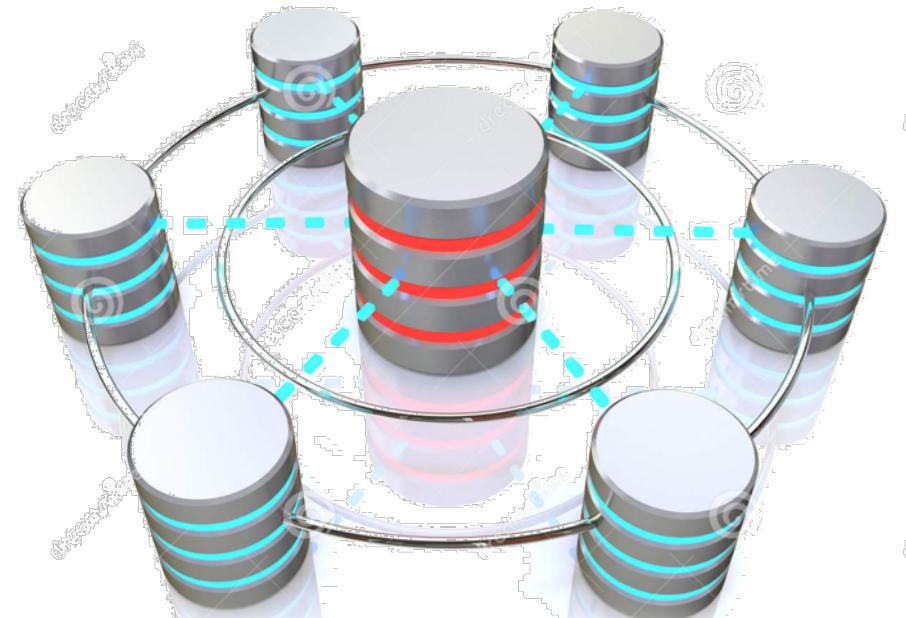




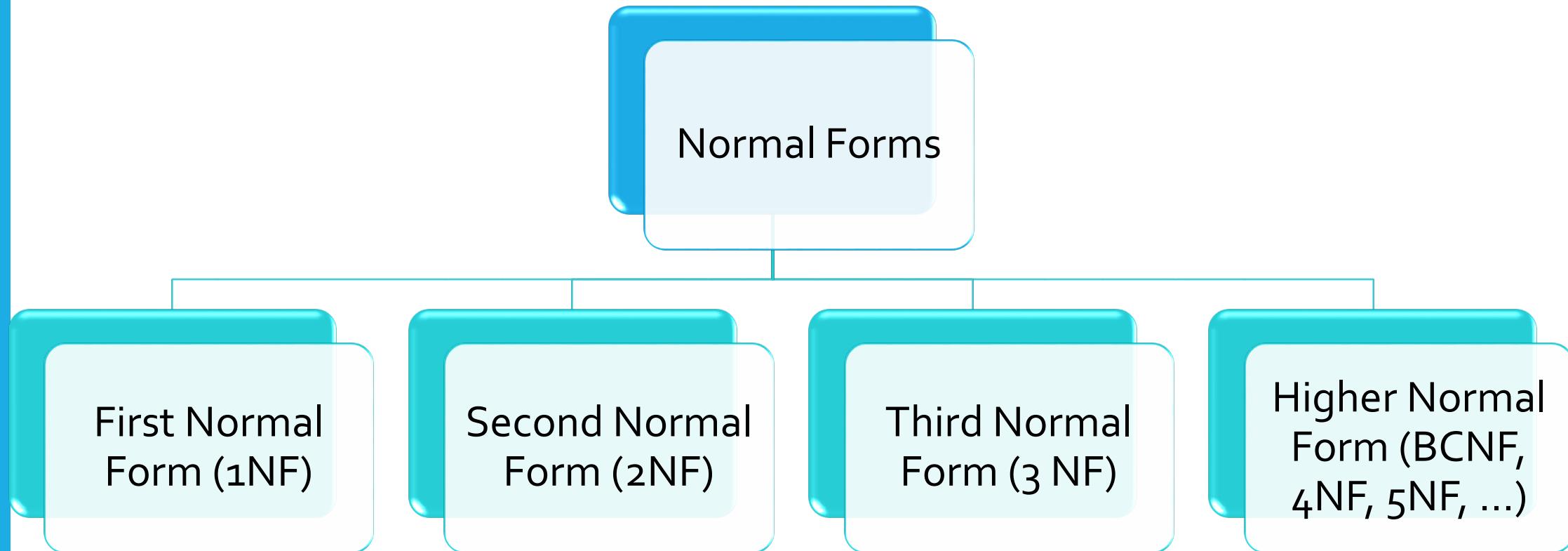
NORMALIZATION

NORMALIZATION AND NORMAL FORMS

- Normalization
- Decompose larger, complex table into simpler and smaller ones
- Moves from lower normal forms to higher normal forms
- Usually done after E-R modeling
 - Should not include computed fields

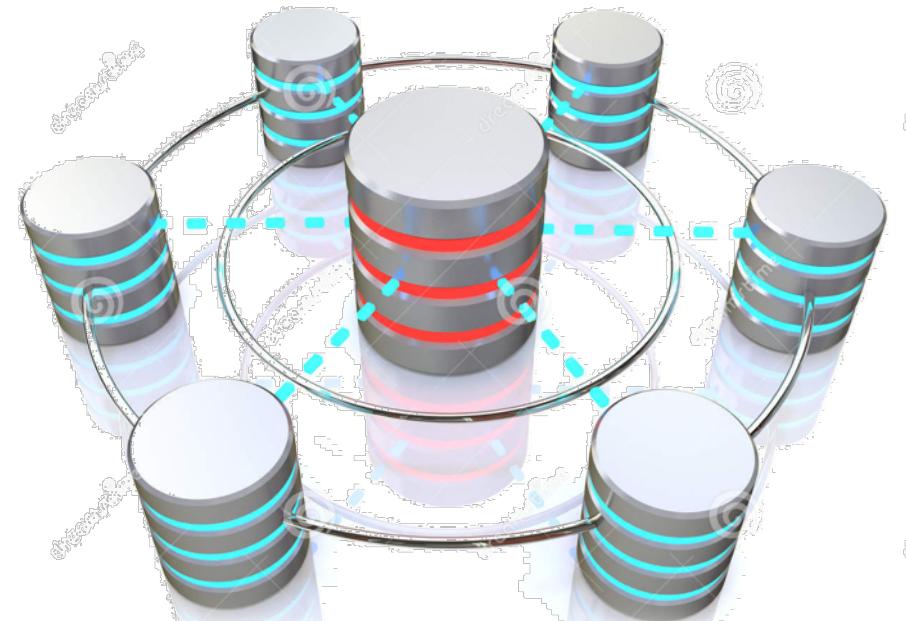


NORMALIZATION AND NORMAL FORMS



NEED FOR NORMALIZATION

- In order to produce good database design, ask questions:
- Does the design ensure all database operations will be efficiently performed?
- Is there any expensive DBMS operations which are avoidable?
- Is the information unnecessarily replicated?



NEED FOR NORMALIZATION

How many rows should be updated if there is a change in student details for Rob? – Update Anomalies

Student_Details	Course_Details	Pre-requisite	Result_Details
0101 Tim 11/4/1985	M1 Physics-Math-Chemistry	7	02/11/2015 82 A
0102 Rob 10/04/1986	M4 Physics-Bio-Chemistry	8	21/11/2015 89 A
0103 Mary 11/07/1985	B3 Botany-Zoology	10	12/11/2015 62 B
0104 Rob 10/04/1986	H6 History-Geo-Eco	9	21/11/2015 89 A
0105 Tom 03/08/1988	C3 Bio-Chemistry	11	12/11/2015 50 C

How many times the student_details should be stored for each of the courses enrolled? – Data Duplication

Is it possible to insert a course which is not enrolled by any student? – Insertion Anomalies

