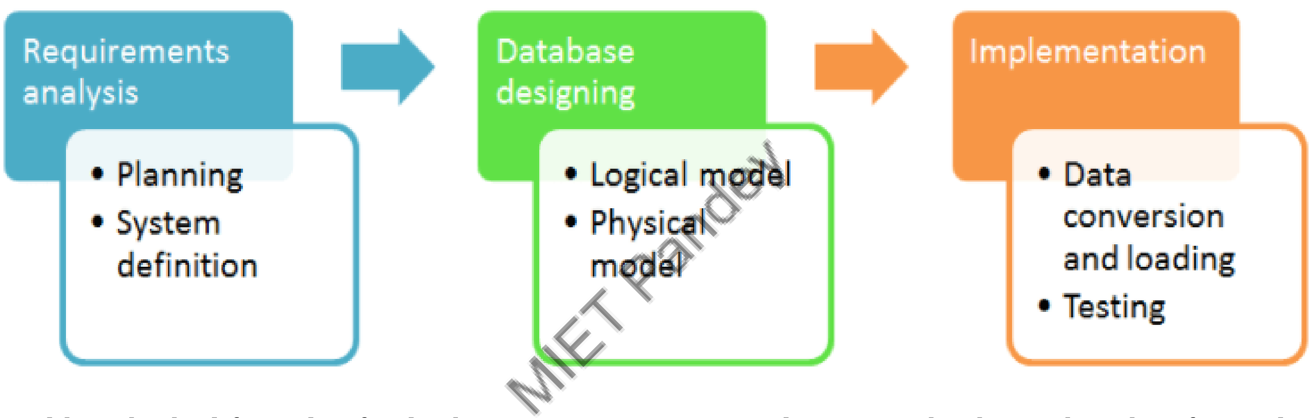
SUBJECT (CODE): DBMS (KCS-501) COURSE:Btech(CSE)

UNIT – III (DATABASE DESIGN & NORMALIZATION)

IMPORTANCEOFDATABASEDESIGN

1. Databasedesignsprovidetheblueprintsofhowthedataisgoingtobestoredinasystem.A proper design of a database highlyaffects the overall performance ofany application.
2. Thedesigningprinciplesdefinedforadatabasegiveaclearideaofthebehaviorofany application and how the requests are processed.
3. Another instanceto emphasizethe databasedesign isthat a proper databasedesign meetsall the requirements of users.
4. Lastly, the processing time ofan application is greatly reduced if the constraints of designing a highly efficient database are properly implemented.

LifeCycle



Although, the life cycle of a database is not an important discussion that has to be taken forward inthis article because we are focused on the database design. But, before jumping directly on the designing models constituting database design it is important to understand the overall workflowand life-cycle of the database.

Databasedesigning

ThemainobjectivesofdatabasedesigninDBMSaretoproducelogicalandphysicaldesignsmodelsof the proposed database system.

- **Logical model** – This stage is concerned with developinga database model based on requirements.Theentiredesignisonpaperwithoutanyphysicalimplementationsorspecific DBMS considerations.
- **Physicalmodel**–Thisstageimplementsthelogicalmodelofthedatabasetakingintoaccount the DBMS and physical implementation factor

Logical

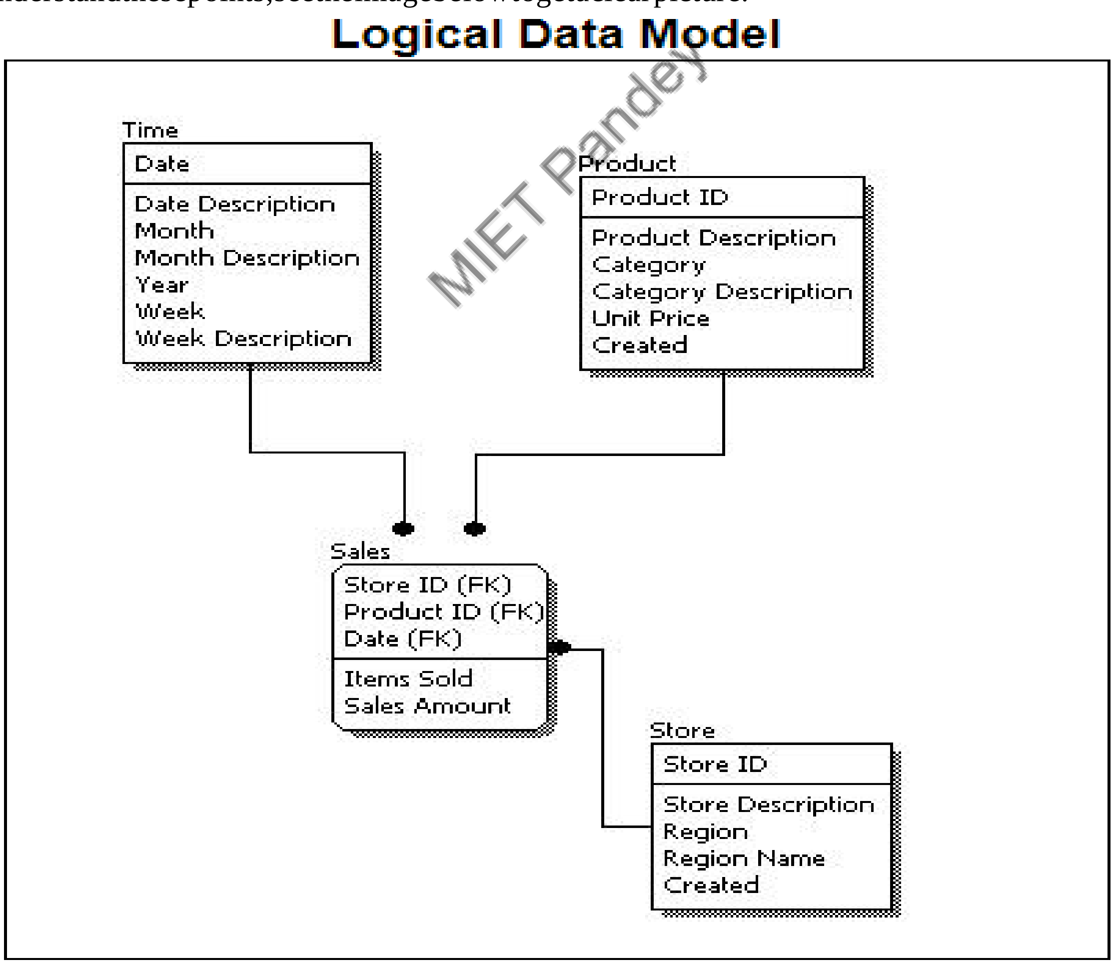
A logical data model generally describes the data in as many details as possible, without having to be concerned about the physical implementations in the database. Features of logical data model might include:

1. Alltheentitiesandrelationshipsamongstthem.
2. Eachentityhaswell-specifiedattributes.
3. Theprimarykeyforeachentityisspecified.
4. Foreignkeyswhichareusedtoidentifyarelationshipbetweendifferententitiesarespecified.
5. Normalizationoccursatthislevel.

Alogicalmodelcanbedesignedusingthefollowingapproach:

1. Specifyalltheentitieswithprimarykeys.
2. Specifyconcurrentrelationshipsbetweendifferententities.
3. Figureouteachentityattributes
4. Resolvemany-to-manyrelationships.
5. Carryouttheprocessofnormalization.

Also, one important factor after following the above approach is to critically examine the design based on requirement gathering. Ifthe above stepsare strictly followed, thereare chances ofcreatinga highly efficient database design that follows the native approach.
Tounderstandthesepoints,seetheimagebelowtogetaclearpicture.



If we compare the logical data model as shown in the figure above with some sample data in the diagram, we can come up withfacts that in aconceptual datamodel there are no presence ofaprimary key whereas a logical data model hasprimary keys for all ofitsattributes. Also, logical data model the cover relationship between different entitiesandcarriesroom for foreignkeys to establishrelationships among them.

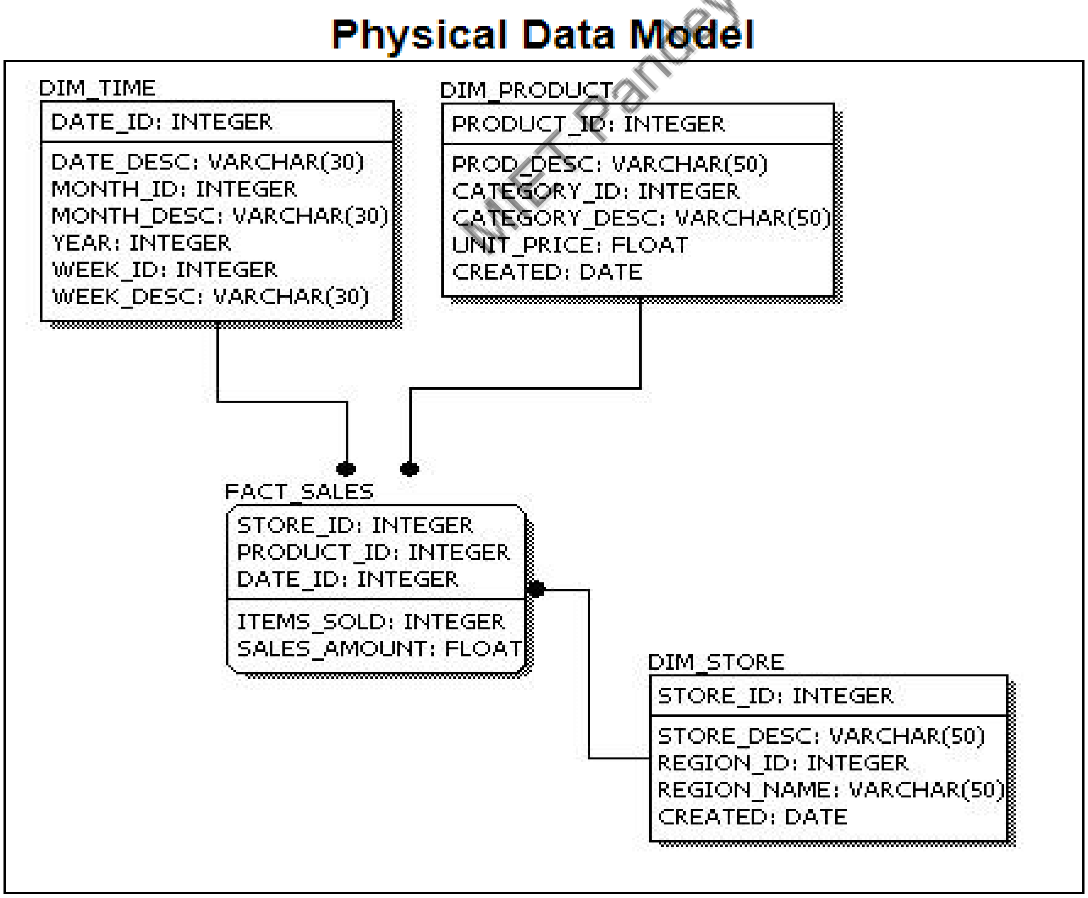
Physical

A Physical data mode generally represents howthe approach or concept of designing the database. The main purpose ofthe physical data model isto showall the**structures** ofthe table including the**column name, column data type, constraints, keys(primary and foreign)**, and the relationship among tables. The following are the features ofa physical data model:

1. Specifiesallthecolumnsandtables.
2. Specifiesforeignkeysthatusuallydefinetherelationshipbetweentables.
3. Basedonuserrequirements,de-normalizationmightoccur.
4. Since the physical consideration is taken into account so there will straightforward reasons for difference than a logical model.
5. Physicalmodelsmight be different fordifferent RDBMS. Forexample, thedatatypecolumn may be different in MySQL and SQL Server.

Whiledesigningaphysicaldatamodel,thefollowingpointsshouldbetakenintoconsideration:

1. Converttheentitiesintotables.
2. Convertthedefinedrelationshipsintoforeignkeys.
3. Convertthedataattributesintocolumns.
4. Modifythedatamodelconstraintsbasedonphysicalrequirements.



Comparingthisphysicaldata modelwiththelogicalwiththepreviouslogicalmodel, wemightconclude thedifferencesthatinaphysicaldatabaseentitynamesareconsideredtablenamesandattributesare consideredcolumnnames.Also,thedatatypeofeachcolumnisdefinedinthephysicalmodel depending on the actual database used.

FUNCTIONALDEPENDENCIES

DeterminestherelationofoneattributetoanotherattributeinaDatabaseManagementSystem(DBMS). FunctionalDependency helpsto maintain the quality ofdata in the database. It playsa vitalrole to find the difference between good and bad database design.
Afunctionaldependencyisdenotedbyanarrow"→".ThefunctionaldependencyofXonYis represented by X → Y.

TypesofFunctionaldependenciesinDBMS:

1. Trivialfunctionaldependency
2. Non-Trivialfunctionaldependency
3. Multivalvefunctionaldependency
4. Transitivefunctionaldependency

1. <u>TrivialFunctionalDependency</u>

InTrivialFunctionalDependency,adependentisalwaysasubsetofthedeterminant.
i.e.IfX→YandYisthesubsetofX,thenitiscalledtrivialfunctionaldependency
For example,

roll_no	name	age
42	abc	17
43	pqr	18
44	xyz	18

Here**,** {roll_no, name} → name is a trivial functional dependency, since the dependentname is a subset of determinant set {roll_no, name}
Similarly,**roll_no→roll_no**isalsoanexampleoftrivialfunctionaldependency.

2. <u>Non-trivialFunctionalDependency</u>

InNon-trivialfunctionaldependency,thedependentisstrictlynotasubsetofthedeterminant.
i.e.If **X→Y**andYisnotasubsetofX,thenitiscalledNon-trivialfunctionaldependency.
For example,

roll_no	name	age
42	abc	17
43	pqr	18
44	xyz	18

Here, **roll_no → name** is a non-trivial functional dependency, since the dependentname is not a subset of determinant roll_no
Similarly, **{roll_no, name} → age** is also a non-trivial functional dependency, since age is not a subset of {roll_no, name}

3. MultivaluedFunctionalDependency

InMultivaluedfunctionaldependency,entitiesofthedependentsetarenotdependentoneachother.
i.e.If **a →{b, c}** andthereexistsnofunctional dependencybetween**band c**, thenit is called a multivalued functional dependency**.**

For example,

roll_no	name	age
42	abc	17
43	pqr	18
44	xyz	18
45	abc	19

Here**,**roll_no→{name,age}isamultivaluedfunctionaldependency,sincethe dependents name & age are not dependent on each other **(i.e.** name → age **or** age → name doesn't exist !**)**

4. TransitiveFunctionalDependency

Intransitivefunctionaldependency,dependentisindirectlydependentondeterminant.
i.e. If **a** → **b** &**b** → **c**, then according to axiom of transitivity, **a** → **c**. This is **a** transitive functional dependency
For example,

enrol_no	name	dept	building_no
42	abc	CO	4
43	pqr	EC	2
44	xyz	IT	1
45	abc	EC	2

Here**,**enrol_no→deptanddept→building_no**,**
Hence,accordingtotheaxiomoftransitivity,**enrol**_no→building_noisavalidfunctional

dependency. This is an indirect functional dependency, hence called Transitive functional dependency.

NORMALIZATION

Normalization is the process of minimizing **redundancy** from a relation or set of relations. Redundancy in relation may cause insertion, deletion, and update anomalies. So, it helps to minimize the redundancy in relations. **Normal forms** are used to eliminate or reduce redundancy in database tables.

- Normalizationistheprocessoforganizingthedatainthedatabase.
- Normalization is used to minimize the redundancyfrom a relation orset of relations. It is also used to eliminateundesirable characteristicslikeInsertion, Update, and Deletion Anomalies.
- Normalizationdividesthelargertableintosmallerandlinksthemusingrelationships.
- Thenormalformisusedtoreduceredundancyfromthedatabasetable.

