Part I

2. Let R be a relation schema, R_1 and R_2 form a decomposition of R. Decomposition is a lossless if for all legal databases instances r of R, $\Pi_{R_1}(r) \bowtie \Pi_{R_2}(r) = r$

5. For evaluating an entire expression tree, ——— passes on tuples to parent operations even as an operation is being executed

6. The time it takes for a disk I/O includes block-transfer time and access time, The latter consists of seek time and rotational latency.

8. A sequence of primitive operations that can be used to evaluate a quary is a query-execution ——— plan.

10. In queryprocessing, the query-execution engine takes the execution plan which contains detailed information on how a particular query or a set of queries will be executed.

11. In deferred database modification scheme, redo? operation is the only operation used in the recovery procedure.

12. Cascading rollbacks can be avoided by applying strict two phase locking protocol to transactions in currency protocols.

13. log is the most widely used structure for recording the modification of database. Since a failure may occur while a update is taking place, it must be written out to non-volatile storage before the actual update to database to be done.

14. A (possibly concurrent) schedule is serializable if it is equal to a serial schedule.

2. Data base design II: Consider a relation schema R(A, B, C, D, E) Given the functional dependencies ABC → D, D → E, E → A.

(a) please find all candidate keys for R ABC, BCD, BCE

(b) please decompose R in 3NF (we only decompose when there is violation of 3NF). If R is already in 3FN just write R(A, B, C, D, E), instead of decomposition. $F_c = \{ABC \to D, D \to E, E \to A\}$ $\{ABCD\} \cup \{DE\} \cup \{EA\}$

(c) Please decompose R in BCNF $\{DE\}$ $\{ABCD\} \Rightarrow \{DE\}, \{EA\}$ $\{ABC\}$

4. Relation R contains 10,000 tuples and has 10 tuples per page. Relation S contains 2000 tuples and also has 10 tuples per page. If only 15 buffers were available, what would be the cost of a sort-merge join?

$(1000 + 200) \cdot (2 + \log_{14}\lceil \frac{1000}{15} \rceil + \log_{14}\lceil \frac{200}{15} \rceil) + 1000 + 200$

5. Describe the process of Cost-Based optimization.

6. Is the following schedule S conflict serializable? Why?

T_1	T_2
read(A)	
write(A)	
	read(A)
	write(A)
read(B)	
write(B)	
	read(B)
	write(B)

$T_1 \to T_2$

Yes