CIS560

Obtaining a Good Database Design – Part 3

1

What we've seen so far:

- We briefly discussed the 3NF and BCNF.
- They are defined using:
  - Functional Dependencies
  - Keys
- We defined functional dependencies.
- We defined closures and how they help
  - Find all functional dependencies
  - Determine whether a dependency violates BCNF

2

What we've seen so far:

- A **superkey** is a set of attributes $A_1, ..., A_n$ s.t. for any other attribute B, we have $A_1, ..., A_n \rightarrow B$
- A **key** is a minimal superkey
  - A set of attributes which is a superkey
  - And for which no subset is a superkey
- We can decompose "bad" relations into BCNF relations.

Boyce-Codd Normal Form

- A relation R is in BCNF if and only if for every functional dependency $X \rightarrow A$:
  - $X \rightarrow A$ is a trivial functional dependency
    or
  - X is a superkey for R

- Equivalently: $\forall$ X, either
  $(X^+ = X)$ or $(X^+ =$ all attributes$)$

A closer look at decompositions

Name	Price	Category
iPad	529	Tablet
iPhone	429	Gadget
iPad	529	Gadget

Name	Price
iPad	529
iPhone	429
~~iPad~~	~~529~~

Name	Category
iPad	Tablet
iPhone	Gadget
iPad	Gadget

Lossless decomposition

A closer look at decompositions

Name	Price	Category
iPad	529	Tablet
iPhone	429	Gadget
iPad	529	Gadget

What's incorrect?

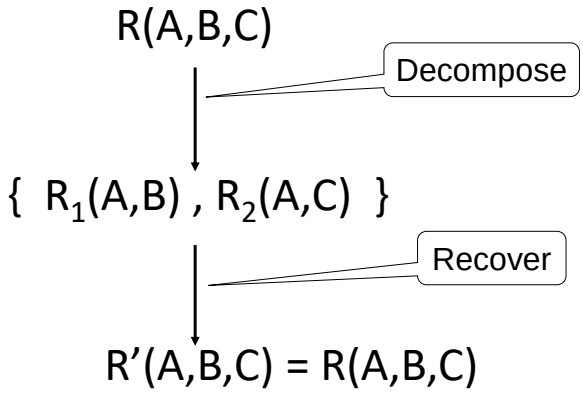
Name	Category
iPad	Tablet
iPhone	Gadget
iPad	Gadget

Price	Category
529	Tablet
429	Gadget
529	Gadget

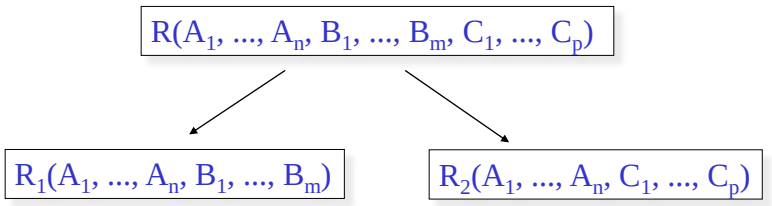
Lossy decomposition

Lossless Decompositions

A decomposition is *lossless* if we can recover the exact information we started with:



Decompositions in General



If $A_1, ..., A_n \rightarrow B_1, ..., B_m$
then the decomposition is lossless

A **BCNF** decomposition is **always lossless**.

A Problem with BCNF?

Professor	Project	Department

FD's: Professor → Department; Project, Department → Professor

So, there is a BCNF violation, and we decompose.

Professor	Department

Professor → Department

Professor	Project

No FDs

In BCNF we lose the FD: Project, Department → Professor

9

So what's the problem?

Professor	Department
Johnson	CIS
Robinson	CIS

Professor	Project
Johnson	Recruitment
Robinson	Recruitment

No problem so far. All *local* FD's are satisfied.

Let's put all the data back into a single table again:

Professor	Department	Project
Johnson	CIS	Recruitment
Robinson	CIS	Recruitment

Violates the dependency: Project, Department → Professor!

10

Preserving Functional Dependencies

- We lose dependencies when a relation with dependency $X \rightarrow Y$ is decomposed and:
  - X ends up in one of the new relations
  - Y ends up only in another
- Such a decomposition is not "dependency-preserving."
- Common form is $AB \rightarrow C$ and $C \rightarrow B$
  - Remember our example?
  - Professor $\rightarrow$ Department
  - Project, Department $\rightarrow$ Professor