TypesofNormalForms

Normal Form	Description
1NF	Arelationisin1NFifit containsanatomicvalue.
2NF	Arelationwillbein2NFifitisin1NFandallnon-keyattributesarefully functional dependent on the primary key.
3NF	Arelationwillbein3NFifitisin2NFandnotransitiondependencyexists.
BCNF	Astrongerdefinitionof3NFisknownasBoyceCodd'snormalform.
4NF	A relation will be in 4NF if it is in Boyce Codd's normal form and has no multi-valued dependency.
5NF	A relation is in 5NF. If it is in 4NF and doesnot contain any join dependency,joining should be lossless.

AdvantagesofNormalization

- Normalizationhelpstominimizedataredundancy.
- Greateroveralldatabaseorganization.
- Dataconsistencywithinthedatabase.
- Muchmoreflexibledatabasedesign.
- Enforcestheconceptofrelationalintegrity.

DisadvantagesofNormalization

- Youcannotstartbuildingthedatabasebeforeknowingwhattheuserneeds.
- Theperformancedegradeswhennormalizingtherelationstohighernormalforms,i.e.,4NF, 5NF.
- Itisverytime-consuminganddifficulttonormalizerelationsofahigherdegree.
- Carelessdecompositionmayleadtoabaddatabasedesign,leadingtoseriousproblems.

FirstNormalForm(1NF)

- Arelationwillbe1NFifitcontainsanatomicvalue.
- It states that an attribute of a table cannot hold multiple values. It must hold only single-valued attribute.
- Firstnormalformdisallowsthemulti-valuedattribute,compositeattribute,andtheir combinations.

Example: RelationEMPLOYEEisnotin1NFbecauseofmulti-valuedattributeEMP_PHONE.

EMP_ID	EMP_NAME	EMP_PHONE	EMP_STATE
14	John	7272826385,9064738238	UP
20	Harry	8574783832	Bihar
12	Sam	7390372389,8589830302	Punjab

ThedecompositionoftheEMPLOYEEtableinto1NFhasbeenshownbelow:

EMP_ID	EMP_NAME	EMP_PHONE	EMP_STATE
14	John	7272826385	UP
14	John	9064738238	UP
20	Harry	8574783832	Bihar
12	Sam	7390372389	Punjab
12	Sam	8589830302	Punjab

SecondNormalForm(2NF)

- Inthe2NF,relationalmustbein1NF.
- In the second normal form, all non-key attributesare fully functional dependent on the primary key

Example: Let'sassume,a schoolcan storethedata ofteachersandthesubjectstheyteach.Ina school,a teacher can teach more than one subject.

TEACHERTABLE

TEACHER_ID	SUBJECT	TEACHER_AGE
25	Chemistry	30
25	Biology	30
47	English	35
83	Math	38
83	Computer	38

Inthegiventable,non-prime attributeTEACHER_AGEisdependentonTEACHER_IDwhichisa proper subset ofa candidate key. That's why it violates the rule for 2NF.
Toconvertthegiventableinto2NF,wedecomposeitintotwotables:

TEACHER_DETAILtable:

TEACHER_ID	TEACHER_AGE
25	30
47	35
83	38

TEACHER_SUBJECTtable:

TEACHER_ID	SUBJECT
25	Chemistry
25	Biology
47	English
83	Math
83	Computer

ThirdNormalForm(3NF)

- Arelationwillbein3NFifitisin2NFandnotcontainanytransitivepartialdependency.
- 3NFisusedtoreducethedataduplication.Itisalsousedtoachievethedataintegrity.
- If there is no transitive dependency fornon-prime attributes, then the relationmustbe inthird normal form.

A relation isin third normal form ifit holdsatleast one ofthe following conditions for every non-trivial function dependency X → Y.

1. Xisasuper key.
2. Yisa primeattribute,i.e.,eachelementofYispartofsomecandidatekey.

EMPLOYEE_DETAILtable:

EMP_ID	EMP_NAME	EMP_ZIP	EMP_STATE	EMP_CITY
222	Harry	201010	UP	Noida
333	Stephan	02228	US	Boston
444	Lan	60007	US	Chicago
555	Katharine	06389	UK	Norwich
666	John	462007	MP	Bhopal

Superkeyinthetableabove:

1. {EMP_ID},{EMP_ID,EMP_NAME},{EMP_ID,EMP_NAME,EMP_ZIP} ....soon

Candidatekey:{EMP_ID}

Non-primeattributes: Inthegiventable,allattributesexceptEMP_IDarenon-prime.

Here, EMP_STATE& EMP_CITY dependent on EMP_ZIP and EMP_ZIP dependent on EMP_ID.Thenon-primeattributes(EMP_STATE,EMP_CITY)transitivelydependentonsuper key(EMP_ID). It violates the rule of third normal form.

That'swhyweneedtomovetheEMP_CITY andEMP_STATEtothenew<EMPLOYEE_ZIP> table, with EMP_ZIP as a Primary key.

