NORMALIZATION

Used in design phase of DDLS or DBLS. (Database Development Life Cycle)

Redundancy

- When same data is stored unnecessarily at different places.
- It has two parts:
  - Data inconsistency:

    When different versions of same data appears in different places.
  - Data Anomalies:

    When all changes in redundant data are not made successfully.
    - Update anomalies:

      Changes in one table need changes in other table.
    - Insertion anomalies:

      Record inserted in one needs record to be inserted in other table.
    - Deletion table:

      Record deleted from one needs record to be deleted from others.

DEPENDENCY

- Functional dependency:
  - Attribute B is functionally dependent on attribute A if for each value of A there is only one value of B.

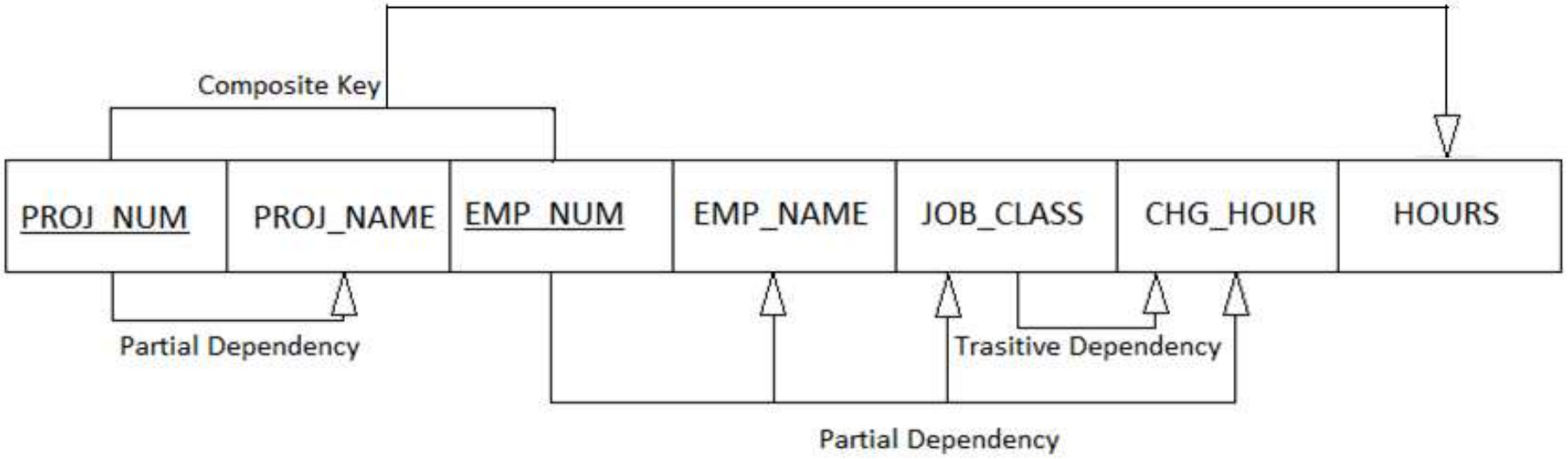
    e.g.

    STU_NUM→STU_NAME,STU_AGE

	STU_NUM	STU_NAME	STU_AGE
	101	Ajay	18
	102	Baldeep	19
	103	Naveen	18
	104	Pankaj	18
	105	Raj	19

Various DEPENDENCIES

- Partial Dependency:
  - Functional dependency based on part of composite primary key.



- Transitive Dependency:
  - Attribute functionally dependent on other attribute, but both are not primary key or at least part of primary key.

Various DEPENDENCIES

- Multivalued Dependency:
  - If one entity has multiple values in table & table has no primary key.

	EMP_NUM	ORG_CODE	ASSIGN_NUM
	10123	RC	1
	10123	RC	3
	10123	UW	4
	10123	UW	5

Various DEPENDENCIES

- Join Dependency:
  - Table cannot be decomposed into two simpler tables.

	AGENT	COMPANY	PRODUCT
	Summet	Abc	Nut
	Summet	Abc	Bolt
	Summet	Cde	Bolt
	Raj	Abc	Nut
	Raj	Abc	Bolt

P1

	AGENT	COMPANY
	Summet	Abc
	Summet	Cde
	Raj	abc

P2

	AGENT	PRODUCT
	Summet	Nut
	Summet	Bolt
	Raj	Nut
	Raj	bolt

Joining P1 & P2

	AGENT	COMPANY	PRODUCT
	Summet	Abc	Nut
	Summet	Abc	Bolt
	Summet	Cde	Nut
	Summet	Cde	Bolt
	Raj	Abc	Nut
	Raj	Abc	Bolt

Various DEPENDENCIES

JOIN DEPENDENCY

But can be decomposed into three

or more than three simpler tables.