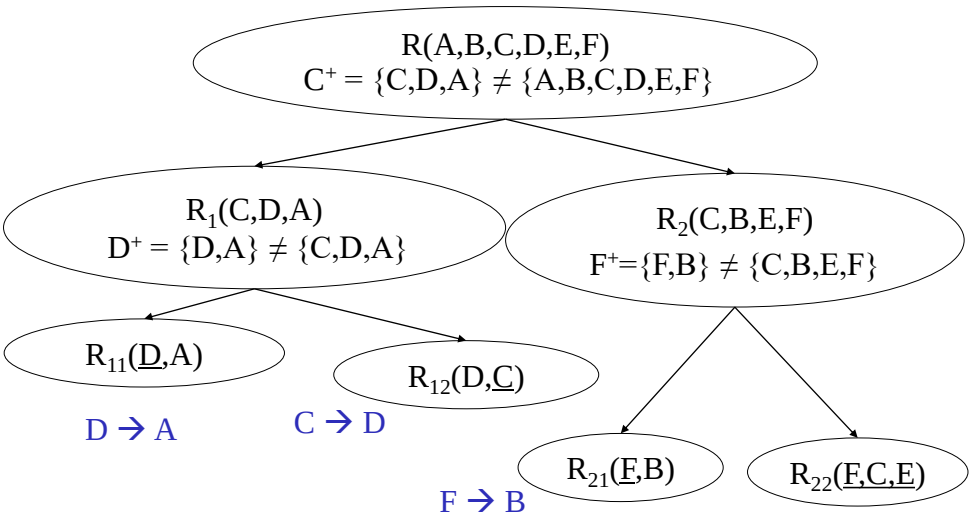
BCNF decomposition does not always preserve dependencies.

Example

R(A,B,C,D,E,F)

AB → C
C → D
F → B
D → A

~~AB → C~~



13

General Decomposition Goals

- Eliminate anomalies
  - Redundancy, update, and delete anomalies
- Recoverability of information
  - Can we get the original relation back?
- Preservation of dependencies
  - Can we enforce the functional dependencies **without** performing joins?

14

BCNF Decompositions

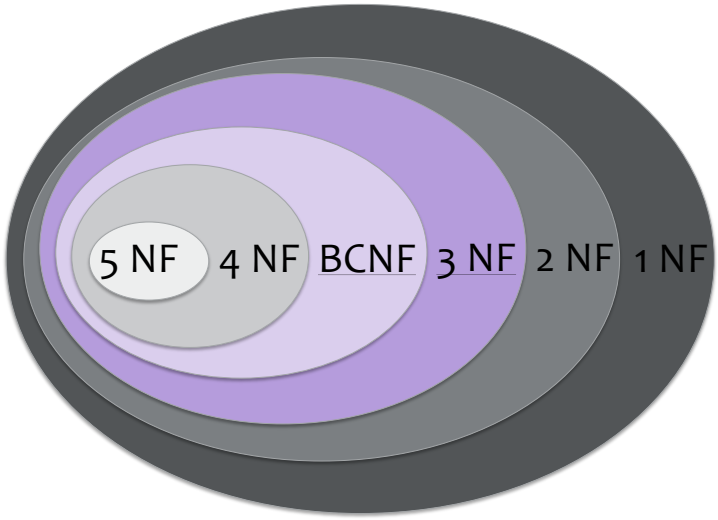
- No anomalies 
- Recoverability of information 
- Sometimes may lose dependencies 

15

What to do?

16

There are other normal forms



17

More Terms

- Candidate Key
  Another name for a minimal superkey
- Prime Attributes
  Attributes of a candidate key
- Non-Prime Attributes
  Do not occur in *ANY* candidate key

18

Normalization 1

Simple attributes

Origin	Country
Liverpool	UK

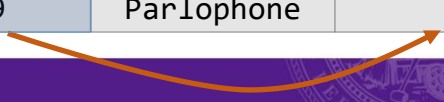
19

Normalization 2

if composite key:

all non-prime attributes depend on the full key

Album	Artist	Label	ArtistCountry
Please Please Me	9	Parlophone	UK



20

Normalization

3

non-prime attributes
not dependent on each other

Album	Artist	Year	Studio	StudioCountry
Please Please Me	9	1963	Abbey Road	UK

21

Every non-key attribute must
provide a fact about the key,
the whole key,
and nothing but the key.

William Kent (1936 – 2005)

22

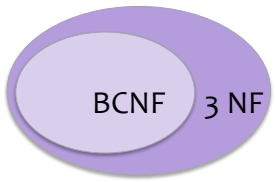
Third Normal Form (3NF)

A relation R is in the third normal form if:

For every nontrivial dependency in R where
$A_1, A_2, ..., A_n \rightarrow B$,
$\{ A_1, A_2, ..., A_n \}$ must be a superkey for R,
or B is part of a key.

3NF vs. BCNF

- R is in BCNF if for every nontrivial FD
  $A_1, A_2, ..., A_n \rightarrow B$,
  then $\{A_1, A_2, ..., A_n\}$ is a superkey.
- BCNF is slightly stronger than 3NF.



- Example: R(A,B,C) with {A,B}→C, C→B
  - 3NF but not BCNF (B is part of the key)

3NF Decompositions

- Recoverability of information 
- Preservation of dependencies 
- May still have anomalies 

Practical advise

Aim for BCNF

Settle for 3NF

In Conclusion of Learning BCNF/3NF

- How can we improve this?
  - Order Number
  - Order Date
  - Customer Name
  - Billing address
  - Product Name
  - SKU
  - Product Category
  - Quantity
  - Unit Price