Normalization

Orlando Karam
okaram@spsu.edu

CC BY SA

SOUTHERN POLYTECHNIC STATE UNIVERSITY

Anomalies

- If we combine information from two tables in the same, we get **anomalies**
- Insert
- Delete
- Update

Id	Name	email
111	Orlando	ok@ok.com
111	Orlando	ok@spsu.edu
222	Lina	lc@spsu.edu

Functional Dependency

- A->B means if two rows have same value of A they necessarily have same value of B
  - –Notice A and B are *sets* of attributes !
- Generalization of candidate key - candidate key says no two rows can have same value
- FDs are determined by **semantics**, you can't say an FD exists just by looking at data
- But can say whether it is NOT true by looking at data

Quick Check

Id	Name	Gender	Age
1	Orlando	Male	35
2	John	Male	35
3	Jane	Female	31
4	Jane	Female	30

Id -> Name ?

Age -> Gender?

Name-> Id?

Name, Age -> Id?

Quick Check

Id	Name	Gender	Age
1	Orlando	Male	35
2	John	Male	35
3	Jane	Female	31
4	Jane	Female	30

Id -> Name ? YES it *could* be true (no two rows with same id)

Age -> Gender? YES it *could* be true (the two rows with same age also have same gender). Do you think it would be true in the real world ?

Name -> Id? NO, 3rd and 4th row, same name (jane) but different Id

Name, Age -> Id? YES, no two rows with same name and same age

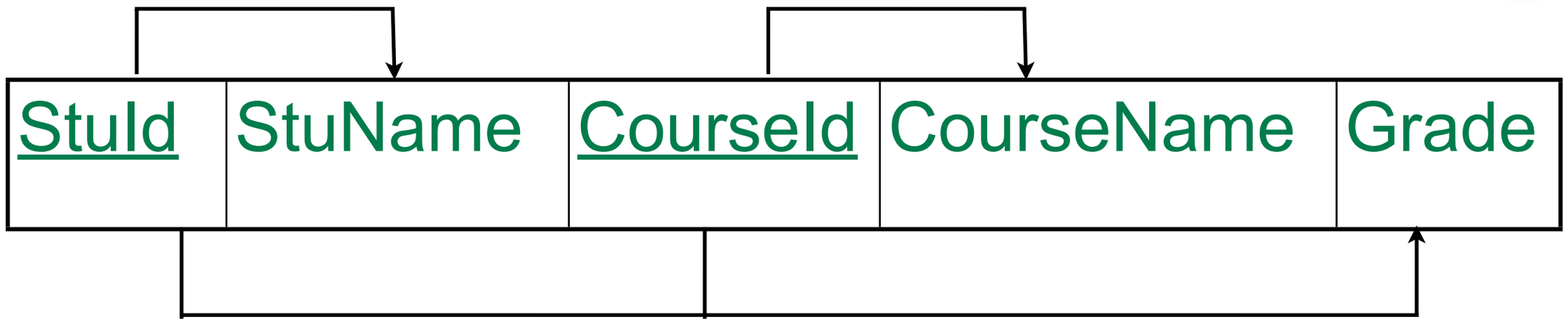
First normal form

- Only atomic attributes (simple, single-valued)
- Given, since we can't represent tables with non-atomic attributes
- Every *relation* is in 1NF by definition

