北京邮电大学 2008——2009 学年第 2 学期

《数据库系统原理》期末考试试题（A）

考试课程				考试时间			2008 年 6 月 16 日					
题号	一	二	三	四	五	六	七	八	九	十	十一	十二	总分
满分	9	10	6	13	8	10	10	12	4	6	6	6	
得分													
阅卷教师													

1. Fill in blanks. (1×9 points)

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

(2) The database system provides users with three levels of data abstraction, the _____ level of abstraction describes only part of the entire database.

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

(4) Let $r_1(R_1)$ and $r_2(R_2)$ be relations with primary keys K_1 and K_2 respectively, the subset α of R_2 is called the _____ referencing K_1 in relation r_1, if for every t_2 in r_2 there must be a tuple t_1 in r_1 such that $t_1[K_1] = t_2[\alpha]$

(5) The _____ takes a query-evaluation plan, executes that plan, and returns the answers to the query.

(6) Query optimization is conducted by two strategies, _____ and heuristic optimization.

(7) when applying "set union, set intersection and set difference" on two relations, the _____ of these two relations must equal, the corresponding attribute value must come from the same domain.

(8) The concurrent control component of a database system supports the atomicity and

_____ of the transaction

(9) The several concurrent control strategies are_____ protocols, timestamp-based protocols, validation-based protocols and multiversion schemes.

2. Choice (1×10 points)

(1)A _____ is a language for specifying the database schema as well as other properties of the data.

A DML   B DDL   C DCL   D DSL

(2)If an entity in the higher-level entity set belongs to no more than one lower-level entity set within a single generalization, then the generation is _____.

A   total generalization   B   partial   generalization
C   disjoint   D   overlapping

(3) Consider the relation R(A, B, C, D ,E , P) and the functional dependencies set F={ A→B, C→P,   E→ A, CE→D }, we decompose R into {R1(A,B,E), R2(C,D,E,P)}, which of the following statement is right? _____

A   It is a lossy and dependency preservation decomposition
B   It is a lossless and dependency preservation decomposition
C   It is a lossy and not dependency preservation decomposition
D   It is a lossless and not dependency preservation decomposition

(4)Consider the following rules about the functional dependencies:
i) If AB→C, then A→C
ii). If A→C, then AB→C
iii). If A→B, then B→C
iv). IF A → B, BC → D , then AC → D
Which are right? _____

A   i) and iv)   B ii) and iii)   C   i) and iii)   D   ii) and iv)

(5) Given a table *Employees* and some SQL queries on it, which queries are right? _____
   Employees(employee-id, company-id, employee-city, age, salary)

   i)   create table *Employees*
             (*employee-id*   char(20),
               company-id   char(20),
               employee-city char(20),
               age   integer,
               salary   integer,

primary key (*employee-id*), primary key (*company-id*))

ii)   select *employee-id*

         from   *Employees*

         where *employee-id*=10522 AND company-id = 'ADB'

         order by *employee-id* desc

iii)   select company-id, sum(salary), employee-city

         from *Employees*

         group by company-id

         having avg(salary)>2000

iv)   alter table *Employees* drop *age*

    A. i) ,   ii)   B. ii),   iii)

    C. ii),   iv)   D. i),   iv)

(6). When using natural join for relation R and S, we require R and S contain one or more

common _____

    A. tuple   B. row   C. record   D. attribute

(7) Which of the following does not belong to the basic steps for query processing ? _____

    A. parsing and translation   B. evaluation   C. sorting   D. optimization

(8) There are relations R (A, B, C) and S (B, C, D, E), which is not hold for the following?

