Lecture 12

Normalization Continued...

Boyce-Codd Normal Form

A relation is in Boyce-Codd normal form (BCNF) if every determinant of a non-trivial FD is a superkey.

The difference between 3NF and BCNF is that 3NF allows a FD to remain in the relation if determinant is non-prime attribute.

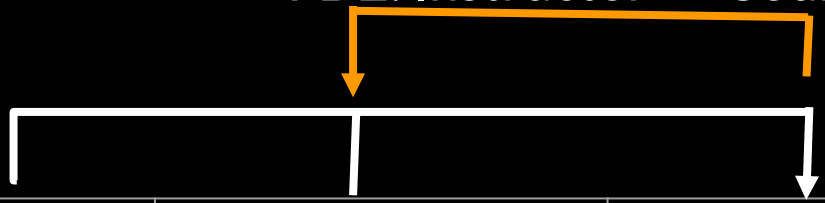
How to Check for BCNF

To test if a relation is in BCNF, we take the determinant of each non-trivial FD in the relation and check if it is a superkey.

Example

FD1: {Student, Course} → Instructor

FD2: Instructor → Course



1NF	Atomicity
2NF	Full F.D.
3NF	Transitive Property
BCNF	?

Student	Course	Instructor
Albert	Database	Mark
Ben	Database	Andrew
Ben	Operating System	Chris
Charlie	Database	Mark
Dan	Operating System	Charles

How to Achieve BCNF?

R1 (Student, Instructor) and R2(Student, Course)

Student	Course
Albert	Database
Ben	Database
Ben	Operating System
Charlie	Database
Dan	Operating System

Student	Instructor
Albert	Mark
Ben	Andrew
Ben	Chris
Charlie	Mark
Dan	Charles

Loss of FD1 !!!

Other Decomposition Possibilities

R1 (Course, Instructor) and R2(Course, Student)

R1 (Instructor, Course) and R2(Instructor, Student)

None of them preserves FD1 !!!

The decision to use 3NF or BCNF depends on the amount of redundancy we are willing to accept and the willingness to lose a functional dependency

BCNF - Point to Remember

With a BCNF decomposition we do not always get dependency preservation, but we can always preserve the lossless-join property (recovery)

Does this Decomposition Preserve Lossless-Join Property ?

R1 (Student, Instructor) and R2(Student, Course)

Student	Course
Albert	Database
Ben	Database
Ben	Operating System
Charlie	Database
Dan	Operating System

Student	Instructor
Albert	Mark
Ben	Andrew
Ben	Chris
Charlie	Mark
Dan	Charles

Do Following Decompositions Preserve Lossless-Join Property

R1 (Course, Instructor) and R2(Course, Student)

R1 (Instructor, Course) and R2(Instructor, Student)

Student	Course	Instructor
Albert	Database	Mark
Ben	Database	Andrew
Ben	Operating System	Chris
Charlie	Database	Mark
Dan	Operating System	Charles

BCNF - Valid Decomposition

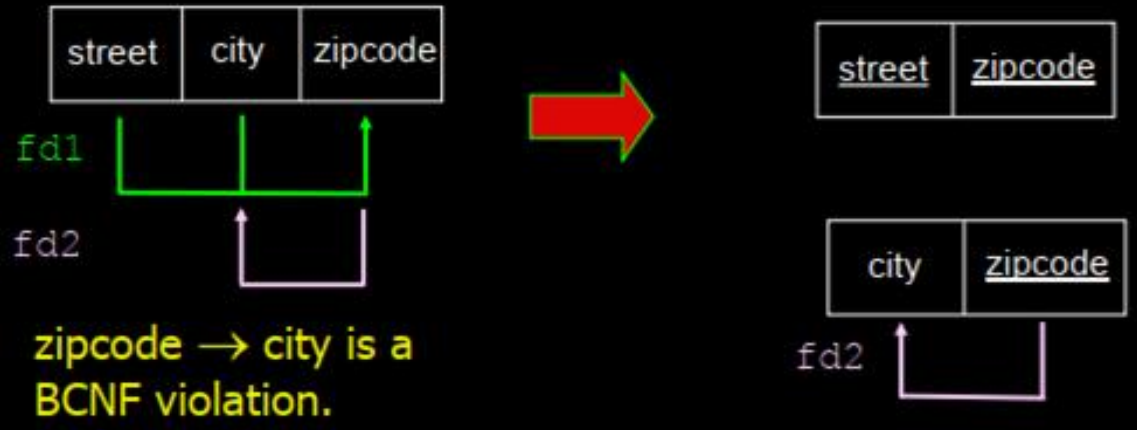
R1 (Instructor, Course)

R2(Instructor, Student)

Another Example

An example of not having dependency preservation with BCNF:

- street,city → zipcode and zipcode → city
- Two keys: {street,city} and {street, zipcode}



More Normalization

Fourth normal form (4NF) and fifth normal formal (5NF) are rarely used in practice.

A relation is in fourth normal form (4NF) if it is in BCNF and contains no non-trivial multi-valued dependencies.

SUPPLY

Sname	Part_name	Proj_name
Smith	Bolt	ProjX
Smith	Nut	ProjY
Adamsky	Bolt	ProjY
Walton	Nut	ProjZ
Adamsky	Nail	ProjX
Adamsky	Bolt	ProjX
Smith	Bolt	ProjY

R_1

Sname	Part_name
Smith	Bolt
Smith	Nut
Adamsky	Bolt
Walton	Nut
Adamsky	Nail

R_2

Sname	Proj_name
Smith	ProjX
Smith	ProjY
Adamsky	ProjY
Walton	ProjZ
Adamsky	ProjX

R_3

Part_name	Proj_name
Bolt	ProjX
Nut	ProjY
Bolt	ProjY
Nut	ProjZ
Nail	ProjX

5th Normal Form

A relation is in fifth normal form (5NF) if the relation has no join dependencies.

What is a join dependency ???

Google !!!

A join dependency implies that spurious tuples are generated when the relations are natural joined.

In Conclusion

What is a functional dependency?
- Trivial Dependency ?
- Prime Attribute?
- Full Functional Dependency?

Normalization
- 1st Normal Form
- 2nd Normal Form
- 3rd Normal Form