Second normal form

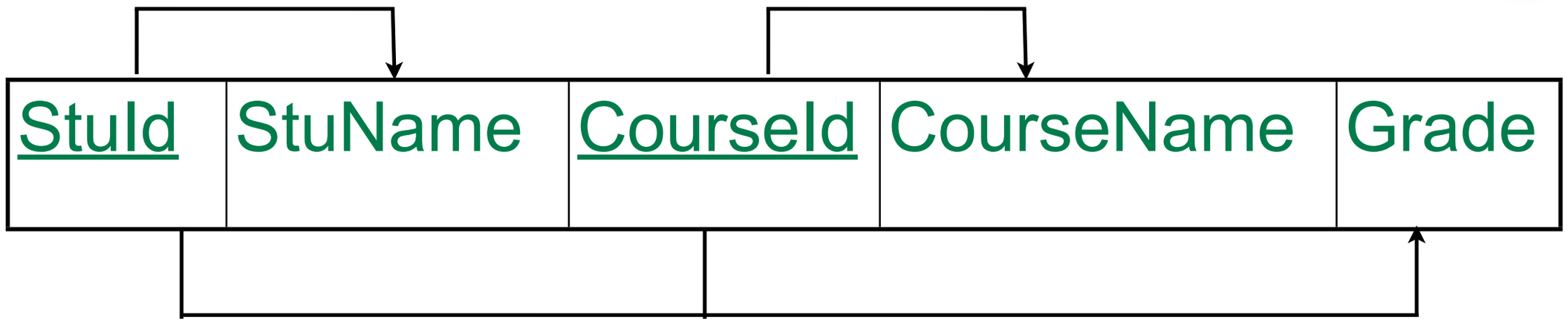
- 1NF and no *partial* (functional) dependencies
- A partial dependency is when a non-key attribute depends on a *part* of the primary key
- Of course, you need a composite PK

2NF Example



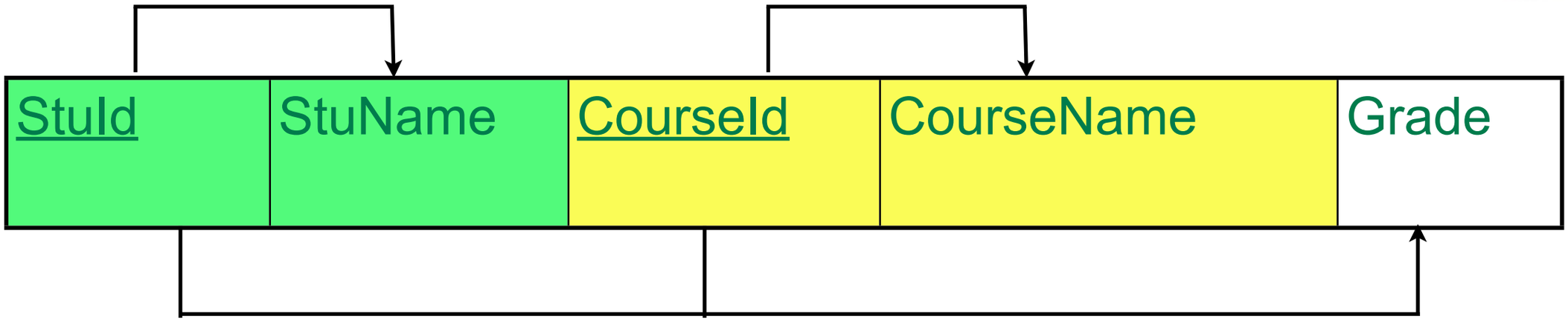
- Can represent FDs with arrow as above or
  - –StuId->StuName,
  - – CourseId->CourseName
  - –StuId,CourseId -> Grade (and StuName, CourseName)
- Any partial FDs ?

2NF Example (again)

