**浙江大学2018年**

**《数据库系统》期末考试题**

**Problem 1：Relational Model and SQL（15 points，3 points per part)**

Following relational schemas represent information of a campus card database in a university.

**card( cno:char(5), name:char(8), depart:char(10), balance:integer )**

**pos( pno:char(4), campus:char(8), location:char(10) )**

**detail( cno:char(5), pno:char(4), cdate:date, ctime:time, amount:integer,**

**remark:char(10) )**

Following tables show an instance of the database:

**card**

|  |  |  |  |
| --- | --- | --- | --- |
| **cno** | **name** | **depart** | **balance** |
| c0001 | 张帅 | CS | 100 |
| c0002 | 李丽 | EN | 200 |
| c0003 | 王浩 | CS | 300 |
| c0004 | 刘萌 | CS | 400 |
| c0005 | 赵亮 | MA | 500 |

**pos**

|  |  |  |
| --- | --- | --- |
| **pno** | **campus** | **location** |
| p001 | 玉泉 | 教育超市 |
| p002 | 玉泉 | 四食堂 |
| p003 | 玉泉 | 四食堂 |
| p004 | 紫金港 | 教育超市 |
| p005 | 紫金港 | 教育超市 |
| p006 | 紫金港 | 一食堂 |

**detail**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **cno** | **pno** | **cdate** | **ctime** | **amount** | **remark** |
| c0001 | p002 | 2018-07-01 | 08:10:10 | 6 | 餐饮 |
| c0001 | p002 | 2018-07-01 | 12:05:12 | 12 | 餐饮 |
| c0001 | p002 | 2018-07-01 | 17:30:20 | 20 | 餐饮 |
| c0001 | p001 | 2018-07-01 | 18:10:10 | 60 | 购物 |
| c0002 | p002 | 2018-07-02 | 08:10:10 | 8 | 餐饮 |
| c0002 | p001 | 2018-07-02 | 08:10:10 | 20 | 购物 |
| c0003 | p003 | 2018-07-02 | 08:10:10 | 25 | 餐饮 |

Given following SQL query on above relations：

**select cno, name**

**from card natural join detail**

**where depart= ‘CS’ and (cdate , pno) in**

**( select cdate, pno**

**from detail**

**where cno=‘c0002’ )**

Please answer following questions：

1. Transform above query to a SQL statement without nested subquery.
2. Transform above query to an equivalent relational algebra expression.
3. Write a SQL statement to find out cards consumed in only one campus in 2018.
4. Write a SQL statement to find out the pos in “紫金港” campus that has the maximum total amount of card consumption in 2018.
5. Write a sequence of SQL statements to complete following transaction:

card “c0002” consumes 20 at pos “p001” at 2018-07-02 08:08:08

**Answers of Problem 1：**

**Problem 2：E-R Model （11 points)**

Please design an ER diagram for a takeaway platform（外卖平台）to capture information about **restaurants**, **customers**, **couriers**（快递员）**,** **orders , payments,** and **deliveries,** etc. A **restaurant** sell **items** on the platform. A **customer** has several **addresses**. A **courier** delivers packages of order to customers.

**Answers of Problem 2：**

**Problem 3: Relational Formalization(12 points, 3 points per part)**

For relation schema R(A, B, C, D, E) with functional dependencies set F={A🡪CD, C🡪B, B🡪D, B🡪E}

1. Compute the Canonical Cover of F.
2. Compute the Closure of attribute set {B}.
3. Decompose the relation R into a collection of BCNF relations, and give out the functional dependency set for each relation.
4. Explain whether above decomposition be dependency preserving or not.

**Answers of Problem 3：**

**Problem 4: XML(12 points, 4 points per part)**

Following DTD depicts information of the campus card database given in **problem 1**:

<!DOCTYPE campus\_cards[

<!ELEMENT campus\_cards( pos+,card+)>

<!ELEMENT pos (campus, location)>

<!ATTLIST pos pno ID #REQUIRED>

<!ELEMENT campus (#PCDATA)>

<!ELEMENT location (#PCDATA)>

<!ELEMENT card (name, depart, balance, detail\*)>

<!ATTLIST card cno ID #REQUIRED>

<!ELEMENT name (#PCDATA)>

<!ELEMENT depart (#PCDATA)>

<!ELEMENT balance (#PCDATA)>

<!ELEMENT detail (cdate, ctime, amount, remark)>

<!ATTLIST detail pno IDREF #REQUIRED>

<!ELEMENT cdate (#PCDATA)>

<!ELEMENT ctime (#PCDATA)>

<!ELEMENT amount (#PCDATA)>

<!ELEMENT remark(#PCDATA)>

]>

1. Base on the instance of the campus card database given in **problem 1**, give an XML document that conforms to above DTD. The document is required to contain the information of **pos “p003”** and **card “c0003”** as well as its consumption details.
2. Give a Path expression to find out the **location** of **pos** where张帅once paid 50.
3. Give an XQuery expression to find outthe **cno** of **card** that once paid at the same **pos** and **date** with 张帅.

