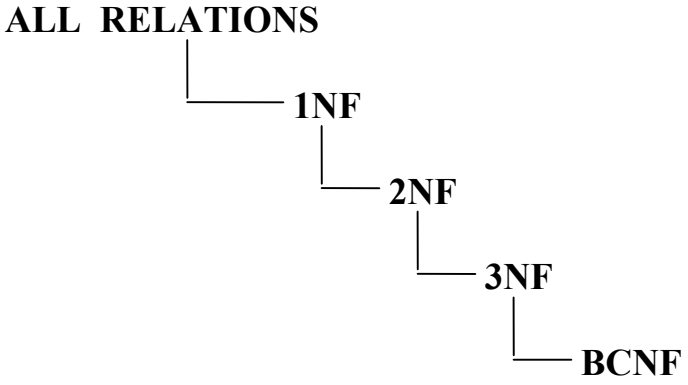
NORMALIZATION

THE FORMALIZATION OF LOGICAL DATABASE DESIGN PRINCIPLES THAT SERVE TO REDUCE UNDESIRABLE PROPERTIES IN RELATIONS

NORMAL FORMS (CONSTRAINTS)



FUNCTIONAL DEPENDENCE

GIVEN A RELATION 'R':

R . X → R .Y

{ X FUNCTIONALLY DETERMINES Y }
OR
{ Y IS FUNCTIONALLY DEPENDENT ON X }

IFF EACH VALUE OF X HAS ASSOCIATED WITH IT A UNIQUE VALUE OF Y

{ 2 TUPLES WITH THE SAME X-VALUE MUST ALSO HAVE THE SAME Y-VALUE }

EXAMPLES :-

faculty (fac_id) → faculty (fac_name)

course (crs_id) → course (description, credit_hours)

class_taken(stu_id, crs_id, sect, term) →
class_taken (grade)

advises (stu_id) → advises (fac_id)

advises (fac_id) --/--> advises (stu_id)

FULL FUNCTIONAL DEPENDENCE

Y is functionally dependent on X and not functionally dependent on any proper subset of X.

EXAMPLE :-

class_taken (stu_id, crs_id, sect, term) → class_taken (grade)

FIRST NORMAL FORM

ALL ATTRIBUTE DOMAINS ARE ATOMIC.

SIMPLIFIES RELATION STRUCTURE.

THEREFORE, SIMPLIFIES OPERATIONS ON RELATIONS.

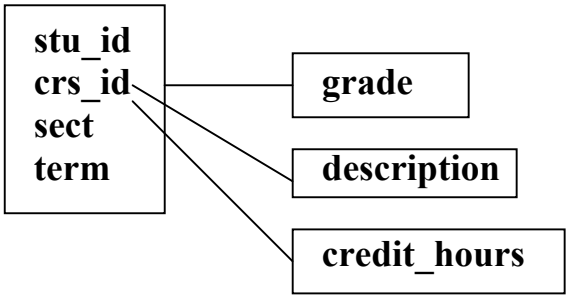
SECOND NORMAL FORM

FIRST NORMAL FORM

PLUS EVERY NONKEY ATTRIBUTE IS FULLY DEPENDENT ON THE PRIMARY KEY.

EXAMPLE :-

class_taken (stu_id, crs_id, sect, term, grade, description, credit_hours)



description AND credit_hours ARE NOT FULLY DEPENDENT ON THE PRIMARY KEY.

class_taken (crs_id) → class_taken (description, credit_hours)

class_taken

stu_id	crs_id	sect	term	gr	desc	cr_hrs
123	CMPS460	1	20001	B	Database	3
887	CMPS460	1	19991	C	Database	3
409	CMPS261	2	19983	A	Data Str	3
.	.	.	.	.	.	.

THE REDUNDANCY IN class_taken LEADS TO "UPDATE ANOMALIES".

INSERT

WE CAN'T HAVE A COURSE TILL SOMEONE TAKES IT.

