北京邮电大学 2020——2021 学年第一学期

《数据库系统原理》期末考试试题（A 卷）

考试注意事项	一、学生参加考试须带学生证或学院证明，未带者不准进入考场。学生必须按照监考教师指定座位就坐。 二、书本、参考资料、书包等物品一律放到考场指定位置。 三、学生不得另行携带、使用稿纸，要遵守《北京邮电大学考场规则》，有考场违纪或作弊行为者，按相应规定严肃处理。 四、学生必须将答题内容做在答题纸上，做在试题及草稿纸上一律无效。

考试课程	数据库系统原理	考试时间	2021 年 2 月 24 日 09:00～ 11:00

题号	一	二	三	四	五	六	七	八	九	十	总分
满分	10	18	10	10	14	6	10	6	6	10	100
得分											
阅卷教师											

1. Choice. (1×10 points)

(1) Application programs are said to exhibit_______independency, if they do not depend on the logical schema, and thus need not be rewritten if the logical schema changes.

A. view   B. logical   D. network   C. physical

(2) With respect to DBS design, a relational table's primary key is defined at the _______Phase.

A. requirement analysis   B. conceptual design

C. logical design   D. physical design

(3) With respect to the relation schema *Branch-schema*(*branch-name, branch-city,* assets) and the relation *branch,* which one is not the metadata stored in system catalog：______

A. names of the relation branch

B. domain and length of attribute assets

C. number of tuples in branch

D. the tuple < Redwood, Brooklyn, 710000>

(4) In SQL language, the statement that can be used for security control is________.

A. insert   B. rollback   C. revoke   D. update

(5) For the entity sets A and B and the relationship set R among them, if the cardinality limit of A is 2….*, and the cardinality limit of B is 0….1, then the mapping cardinality from A to B is _______.

A. one to one   B. many to one   C. many to many   D. one to may

(6) In the_______file organization, records of several different relations can be stored in the same file.

A. sequential B. multitable clustering C. hashing D. heap

(7) The functional dependency set F={ A→D, E→D, D→B, BC→D, CD→A } holds on the relation schema R = (A, B, C, D, E) （ADE）$^+$ = _______

A．{A，D，E }

B．{A，D，B}

C． {A，D，B，E}

D． { A, B, C, D, E }

(8) Which of the following statements is correct:_______

A. the two-phase locking protocol ensures conflict serializability and freedom from deadlock;

B. the two-phase locking protocol ensures conflict serializability，but does not ensure freedom from deadlock;

C. the two-phase locking protocol does not ensure conflict serializability，but ensure freedom from deadlock;

D. the two-phase locking protocol does not ensure conflict serializability and does not ensure freedom from deadlock;

(9) Transactions are required to have the ACID properties._______ ensures that either all operations of the transaction are properly reflected in the database or none are.

A. Atomicity B. Consistency C. Isolation D. Durability

(10)As for the following equivalence rules for transformation of relational expressions, which one is not right：

____

A. $\prod_L(E1 \cup E2) = (\prod_L(E1)) \cup E2$

B. $\sigma_\theta (E1 - E2) = \sigma_\theta (E1) - \sigma_\theta(E2)$

C. $E1 \cap E2 = E2 \cap E1$

D. $\sigma_\theta (E1 \times E2) = E1 \infty_\theta E2$

2. (18 points) Consider the following relations in an enterprise database, where the primary keys are underlined.

Employee(*<u>employeeID</u>*, *employeename*, *age*, *address*, *sex*, *salary*, *deptID*)

Department(*<u>deptID</u>*, *deptname*, *managerID*, *managername*)

DepartLocations(*<u>deptID</u>*, *<u>deptlocation</u>*)

Project (*<u>projectID</u>*, *projectname*, *projectlocation*, *deptID*)

Workson(*<u>employeeID</u>*, *<u>projectID</u>*, *hours*)

Dependent(*<u>employeeID</u>*, *<u>dependentname</u>*, *sex*, *Birthdate*, *Relationship*)

(1) Use a ***relational algebras expression*** to delete from the table *Department* all departments that have some projects at Beijing. (4 points).

