

CS 589 Principles of Databases, Unit 2, Part 5

Principles of Databases

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

Unit 2, Part 5

David Maier

1

Principles of Databases

Example

Q D C N R P T

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

$QRT^+ = QRTDCNP$

$QR^+ =$

$QT^+ =$

$RT^+ =$

$Q^+ =$

$D^+ =$

$T^+ =$

Unit 2, Part 5

David Maier

2

CS 589 Principles of Databases, Unit 2, Part 5

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

CS 589 Principles of Databases, Unit 2, Part 5

Principles of Databases

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 $L \rightarrow G$

Unit 2, Part 5

David Maier

5

Principles of Databases

Another Example

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

Unit 2, Part 5

David Maier

6

CS 589 Principles of Databases, Unit 2, Part 5

Principles of Databases

Aside: Boyce-Codd Normal Form (BCNF)

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

Unit 2, Part 5

David Maier

7

Principles of Databases

BCNF Example

Airport Company BillTo

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

This schema is in 3NF

This schema is not in BCNF

Unit 2, Part 5

David Maier

8

CS 589 Principles of Databases, Unit 2, Part 5

Principles of Databases

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 → IN, I → N, N → I`}

Unit 2, Part 5

David Maier

9

Principles of Databases

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

Unit 2, Part 5

David Maier

10

CS 589 Principles of Databases, Unit 2, Part 5

Principles of Databases

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 ...

Unit 2, Part 5

David Maier

11

Principles of Databases

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 $CE \rightarrow B$?

Unit 2, Part 5

David Maier

12

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

$$G'' =$$

CS 589 Principles of Databases, Unit 2, Part 5

Principles of Databases

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

Unit 2, Part 5

David Maier

19

Principles of Databases

Synthesis Example

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

Unit 2, Part 5

David Maier

20