

<epam>

# Normal Forms — Part 2

Relational Databases Basics

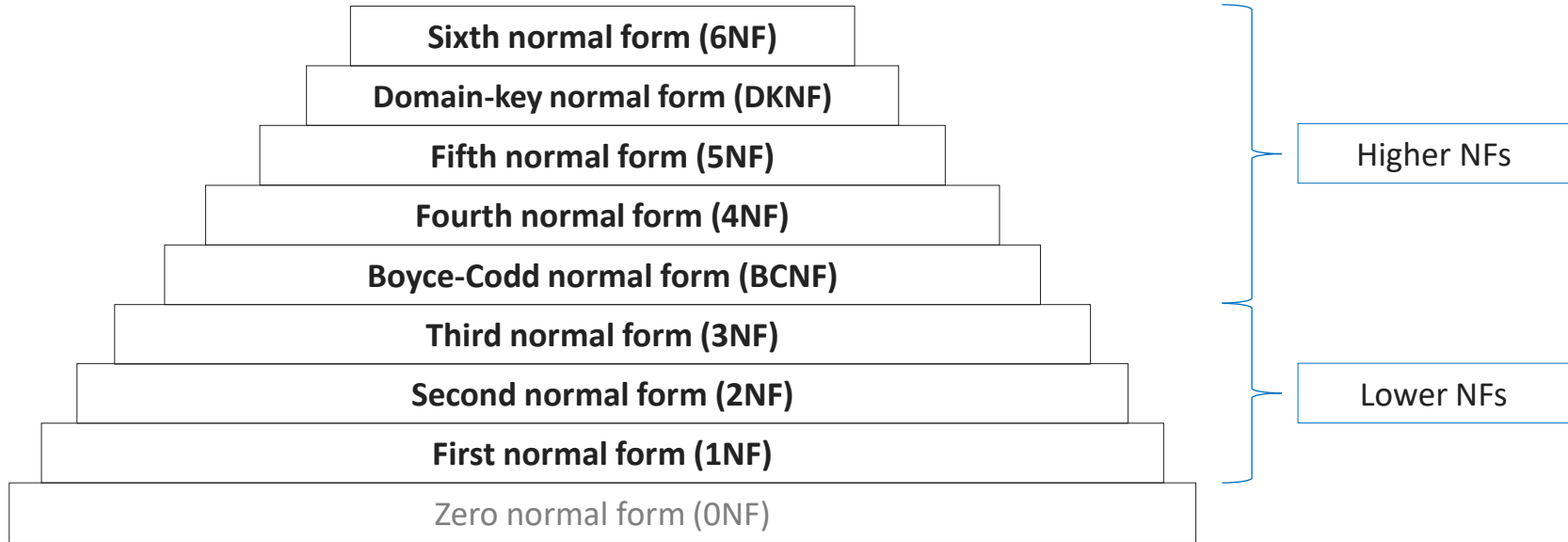


**TRAINING**  
CENTER

— <epam> —

# Normal forms hierarchy

---



A relvar is in **3NF** if it satisfies 2NF and has no nonprime attribute that is transitively dependent on the primary key.

*In other words: there must not be an attribute that is not a part of PK and transitively dependent on PK.*

### 3NF: third normal form (violation example)

Nonprime attributes “g\_dean”, “g\_dean\_dob” are transitively dependent on candidate keys and thus should be moved to another relation

group

<u>g_id</u> PK	<u>g_number</u> Alternate key	<u>g_start_year</u>	<u>g_faculty</u>	<u>g_dean</u>	<u>g_dean_dob</u>
1	1	1998	Physics	John Smith	1974-01-12
2	1	1999	Physics	John Smith	1974-01-12
3	1	2000	Mathematics	Joe Black	1984-12-23
4	2	2000	IT	Jane Dow	1985-05-28
5	1	2008	IT	Jane Dow	1985-05-28
6	2	2008	Chemistry	Walter White	1969-02-20

$\{g\_id\} \rightarrow \{g\_faculty\} \rightarrow \{g\_dean\} \rightarrow \{g\_dean\_dob\}$   
 $\{g\_number, g\_start\_year\} \rightarrow \{g\_faculty\} \rightarrow \{g\_dean\} \rightarrow \{g\_dean\_dob\}$