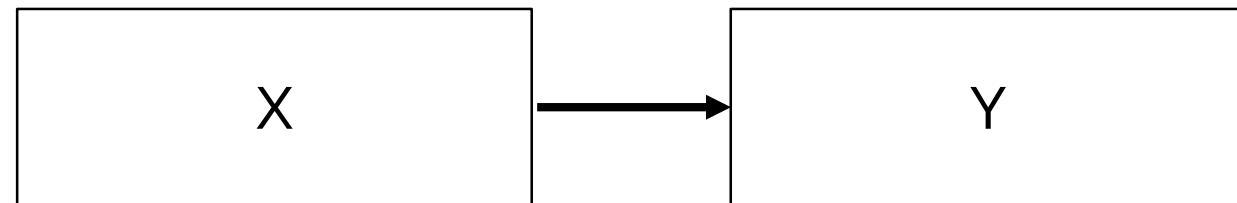
Can we delete the course_details for a student Tom? – Delete Anomalies

FUNCTIONAL DEPENDENCY

- Consider the relation
 - Result (student#, course#, CourseName, InstructorName Marks, Grade)
 - Student# and course# together defines exactly one value of marks. Student#, course# ↳ Marks
 - Student# and course# determines Marks or Marks is functionally dependent on student# and course#
- Other functional dependencies in the relation:
 - Course# - CourseName
 - Course# - InstructorName
 - Marks - Grade

FUNCTIONAL DEPENDENCY

- In a given relation R, X and Y are attributes. Attribute Y is functionally dependent on attribute X if each value of X determines exactly one value of Y.



FD TYPES

Full Functional Dependency

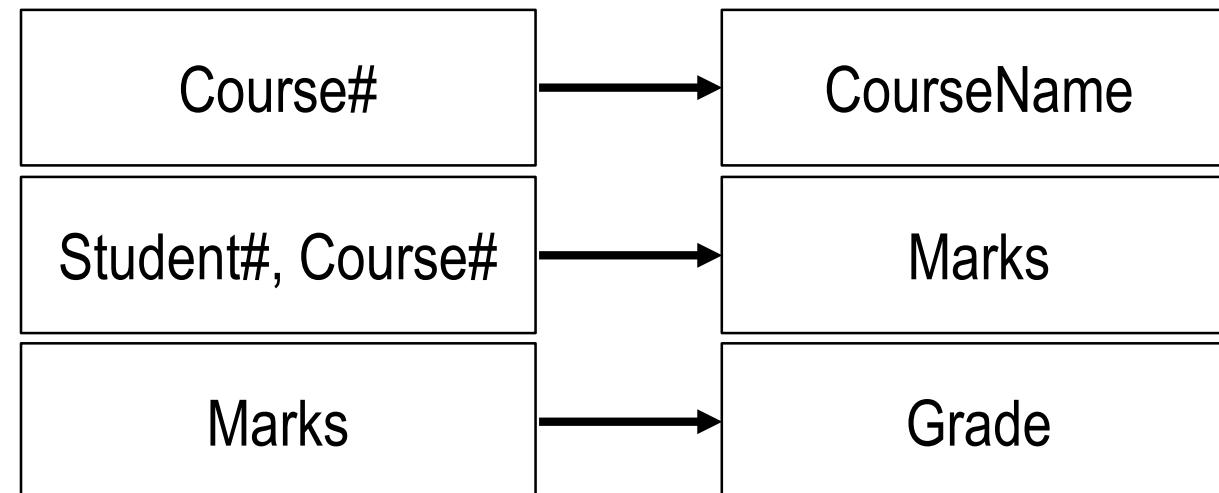
Partial Functional Dependency

Transitive Dependency

FD TYPES

Full Functional Dependency

- Attribute Y is functionally dependent on attribute X, if and only if it is not functionally dependent on the subset of X
- Result(student#, course#, CourseName, InstructorName, Marks, Grade)

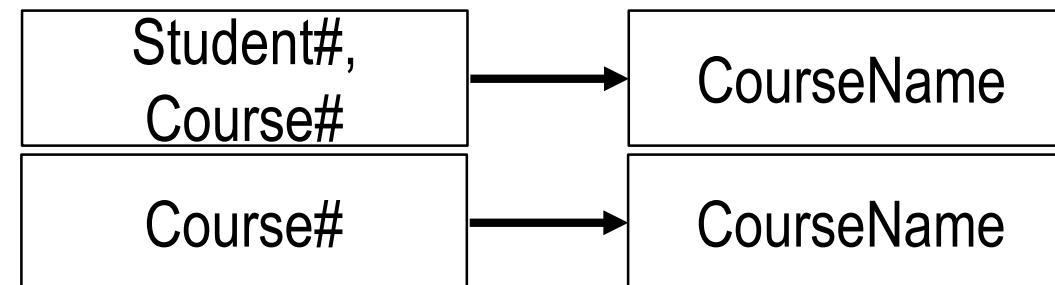


Marks is dependent on {Student#, Course#} together and not on the subset of {Student#, Course#}

