细则：**block transfers** 的数量 **3** 分，**seeks** 的数量 **2** 分，全对给 **5** 分

Problem 7: Concurrency Control (13 points)

Given the following schedule, please answer the following questions:

(1) Draw the precedence graph for the schedule. (3 points)

(2) Is the schedule conflict serializable? Why? (2 points)

(3) Is it possible that the schedule is generated by the 2PL protocol with lock conversions? Explain. (5 points)

(4) Which conditions should be satisfied if we want the schedule to be recoverable? (3 points)

T1	T2	T3
	read C	
read B		
	write C	
	read A	
		read C
	write A	
read A		
		write C
write B		
		read B

Answers of Problem 7:

(1)
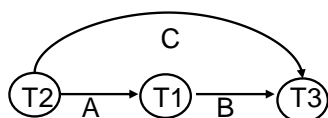
评分细则：图中结点错误或连接错误，每错一处扣 **1** 分，全对给 **3** 分

(2)

The schedule is conflict serializable, as the precedence graph is acyclic.

评分细则：结论正确给 **1** 分，解释正确给 **1** 分，结论错误不给分

(3)

It is possible that the schedule is generated by the 2PL protocol with lock conversions.

T1	T2	T3
	Lock-S (C) read C	
Lock-S (B) read B		
	Upgrade (C) write C	
	Lock-X (A) read A UL (C)	
		Lock-S (C) read C
	write A UL (A)	
Lock-S (A) read A		
		Upgrade (C) write C
Upgrade (B) write B UL (A, B)		
		Lock-S (B) read B UL (B, C)

评分细则：结论正确给 **1** 分，流程图正确给 **4** 分，或者文字解释，表明题意也可给

分，结论错误不给分

(4)
T1 must commit before T3 does.
T2 must commit before T1 does.
T2 must commit before T3 does.
评分细则：每条执行顺序正确给 **1 分**，全对给 **3 分**

Problem 8: Recovery (14 points)

Given the following log file that supports logical undo, please answer the following questions:

(1) The system crashes just after the last log record. What are the values of B and C in the database after system crash? (3 points)
(2) Which transactions should redo and undo, respectively? (3 points)
(3) What are the start and end points for redo and undo, respectively? (3 points)
(4) What are the log records added during recovery? (5 points)

1    $<T_0 \text{ start}>$
2    $<T_0, B, 2000, 2050>$
3    $<T_1 \text{ start}>$
4    $<T_1, B, 2050, 2100>$
5    $<T_1, O_1, \text{operation-begin}>$
6    $<\text{checkpoint } \{T_0, T_1\}>$
7    $<T_1, C, 700, 400>$
8    $<T_0 \text{ commit}>$
9    $<T_1, O_1, \text{operation-end}, (C, +300)>$
10   $<T_2 \text{ start}>$
11   $<T_2, O_2, \text{operation-begin}>$
12   $<T_2, C, 400, 300>$
13   $<T_2, O_2, \text{operation-end}, (C, +100)>$
14   $<T_2, \text{commit}>$

Answers of Problem 8:

(1)
B=2100
C= 300 or 400 or 700
评分细则：**B 值正确给 1 分，C 值漏写扣 1 分，三个值完整写出给 2 分**

(2)
redo: T_0 and T_2 undo: T_1
评分细则：**redo 完整写出给 2 分，漏写扣 1 分，undo 写对给 1 分**

(3)

redo: 7-14   undo: 14-3

评分细则：redo、undo 顺序错写一个扣 1 分，redo 顺序写成 6-14 也可给分，全对给 3 分

(4)

$<T_1, C, 600>$

$<T_1, O_1, operation-abort>$

$<T_1, B, 2050>$

$<T_1, abort>$

评分细则：错一个扣 1 分，写出一个给 1 分，全对给 5 分