EMPLOYEEtable:

EMP_ID	EMP_NAME	EMP_ZIP
222	Harry	201010
333	Stephan	02228
444	Lan	60007
555	Katharine	06389
666	John	462007

EMPLOYEE_ZIPtable:

EMP_ZIP	EMP_STATE	EMP_CITY
201010	UP	Noida
02228	US	Boston
60007	US	Chicago
06389	UK	Norwich
462007	MP	Bhopal

BoyceCoddnormalform(BCNF)

- BCNFistheadvanceversionof3NF.Itisstricterthan 3NF.
- AtableisinBCNFifeveryfunctionaldependencyX→Y,Xisthesuperkeyofthetable.
- ForBCNF,thetableshouldbein3NF,andforeveryFD,LHSissuperkey.

Example: Let'sassumethereisacompanywhereemployeesworkinmorethanonedepartment.
EMPLOYEEtable:

EMP_ID	EMP_COUNTRY	EMP_DEPT	DEPT_TYPE	EMP_DEPT_NO
264	India	Designing	D394	283
264	India	Testing	D394	300
364	UK	Stores	D283	232
364	UK	Developing	D283	549

IntheabovetableFunctionaldependenciesareasfollows:

1. EMP_ID→EMP_COUNTRY
2. EMP_DEPT→{DEPT_TYPE,EMP_DEPT_NO}

Candidatekey:{EMP-ID,EMP-DEPT}
ThetableisnotinBCNFbecauseneitherEMP_DEPT norEMP_IDalonearekeys.
Toconvert the giventableinto BCNF, we decompose it into three tables:

EMP_COUNTRYtable:

EMP_ID	EMP_COUNTRY
264	India
264	India

EMP_DEPTtable:

EMP_DEPT	DEPT_TYPE	EMP_DEPT_NO
Designing	D394	283
Testing	D394	300
Stores	D283	232
Developing	D283	549

EMP_DEPT_MAPPINGtable:

EMP_ID	EMP_DEPT
D394	283
D394	300
D283	232
D283	549

Functionaldependencies:

1. EMP_ID→ EMP_COUNTRY
2. EMP_DEPT→{DEPT_TYPE,EMP_DEPT_NO}

Candidate keys:

Forthefirsttable:EMP_IDEorthesecondtableforthethirdtable:{EMP_ID,EMP_DEPT} Now, thisis in BCNF becauseleft side part of both the functional dependenciesis a key.

Fourthnormalform(4NF)

- Arelationwillbein4NFifitisinBoyceCoddnormalformandhasnomulti-valued dependency.
- Fora dependency A→B,iffora singlevalueofA, multiple valuesofBexists,thentherelation will be a multi-valued dependency.

STUDENT

STU_ID	COURSE	HOBBY
21	Computer	Dancing
21	Math	Singing
34	Chemistry	Dancing
74	Biology	Cricket
59	Physics	Hockey

The given STUDENT table is in 3NF, but the COURSE and HOBBY are two independent entity. Hence, there is no relationship between COURSE and HOBBY.
In the STUDENT relation, a student with STU_ID,**21** contains two courses,**Computer**and**Math**and two hobbies,**Dancing**and**Singing**. So there is a Multi-valued dependency on STU_ID, which leads to unnecessary repetition of data.
Soto maketheabovetableinto4NF,wecandecomposeitintotwotables:

STUDENT_COURSE

STU_ID	COURSE
21	Computer
21	Math
34	Chemistry
74	Biology
59	Physics

STUDENT_HOBBY

STU_ID	HOBBY
21	Dancing
21	Singing
34	Dancing
74	Cricket
59	Hockey

Fifthnormalform(5NF)

- A relation is in 5NF if it is in 4NF and not contains any join dependency and joining should be lossless.
- 5NF is satisfied when all the tables are broken into as many tables as possible in order to avoid redundancy.
- 5NFisalsoknownasProject-joinnormalform(PJ/NF).

SUBJECT	LECTURER	SEMESTER
Computer	Anshika	Semester1
Computer	John	Semester1
Math	John	Semester1
Math	Akash	Semester2
Chemistry	Praveen	Semester1

In the above table, John takes both Computer and Math class for Semester 1 but he doesn't take Math class for Semester 2.Inthis case, combination ofall these fieldsrequiredto identifya valid data.

Supposeweadda newSemesterasSemester3 butdonotknowaboutthesubjectand whowillbetaking that subject so weleaveLecturer andSubjectasNULL. Butallthree columnstogetheractsasa primary key, so we can't leave other two columns blank.

Soto maketheabovetableinto5NF,wecandecomposeitintothreerelationsP1,P2 &P3:

SEMESTER	SUBJECT
Semester1	Computer
Semester1	Math
Semester1	Chemistry
Semester2	Math

P2

SUBJECT	LECTURER
Computer	Anshika
Computer	John
Math	John
Math	Akash
Chemistry	Praveen

P3

SEMSTER	**LECTURER**
Semester1	Anshika
Semester1	John
Semester1	John
Semester2	Akash
Semester1	Praveen

Astatementin whichsomecolumnsofanyrelationarecontainedinothercolumnsisknownas an**InclusionDependency**.Inclusiondependencies,likefunctionaldependencies,representone-to-many relationships. However, inclusion dependenciesare more commonlyused to represent relationships betweenrelations. **Aforeign key is an example of inclusion dependency.** Therelation whichitis referring is contained in the column of primary key.

InclusionDependencyExample

Let's say we take tworelations, namely R and S thatare created byusing two entity setsin a way that everyentityinRisalsoSentity.Inclusiondependenceoccurswhenprojecting R'skeyattributesgivesa relation that is contained in the relation acquired by projecting S'skey attributes.