P3 ➔

	COMPANY	PRODUCT
	Abc	Nut
	Abc	Bolt
	Cde	bolt

	AGENT	COMPANY	PRODUCT
	Summet	Abc	Nut
	Summet	Abc	Bolt
	Summet	Cde	Nut
	Summet	Cde	Bolt
	Raj	Abc	Nut
	Raj	Abc	Bolt

Joining with P3➔

	AGENT	COMPANY	PRODUCT
	Summet	Abc	Nut
	Summet	Cde	Bolt
	Summet	Abc	Bolt
	Raj	Abc	Nut
	Raj	Abc	Bolt

NORMALIZATION

- Process of correcting & Evaluating tables to minimize data redundancies, thereby reducing data anomalies.
- Steps towards normalization are:
  - Convert ER model to tables.
  - Examine tables for redundancy.
  - Convert to non-redundant forms.
  - Non-redundant forms then converted to database defination to achive objectives of database design.

NORMALIZATION

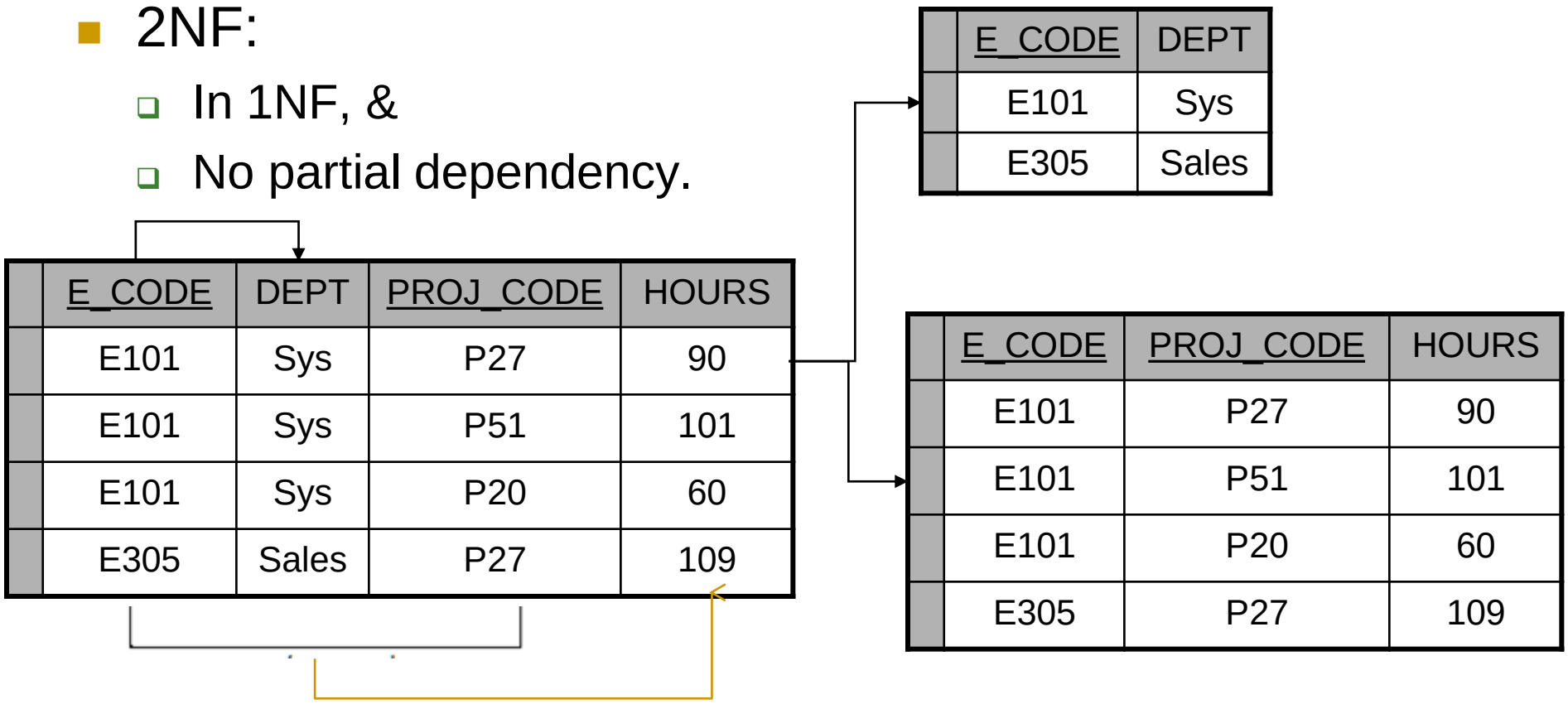
- 1NF:
  - No repeating groups

	E_CODE	DEPT	PROJ_CODE	HOURS
	E101	Sys	P27	90
			P51	101
			P20	60
	E305	Sales	P27	109

	E_CODE	DEPT	PROJ_CODE	HOURS
	E101	Sys	P27	90
	E101	Sys	P51	101
	E101	Sys	P20	60
	E305	Sales	P27	109

NORMALIZATION

- 2NF:
  - In 1NF, &
  - No partial dependency.



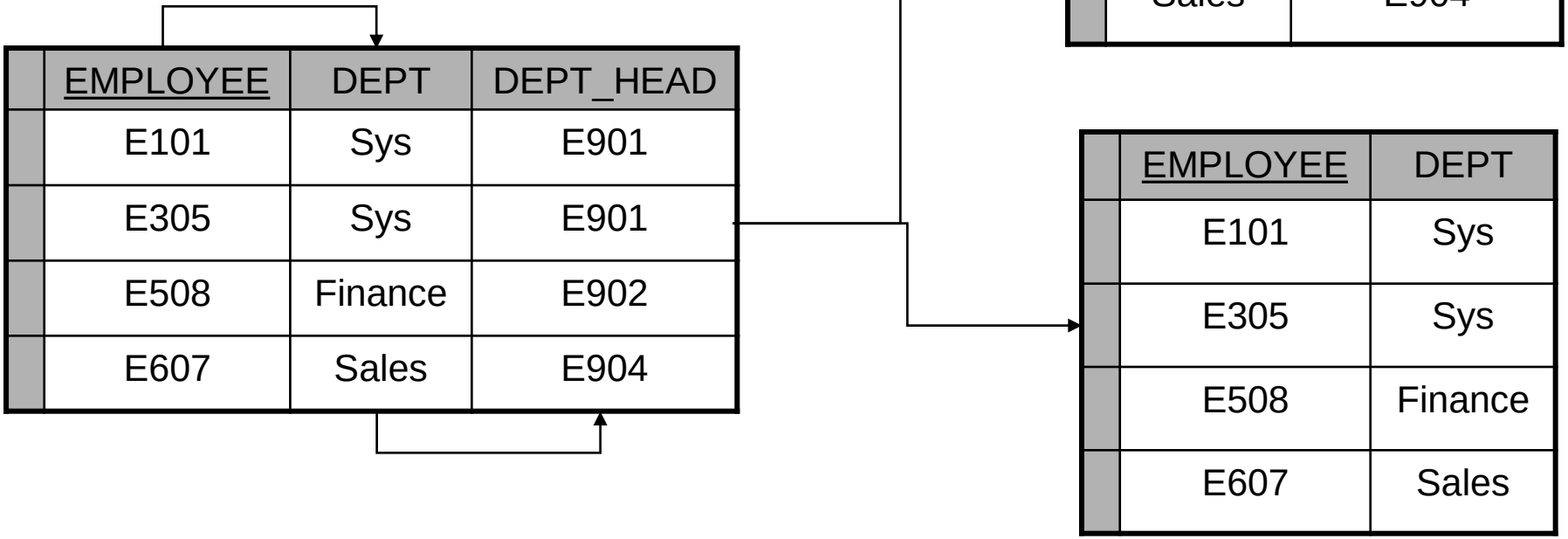
E_CODE	DEPT	PROJ_CODE	HOURS
E101	Sys	P27	90
E101	Sys	P51	101
E101	Sys	P20	60
E305	Sales	P27	109

E_CODE	DEPT
E101	Sys
E305	Sales

E_CODE	PROJ_CODE	HOURS
E101	P27	90
E101	P51	101
E101	P20	60
E305	P27	109

NORMALIZATION

- 3NF:
  - In 2NF, &
  - No transitive dependency.



