**诚信应考,考试作弊将带来严重后果！**

姓名 学号 学院 专业 座位号



( 密 封 线 内 不 答 题 )

……………………………………………………密………………………………………………封………………………………………线……………………………………线………………………………………

**华南理工大学期末考试**

**《数据库系统》试卷A**

**注意事项：1. 考前请将密封线内填写清楚；**

**2. 所有答案请直接答在试卷上；**

**3．考试形式： 闭 卷；**

**4. 本试卷共 大题，满分100分， 考试时间120分钟**。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **题 号** | **Part I** | **Part II** | | | | | **总分** |
| **1** | **2** | **3** | **4** | **5** |
| **得 分** |  |  |  |  |  |  |  |
| **评卷人** |  |  |  |  |  |  |  |

**Part I [20 pts.] (1pt each) Fill in the blanks with the best answer**

1. The collection of information stored in the database at a particular moment is called an instance of the database. The overall design of the database is called the database schema.
2. A relation schema R is in third/3 normal form if for all α → *β* in F+,at least one of the following holds: α → *β* is trivial; α is a superkey for R;Each attribute A in α → *β* is contained in a candidate key for R
3. Let R be a relation schema, R1 and R2 form a decomposition of R. Decomposition is a lossless\_ if for all legal databases instances r of R, .
4. In E-R model, an entity is represented by a set of \_\_attributes\_\_\_\_. A \_\_relationship\_\_\_ is an association among several entities.
5. Assume relation *r* has *br* blocks and relation *s* has *bs* blocks, therefore, in the best case, only *br* + *bs* block transfers would be required for .
6. An ideal hash function is uniform and random, the former require that each bucket is assigned the same number of search-key values from the set of all possible values.
7. To generate query-evaluation plans for an expression, we have to generate logically equivalent expressions using equivalence rules.
8. Consider a B+-tree of order n, if there are K search-key values in the file, the path from the root to the leaf node is no longer than \_⎡log⎡n/2⎤K⎤\_ .
9. A transaction has the following properties: atomicity, consistency, isolation and durability.
10. When the final statement of a transaction has been executed, the transaction enters the partially committed state. After a transaction has been rolled back and the database has been restored to its previous state, the transaction enter the aborted state.
11. A schedule S is recoverable if a transaction Tj in S reads a data item previously written by a transaction Ti , then the commit operation of Ti appears before the commit operation of Tj.
12. Multivalued attribute values or composite attribute values are not atomic.
13. A relation schema may have an attribute that corresponds to the primary key of another relation. The attribute is called a foreign key.

,

Answer:

1. \_\_\_\_\_\_\_\_\_\_\_,\_\_\_\_\_\_\_\_\_\_\_\_\_
2. \_\_\_\_\_\_\_\_\_\_\_,\_\_\_\_\_\_\_\_\_\_\_\_\_
3. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
4. \_\_\_\_\_\_\_\_\_\_\_,\_\_\_\_\_\_\_\_\_\_\_\_\_
5. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
6. \_\_\_\_\_\_\_\_\_\_\_,\_\_\_\_\_\_\_\_\_\_\_\_\_
7. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
8. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
9. \_\_\_\_\_\_\_\_\_\_\_,\_\_\_\_\_\_\_\_\_\_\_\_\_\_
10. \_\_\_\_\_\_\_\_\_\_\_,\_\_\_\_\_\_\_\_\_\_\_\_\_\_
11. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
12. \_\_\_\_\_\_\_\_\_\_\_,\_\_\_\_\_\_\_\_\_\_\_\_\_\_
13. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Part II [80 pts.] Answer the following questions**

1. [16points]Database design I: Consider the following conditions
   * 1. The STUDENT may be taught by one and only one teacher. The TEACHER may be instructor of one or more STUDENTS.
     2. The TEACHER may be responsible for one and only one CLASS. The CLASS may be the responsibility of one and only one TEACHER.
     3. The CLASS may be made of one or more STUDENTS. The STUDENT must be a member of one and only one CLASS.
     4. The CLASS must have one and only one ROOM. The ROOM may belong to one or more CLASS.

