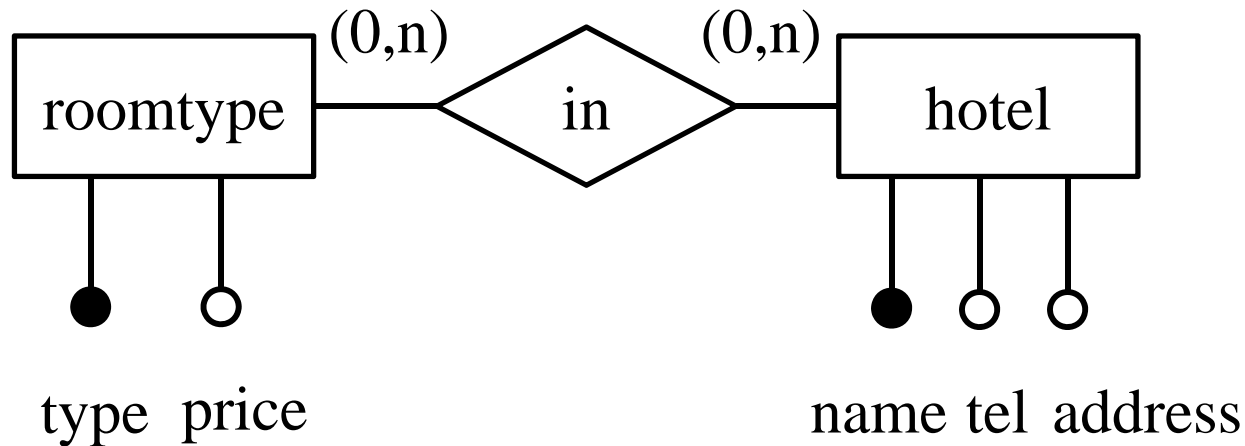


Normal Forms

Stéphane Bressan





We have the following functional dependencies:

$\{\text{type}\} \rightarrow \{\text{price}\}$

and

$\{\text{name}\} \rightarrow \{\text{tel}, \text{address}\}$

One Table

type	price	name	tel	address
superior	145	Sangria Clarke Quay	65166516	1 Clarke Quay
standard	75	Sangria Clarke Quay	65166516	1 Clarke Quay
suite	250	Sangria Clarke Quay	65166516	1 Clarke Quay
superior	145	Sangria Holland V	65166516	13 Holland Drive
standard	76	Sangria Holland V	65166516	13 Holland Drive
suite	250	Sangria Holland V	65165555	13 Holland Drive
junior suite	200	Sangria Holland V	65165555	13 Holland Drive
executive	175			

We have the following functional dependencies:

$\{\text{type}\} \rightarrow \{\text{price}\}$

and

$\{\text{name}\} \rightarrow \{\text{tel}, \text{address}\}$

Redundant Storage

type	price	name	tel	address
superior	145	Sangria Clarke Quay	65166516	1 Clarke Quay
standard	75	Sangria Clarke Quay	65166516	1 Clarke Quay
suite	250	Sangria Clarke Quay	65166516	1 Clarke Quay
superior	145	Sangria Holland V	65166516	13 Holland Drive
standard	76	Sangria Holland V	65166516	13 Holland Drive
suite	250	Sangria Holland V	65165555	13 Holland Drive
junior suite	200	Sangria Holland V	65165555	13 Holland Drive
executive	175			

The same information is repeated, possibly unnecessarily.

Update Anomaly

type	price	name	tel	address
superior	145	Sangria Clarke Quay	65166516	1 Clarke Quay
standard	75	Sangria Clarke Quay	65166516	1 Clarke Quay
suite	250	Sangria Clarke Quay	65166516	1 Clarke Quay
superior	145	Sangria Holland V	65166516	13 Holland Drive
standard	76	Sangria Holland V	65166516	13 Holland Drive
suite	250	Sangria Holland V	65165555	13 Holland Drive
junior suite	200	Sangria Holland V	65165555	13 Holland Drive
executive	175			

The price of standard rooms is replicated and was wrongly entered in one of the replicas.

Deletion Anomaly

type	price	name	tel	address
superior	145	Sangria Clarke Quay	65166516	1 Clarke Quay
standard	75	Sangria Clarke Quay	65166516	1 Clarke Quay
suite	250	Sangria Clarke Quay	65166516	1 Clarke Quay
superior	145	Sangria Holland V	65166516	13 Holland Drive
standard	76	Sangria Holland V	65166516	13 Holland Drive
suite	250	Sangria Holland V	65165555	13 Holland Drive
junior suite	200	Sangria Holland V	65165555	13 Holland Drive
executive	175			

If Sangria Holland V stops offering junior suites, then their price disappears from the database.

Insertion Anomaly

type	price	name	tel	address
superior	145	Sangria Clarke Quay	65166516	1 Clarke Quay
standard	75	Sangria Clarke Quay	65166516	1 Clarke Quay
suite	250	Sangria Clarke Quay	65166516	1 Clarke Quay
superior	145	Sangria Holland V	65166516	13 Holland Drive
standard	76	Sangria Holland V	65166516	13 Holland Drive
suite	250	Sangria Holland V	65165555	13 Holland Drive
junior suite	200	Sangria Holland V	65165555	13 Holland Drive
executive	175			

No hotels offers executive rooms yet. We cannot store their price.

