M T W T F S S

Page No.:
Date:
YOUVA

Unit-2

Database design

Topics to cover:

Entity Relationship Model - ER Diagrams - Enhanced ER model - ER to Relational Mapping - functional dependencies - Non-loss decomposition - First, Second, Third Normal forms dependency preservation - Boyce Codd Normal Form - Multi Valued Dependencies and Fourth Normal Form - Join Dependencies - Recovery based on deferred.

Part-A

1) ER model: It is a map of entities and their relation, Entity is a real world object
2) Weak entity set has no primary key, Strong entity sets have primary key
3) Limitation of E-R: does not show complex constraints clearly, can not represent procedures, limited in modelling, overcomed by EER model
4) Specialization: Top-down approach divides an entity to sub entities based on characteristics
5) Aggregation: Treats a relationship as a higher-level entity to form another relationship shows whole part relationship
6) Characteristics of relational database: Controls redundancy, secured, multiple users are allowed
7) Primary key - Unique, Candidate keys Attributes that are primary, Superkey - primary key
8) Functional dependency: $X \rightarrow Y$ X uniquely determines Y. If two tuples have same X values, they must have the Y values
9) Entity: It is a real world object / conceptual object
10) Entity set: A group of similar type of entities having same attributes
11) Attributes: Columns / properties of the entity
12) Derive Attributes: It is created from another attribute
13) Recursive relationship: Entities relates to itself
14) Cardinality: Specifies how many times one entity can be associated with another
15) Key: used to uniquely identify elements

16) Relational schema: Definition of relation attributes and primary keys
17) Use of functional dependency: Test legal relations, specify constraints, helps in normalization
18) Normalization: Process of decomposing tables
19) Functional dependency rule: Reflexive, Augmentation, Transitive
20) Data Anomalies: problems caused by redundancy like repeated insertion
21) 1NF: no repeating groups, Atomic values only, each table has a primary key
22) 2NF: 1NF + no partial dependency
23) 3NF: 2NF + no transitive dependency
24) BCNF: A relation is in BCNF if every determinant is a candidate key
25) Decomposition properties: no loss of data, all dependencies preserved
26) MVD: One attribute determine multiple independent values is called Multivalued dependency (MVD)
27) 4NF: BCNF + no non trival Multivalued dependencies
28) 5NF: Every Join dependency is implied by the candidate keys
30) Object Based Logical Models: ER Model, Object Oriented, Semantic, Functional
31) Record Based Logical Models: Relation, network, Hierarchical
32) Domain: Set of valid values for an attributes
33) Candidate key: Attributes that qualifies to become primary key
34) Primary key: unique values only
35) Foreign key: Field in one table refers to another from other table
36) Alternate key: Candidate key other than choosen key
37) Composite key: primary key made of more than one attribute
38) Steps in query processing: Parsing & Translation, optimization and Evaluation
40) Trival dependencies: $\alpha \rightarrow \beta$ is trival if β is contained in α

Normalization

Normalization is a step by step process used to design a good database structure by removing redundancy, avoiding anomalies and ensuring data integrity

1NF:

*) Each row has atleast one unique value
*) No repeating or multivalued attributes
*) Each column contains atomic values

Ex:

before 1NF

Student	Subject
A	M, E

1NF →

Student	Subject
A	M
A	E

2NF:

*) It is 1NF
*) It has no partial dependencies

Ex: before 2NF

Student	Course	Instructor
A	B	C

2NF →

Student table

Student	Course
A	B

Instructor table

Course	Instructor
B	C

Student is dependent on course and Instructor is not dependent on Student directly & called partial dependencie

3NF

*) It is 2NF
*) It has no transitive dependencies

Example: | EMP ID | EmpName | Dept ID | Dept Name |
Dept Name depends on Dept ID

2NF →

Employee Table

EMP ID	EmpName	Dept ID

Dept table

Dept ID	Dept Name

BCNF Boyce codd Normal form:

*) It is 3NF
*) For each functional dependencie
    Ex: A → B
        A Must be Superkey
*) In 3NF Non key → Non key is allowed
    but BCNF removes that feature