(2) For each department which has at least 10 employees, list its *deptID*, and calculate the total number of the employees who work in this department and whose salaries are more than 4000. Give one or more SQL statements to list the query result in **descending order** of the attribute *deptID*. (5 points)

(3) Use SQL statements to 1) add a new attribute *city* into the table *DepartLocations*. It is assumed that the data type of *city* is varchar(50); 2) and then define a nonclustered index on attribute *city*. (4 points)

(4) A new project is started, and its information is as follows: *projectID*=1021, *projectname*='Building', *projectlocation*='Shanghai', *deptID*='201'. Find out from the department '201' all employees who currently do not work for any projects, and assign to them this new project '1021'. It is assumed that these selected employees will work for the new project for 500 hours. (5 points)

Give SQL statements to modify the tables in the database, so as to record all the information about the new project '1021' and the employees working on this project.

3. (10 points) Convert the following E-R diagram to the proper relation schemas and identify the primary key of each relation.

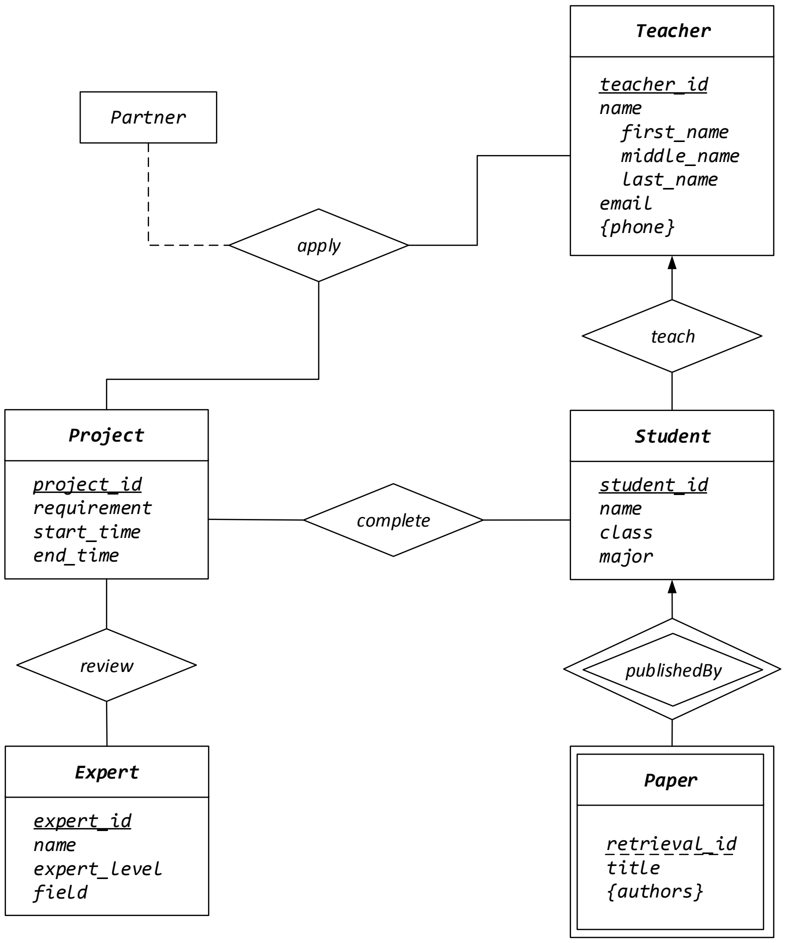


Fig. 1 E-R diagram

4. (10 points) A hospital database needs to store information about doctors, patients, sickroom (病房), and departments (科室). Consider the following information:

- Each doctor has descriptive attributes of identifier number, name, age, and technical title.
- Each patient has descriptive attributes of the number of medical records(病历) , name, age, and sex.
- Each sickroom has descriptive attributes of the number of sickroom, and address.
- Each department has descriptive attributes of name, address, and telephone-number.
- Integrity constraints:

a. Each doctor must belong to one (and only one) department; and for each department, there are more than one doctors belonging to it.

b. Each patient is taken care of by one and only one responsible doctor; a doctor may be responsible for no patients, or only one patients, or more than one patients.

c. Each patient lives in one and only one sickroom; a sickroom may contain more than one patient.

d. Each sickroom can be managed by more than one department; but for some departments, there are no sickrooms managed by them, while for other departments, there are more than one managed sickroom.

Design the E/R diagram for hospital database on basis of the information mentioned above.

Note: mapping cardinality of each relationship and participation of each entity to the relationship should be described in the diagram.