One Table

type	price	name	tel	address
superior	145	Sangria Clarke Quay	65166516	1 Clarke Quay
standard	75	Sangria Clarke Quay	65166516	1 Clarke Quay
suite	250	Sangria Clarke Quay	65166516	1 Clarke Quay
superior	145	Sangria Holland V	65166516	13 Holland Drive
standard	76	Sangria Holland V	65166516	13 Holland Drive
suite	250	Sangria Holland V	65165555	13 Holland Drive
junior suite	200	Sangria Holland V	65165555	13 Holland Drive
executive	175			

All the problems are caused by the functional dependencies

Learning Objectives

Understand the rationale (anomalies) and definition of each of the main normal forms based on functional dependencies (2NF, BCNF, 3NF)

An Example

$R = \{\text{type}, \text{price}, \text{name}, \text{tel}, \text{address}\}$

We have the following functional dependencies:

$\{\text{type}\} \rightarrow \{\text{price}\}$

and

$\{\text{name}\} \rightarrow \{\text{tel}, \text{address}\}$

The candidate and primary key is $\{\text{name}, \text{type}\}$

The left-hand side of the two functional dependencies in the minimal cover above are **proper subsets of the candidate key**. This suggests that we are dealing with **more than one entity**.

Let us forbid this ... and create **a normal form**.

(Cont.)

$\{\text{type}\} \rightarrow \{\text{type}\}$ holds on R

The candidate and primary key is $\{\text{name}, \text{type}\}$

The left-hand side of $\{\text{type}\} \rightarrow \{\text{type}\}$, $\{\text{type}\}$, is a **proper subsets of the candidate key**. But this is not a problem because $\{\text{type}\} \rightarrow \{\text{type}\}$ is a **trivial functional dependency**.

Let us discount this case...

(Cont.)

$R = \{A, B, C, D\}$

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

$\{B\} \rightarrow \{A\}$

Candidate keys $\{A, D\}$ and $\{B, D\}$

The left-hand side of $\{B\} \rightarrow \{A\}$, $\{B\}$, is a proper subset of a candidate key but there is nothing we can do about it because $\{A\}$ is a **prime attribute** (part of a candidate key).

Let us just discount this case...

Second Normal Form (2NF)

R is a relation schema, with the set F of FDs

R is in 2NF if and only if

For all X : $X \subset R$

and, for all $A \in R$

such that there exists a FD: $X \rightarrow \{A\}$ in F^+

Then

$A \in X$ ($X \rightarrow \{A\}$ is **trivial**), or

X is **NOT** a proper subset of a candidate key for R, or

A is **part of some candidate key** for R (A is a prime attribute)

Another Example

$R = \{A, B, C, D\}$

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

$\{C\} \rightarrow \{D\}$

Candidate key $\{A, B\}$

The left-hand side of $\{C\} \rightarrow \{D\}$, $\{D\}$, is **not a proper subset of a candidate key**. It is a problem because $\{D\}$ does not directly depends on $\{A, B\}$. $\{D\}$ depends **transitively** on $\{A, B\}$.

Let us forbid this ... and create **a new normal form**.

Let us make sure every non-trivial functional dependency corresponds to a candidate key (superkey, actually).

Boyce-Codd Normal Form (BCNF)

R is a relation schema, with the set F of FDs

R is in BCNF if and only if

For all X : $X \subset R$

And, for all $A \in R$

such that there exists a FD: $X \rightarrow \{A\}$ in F^+

Then

$A \in X$ ($X \rightarrow \{A\}$ is **trivial**), or

X is a **superkey** for R

Another Example

$R = \{A, B, C, D\}$

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

$\{B\} \rightarrow \{A\}$

Candidate keys $\{A, D\}$ and $\{B, D\}$

in $\{B\} \rightarrow \{A\}$ $\{B\} \rightarrow \{A\}$ $\{B\}$ is not a **superkey**. $\{A\}$ is a **prime attribute** (part of a candidate key).

Let us discount this case... and **create a new normal form**.

Because we cannot untangle the functional dependencies.

Third Normal Form (3NF)

R is a relation schema, with the set F of FDs

R is in 3NF if and only if

For all $X: X \subset R$

And, for all $A \in R$

such that there exists a FD: $X \rightarrow \{A\}$ in F^+

Then

$A \in X$ ($X \rightarrow \{A\}$ is **trivial**), or

X is a **superkey** for R, or

A is **part of some candidate key** for R

(A is called a prime attribute)

BCNF \subset 3NF \subset 2NF

BCNF:

Trivial, or

X is a superkey for R

3NF:

Trivial, or

X is a superkey for R, or

A is part of some candidate key for R

2NF:

Trivial, or

X is not a proper subset of a candidate key for R, or

A is part of some candidate key for R

Decomposition

name	tel	address
Sangria Clarke Quay	65166516	1 Clarke Quay
Sangria Holland V	65165555	13 Holland Drive

type	price
superior	145
standard	75
suite	250
junior suite	200
executive	175

type	name
superior	Sangria Clarke Quay
standard	Sangria Clarke Quay
suite	Sangria Clarke Quay
superior	Sangria Holland V
standard	Sangria Holland V
suite	Sangria Holland V
junior suite	Sangria Holland V

Credits

Copyright © 2018 by Stéphane Bressan

The content of this lecture is based
on the book “Introduction to
database Systems”

By
S. Bressan and B. Catania, McGraw
Hill publisher

Images and clips used in this
presentation are licensed from
Microsoft Office Online Clipart and
Media

For questions about the content of
this course and about copyrights,
please contact Stéphane Bressan

steph@nus.edu.sg