FD TYPES

Partial Functional Dependency

- Attribute Y is partially dependent on attribute X, if and only if it is dependent on the subset of attribute X.
- REPORT (Student#, Course#, StudName, CourseName, Marks, Grade)

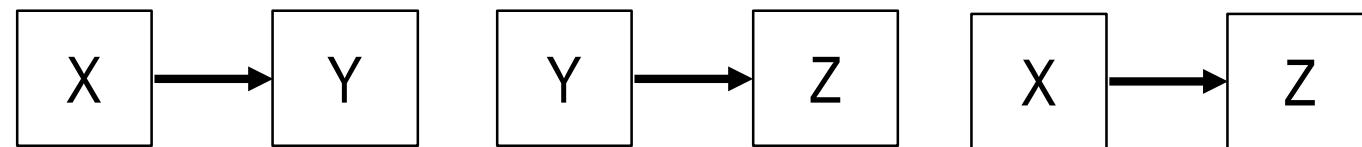


COURSENAME IS DEPENDENT ON {STUDENT#, COURSE#} AND ON THE SUBSET {COURSE#}

FD TYPES

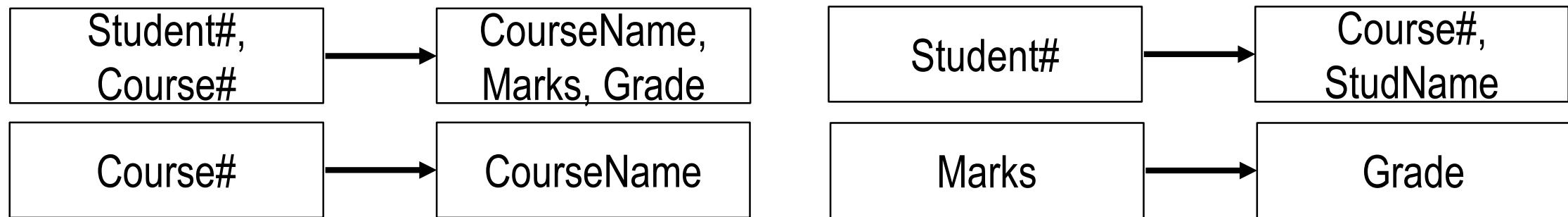
Transitive Dependency

X, Y, Z are three attributes



KEY ATTRIBUTE - ALTERNATE DEFINITION

- In a Given relation R, if an attribute X determines all other attributes, then X is said to be a Key attribute
- REPORT (Student#, Course#, StudName, CourseName, Marks, Grade)



{STUDENT#, COURSE#} TOGETHER DETERMINES ALL THE ATTRIBUTES IN THE RELATION AND HENCE IS THE KEY ATTRIBUTE OF THE RELATION

NORMALIZATION

Types of Normal Form

First Normal Form
(1NF)

Second Normal Form
(2NF)

Third Normal Form
(3NF)

FIRST NORMAL FORM

Types of Normal Form

First Normal Form
(1NF)



A relation schema is in 1NF, if and only if:
All attributes in the relation are atomic in nature and
there are no repeating elements or group of elements.

Second Normal Form
(2NF)

Third Normal Form
(3NF)

FIRST NORMAL FORM (1NF)

Student Marks Table

Student_Details	Course_details		Pre-requisite	Result_details
0101 Tim 11/4/1985	M1 Physics-Math-Chemistry	7	Basic Math	02/11/2015 82 A
0102 Rob 10/04/1986	M4 Physics-Bio-Chemistry	8	Basic Physics	21/11/2015 89 A
0103 Mary 11/07/1985	B3 Botany-Zoology	10	Basic Biology	12/11/2015 62 B
0104 Rob 10/04/1986	H6 History-Geo-Eco	9	Basic History	21/11/2015 89 A
0105 Tom 03/08/1988	C3 Bio-Chemistry	11	Basic Biology	12/11/2015 50 C