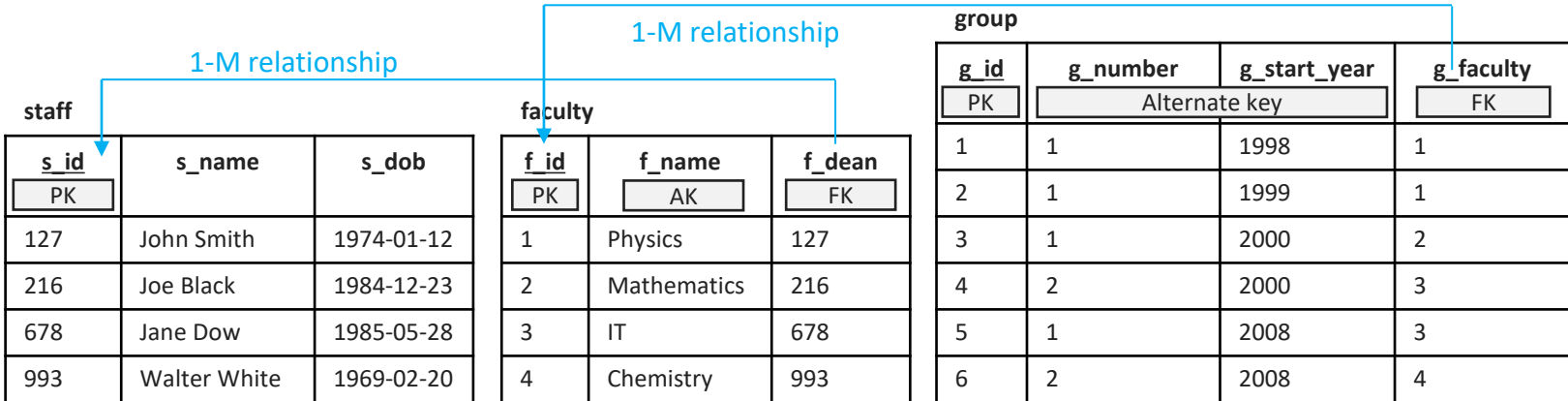
# 3NF: third normal form (dealing with violations)

Before

group

<u>g_id</u> PK	g_number	g_start_year	g_faculty	g_dean	g_dean_dob
Alternate key					
1	1	1998	Physics	John Smith	1974-01-12
2	1	1999	Physics	John Smith	1974-01-12
3	1	2000	Mathematics	Joe Black	1984-12-23
4	2	2000	IT	Jane Dow	1985-05-28
5	1	2008	IT	Jane Dow	1985-05-28
6	2	2008	Chemistry	Walter White	1969-02-20

After



## 3NF: third normal form (conclusion)

---



3NF doesn't protect you from “local mistakes”, but it keeps you from violating “global rules”.



One of the most obvious evidence of 3NF violation is that a relation has some attributes with multiple values duplication (for many tuples). Such attributes are the first candidates to moving to another relation.

A relvar is in **BCNF** if it satisfies 2NF and has no attribute that is transitively dependent on the primary key.

*In other words: there must not be an attribute that is transitively dependent on PK.*

Compare to 3NF:

A relvar is in **3NF** if it satisfies 2NF and has no ~~nonprime~~ attribute that is transitively dependent on the primary key.

*In other words: there must not be an attribute that is ~~not a part of PK and~~ transitively dependent on PK.*

## BCNF: Boyce-Codd normal form (violation example)

**gift**

PK		
<u>g_set</u>	<u>g_type</u>	<u>g_name</u>
AK...		...AK
Set 1	Toys	Teddy Bear
Set 1	Sweets	Marmalade
Set 1	Fruits	Oranges
Set 2	Toys	Teddy Bear
Set 2	Sweets	Marmalade
Set 3	Toys	Teddy Bear

“g\_name” attribute is dependent on “g\_type” attribute  
(while “g\_type” itself is not a candidate key)

This statement holds due to the following subject matter rules:

- No set may contain two or more items of the same type.
- No set may contain two or more items of the same name.
- No item type may contain two or more items.