EMPLOYEE	DEPT	DEPT_HEAD
E101	Sys	E901
E305	Sys	E901
E508	Finance	E902
E607	Sales	E904

DEPT	DEPT_HEAD
Sys	E901
Finance	E902
Sales	E904

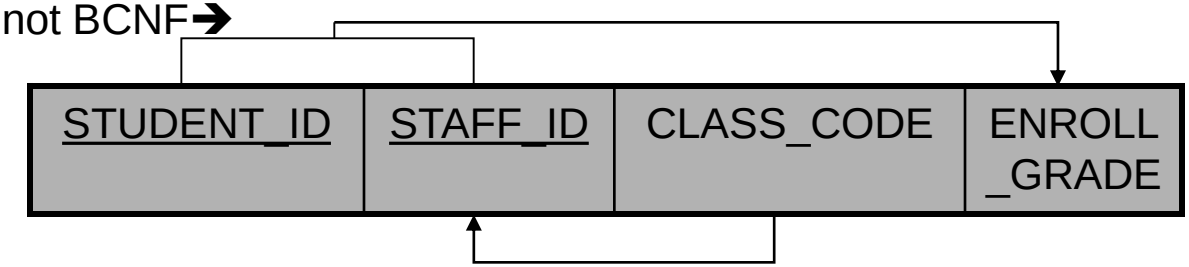
EMPLOYEE	DEPT
E101	Sys
E305	Sys
E508	Finance
E607	Sales

NORMALIZATION

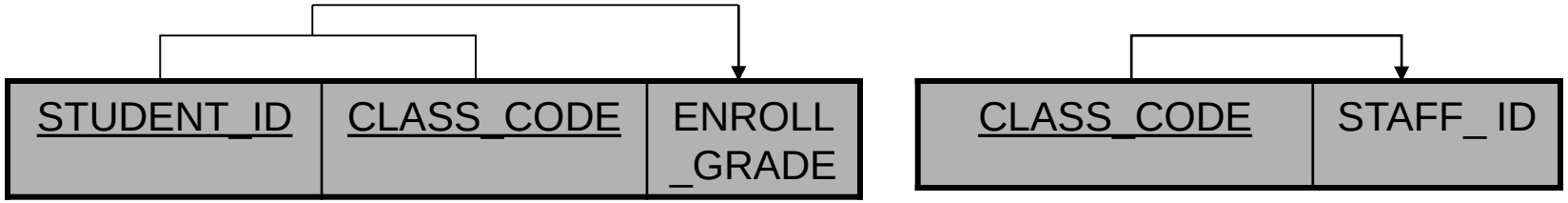
- BCNF:
  - Every determinant is candidate key.

STUDENT_ID	STAFF_ID	CLASS_CODE	ENROLL _GRADE

In 3NF, not BCNF➔



In 3NF, & BCNF➔



NORMALIZATION

- 4NF:
  - In 3NF or BCNF, &
  - At most one multivalued dependency.

	STUDENT	COURSE	TEXTBOOK
	Ankit	Physics	Mechanics
	Ankit	Physics	Optics
	Raj	Physics	Mechanics
	Raj	Physics	Optics
	Ankit	Chemistry	Organic
	Ankit	Chemistry	Inorganic
	Ajay	English	Grammer
	Ajay	English	literature

	STUDENT	COURSE
	Ankit	Physics
	Raj	Physics
	Ankit	Chemistry
	Ajay	English

	COURSE	TEXTBOOK
	Physics	Mechanics
	Physics	Optics
	Chemistry	Organic
	Chemistry	Inorganic
	English	Grammer
	English	literature

NORMALIZATION

- 5NF: (PJNF)
  - In 4NF, &
  - No join dependency.

	AGENT	COMPANY	PRODUCT
	Summet	Abc	Nut
	Summet	Abc	Bolt
	Summet	Cde	Bolt
	Raj	Abc	Nut
	Raj	Abc	Bolt

P1

	AGENT	COMPANY
	Summet	Abc
	Summet	Cde
	Raj	abc

P2

	AGENT	PRODUCT
	Summet	Nut
	Sunnet	Bolt
	Raj	Nut
	Raj	bolt

P3

	COMPANY	PRODUCT
	Abc	Nut
	Abc	Bolt
	Cde	bolt

Denormalization

- Tables are decomposed to attain normalized form. But, joining larger tables during query processing takes more time & efforts. So, tables are denormalized to lower normalization forms to attain processing speed.
- "*Normalized Relations*" is one of the design goal of database, but it is not more important than "*Processing Speed*" (i.e.Another Goal).