FIRST NORMAL FORM (1NF)

Student Marks Table in 1NF

Student#	Student_Name	Date Of Birth	Course#	CourseName	Pre-Requisite	Duration in days	Date Of Exam	Marks	Grade
0101	Tim	11/4/1985	M1	Physics-Math - Chemistry	Basic Math	7	02/11/2015	82	A
0102	Rob	10/04/1986	M4	Physics-Bio - Chemistry	Basic Physics	8	21/11/2015	89	A
0103	Mary	11/07/1985	B3	Botany - Zoology	Basic Biology	10	12/11/2015	62	B
...

SECOND NORMAL FORM (2NF)

Types of Normal Form

First Normal Form
(1NF)

Second Normal Form
(2NF)

Third Normal Form
(3NF)

A relation is said to be in 2NF, if and only if:

- It is in 1NF.
- No partial dependency exists between non-key attributes and key attributes.

SECOND NORMAL FORM (2NF)

Student Marks Table in 1NF

Student#	Student_Name	DOB	Course#	CourseName	Pre Requisite	Duration in days	Date of Exam	Marks	Grade
0101	Tim	11/4/1985	M1	Physics-Math - Chemistry	Basic Math	7	02/11/2015	82	A
0102	Rob	10/04/1986	M4	Physics-Bio-Chemistry	Basic Physics	8	21/11/2015	89	A
0103	Mary	11/07/1985	B3	Botany-Zoology	Basic Biology	10	12/11/2015	62	B

SECOND NORMAL FORM (2NF)

- Key and Non Key Attributes of Student_marks table
- {Student#, Course#} : Candidate key
- Key Attributes: Student#, Course#
- Non Key Attributes: StudentName, DOB, CourseName, Pre-Requisite, DurationDays, Dateof exam, Marks, Grade
- Functional dependencies:

STUDENT#, COURSE# → MARKS

STUDENT#, COURSE# → GRADE

MARKS → GRADE

STUDENT# → STUDENTNAME, DOB

COURSE# → COURSENAME, PRE-REQUISITE, DURATIONDAYS,

DATEOF EXAM

Partial
Dependency
with the Key
attribute

Split/Decompose the tables
to remove partial
dependencies

SECOND NORMAL FORM (2NF)

Student Table

<u>Student#</u>	<u>Student_Name</u>	Date Of Birth
0101	Tim	11/4/1985
0102	Rob	10/04/1986
0103	Mary	11/07/1985

Result Table

<u>Student#</u>	<u>Course#</u>	Marks	Grade
0101	M1	82	A
0102	M4	89	A
0103	B3	62	B

Course Table

<u>Course#</u>	<u>CourseName</u>	Prerequisite	Durationin days	Date Of Exam
M1	Physics-Math-Chemistry	Basic Math	7	02/11/2015
M4	Physics-Bio-Chemistry	Basic Physics	8	21/11/2015
B3	Botany-Zoology	Basic Biology	10	12/11/2015

THIRD NORMAL FORM (3NF)

Types of Normal Form

First Normal Form
(1NF)

Second Normal Form
(2NF)

Third Normal Form
(3NF)

A relation R is said to be in 3NF if and only if:

- It is in 2NF.
- No transitive dependency exists between non-key attributes and key attributes through another non-key attribute.

THIRD NORMAL FORM (3NF)

Result_table

Student#	Course#	Marks	Grade
0101	M1	82	A
0102	M4	89	A
0103	B3	62	B

STUDENT#, COURSE# → MARKS

STUDENT#, COURSE# → GRADE

MARKS → GRADE

STUDENT#, COURSE# → MARKS → GRADE:TD

Remove

THIRD NORMAL FORM (3NF)