4NF:

*) It is in BCNF
*) Has no Multivalued dependency

Example:

| Book | Author | Language |

In that a Author can write more than in one Language and A Book can have more than one author, then

| Book | Author | is a table also | Author | Language | is another

And repeated values are made as once

5NF:

*) It is in 4NF
*) Every Join dependencies can be reconstructed without loss
*) The table cannot be further decomposed without losing information

Ex: Book | Author | Language —5NF→

Table 1:
Book | Language

Table 2:
Book | Author

Table 3:
Author | Language

And same as 4NF rules

M T W T F S S

Page No.: YOUVA

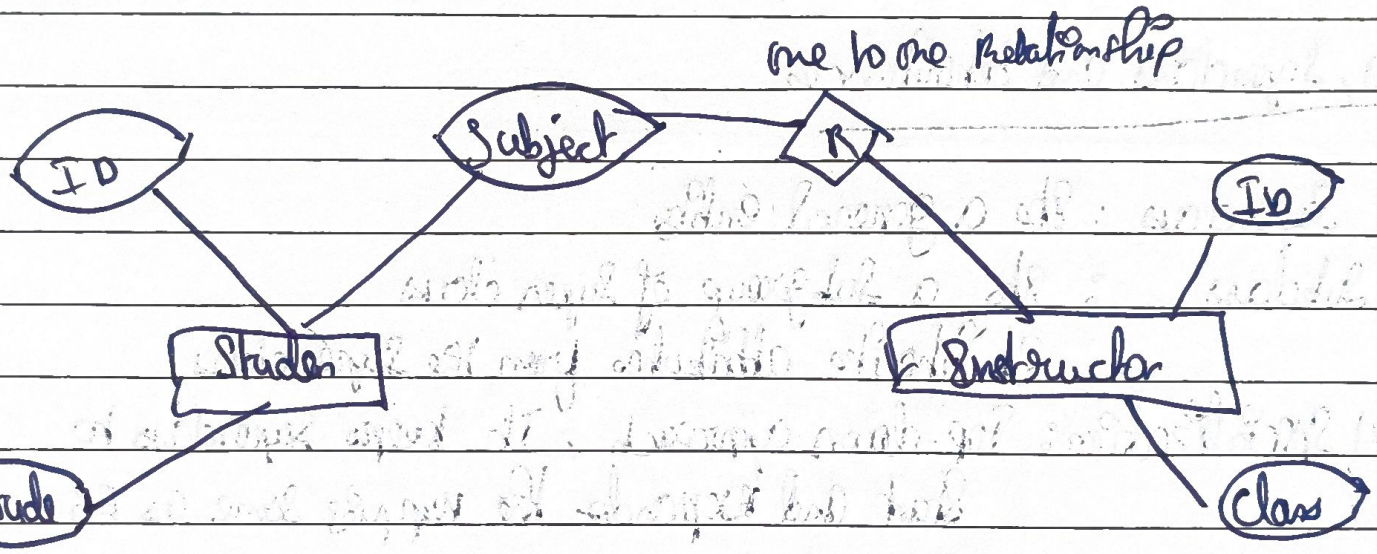
Date:

Components of ER model

E-R diagram is a visual representation of real world data.
It shows Entities ▭, Attributes ⬭, relation ◇

↓	↓	↓
Real world objects	Properties of Entities Columns	Relationship between Entity and Attributes

Ex:



Relationships have few types
- one to one
- one to many
- many to one
- many to many

M T W T F S S

Page No.: YOUVA

Date:

~~Eng~~

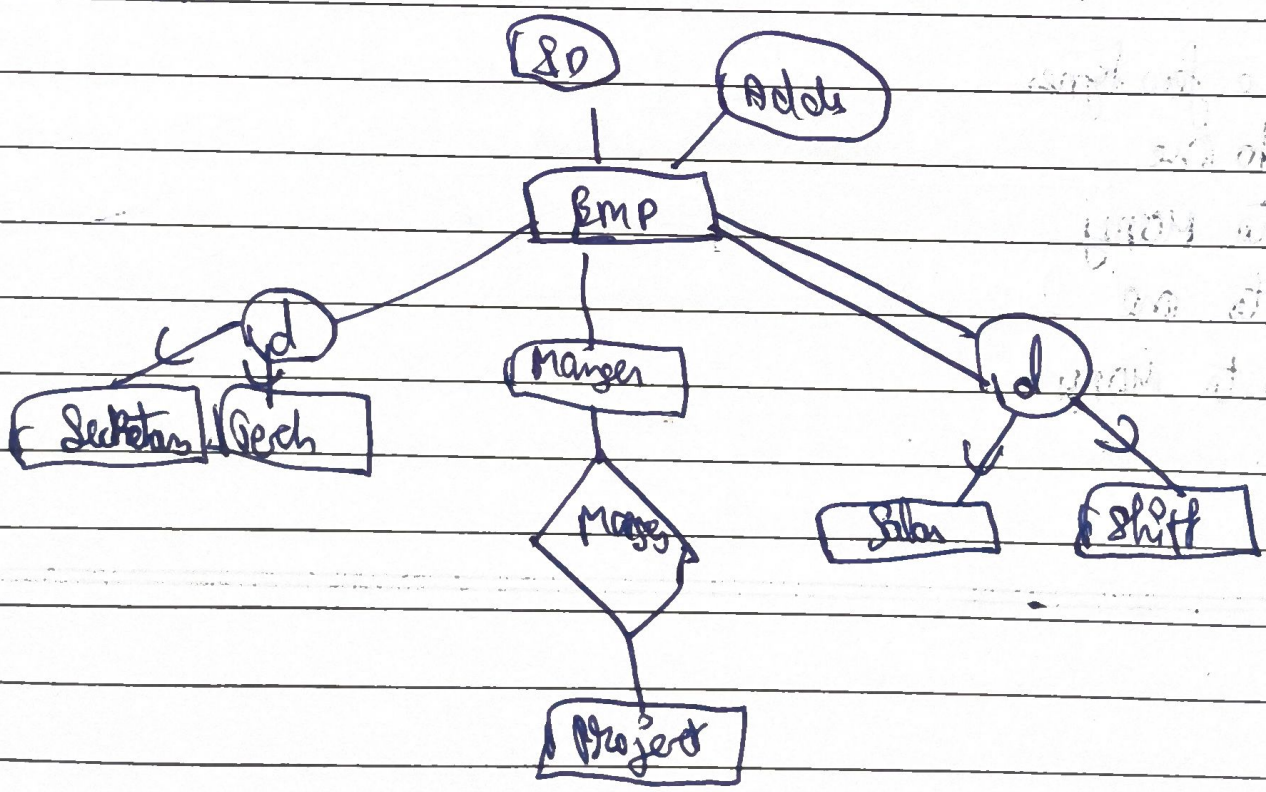
Enhanced Entity-Relationship Model

*| It is same as ER
*| It adds the Subclass -> divides entities into meaningful subgroups
Specialization → Top-down approach for subclasses
Generalization → Bottom to top approach for combining subclasses
Categorization → when subclass represents a union of types different entity types

Subclass, superclass and Inheritances

*| Superclass : It's a general entity
*| Subclass : It's a subgroup of super class
It inherits attributes from the super class
*| Specialization: Top-down approach, It keeps superclass to start and expands the mapping same as ER

ER



M T W T F S S

Page No.:	YOUVA
Date:	

Generalization:

Combines two similar / simple ER model together with common entity and attributes

Constraints in specialization & generalization:

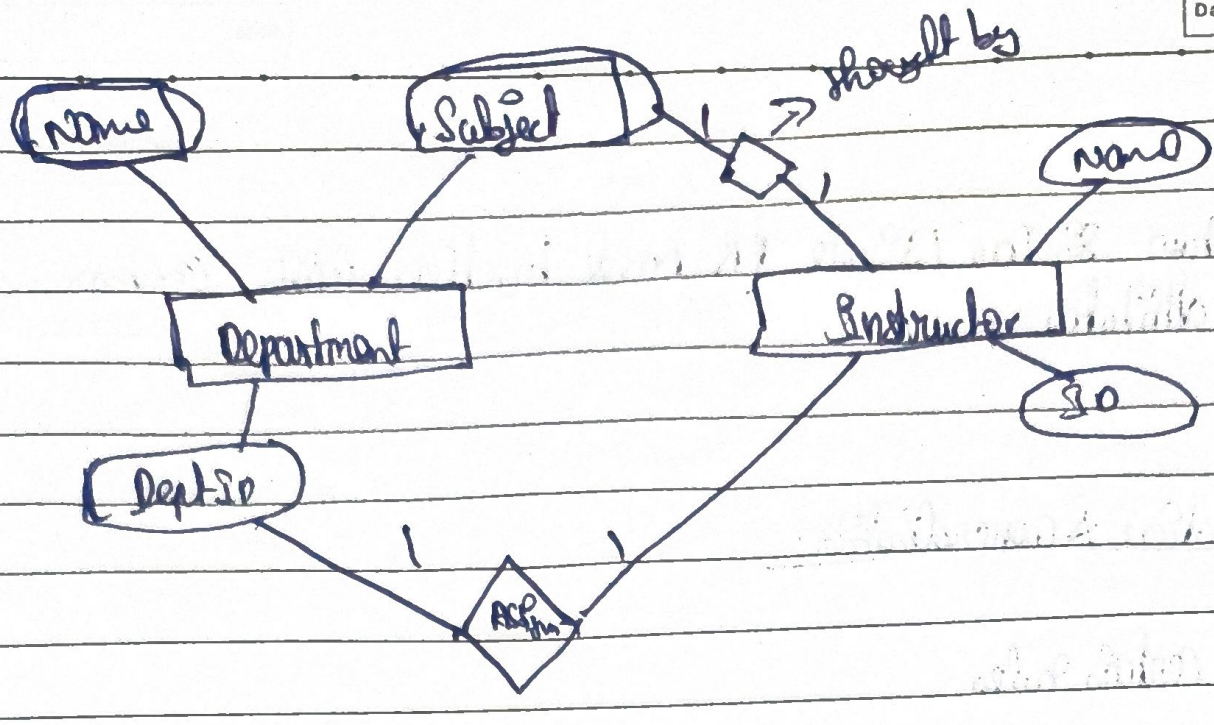
Subclass follow certain rules

1) predicate defined : If all subclass uses same superclass attributes, that is called attributed defined specialization
2) Disjointness constraint: Entity belongs to only one subclass, overlapping opposite
3) completeness constraint: Defines whether every superclass entity must belong to a any subclass
   Total : Every entity must be in some subclass
   Partial : Some entities may not belong to subclass

ER to Relational mapping

*) It is method of representing tables with their columns with relations and connection
*) Tables -> Entities (Real world Objects)
   Column -> Attributes
   Relation -> How it is connected

Step 1 : Identify Entities
Step 2 : Identify Relationships + Cardinality -> 1-1 or 1 to N
Step 3 : Identify key Attributes Ex: Regd (entity), Regd name is the attribute
Step 4 : Identify other Attributes
Step 5 : Draws ER diagram



Its a simple E-R design.

Data Model and Its types

A Data Model is a Conceptual tool used to describe

- How data is stored
- How data is connected
- What rules apply
- Data meaning

A Data Model describes database as

1) Physical level
2) Conceptual level / Logical level
3) View level

Type of Data Models:

1) Relational Data Model
   - *) Most widely used
   - *) Represents data in tables
   - *) Each tables : rows x columns
   - *) Record based model
   - *) Supports primary key, foreign keys
   - *) Used in DBMS like MSSQL, Oracle

2) Entity - Relationship model : Just same point from E-R
3) object based data model : OOPs, encapsulation, Inheritance
4) Semi-structured data model : It does not follow fixed structure across columns, Each data can have different attributes
XML, JSON
5) Other models : Hierarchical ; network model
↓ ↓
Tree graph