

CMPSCI 645: Homework 4 - Part B

Due: Monday, April 9 2018, 11:59 pm

Prof. Yanlei Diao

Name:

Question	Points	Score
5	16	
6	24	
Total	40	

Please turn in this homework electronically, as a PDF, through Gradescope. You may handwrite your solutions, and then scan the document, or type directly into the PDF form. Make sure the PDF you upload includes all pages (including this front page).

Question 5 [3 parts, 16 points] Query Optimization

Consider the following schema

Sailors(sid, sname, rating, age)

Boats(bid, bname, size)

Reserves(sid, bid, day)

Reserves.sid is a foreign key to Sailors and Reserves.bid is a foreign key to Boats.bid.

We are given the following information about the database:

Sailors contains 50 pages with 20 records per page, so 1000 records in total.

Boats contains 10 pages with 10 records per page, so 100 records in total.

Reserves contains 250 pages with 40 records per page, so 10,000 records in total.

There are 100 values for Reserves.sid.

There are 50 values for Reserves.bid.

There are 1000 values for Reserves.day

In the following queries, assume that a System R style optimizer is used.

Consider Query 1:

```
SELECT S.sid, S.sname, B.bname
FROM   Sailors S, Reserves R, Boats B
WHERE  S.sid = R.sid AND R.bid = B.bid AND R.day = 'July 4, 2003';
```

(a) [4 points] Assuming uniform distribution of values and column independence, estimate the number of tuples returned by this query.

Consider Query 2:

```
SELECT    S.sid, S.sname, B.bname
FROM      Sailors S, Reserves R, Boats B
WHERE     S.sid = R.sid AND R.bid = B.bid
```


(b) [4 points] Draw all possible left-deep query plans for this query:

(c) [8 points] List all the one-relation, two-relation, and three-relation subplans that a System-R optimizer will consider in optimizing Query 2 based on the dynamic programming algorithm that we learned in class. There is no need to consider the particular join methods in this question.

Question 6 [24 points] Transaction Management

(a) [6 points] Serializable Schedules

Consider the following schedule involving three transactions T1, T2 and T3:

	1	2	3	4	5	6	7	8	9	10
	<hr/>									
										
T1:	W(A)	R(B)	R(C)		W(B)					
T2:							R(C)	W(B)	W(C)	
T3:				W(B)		W(C)				R(A)

(i) Draw the precedence graph for this schedule.

(ii) Judge if this schedule is conflict serializable or not. If it is conflict serializable, give the equivalent serial schedule (just write the order of the transactions). Otherwise, briefly explain why the schedule is not conflict serializable.

Consider three transactions, T1, T2, and T3, with timestamps 1, 2, 3, respectively. Now consider the following sequences of actions, listed in the order that they are produced from user queries:

Explain how this sequence of actions will be executed using the **Strict Two Phase Locking** protocol (2PL). The following table is designed for you to specify in time order the activities that take place in the lock manager and in access to the database. The activities include:

- The first few actions are already given below. Please complete the other actions for this schedule. List only one activity in each row, and use as many rows as needed for this schedule.

Time stamp	Action: please list one action at a time
1	T1: S(A)
2	T1: R(A)
3	T2: X(A)
4	T2: blocked
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

(c) [12 points] The ARIES Recovery Algorithm

Consider the execution shown in the following figure.

LSN	Log Record
00	Update: T1 writes P2
10	Update: T1 writes P1
20	Update: T2 writes P5
30	Update: T3 writes P3
40	T3 commits
50	Update: T2 writes P5
60	Update: T2 writes P5
70	T2 aborts

(i) [3 points] Describe the actions taken to rollback transaction T2.

(ii) [9 points] Assume that the system crashed right after generating the log record 70. Now the system wakes up and runs the ARIES recovery algorithm. Please show all new log entries generated in the recovery process, where each log record includes the LSN, type of log record, transaction ID, page ID, prevLSN, and undonextLSN.

LSN	Type	Xact ID	Page ID	prevLSN	undonextLSN
80					
90					
100					
110					
120					
130					
140					