

# ***Normalization 2: Boyce Codd Normal Form***

---

# Example

---

Suppose instructors can have multiple offices.  
We might have this relation:

inst\_offices(ID, name, office)

For example:

ID	name	office
gill1992	Harman	DH282
gill1992	Harman	MH999
benr9431	Ben	MH213

Is there redundancy here?

# Identifying the redundancy

---

ID	name	office
gill1992	Harman	DH282
gill1992	Harman	MH999
benr9431	Ben	MH213

We can see 'Harman' multiple times, and know that ID determines the name.

**Question:** Would the fact that ID  $\rightarrow$  name cause redundancy if ID were a key in this table?

# Another Example

---

Let's look at students and their advisers:

advises(inst\_ID, student\_ID, student\_name)

For example:

inst_ID	student_ID	student_name
gill1992	john4320	John Harper
gill1992	vlim5230	Victor Lim
benr9431	john4320	John Harper

Is there redundancy here?

# Identifying the redundancy

---

inst_ID	student_ID	student_name
gill1992	john4320	John Harper
gill1992	vlim5230	Victor Lim
benr9431	john4320	John Harper

Redundancy can exist if two things are present.

**Question:** What are they?

1. a functional dependency exists
2. the possibility of multiple rows involving the source of the multiple dependencies

# Conditions for redundancy

---

inst_ID	student_ID	student_name
gill1992	john4320	John Harper
gill1992	vlim5230	Victor Lim
benr9431	john4320	John Harper

Two conditions are present:

1. a functional dependency:

$\text{student\_ID} \rightarrow \text{student\_name}$

2. the possibility of multiple rows:

student\_ID is not a superkey

Remember:  
a functional  
dependency  
involves two  
sets of  
attributes of  
a single  
table

# Boyce Codd Normal Form (BCNF)

---

A relation schema  $R$  is in **Boyce Codd Normal Form** if:

- whenever there's a (non-trivial) functional dependency  $X \rightarrow Y$  for  $R$
- then  $X$  is be a superkey for  $R$

This is exactly our idea from the examples.

To show a relation schema is not in BCNF:

- just show a FD  $X \rightarrow Y$  such that  $X$  is not a superkey

# Question

---

Here is the “instructor offices” schema

inst\_offices(ID, name, office)

Is it in Boyce Codd Normal Form?

- a) yes
- b) no, because of FD  $ID \rightarrow name$
- c) no, because of FD  $ID, office \rightarrow name$
- d) no, because of FD  $ID \rightarrow name$  and because  $ID$  isn't a superkey



# Fixing redundancy by “splitting”

---

This schema isn't in BCNF:

ID	name	office
gill1992	Harman	DH282
gill1992	Harman	MH999
benr9431	Ben	MH213

So,  
problem  
solved?

These schemas are in BCNF:

ID	name
gill1992	Harman
benr9431	Ben

ID	office
gill1992	DH282
gill1992	MH999
benr9431	MH213

# How is table-splitting a solution?

---

## Idea:

Splitting a table into two seems okay if the original table could be reconstructed from the two new tables

## Problem:

Can you reconstruct any table no matter how you split it up?

# “Reconstructing” a table

ID	name
gill1992	Harman
benr9431	Ben

natural  
join

ID	office
gill1992	DH282
gill1992	MH999
benr9431	MH213

ID	name	office
gill1992	Harman	DH282
gill1992	Harman	MH999
benr9431	Ben	MH213

This is okay, but can you always reconstruct after splitting?

# “Reconstructing” a table, part 2

split

a	b	c
1	2	3
2	2	4

a	b
1	2
2	2

b	c
2	3
2	4

natural  
join

a	b	c
1	2	3
2	2	3
1	2	4
2	2	4

This is not okay.

If we split a table, we need to make sure we can reconstruct it perfectly.

“Lossless decomposition”