$\{g\_set, g\_type\} \rightarrow \{g\_name\}$

$\{g\_type\} \rightarrow \{g\_name\}$



# BCNF: Boyce-Codd normal form (dealing with violations)

**Before**

**gift**

PK		g_name ...AK
<u>g_set</u>	<u>g_type</u>	
AK...		
Set 1	Toys	Teddy Bear
Set 1	Sweets	Marmalade
Set 1	Fruits	Oranges
Set 2	Toys	Teddy Bear
Set 2	Sweets	Marmalade
Set 3	Toys	Teddy Bear

**After**

**set**

PK	
<u>s_name</u>	<u>s_o_type</u>
	FK
Set 1	Toys
Set 1	Sweets
Set 1	Fruits
Set 2	Toys
Set 2	Sweets
Set 3	Toys

1-M relationship

**object**

PK	g_name
<u>g_type</u>	
Toys	Teddy Bear
Sweets	Marmalade
Fruits	Oranges

## BCNF: Boyce-Codd normal form (conclusion)

---



BCNF (like 3NF) doesn't protect you from “local mistakes”, but it keeps you from violating “global rules”.



One of the most obvious evidence of BCNF violation is that a relation has candidate keys intersection.



Most relations that satisfy 3NF also satisfy BCNF, because BCNF violation require specific subject matter rules.

A relvar is in **4NF** if it satisfies BCNF and any multi-valued dependency that holds in this relation is implied by some superkey.

*In other words: there must not be non-trivial multi-valued dependencies.*

## 4NF: fourth normal form (violation example)

One applicant may apply to several faculties, i.e.:

$\{ua\_applicant\} \twoheadrightarrow \{ua\_faculty\}$

One faculty has several entrance exams, i.e.:

$\{ua\_faculty\} \twoheadrightarrow \{ua\_exam\}$

**university\_application**

<u>ua_applicant</u>	<u>ua_faculty</u>	<u>ua_exam</u>
PK		
Ivanov I.I.	Math. faculty	Computer science
Ivanov I.I.	Math. faculty	Mathematics
Ivanov I.I.	Phys. faculty	Computer science
Ivanov I.I.	Phys. faculty	Physics
Petrov P.P.	Math. faculty	Computer science
Petrov P.P.	Math. faculty	Mathematics
Sidorov S.S.	Phys. faculty	Computer science
Sidorov S.S.	Phys. faculty	Physics

For each applicant we have to add as many rows, as many exams there are on a faculty, this applicant applies to.

# 4NF: fourth normal form (dealing with violations)

## Before

university\_application

<u>ua_applicant</u>	<u>ua_faculty</u>	<u>ua_exam</u>
PK		
Ivanov I.I.	Math. faculty	Computer science
Ivanov I.I.	Math. faculty	Mathematics
Ivanov I.I.	Phys. faculty	Computer science
Petrov P.P.	Phys. faculty	Physics
Petrov P.P.	Math. faculty	Computer science
Petrov P.P.	Math. faculty	Mathematics
Sidorov S.S.	Phys. faculty	Computer science
Sidorov S.S.	Phys. faculty	Physics

For each applicant we have to add as many rows, as many exams there are on a faculty, this applicant applies to.

## After

application

<u>a_applicant</u>	PK	<u>a_faculty</u>
Ivanov I.I.		Math. faculty
Ivanov I.I.		Phys. faculty
Petrov P.P.		Phys. faculty
Petrov P.P.		Math. faculty
Sidorov S.S.		Phys. faculty

exam

<u>e_faculty</u>	PK	<u>e_exam</u>
Math. faculty		Computer science
Math. faculty		Mathematics
Phys. faculty		Computer science
Phys. faculty		Physics

No relationships here!

## 4NF: fourth normal form (dealing with violations, more realistic)

### After

**applicant**

<u>a_id</u> PK	a_name
1	Ivanov I.I.
2	Petrov P.P.
3	Sidorov S.S.

**faculty**

<u>f_id</u> PK	f_name
1	Math. faculty
2	Phys. faculty

**exam**

<u>e_id</u> PK	e_name
1	Computer science
2	Mathematics
3	Physics

1-M

**m2m\_applicant\_faculty**

<u>a_id</u> PK	<u>f_id</u>
1	1
1	2
2	1
2	2
3	2

1-M

1-M

**m2m\_faculty\_exam**

<u>f_id</u> PK	<u>e_id</u>
1	1
1	2
2	1
2	3

1-M

## 4NF: fourth normal form (conclusion)

---



4NF (like BCNF) doesn't protect you from “local mistakes”, but it keeps you from violating most “global rules”.



One of the most obvious evidence of 4NF violation is that a relation has a group of three or more attributes with pairwise dependencies inside such group.



Most relations that satisfy BCNF also satisfy 4NF, because 4NF violation require specific subject matter rules.

<epam>

# Normal Forms — Part 2

Relational Databases Basics



**TRAINING**  
CENTER

— <epam> —