DELETE

IF WE DELETE THE LAST CLASS_TAKEN TUPLE REFERENCING A COURSE, WE LOSE THE COURSE.

UPDATE

IF A COURSE DESCRIPTION CHANGES, EITHER

A) WE UPDATE THEM ALL

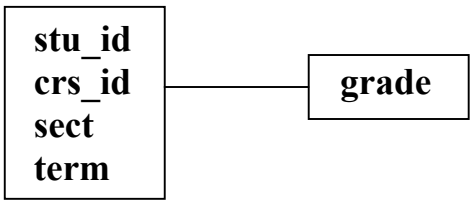
OR

B) THE DATABASE IS INCONSISTENT

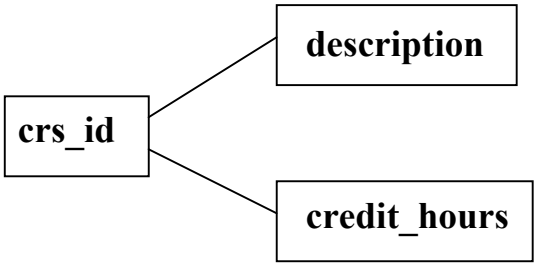
SOLUTION :-

DECOMPOSE CLASS_TAKEN INTO TWO RELATIONS THAT ARE IN SECOND NORMAL FORM.

class_taken (stu_id, crs_id, sect, term, grade)



course (crs_id, description, credit_hours)



THIRD NORMAL FORM

1NF + 2NF + EVERY NONKEY ATTRIBUTE IS NONTRANSITIVELY DEPENDENT ON THE PRIMARY KEY.

TRANSITIVITY

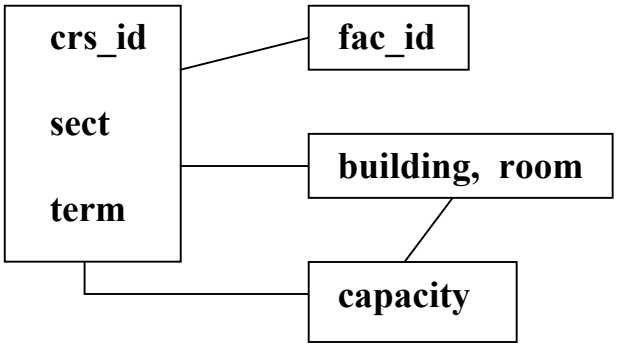
If R . X → R . Y and R . Y → R . Z

THEN

R . X → R . Z (TRANSITIVELY)

EXAMPLE :-

class_offered (crs_id, sect, term, fac_id, building, room, capacity)



class_offered is in 1NF (OBVIOUS)

class_offered is in 2NF

EVERY NONKEY ATTRIBUTE IS FULLY DEPENDENT ON P.K.

class_offered is not in 3NF

CAPACITY IS TRANSITIVELY DEPENDENT ON PRIMARY KEY

UPDATE ANOMALIES :-

<u>INSERT</u> - WE CAN'T HAVE A BUILDING / ROOM UNTIL A COURSE MEETS IN IT (NOR CAN WE RECORD A CAPACITY).

<u>DELETE</u> - IF WE DELETE THE ONLY COURSE OFFERED IN A ROOM, WE LOSE THE ROOM.

<u>UPDATE</u> - IF A ROOM'S CAPACITY CHANGES,

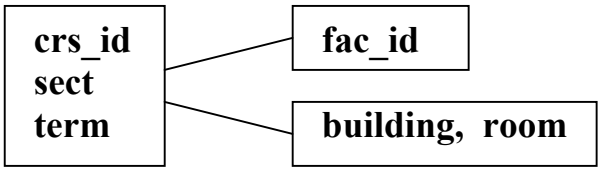
A) MASS UPDATES

B) INCONSISTENCY

SOLUTION :

DECOMPOSE class_Offered INTO TWO 3NF RELATIONS.

class_offered (crs_id, sect, term, fac_id, building, room)



room (building, room, capacity)