Data Management – exam of 15/01/2009

COGNOME: _____

NOME: _____

MATRICOLA: _____

Autorizzo la pubblicazione del mio voto di questo esame sul sito web `http://www.dis.uniroma1.it/~lenzerini`, secondo quanto prevede il decreto legislativo 196/2003 (codice in materia di protezione dei dati personali) che dichiaro di conoscere. In fede,

_____

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

$$S = r_1(A)\, r_2(B)\, w_1(A)\, w_1(C)\, c_1\, r_2(C)\, w_2(B)\, w_3(B)\, r_2(A)\, w_4(D)\, c_2\, w_4(B)\, r_3(A)\, c_3\, c_4.$$

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)*** Prove or disprove the following statement: every schedule accepted by the "multiversion timestamp" method is strict.

Problem 3 ***(part 2)*** Let `Citizen(Cod,Name,BirthDate,BirthCity,Address)` be a relation (whose key is `Cod`) that is frequently updated, and is queried by a single type of query, asking for name and address of all citizens who are born in a given city (every query of the above type specifies such a city). We know that every page contains 100 records of the relation `Citizen`. Currently, there are about 670.000 citizens and about 100 cities in our database, and the births of the citizens distribute uniformly on the various cities.

1. Tell which file organization you would choose for the relation `Citizen`, and motivate your choice.
2. On the basis of the choice done for the previous question, tell how many page accesses are needed in the current situation for a query of the above type, and explain the answer.
3. If the relation *Citizen* were static (never updated), tell which file organization you would have chosen, and motivate your choice.
4. On the basis of the choice done for the previous question, tell how many page accesses are needed in the current situation for a query of the above type, and explain the answer.

Problem 4 ***(part 2)*** Consider a relation `R` represented through an extendible hashing scheme, based on a hash function h defined on the search key K. Let S_1 denote the current situation, in which the value of the global depth g is 1, every page contains at most 4 values of the search key, and the following values of the search key appear in the index: 4^*, 12^*, 32^*, 16^*, 1^*, 5^*, 21^*, 13^*. Note that, in our notation, x^* denotes the value v of the search key K such that $h(v) = x$.

1. Draw a picture illustrating the pages in above described current situation S_1.
2. Explain what happens if, starting from situation S_1, we insert the value 10^* of the search key, drawing a picture illustrating the pages of the resulting situation S_2.
3. Explain what happens if, starting from situation S_2, we insert the value 20^* of the search key, drawing a picture illustrating the pages in the resulting situation S_3.