

Normalization Process

Fix normal form violations by *decomposing* relations.

- for a functional dependency $X \rightarrow Y$, split R into relations XY and $R-Y$

When decomposing a relation R to satisfy a normal form, must ensure

- lossless join
 - cannot get spurious tuples
- dependency preservation
 - each functional dependency of R is represented in some relation or can be derived from those that are

for FD $X \rightarrow Y$, splitting R into XY and $R-Y$ is lossless as long as $X \cap Y = \emptyset$ i.e. X and Y do not have any attributes in common

more generally, splitting R into $R1$ and $R2$ is lossless as long as $R1 \cap R2$ is a superkey for $R1$ or $R2$

Preserving Dependencies

TASK(project, manager, task, deadline, budget)

- project \rightarrow manager [every project has a single manager]
- project, task \rightarrow deadline [each of the project's tasks has its own deadline]
- manager, deadline \rightarrow budget [budget specific to manager and deadline]

Is this in 2NF?

- no, because non-key manager depends on project, a subset of a key

Decompose to satisfy 2NF.

TASK(project, task, deadline, budget)
MANAGER(project, manager)

...but this is not dependency-preserving because manager, deadline \rightarrow budget is not represented in either relation (or derivable from ones that are)

Preserving Dependencies

TASK(project, manager, task, deadline, budget)

- project \rightarrow manager [every project has a single manager]
- project, task \rightarrow deadline [each of the project's tasks has its own deadline]
- manager, deadline \rightarrow budget [budget specific to manager and deadline]

Try a different decomposition. Observe:

- project \rightarrow manager violates 2NF because non-key manager depends on project, a subset of a key
- manager, deadline \rightarrow budget violates 3NF because non-key budget depends on non-key manager, deadline

Decompose.

TASK(project, task, deadline)
MANAGER(project, manager)
BUDGET(manager, deadline, budget)

This is dependency-preserving, lossless, and BCNF.

- TASK \cap MANAGER = { project }, a superkey for MANAGER
- (TASK \cap MANAGER) \cap BUDGET = { manager, deadline }, a super key for BUDGET