5. (14 points) Consider the following relation schema Employee and the functional dependency set F hold on Employee:

Employee (ID, Name, Department, Dependent_name, Dependent_birthday, Dependent_sex)

F={ ID→Name, ID→Department,

(ID, Dependent_name) →Dependent_birthday, (ID, Dependent_name) →Dependent_sex }

(1) List all the candidate keys of Employee. (2 points)

(2) What is the highest normal form of Employee? Why? (4 points)

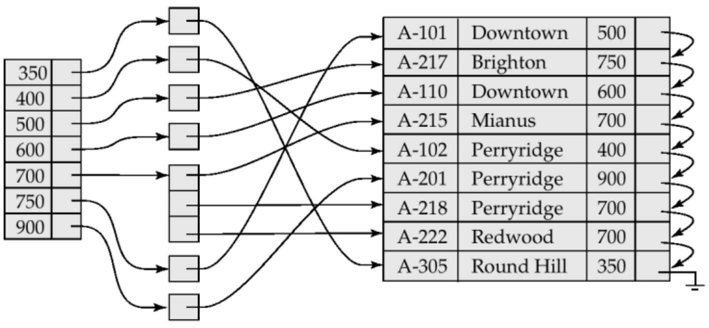
(3) Is the decomposition ρ={R1(ID, Name, Department), R2(Dependent_name, Dependent_birthday, Dependent_sex) } on Employee dependency-preserving? Why? (4 points)

(4) Give a lossless-join and dependency-preserving decomposition of Employee into 3NF. (4 points)

6. (6 points) Given the data file ***account***(*account_number*, *branch_name*, *balance*) and the index defined on *balance*, as shown in the following figure,

(1) Is the index a dense or sparse index, why? (3 points)

(2) Is the index a clustering or non- clustering index, why? (3 points)



7. (10 points) There are three relations in the database, describing the employees and the companies that the employees work at,

Employee(employee-id, name, age, sex, e-city)

Company(company-id, company-name, company-city)

Works(employee-id, company-id, salary)

(1) Give an SQL statement to find out the *name* and *id* of all the *employees*, who are male, work at the company named SHOUGANG, and whose *salaries* are below $2000. (3 points)

(2) Translate the SQL statement in (1) into an initial query tree, and give an optimized query tree for it, by means of heuristic query optimization.(7 points)

8. （6 points）Consider the concurrent transactions T1, T2 , and T3 under the schedule S1：

（1） Is the concurrent schedule S1 a recoverable schedule, and why?（3 points）

（2） Is S1 a Cascading or Cascadeless schedule? and why?(3 points)

S1:

T1	T2	T3

read(X)
X: = X +50
write(X)		

	read(Y)	
commit		

		read(X)
		X:= X-100
		write(X)

	Y:= Y -1
	write(Y)	

	read(X)	
		commit

	X:= X *2
	write(X)
	commit	

9.（6 points）Given a concurrent schedule S2, as shown below,

(1) Construct the precedence graph for it. (3 points)

(2) Is S2 a serializable schedule ? If not, give the reason. If it is, give a serial schedule equivalent to S2. (3 points)

S2:

T_1	T_2	T_3	T_4	T_5
				read(E) E: = E +25

		read(B)
		B: = B+90
		Write(B)		

read(A) A: = A*5				
				write(E)
Write(A)				
	read(A) A: = A+1			
				read(D)
	Write(A)			
		read(A) A: = A/3		
			read(C)	
		Write(A)		
			C: = C+100	
				D: =D*5 write(D)
			write(C)	
read(D) D: =D+2				
			read(B) B: = B *100	
write(D)				
			Write(B)	

	read(W)
	W: =W+30
	Write(W)			

10. (10 points) Considering the concurrent schedule S3 for transactions T_1, T_2, T_3 and T_4, and the data items A, B, C, D modified by these transactions. It is assumed that the initial values of these data items are A=4, B=0, C=30, D=10. When a failure occurs, the log-based recovery scheme consults the log to determine the recovery operations (i.e. redo, undo, ignore) done on T_1, T_2, T_3 and T_4.

(1) Write the log file records for schedule S. (4 points)

(2) After recovery operations on T_1, T_2, T_3 and T_4 are completed, what are the values of the data items A, B, C, D in the database? (4 points)

(3) Write down the redo, undo, ignore list of transactions respectively. (2 points)

S3：

T_1	T_2	T_3	T_4	recovery
begin transaction				

read(A)
A:=A+6
write(A)				

	begin transaction
	read(D)			

commit				

	D:=D+10
	write(D)			

		begin transaction
		read(B)
		B:=B+5
		write(B)		

				checkpoint

			begin transaction
			read(C)	

	read(A)
	A:=A+10
	write(A)			

	commit			

			C:=C+10
			write(C)	

		read(A)
		A:=A+30
		write(A)		

		commit		

			read(D)
			D:=D-10
			write(D)	

****Failure****				