Let'snametherelationsRasteacherandSasstudent,sotaketheattributeasteacher_id, so wecan write:

- teacher.teacher_id-->student.teacher_id

teacher:

teacher_id(primarykey)	**name**	**department**
1	RamKumar	DBMS

student:

student_1	**name**	**teached_id(foreign key)**	**age**
1	RahulSingh	1	18

Lossless-join decomposition is a process in which a relation is decomposed into two or more relations. This property guarantees that the extra or less tuple generation problem does not occur and no information islost fromtheoriginal relation during the decomposition.It isalsoknownasnon-additive join decomposition.

When the sub relations combine again then the new relation must be the same as the original relation was before decomposition.

ConsiderarelationRifwedecomposeditintosub-partsrelationR1andrelationR2. The decomposition is lossless whenit satisfies the following statement −

- If weunionthesubRelationR1 andR2thenit mustcontainalltheattributesthatareavailablein the original relation R before decomposition.
- Intersections of R1and R2 cannotbe Null. The sub relation mustcontain a common attribute. The common attribute must contain unique data.

Thecommonattributemustbea superkeyofsubrelationseitherR1 orR2. Here,

R=(A,B,C)
R1=(A,B)
R2=(B,C)
The relation R has threeattributes A, B,andC.The relation R is decomposed into tworelation R1 and R2. . R1 and R2 both have 2-2 attributes.The common attributes are B.

TheValueinColumnBmustbeunique.ifitcontainsaduplicatevaluethentheLossless-join decomposition is not possible.

DrawatableofRelationRwithRawData−

R(A,B,C)

A	B	C
12	25	34
10	36	09
12	42	30

Itdecomposesintothetwosubrelations−

R1(A,B)

A	B
12	25
10	36
12	42

R2(B,C)

B	C
25	34
36	09
42	30

Now,wecancheckthefirstconditionforLossless-joindecomposition.

TheunionofsubrelationR1 andR2isthesameasrelationR.

$R_1 U R_2$=R

Wegetthefollowingresult−

A	B	C
12	25	34
10	36	09
12	42	30

The relation is the same as the original relation R. Hence, the above decomposition is Lossless-join decomposition.

FunctionalDependency

Functionaldependencyreferstotherelation ofoneattribute ofthe databasetoanother.Withthehelpof functional dependency, the quality of the data in the database can be maintained.

Thesymbolforrepresentingfunctionaldependencyis->(arrow).

ExampleofFunctionalDependency

Considerthefollowingtable.

EmployeeNumber	Name	City	Salary
1	bob	Bangalore	25000
2	Lucky	Delhi	40000

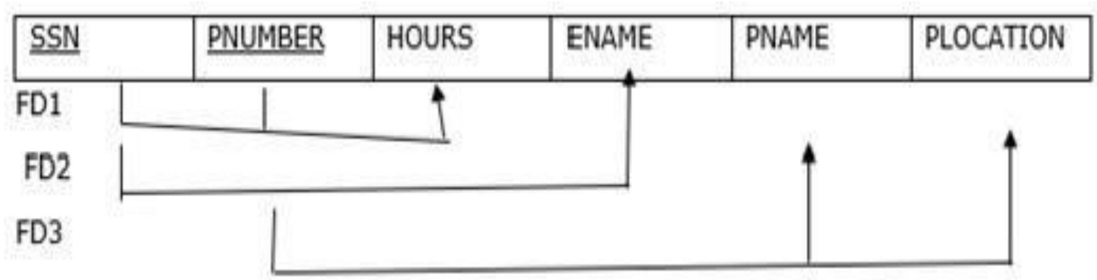
The details of the name of the employee, salary and city are obtained by the value of the number of Employee (or id of an employee). So, it can be said that the city, salary and the name attributes are functionally dependent on the attribute Employee Number.

Example

SSN->ENAMEreadasSSNfunctionallydependentonENAMEorSSN determines ENAME.

PNUMBER->{PNAME,PLOCATION}(PNUMBERdeterminesPNAMEandPLOCATION)

{SSN,PNUMBER}->HOURS(SSNandPNUMBERcombineddeterminesHOURS)



TransitiveDependency

Thetransitivedependencyisbeing obtainedbyusingtherelationofmorethanthreeattributes. These dependencies are being used to normalize the database in 3NF.

Example of Transitive Dependency

Consider the following table −

Book	Book_Author	Age_of_Author
ABC	Hari	45
PQR	James	60

Thedependenciesareasfollows−

{Book}->{Book_Author}

{Book_Author}doesnot->{Book}

{Book_Author}->{Age_of_Author}

Hence, as per thetransitivity, the{Book}-> {Age_of_Author}. Therefore, it oneknowsthe book then it must know the age of the Author.

Problem

ArelationR(ABCDEF)andF:{AB->C,C->A,B->DE,ABD->F}.Findthetransitivedependency.

Solution

AB+=ABCDE=>ABiscandidatekey

C+=CA

B+=BDE

ABD+=ABDFCE=>ABDisnotcandidatekey[sinceABiscandidatekey].

=>keyattribute={A,B}andnon-keyattributes={C,D,E}

AB-> C is not a transitive dependency.

C->Aisnotatransitivedependency.

B->DEisatransitivedependency[SinceBisnotacandidatekey/superkeyandDEisanon-key attribute}.

ABD->Fisnotatransitivedependency.[Since,ABDisasuperkey].

