Data Management – exam of 18/02/2009

LAST NAME: _____

NAME: _____

STUDENT CODE: _____

I allow the publication of my grade on the Web site `http://www.dis.uniroma1.it/~lenzerini`, according to the Italian regulation (decreto legislativo 196/2003), which I hereby declare to know.
Faithfully,

_____

Problem 1 ***(part 1)*** Consider the following schedule

$$S = r_2(T)\, r_1(P)\, w_1(S)\, w_1(P)\, c_1\, r_2(S)\, w_2(T)\, r_2(P)\, w_4(Q)\, w_3(T)\, c_2\, w_4(T)\, r_3(P)\, c_3\, w_4(P) c_4.$$

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.
2. Tell whether S is strict or not, and explain the answer.
3. Tell whether S is ACR (Avoid Cascading Rollback) or not, and explain the answer.
4. 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 ***(part 1)*** A schedule constituted by only "write" actions is called a *write-only* schedule. Prove or disprove the following statement: every write-only schedule that is view serializable is also conflict serializable.

Problem 3 ***(part 2)*** Let `PriceList(ProductCod,Description,ProductionYear,Cost)` be a relation (whose key is `ProductCod`) that is never updated, and is queried by a single type of query, asking for the code, the description and the cost of all products produced in a given interval of years.

1. Which file organization would you choose for the relation `PriceList`? Explain in detail your answer.
2. On the basis of the choice done for the previous question, tell which is the function characterizing the cost of answering the queries of the above type, with respect to the number of pages needed to store the relation. Explain in detail your answer.
3. If the relation *PriceList* were frequently updated, which file organization would you have chosen? Explain in detail your answer.

Problem 4 ***(part 2)***

Consider the relations R(A,B) and S(B,C), where R is constituted by 3000 pages and S is constituted by 1000 pages.

What is the cost (in terms of number of pages that are either read or written) of the natural join between R and S if we use the nested loop algorithm? Explain in detail your answer.

Suppose the buffer has currently 502 pages available. What is the cost (in terms of number of pages that are either read or written) of the natural join between R and S if we use the block nested loop algorithm? Explain in detail your answer.