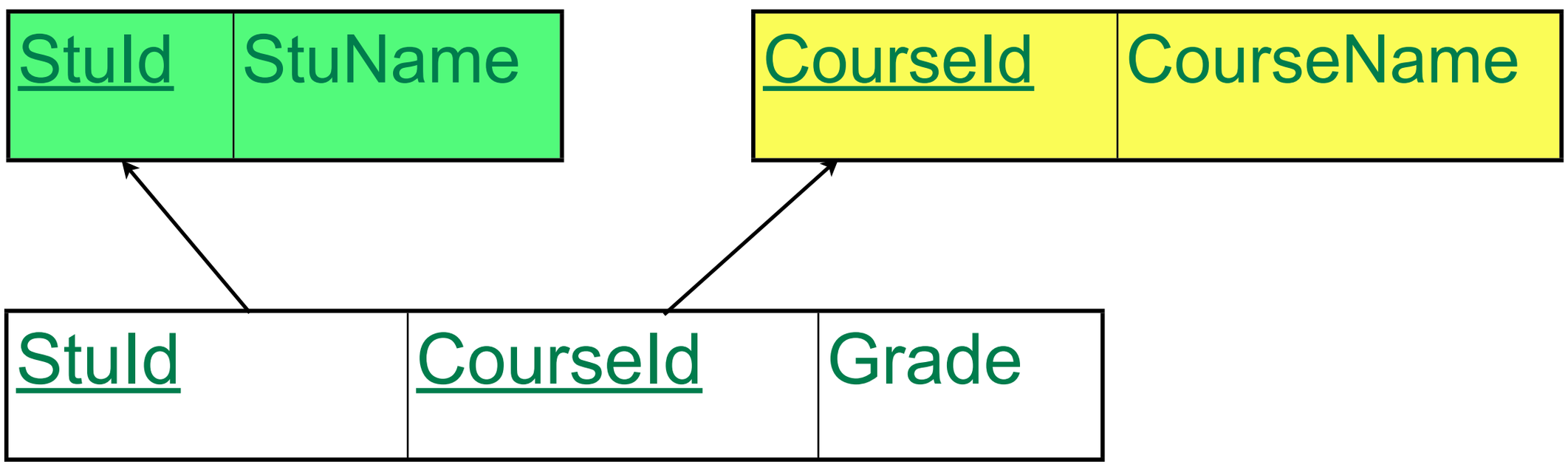


- Can represent FDs with arrow as above or
  - **StuId->StuName** **PARTIAL**
  - **CourseId->CourseName** **PARTIAL**
  - StuId,CourseId -> Grade (and StuName, CourseName)
- How do we convert the partial dependencies into normal ones ? By breaking into more tables

2NF: Normalizing



Becomes ... (notice above arrows mean functional dependency, below they represent foreign key constraints)



You try it now

SeriesId	EpisodeId	SeriesTitle	EpisodeTitle	AiringDate

- FDs
  - SeriesId -> SeriesTitle
  - SeriesId,EpisodeId -> SeriesTitle, EpisodeTitle, AiringDate
- Any partial dependencies ? How do we eliminate them ?

Solution

SeriesId	EpisodeId	SeriesTitle	EpisodeTitle	AiringDate

- FDs
  - –SeriesId -> SeriesTitle PARTIAL
  - –SeriesId,EpisodeId -> SeriesTitle, EpisodeTitle, AiringDate
- How do we eliminate the partial FD? Make new table !

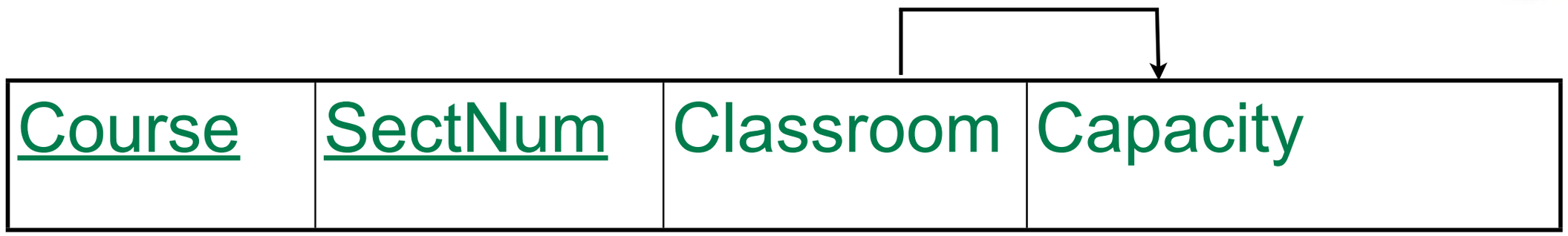
SeriesId	EpisodeId	EpisodeTitle	AiringDate

