Data Management – exam of 3/6/2015

Problem 1 Consider the following schedule

$$S = r_1(A)\, r_1(B)\, w_2(A)\, r_3(C)\, w_1(C)\, w_2(B)\, w_2(C).$$

1. Tell whether S is accepted by the 2PL scheduler with exclusive and shared locks. If the answer is yes, then show the schedule obtained from S by adding suitable lock and unlock commands. If the answer is no, then explain the answer, and then tell whether there is a single action in the schedule whose removal brings the schedule in the class of 2PL with exclusive and shared locks.
2. Tell whether S is conflict-serializable. If the answer is yes, then show a serial schedule that is conflict-equivalent to S. If the answer is no, then explain the answer.

Problem 2 Answer the following three questions.

1. Describe in detail an algorithm that, given two schedules on the same set of transactions, checks whether the two schedules are view-equivalent.
2. Describe in detail an algorithm that, given a schedule, checks whether the schedule is view-serializable, by using the algorithm illustrated above for item 1.
3. Discuss in detail the computational complexity of the two algorithms illustrated for items 1 and 2 above.

Problem 3
Consider the following schedule S:

$$w_3(y)\, r_3(t)\, r_1(x)\, w_1(z)\, w_2(x)\, r_3(z)\, w_2(y)\, w_4(x)$$

and tell whether S is accepted by the timestamp-based scheduler, explaining your answer in detail.

Problem 4
A relation R has 200.000 pages, and our DBMS has 200 buffer frames free. Describe in detail the algorithm you would use to sort R under this condition, and tell which is the cost of the algorithm in terms of page accesses.

Problem 5
Consider the relation `MEETING(code,year,nation,cost,number)` that stores information about meetings, with the code of meeting, the year when the meeting has taken place, the nation where the meeting has taken place, the cost of the meeting, and the number of people that participated in the meeting. The relation occupies 600.000 pages, each of 1600 KB. We assume that all fields in every record have the same size of 20 KB, independently of the field type. There is a sparse B^+-tree index on `MEETING` with search key $\langle$`code,year,nation`$\rangle$, using alternative 2. Consider the query

```
select code,nation from MEETING;
```

that asks for the code and the nation of all the meetings, and tell which algorithm you would use for executing the query, and how many page accesses such algorithm needs for computing the answer.

Problem 6
Describe in detail the one-pass algorithm for eliminating duplicates from a relation R, in the case where R is not sorted, and there is no index defined on it. Also, tell which are the conditions under which such one-pass algorithm can be used.