_____

    A. $\Pi_A(R) \Join \Pi_D(S)$ B. $R \cup S$ C. $\Pi_B(R) \cap \Pi_B(S)$ D. $R \Join S$

(9) Considering the following statements,

i) A recoverable schedule is a schedule, where for each pair of T_i and T_j , if T_j reads a data
   items previously written by T_i , the *commit* operation of T_j appears before the *commit*
   operation of T_i.

ii) Every cascadeless schedule is also recoverable, but a recoverable may not be a cascadeless
   schedule.

iii) In the strict 2PL protocol, all *X*-locks and *S*-locks taken by a transaction are held until that
   transaction commits.

iv) With respect to the deferred database modification scheme, before T_i executes write(*X*), a
   record <T_i , *X, V2*> is written into the log, *V2* is the value to be written to *X*.

  the correct descriptions are _____ :

    A. i) , ii)   B. ii),   iii)

    C. ii),   iv)   D. i),   iv)

(10) With respect to three types of failures in DBS, if the system enters the deadlock state, then the _____ occurs.

A. transaction failure   B. logical error   C. system crash   D. disk failure

3（6 points）There is a database SPJ，including S，P，J，SPJ four relational schemes：
S（SNO，SNAME，STATUS，CITY）
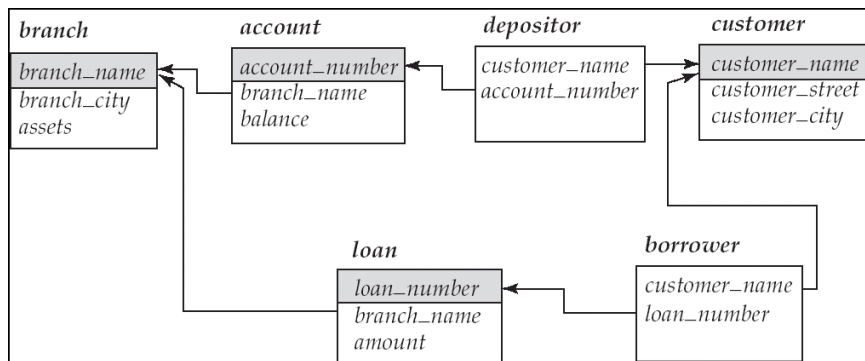P（PNO，PNAME，COLOR，WEIGHT）
J（JNO，JNAME，CITY）
SPJ（SNO，PNO，JNO，QTY）
where：supplier table S includes supplier number SNO, supplier name SNAME，supplier status STATUS, supplier city CITY；part table P includes part number PNO, part name PNAME, color COLOR, weight WEIGHT；project table J includes project number JNO, project name JNAME, project city CITY；supply condition table SPJ includes supplier number SNO, part number PNO, project number JNO, supply quantity QTY，which indicates the quantity of the part that a supplier supplies to a certain project is QTY. Please use relational algebra to write the following queries.
(1) find supplier number SNO which supply part with color 红色 for project J1.   (2 point)
(2) find the project number JNO which does not use part with color 红色 and supplier city is 天津.   (2 point)
(3) find the project number JNO that at least uses the whole parts that supplier number S1 supplies.   (2 point)

4. **(13 points).** Here is the schema diagram for the bank-customer database.

branch
branch_name
branch_city
assets

account
account_number
branch_name
balance

depositor
customer_name
account_number

customer
customer_name
customer_street
customer_city

loan
loan_number
branch_name
amount

borrower
customer_name
loan_number

Use SQL statements to implement the following operations:

(1) Define the table *account*, it is assumed that the *null* value is inappropriate for the attribute *branch_name* and the attribute *balance* ranges from 5000 to 100,000.　　　　(3 points)

(2) Find out the *name* of each *customer* who has *accounts* in only one *branch*.　　(5 points)

(3) Modify the structure of the table *branch*, add a new attribute *account_amount* into it, which describes the total number of *accounts* in a *branch*. And then assign the correct value of *account_amount* for each *branch* in the table *branch*.　　　　　　(5 points)

5. (8 Points)A school is going to arrange a sports day for the students. A database to keep track of participants and activities during the sports day is to be created. Consider the following information:

(1). Participating persons. Each person has a number and a name. The person is identified by the number.