SeriesId	SeriesTitle

Third normal form

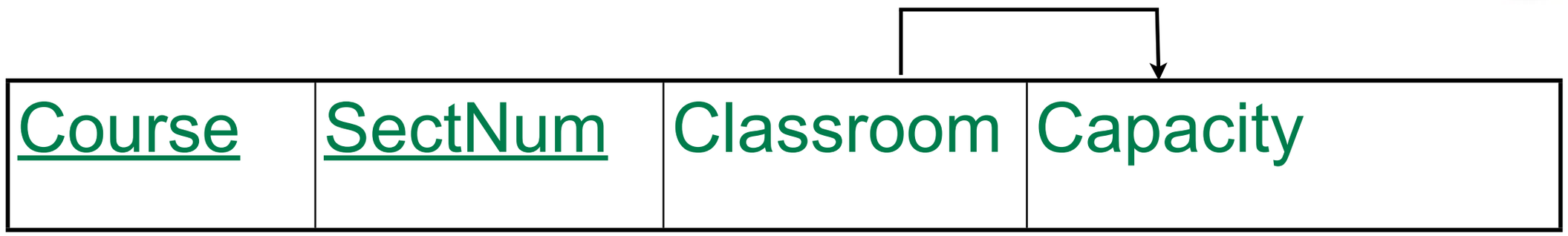
- 2NF and no *transitive* (functional) dependencies
- A *transitive* dependency is when a non-key attribute depends on another non-key attribute

3NF Example



- Classroom->Capacity
- Any partial FDs ?
- Any transitive FDs ?

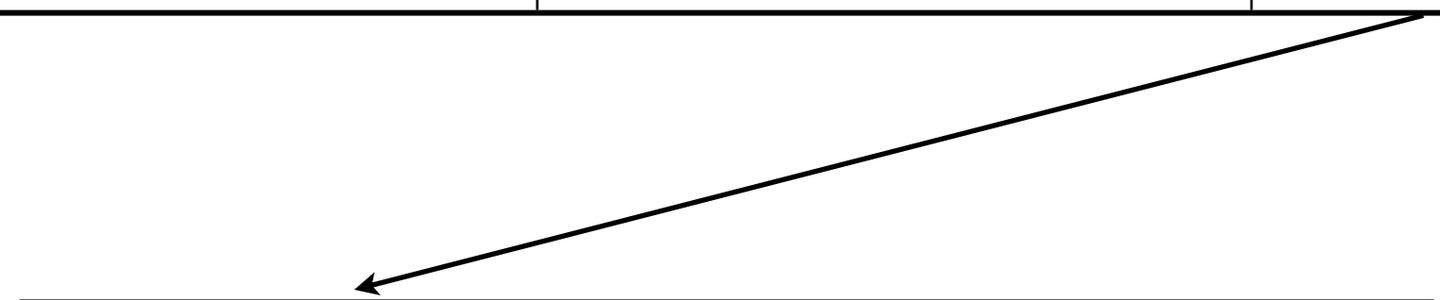
3NF Example



- Classroom->Capacity TRANSITIVE
- Any partial FDs ? NO
- Any transitive FDs ? Yes ! how do we eliminate ? by breaking into its own table

3NF Normalization

Course	SectNum	Classroom



Classroom	Capacity

You try it

StudentId	ProgramId	StudentName	ProgramName

- ProgramId -> ProgramName
- Any Partial FDs ?
- Any Transitive FDs?

Solution

StudentId	ProgramId	StudentName	ProgramName

- ProgramId -> ProgramName TRANSITIVE
- Any Partial FDs ? NO
- How to eliminate transitive FD ? New table !

StudentId	ProgramId	StudentName

StudentId	ProgramName

Further Normalization

- Boyce-Codd Normal form (BCNF)
  - Slight difference with 3NF
  - To be in 3NF but not in BNF, needs two composite candidate keys, with one attribute of one key depending on one attribute of the other
  - Not very common :)
- Fourth Normal Form (4NF)
  - To break it, need to have two independent multivalued attributes in the same table
- Usually, if you're in 3NF you're in BCNF, 4NF, …