Multivalued dependency (MVD) is having the presence of one or more rows in a table. It implies the presence of one or more other rows in that same table. A multivalued dependency prevents fourth normal form. A multivalued dependency involvesat least three attributes of a table.

When existence of one or more rows in a table implies one or more other rows in the same table, then the Multi-valued dependencies occur.

IfatablehasattributesP,QandR,thenQandRaremulti-valuedfactsofP.

Itisrepresentedbydoublearrow−

->->

Forourexample:

P->->QP->->R

Intheabovecase,MultivaluedDependencyexistsonlyifQandRareindependentattributes. A table with multivalued dependency violates the 4NF.

Example

Letusseeanexample&mins;

<Student>

StudentName	**CourseDiscipline**	**Activities**
Amit	Mathematics	Singing
Amit	Mathematics	Dancing
Yuvraj	Computers	Cricket
Akash	Literature	Dancing
Akash	Literature	Cricket
Akash	Literature	Singing

Intheabovetable,wecanseeStudents**Amit**and**Akash**haveinterestinmorethanoneactivity.
This is multivalued dependency because**CourseDiscipline**of astudent are independent of Activities, but are dependent on the student.
Therefore,multivalueddependency−

StudentName->->CourseDisciplineStudentName->->Activities

TheaboverelationviolatesFourthNormalForminNormalization.

Tocorrectit,dividethetableintotwoseparatetablesandbreakMultivalued Dependency −**<StudentCourse>**

StudentName	**CourseDiscipline**
Amit	Mathematics
Amit	Mathematics
Yuvraj	Computers
Akash	Literature

Akash	Literature
Akash	Literature

<StudentActivities>

StudentName	**Activities**
Amit	Singing
Amit	Dancing
Yuvraj	Cricket
Akash	Dancing
Akash	Cricket
Akash	Singing

Thisbreaksthemultivalueddependencyandnowwehavetwofunctionaldependencies−

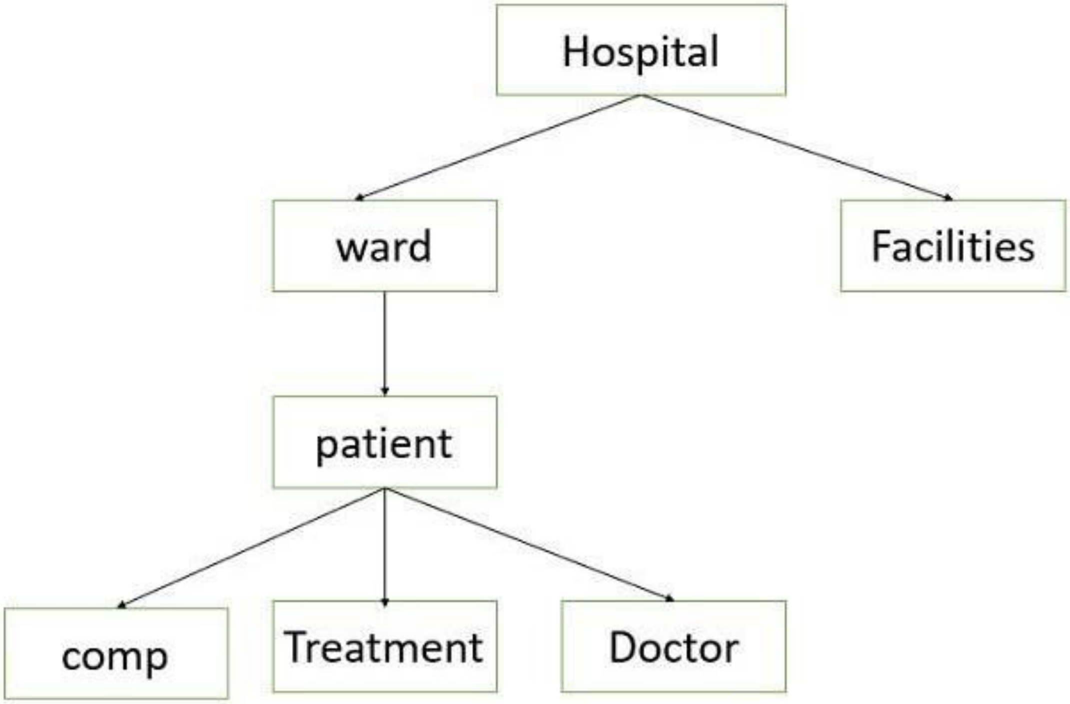
StudentName->CourseDisciplineStudentName->Activities

Joindependency

Join dependency isa constraint which is similar to functional dependency or multivalueddependency. It is satisfied if and only if the relation concerned is the join of a certain number of projections. Such type of constraint is called join dependency.

Let'sconsidera special classofjoin dependencies which helpto capturedatadependenciespresent in a hierarchical data structure.

Example1



The above hierarchical organization informs regarding ward and patients currently admitted to a ward depend only on the hospital but not the facilities present in that hospital. Since hospitals have multiple wards, functional dependencies are not adequate to describe the data dependency among hospitals and wards or facilities.

Inthiscase,multivalueddependencies,

Hospital->-> ward or
Hospital->->facilitieshold.

Using first order hierarchical decomposition wouldenableustorepresent data dependencies present in hierarchical data structure in a more natural way.

Thuswecanstorehospitaldatabaseaslosslessjoinofthefollowing− Hospital_facility(hospital, facilities),
Hospital_ward(hospital,ward,patient,complaints,treatment,doctor) Example
2
A relation R satisfies join dependency if R is equal to the join of R1,R2,.....Rn where Ri is a subset of the set of attributes of R.