(2). Teams. Each team has a number and a name. The team is identified by the number.

(3). A team consists of several persons, and a person can be a member of several teams.A team must have at least one person, and a person may not participate in any team.

(4). Activities, such as baseball or long jump. Each activity has a number, a name, and a starting time, and is identified by the number.

(5). Participation in the activities. Both teams and individual persons can participate in more than one activities. Each person and each team must participate in at least one activity. Each activity can have many teams and many persons.

Draw the E/R diagram for the database based on the information mentioned above.(8 Points)

6 (10points)Consider the relation schema about the sales information :

Sales(Sales-TransactionNo, Item-No, Item-Price, Item-QuantitySold, Seller, SellerDistrict)
The functional dependencies on Sales are shown below:

> Sales-TransactionNo, Item-No → Item-QuantitySold
>
> Item-No → Item-Price
>
> Sales-TransactionNo →Seller
>
> Seller→SellerDistrict

（1）Compute (Sales-TransactionNo, Item-Price)$^+$(2points)

（2）What is the candidate key of this relation? (2points)

（3）What is the highest normal form of R? (2points)

(4)Give the lossless-join, dependency-preserving decomposition of this schema into 3NF. (4points)

7.(10 points)Consider the following E/R diagram. Create the relational schema that captures this E/R diagram. For every relation in your schema, specify the primary key of that relation.
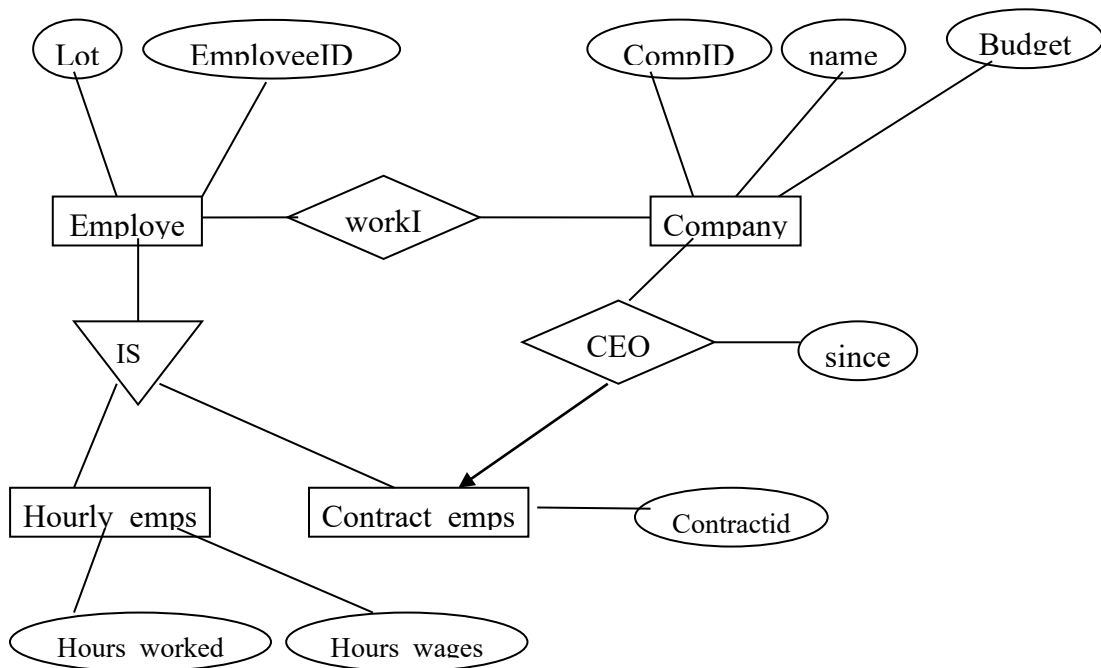
8.(12 points) There are three relations in the database, describing the teachers and the universities that the teachers work at,

    teacher(teacher-id, t-name, age, sex, t-telephone, t-city)
    university(university-id, u-name, u-telephone, u-city)
    works(teacher-id, university-id, salary)
(1) Give an SQL statement to find out the *name* and id of all the teachers, who are female, work at BUPT university, who are older than 40 and whose *salaries* are more than 6000 RMB.   (3 points)

(2) Translate this SQL statement into an initial query tree, and give an optimized query tree for it by means of heuristic query optimization. (9 points)

9. (4 points) Given the relation student(Sno, Sdept, Sname, Ssocre) as shown below, which is

organized as a sequential file.

