

**Q 1** *(3+3+3=9 points)***:** You are given the below the set F of functional dependencies for relation R(A,B,C,D,E),

F = {CE → D, D → B, C → A}.

1. Find all keys (candidate keys).
2. Identify the best normal form that R satisﬁes (1NF, 2NF, 3NF, or BCNF). Justify your answer.
3. If the relation is not in BCNF, decompose it until it becomes BCNF. At each step, identify a new relation, decompose and re-compute the keys and the normal forms they satisfy.

**Q 2** *(2+2+2=6 points)***:** You are given the below the set F of functional dependencies for relation R(A,B,C,D,E),

F = {AB → C, AB → D, D → A, BC → D, BC → E}.

1. Is this relation is in BCNF? If not, show all dependencies that violate it.
2. Is this relation in 3NF? If not, show all dependencies that violate it
3. Is the following dependency implied by the above set of dependencies? If so, show how using the Inference Rules (Armstrong’s Axioms) : ABC → AE

**Q 3** *(5 points)***:** You are given the relation below for a relation schema R(A,B,C,D,E). You do not know the functional dependencies for this relation. This question is independent of Question 3 above.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A | B | C | D | E |
| ’a’ | 122 | 1 | ’s1’ | ’a’ |
| ’e’ | 236 | 4 | ’e2’ | ’b’ |
| ’a’ | 199 | 1 | ’b5’ | ’c’ |
| ’b’ | 213 | 2 | ’z8’ | ’d’ |

Suppose this relation is decomposed into the following two relations: R1(A,B,C,D) and R2(A,C,E).

Is this decomposition lossless? Explain your reasoning.

**Q 4** *(5 points)***:** Consider relation R(X, Y, Z). Relation R currently has three tuples: (6, 4, 2), (6, 6, 8) and (6, 4, 8). Which of the following three functional dependencies may hold in the relation R? If the dependency cannot hold , explain why by specifying the tuples that cause the violation.

**1.** Y → X **2.** Z → Y **3.** XY → Z

**Q 5** *(5 points)***:** You are given the below set of functional dependencies for a relation R(A,B,C,D,E,F,G),

F = {AD → BF, CD → EGC, BD → F, E → D, F → C, D → F}.

Find the minimal cover for the above set of functional dependencies using the algorithm described in class.

**Q 6** *(5+5+5=15 points)***:** Consider following relational schema for database to maintain Documents of an organization.

DOCUMENT(doc-id, title, size, doc-type) ABOUT(doc-id , top-id *,* rev-id)

TOPIC(top-id, top-name, parent-top-id) REVIEWER(rev-id, rev-name)

SSYSTEM(sys-id, sys-name, capacity, sys-type) STORED(doc-id , sys-id , timestamp)

A DOCUMENT is uniquely identified by its doc-id and has a title, a size (in bytes), and a doc-type (e.g. “.ppt”, “.pdf”, “.html”, etc.). A document can be ABOUT one or more TOPICs, as identified by a REVIEWER. But different reviewers may disagree in their assessment: e.g., Reviewer 1 might say that document 17 is about both topics X and Y (so there will be two corresponding tuples in ABOUT), while Reviewer 2 says that this document is about Z. A TOPIC is identified by its top-id, has a name (e.g., “Query Optimization”), and a unique parent topic, identified by the parent-top-id. For example, topic X (say “Query Optimization”) might have as unique parent the topic Z (say “Databases”). A REVIEWER has a unique rev-id and a name. A SSYSTEM (Storage System) is uniquely identified by its sys-id, has a system name sys-name (e.g. “MegaStore 3000”), a storage capacity (in bytes), and a system type sys-type. Finally, documents are STORED on storage systems. A timestamp is used to record when a particular document doc-id was stored in system sys-id.

Write SQL queries for each statement given below.

1. Give the doc-ids of documents for which exactly one ABOUT entry exists.
2. What is the maximum size of documents that are about the topic “Data Mining”?
3. Give the id and name of the topic with the most entries in ABOUT.