**Answers of Problem 4：**

**Problem 5: B+ -Tree (12 points, 4 points per part)**

For the following B+-tree (n=4 )：

158

228

48 98

228 258

8 38

48 68

158 218

98 128 138

1. Draw the B+ -tree after inserting four entries 28,168, 58,18 sequentially.
2. Draw the B+ -tree after deleting four entries 8, 38, 98,128 from the original B+-tree sequentially.
3. Assuming (a) there are 3 blocks in main memory buffer for B+-tree operation, at first these blocks are empty. (b) Least Recently Used (LRU strategy) is used for buffer replacement. (c) each node of B+-tree occupies a block.

Please count the number of B+-tree blocks transferred to buffer in order to complete the operation in 1).

**Answers of Problem 5：**

**Problem 6: Query Processing (16 points, 4 points per part)**

For the relational schemas of the campus card database given in **problem 1**, there are following assumptions:

* ncard=10,000 , npos=100, ndetail=10,000,000
* lcard=25, lpos=22, ldetail=29
* V(campus, pos) = 6, V(location, pos) = 20
* V(depart, card) = 100, V(name, card) = 5000
* Thevalue ofattribute **cdate** in **detail** table is uniformly distributed between ‘2017-01-01’ and ‘2017-12-31’.
* block size is 4K bytes.
* size of B+-tree pointer is 4 bytes.
* **card** and **detail** tablesare stored as sequential files based on search key **cno**.
* there is a B+-tree index on **detail(cno)**.

1. Estimate the size (i.e. number of records) returned by following SQL statement :

**select** d1.cno, d2.cno

**from** detail d1, detail d2

**where** d1.pno=d2.pno and d1.cdate=d2.cdate and

d1.cdate between ‘2017-05-01’ and “2017-07-31’

1. Estimate the number of blocks of **card** and **detail** tables respectively.
2. Estimate the height of the B+-tree index on **detail(cno)**.
3. Estimate the cost for evaluating expression “**σ name=’张帅’ (card) detail** ” using file scan for **σ** operationfollowed by indexed-loop join method for operation.

(Hind: the cost is measured by number of blocks transferred to main memory and times to seek disk.)

**Answers of Problem 6：**

**Problem 7: Concurrency Control (12 points，3 points per part)**

Consider following schedule S with five transactions T1,T2,T3,T4 and T5:

***S: r1(A ) w2(A) r2 (B) w3(A) w3(B) w4(C) w4(B) w5(C) w3(C)***

where: ***ri(X)*** means transaction ***Ti*** read data ***X***.

***wi(X)*** means transaction ***Ti*** write data ***X***.

1. Draw the precedence graph of S.
2. Explain whether S is serializable.
3. Explain whether S be generated by the two–phase locking protocol.
4. If S is serializable, give out all serial schedules to which S is equivalent. If S is not serializable, figure out an operation in S such that if this operation was removed from S，S would become serializable.

**Answers of Problem 7：**

**Problem 8: Aries Recovery Method (10 points, 2 points per part)**

A DBMS uses **Aries** algorithm for system recovery. Following figure is a log file just after system crashes. The log file consists of 16 log records with LSN from 1001 to 1016. The figure does not show PrevLSN and UndoNextLSN in log records. Assuming that last completed checkpoint is the log record with LSN 1012.

1010

1006

T3

1011： <T2, 8002.2, 88, 99>

PageLSN

PageID

1006

1011

8002

1010

1010

8001

T4

1011

T2

LastLSN

Txn

RecLSN

1012： checkpoint

1010： <T4, 8001.1, 22, 77 >

1009： <T4 begin>

1008： <T1 commit>

1004： <T2 , 8001.2, 33, 44>

1006： <T3, 8002.1, 55, 66 >

1001： <T1 begin>

1005： <T3 begin>

1003： <T2 begin>

1002： <T1 , 8001.1, 11, 22>

1013： <T2 commit>

1014： <T3, 8002.1, 66, 77>

1016： <T4 commit>

1015： <T4, 8003.1, 11, 22>

Please answer following questions：

1. Which log record is the start point of Redo Pass?
2. Which log record is the end point of Undo Pass?
3. After Analysis Pass，what content is the dirty page table?
4. After recovery, what is the value of data items identified by “8002.1” and “8002.2” respectively?
5. After recovery, what additional log records appended to log fie?

**Answers of Problem 8:**