Taking the attribute Sscore as the search key, define a *dense* and secondary index for the indexed file *student*. The index file and index entries in the index file should be drawn. (4 points)

Sno	Sdept	Sname	Ssocre
S1	Automation Dep.	Zhang ming	88
S2	Automation Dep.	Wang gang	90
S3	Computer Dep.	Yu feng	60
S4	Computer Dep.	Li yong	67
S5	Economy Dep.	Wang jin	98
S6	Economy Dep.	Xin xin	67
S7	Finance Dep.	Li ya	67
S8	Finance Dep.	Li dong	82
S9	History Dep.	Wang ying	70

10. (6 points). Given a concurrent schedule S_1, as shown below,

(1) Construct the precedence graph for it.

(2) Is S_1 a serializable schedule ? and why ?

S_1

T1	T2	T3	T4	T5
				read(Z)
read(Y)				
	write(X)			
	read(Y)			
read(Z)				
		read(Z)		
			write(Z)	
				write(X)
		read(Z)		

		write(Y)		

11. (6 points) Considering the concurrent schedule S on the transaction set {T1, T2, T3} that is

under the timestamp protocol, it is assumed that the timestamps of T1, T2 and T3 are 1, 2, and 3 respectively, the initial values of R-timestamp(X), W-timestamp(X), R-timestamp(Y), W-timestamp(Y), R-timestamp(Z), W-timestamp(Z) are all 0,

(1) What are the values of R-timestamp(X) and W-timestamp(Y) when T1 commits?

(2 points)

(2) What are the values of R-timestamp(Y) and W-timestamp(Z) when T2 commits?

(2 points)

(3) What are the values of R-timestamp(X) and W-timestamp(Z) when T3 commits?

(2 points)

S

T1	T2	T3
write(Y)		
	write(X)	
		read(Y)
write(Z)		
	read(Y)	
		write(Z)
commit		
	commit	
		write(X)
		commit

12. (6 points) Considering the concurrent transactions T_1, T_2, T_3, T_4 and T_5, and the data items A, B, C and D modified by these transactions. It is assumed that

- the initial values of these data items are A=100, B=0, C=200, D=300.
- immediate database modification and checkpoint techniques are employed

 With respect to the log in the following figure that describes the concurrent executing of T_1, T_2, T_3, T_4 and T_5, 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, T_4 and T_5.

 After recovery operations on T_1, T_2, T_3, T_4 and T_5 are completed,

(1) what are the values of the data items **A**, **B**, **C** and **D** in the database?   (4 points)

(2) which transaction successfully updates the data item **A** and determines the final value of **A**.

$<T_1$ start$>$

$<T_1$, A, 100, 300$>$

$<T_3$ start$>$

$<T_3$, B, 0, 30$>$

$<T_2$ start$>$

$<T_2$, A, 300, 500$>$

$<T_3$, C, 200, 400$>$

$<T_4$ start$>$

$<T_1$, D, 300, 100$>$

$<T_4$, D, 100, 200$>$

$<T_1$ commit$>$

$<$checkpoint $\{T_2, T_3, T_4\}>$

$<T_5$ start$>$

$<T_4$, B, 30, 50$>$

$<T_3$ commit$>$

$<T_5$, C, 400, 600$>$

$<T_2$ commit$>$

$<T_5$, A, 500, 800$>$

System crash