Notes: Assume entity CLASS has the following attributes: CID and CNAME, entity ROOM has the following attributes: RID and LOCATION, entity STUDENT has the following attributes: SID, LASTNAME, and FIRSTNAME, entity TEACHER has the following attributes: TID, TEACHERNAME, and TITLE.

1. [8points]Construct an ER diagram showing these relationships.



1. 4points] construct appropriate relation schemas for the above ER diagrams.

CLASS (CID, CNAME);

ROOM(RID, LOCATION)

STUDENT(SID, LASTNAME, FIRSTNAME)

TEACHER (TID, TEACHERNAME, TITLE).

S- T(SID,TID); T-C(CID,TID); C-R(CID,RID); S- C(SID,CID);

1. [4points] Create an index *std\_index* on the *Student* relation with *SID* as the search\_key.

Create Index std\_index on STUDENT(SID)

1. [6points] In database design, how to represent relationship sets as relational schemas?

A many-to-many relationship set :



Many-to-one /one-to-many relationship sets:



1. [14points]: Let *R* =(*A,B, C, D, E, F* ) be a relation withfunctional dependency *F* ={*A* → *CB, E* → *F A* }.
   * 1. Compute the candidate keys for R. [2pts]

Let us compute . .

Let us compute . 

Thus , ED is a superkey.

It is easy to see that , 

Thus, ED is a candidate key.

Note that ED is not implied by any other attributes, thus any candidate key of R must contain ED. Further, ED is the unique candidate key of R.

1. Is *R* in 3NF? If it is, justify your answer. If not, produce a decomposition of *R* into 3NF. [6pts]

Not. In *A* → *CB,*  A is a not a superkey and is not contained in the candidate key.

Compute the canonical cover, we have .

According to , we get *R*1 = (*A,B, C* ), *R*2 = (*A, E, F* ).

Note none of schemas contains ED, we generate *R*3 = (*D, E* ).

Thus, decomposition of *R:*

*R*1 = (*A,B, C* ), *R*2 = (*A, E, F* ), *R*3 = (*D, E* ).

1. Is *R* in BCNF? If it is, justify your answer. If not, produce a decomposition of *R* into BCNF. [6pts]

*否*

*1. A* → *CB* 违反BCNF定义

*R*1 = (*A,B, C* )

*R*2 = (*A, D, E, F* )

*2. R*2中*E* → *F A*违反BCNF定义,对*R*2分解

*R*3 = (*A, E, F* )

*R*4 = (*D, E* )

Final decomposition: *R*1*, R*3*, R*4

1. [28points] BOOK (Bookid,Title,Publishername)

BOOK\_AUTHORS(Bookid,Authorname)

PUBLISHER(Publishername, Address, Phone)

BOOK\_COPIES(Bookid,Branchid,No\_Of\_Copies)

LIBRARY\_BRANCH(Branchid,Brachname,Address)

BOOK\_LOANS(Bookid, Branchid, Cardno,DateOut,Duedate)

BORROWER(Cardno,Name, Address,Phone)

1. [3points]Write appropriate SQL DDL statements for declaring the BOOK\_AUTHORS relation.

**create table** BOOK\_AUTHORS (Bookid **char** (20), Authorname**char** (200));

1. [6points]Give an expressions in **relational algebra** to express the following queries:

Q1:Retrieve the names of all borrowers who do not have any books checked out.

Answer: (Note: We will use S for SELECT, P for PROJECT, \* for NATURAL JOIN, - for SET DIFFERENCE, F for AGGREGATE FUNCTION)



Q2:For each book that is loaned out from the "Sharpstown" branch and whose DueDate is today, retrieve the book title, the borrower's name, and the borrower's address.



1. [16points]Give an expressions in **SQL** to express the following queries:

Q1:How many copies of the book titled The Lost Tribe are owned by the library branch whose name is "Sharpstown"?

SELECT NoOfCopies

FROM ( (BOOK NATURAL JOIN BOOK\_COPIES ) NATURAL JOIN

LIBRARY\_BRANCH )

WHERE Title='The Lost Tribe' AND BranchName='Sharpstown'

Q2:For each library branch, retrieve the branch name and the total number of books loaned out from that branch.

SELECT L.BranchName, COUNT(\*)

FROM BOOK\_COPIES B, LIBRARY\_BRANCH L

WHERE B.BranchId = L.BranchId

GROUP BY L.BranchName

Q3: Retrieve the names, addresses, and number of books checked out for all borrowers who have more than five books checked out.