Result Table

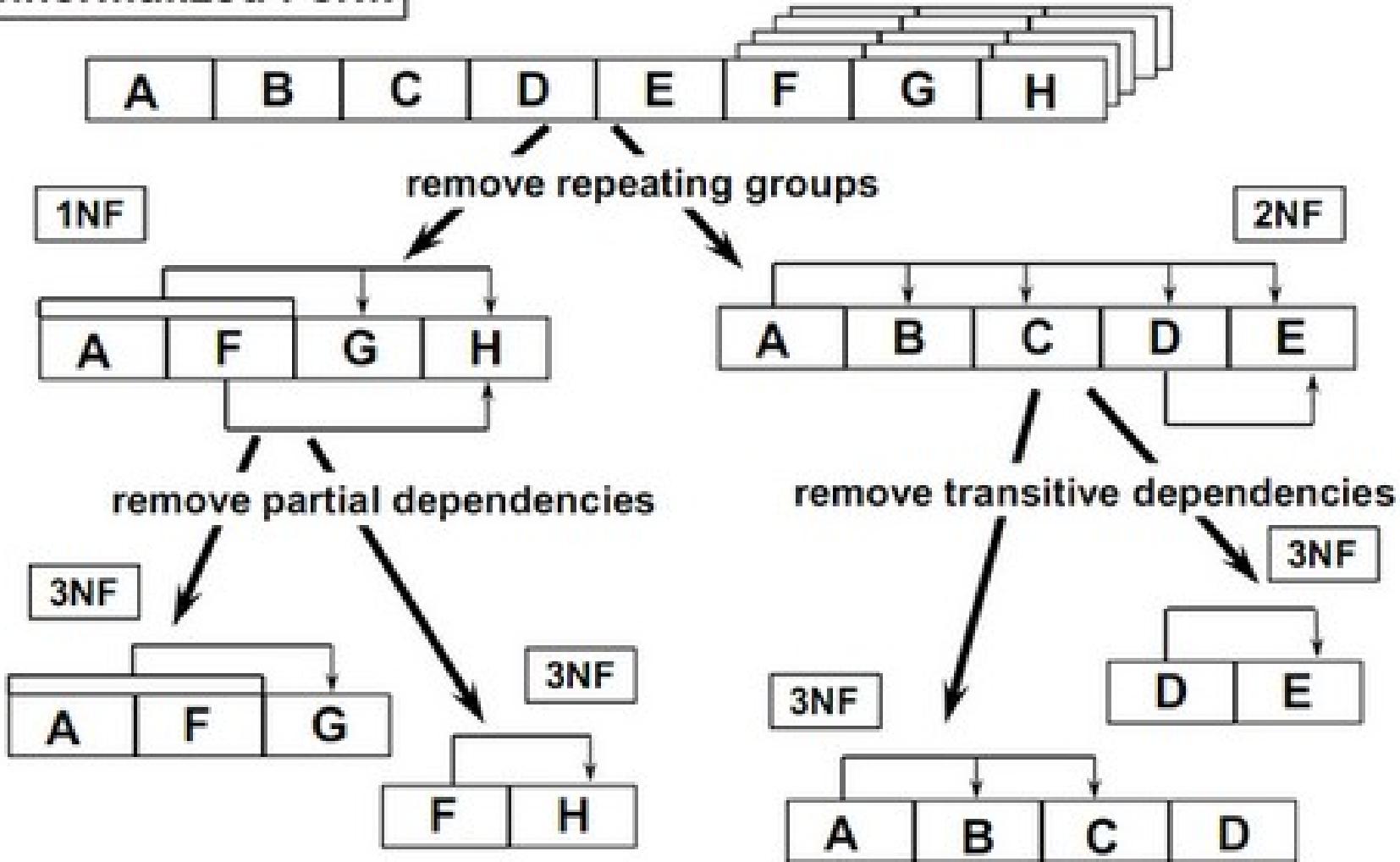
<u>Student#</u>	<u>Course#</u>	Marks
0101	M1	82
0102	M4	89
0103	B3	62

Marks Grade Table

<u>Marks</u>	<u>Grade</u>
82	A
89	A
62	B

NORMALIZATION IN A NUTSHELL

Unnormalized Form



PROS AND CONS OF NORMALIZATION



- Based on mathematical foundation
- Removes the redundancy to a large extent
- After 3NF, data redundancy is minimized to the extent of foreign keys
- Removes the anomalies present in INSERTs, UPDATEs and DELETEs



- Data retrieval or SELECT operation performance will be severely affected
- Normalization might not always represent real world scenarios

THANK YOU