

Normal Form Review

Key

FDs F on schema R , $K \subseteq R$

1. $K^+ = R$
2. No proper subset of K has this property

Can be multiple keys

An attribute A is prime if it is in some key K

Superkey: Just Part 1. above

Example

$Q D C N R P T$

$\{DC \rightarrow N, DN \rightarrow C, QDCT \rightarrow RP, P \rightarrow D, QRT \rightarrow DC\}$

$QRT^+ = QRTDCNP$

$QR^+ =$

$QT^+ =$

$RT^+ =$

$Q^+ =$

$R^+ =$

$T^+ =$

CS 589 Principles of Databases, Lecture 2-4

Principles of Databases

More Examples

$\{DC \rightarrow N, DN \rightarrow C, QDCT \rightarrow RP, P \rightarrow D, QRT \rightarrow DC\}$

$QCPT^+ =$

$CPT^+ =$

$QPT^+ =$

$QCT^+ =$

$QCP^+ =$

What about QRPT?

Unit 2, Part 5

David Maier

3

Principles of Databases

Different Example

Flight **D**ate **P**ilot **G**ate

$\{LD \rightarrow PG, L \rightarrow G\}$

LD is a key, so L, D are prime

P, G are not prime -

Unit 2, Part 5

David Maier

4

Third Normal Form (3NF)

Schema R with FDs F is in third normal form if for any nontrivial FD $X \rightarrow A$,

Last example was not in 3NF

Consider $L \rightarrow G$

Another Example

Flight **D**ate **P**ilot**I**D **P**ilot**N**ame
{ $LD \rightarrow IN$, $I \rightarrow N$, $N \rightarrow I$ }

Aside: Boyce-Codd Normal Form (BCNF)

Schema R with FDs F is in Boyce-Codd normal form if for any nontrivial FD $X \rightarrow A$,

BCNF Example

Airport **C**ompany **B**ill**T**o

$\{A \rightarrow C, C \rightarrow B, B \rightarrow A\}$

This schema is in 3NF

This schema is not in BCNF

Normalization through Decomposition

If R not in 3NF for FDs F, find an FD
 $Y \rightarrow Z$ in F^+ where Y is not a key; split
into

FLight Date PilotID PilotName
{LD $\rightarrow$ IN, I $\rightarrow$ N, N $\rightarrow$ I}

Problems with Decomposition

1. Figuring out non-prime attributes can be hard
2. Might end up with more relations than you need
3. Might not be able to enforce FDs

Decomposition Example

$R = A \ B \ C \ D \ E$

$\{AB \rightarrow CDE, AC \rightarrow BDE, B \rightarrow C, C \rightarrow B, C \rightarrow D, B \rightarrow E\}$

AB, AC are keys, so $C \rightarrow D$ is a problem

Decompose on $C \rightarrow D$

$B \rightarrow E$ is a problem, decompose again

But ...

Enforcing FDs Example

$R = A \ B \ C \ D \ E$

$\{A \rightarrow BCDE, CD \rightarrow E, CE \rightarrow B\}$

A is the only key, so $CD \rightarrow E$ is a problem

Decompose on $CD \rightarrow E$

How do you enforce $CE \rightarrow B$?

Normalization through Synthesis

Start with schema R and FDs F

Generate database schema

$R_1, R_2, \dots, R_n$ such that

- Every relation schema is in 3NF
- Can enforce F
- Avoid extra relation schemas
- Have a lossless decomposition from R onto $R_1, R_2, \dots, R_n$.

Depends on a cover for F of a certain form

What's a Cover?

A set of FDs G is a cover for a set of FDs F if

G is a non-redundant cover if you cannot eliminate FDs and still be a cover

$\{A \rightarrow B, B \rightarrow C, A \rightarrow C\}$

Can test FDs one by one to see if any can be left out

Minimum Cover

A minimum cover is one with fewest possible FDs

Non-redundant, but not minimum:

$\{A \rightarrow BC, B \rightarrow A, AD \rightarrow E, BD \rightarrow I\}$

Finding a Minimum Cover

Replace every FD $X \rightarrow Y$ in F with $X \rightarrow X^+$, then find a non-redundant cover

$\{A \rightarrow BC, B \rightarrow A, AD \rightarrow E, BD \rightarrow I\}$

Reduced Covers

Are there extraneous attributes?

Can appear on the right side or the left

$\{AB \rightarrow C, A \rightarrow B\}$

$\{A \rightarrow BC, B \rightarrow C\}$

How to remove?

For $XA \rightarrow Y$, see if

For $X \rightarrow YB$, see if

Reduce Left Sides First

$G = \{A \rightarrow C, AB \rightarrow DE, AB \rightarrow CDI, AC \rightarrow J\}$

Reduce left sides

$G' =$

Reduce right sides

$G'' =$

Synthesis Algorithm

Start with FDs F

Find G = minimum cover for F

Find H = reduced cover for G

For each $X \rightarrow Y$ in H , add a schema XY

Synthesis Example

$$F = \{A \rightarrow BC, B \rightarrow A, \\ AD \rightarrow E, BD \rightarrow I\}$$

Lossless Decomposition

If none of the FDs represents a key for all the attributes, need to add one more relation to get a lossless decomposition

$R = ABCDE$

$F = \{A \rightarrow B, C \rightarrow D\}$