RelationR

Dept	Subject	Name
CSE	C	Ammu
CSE	C	Amar
CSE	Java	Amar
IT	C	bhanu

Here,

dept->->subject
dept->-> name
The above relation is in 4NF. Anomalies can occur in relation in 4NF if the primary key has three or more fields. The primary key is(dept, subject, name). Sometimes decomposition ofa relation into two smaller relations does not remove redundancy. In such cases it may be possible to decompose the relation in three or more relations using 5NF.

Theaboverelation saysthatdeptoffersmany electivesubjectswhicharetaken byavariety ofstudents. Students have the opinion to choose subjects. Therefore all three fields are needed to represent the information.

The above relation does not show non-trivial MVDs since the attributes subject and name are dependent; they are related to each other (A FD subject->name exists). The relation cannot be decomposed in two relations (dept, subject) and (dept,sname).

Thereforetherelationcanbedecomposedintofollowingthreerelations− R1(dept,

subject)
R2(dept,name)and
R3(subject,name)anditcanbeshownthatdecompositionislossless.

R1

Dept	Subject
CSE	C
CSE	Java
IT	C

R2

Dept	Name
CSE	Ammu
CSE	Amar
IT	bhanu

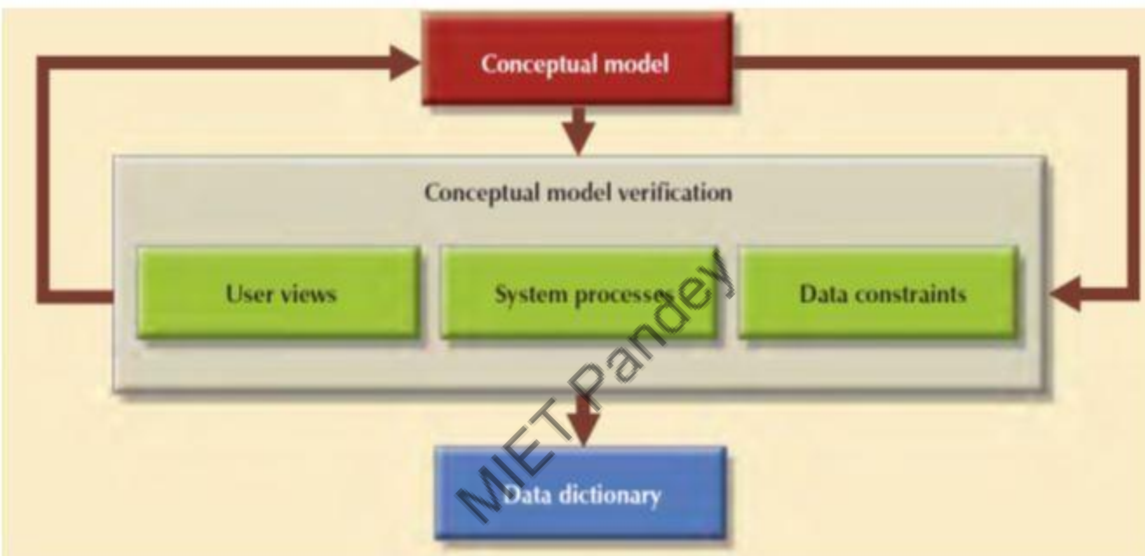
R3

Subject	Name
C	Ammu
C	Amar

Subject	Name
Java	Amar
C	bhanu

OTHERDESIGNINGAPPROACHES

Centralizeddesign:
Decentralized design might be used when the data component ofthe system hasa considerable number ofentitiesandcomplexrelationsonwhichverycomplexoperationsareperformed.Decentralizeddesign is also likely to be employed whenthe problem itself is spreadacross severaloperational sitesand each element isa subset of the entire data set.Consider the following diagram.



Centralized design

Decentralizeddesign:

Inlargeandcomplexprojects,thedatabasedesigntypicallycannotbedonebyonlyoneperson.Instead, acarefullyselectedteamofdatabasedesignersisemployedtotackleacomplexdatabaseproject.Within the decentralized design framework,the database design task is divided into several modules.Oncethe design criteria have been established, the lead designer assigns design subsets or modules to design groupswithin the team.Each design group createsa conceptual data model correspondingto the subset being modeled. Each conceptual modelisthen verified individually against the user views, processes, and constraints for each of the modules. After the verification process has been completed,all modules areintegratedintooneconceptual model.Naturally,after the subsetshavebeenaggregatedintoa larger conceptual model, the lead designer must verify thatthe combined conceptual modelis stillable to support all of the required transactions.

Keepinmindthattheaggregationprocessrequiresthedesignertocreateasinglemodelinwhich various aggregation problems must be addressed.

• Synonymsand homonyms. Various departments might know the same object by different names (synonyms),ortheymightusethesamenametoaddressdifferentobjects(homonyms).Theobjectcan be an entity, an attribute, or a relationship.

• Entityandentitysubtypes. Anentitysubtypemightbeviewed asaseparateentitybyoneormore departments. The designer must integrate such subtypes into a higher-level entity.

• Conflictingobjectdefinitions.Attributescanberecordedasdifferenttypes(character,numeric),or different domains can be defined for the sameattribute. Constraint definitions, too, can vary.The designer must remove such conflicts from the model.