SELECT B.CardNo, B.Name, B.Address, COUNT(\*)

FROM BORROWER B, BOOK\_LOANS L

WHERE B.CardNo = L.CardNo

GROUP BY B.CardNo

HAVING COUNT(\*) > 5

Q4: For each book authored (or co-authored) by "Stephen King", retrieve the title and the number of copies owned by the library branch whose name is "Central".

SELECT TItle, NoOfCopies

FROM ( ( (BOOK\_AUTHORS NATURAL JOIN BOOK) NATURAL JOIN

BOOK\_COPIES)

NATURAL JOIN LIBRARY\_BRANCH)

WHERE Author\_Name = 'Stephen King' and BranchName = 'Central'

d.[3points] Record the fact that the manager didn't maintain information about the book named “T&G”, i.e. remove information about “T&G”.

**delete from** *BOOK\_AUTHORS***where** *Bookid* **in** (**select** Bookid**from** BOOK

**where** Title *=* “T&G”);

**delete from**  *BOOK\_COPIES***where** *Bookid* **in** (**select** Bookid**from** BOOK

**where** Title *=* “T&G”);

**delete from**  *BOOK\_LOANS***where** *Bookid* **in** (**select** Bookid**from** BOOK

**where** Title *=* “T&G”);

**delete from**  *BOOK***where** Title *=* “T&G”;

1. [16pts]Query Processing, Optimization and Transaction
   1. [4points] Please describe the implementation process of selection operation , where r is a relation, A is an attribute and is not a candidate key, r has a primary index on A. If there are n matching records, the B+ tree index is of height h, and each disk block contains at most d records, please analyze the overhead in the best case.

Process:

1. The algorithm walks from the root of the B+ tree to the leaf containing valuec.

2. According the address given by the leaf, the algorithm retrieves the n matching records.

Overhead:

1. It accesses disk h times. in each time, it needs a seek and a block transfer.

2. In the best case, these n records are distributed on n/d disk blocks. Thus it seeks disk 1 time and transfer block n/d times.

* 1. [4points] Describe the process of Indexed nested-loop join



Preconditions of usage:

* + join is an equi-join or natural join and
  + an index is available on the inner relation’s join attribute
    - Can construct an index just to compute a join.
* Algorithm: For each tuple *tr* in the outer relation *r,* use the index to look up tuples in *s* that satisfy the join condition with tuple *tr.*
  1. [4points] Please describe the two-phase locking protocol and prove that it ensures conflict-serializable schedules and does not ensure freedom from deadlocks.

The protocol:

Phase 1: Growing Phase

* 1. transaction may obtain locks
  2. transaction may not release locks

Phase 2: Shrinking Phase

* 1. transaction may release locks
  2. transaction may not obtain locks

The protocol assures serializability.

Proof. It can be proved that the transactions can be serialized in the order of their **lock points**(i.e. the point where a transaction acquired its final lock).

The protocol does not ensure freedom from deadlocks.

Proof. T1 lock-x (A)

T2 lock-x (B)

T1 lock-x (B)

T2 lock-x (A)

* 1. [4points] Below we show some log of a DBMS, please describe the recovery procedure using immediate database modification.

<*T*0**start**> <*T*0**start**> <*T*0**start**>

<*T0,* A, 1000, 950> <*T0,* A, 1000, 950> <*T0,* A, 1000, 950>

*<T*o*,* B, 2000, 2050> *<T*o*,* B, 2000, 2050> *<T*o*,* B, 2000, 2050>

<*T*0 **commit**> <*T*0 **commit**>

<*T*1 **start**> <*T*1 **start**>

<*T*1, C, 700, 600> <*T*1, C, 700, 600>

<*T*1 **commit**>

(a) (b) (c)

Recovery actions in each case above are:

(a) redo (*T*0) and redo (*T*1): A and B are set to 950 and 2050 respectively. Then *C* is set to 600

(b) undo (*T*1) and redo (*T*0): C is restored to 700, and then *A* and *B* are

set to 950 and 2050 respectively.

(c) undo (*T*0): B is restored to 2000 and A to 1000.