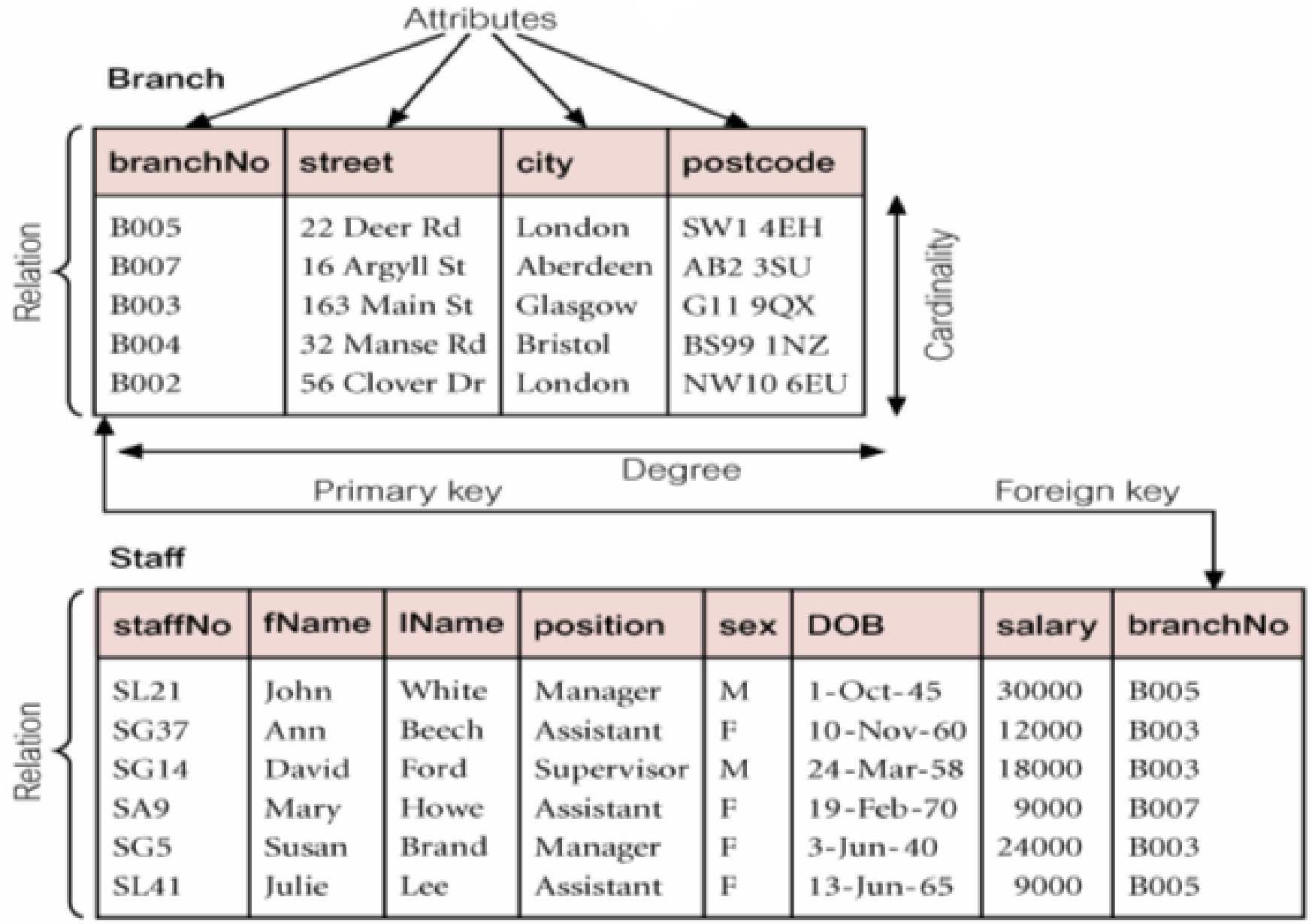
UNIT 2

Lecture 10

Relational Model

RDBMS Terminology

Relational terms	Previously used terms
Relation	Table
Tuple	Row, Record
Cardinality	Number of rows
Attribute	Column, Field
Degree	Number of columns
Primary Key	Unique Identifier
Domain	Set of Legal Values


Attributes
Branch
branchNo
street
city
postcode
B005
22 Deer Rd
London
SW1 4EH
B007
16 Argyll St
Aberdeen
AB2 3SU
B003
163 Main St
Glasgow
G11 9QX
B004
32 Manse Rd
Bristol
BS99 1NZ
B002
56 Clover Dr
London
NW10 6EU
Relation
Cardinality
Degree
Primary key
Foreign key
Staff
staffNo
fName
lName
position
sex
DOB
salary
branchNo
SL21
John
White
Manager
M
1-Oct-45
30000
B005
SG37
Ann
Beech
Assistant
F
10-Nov-60
12000
B003
SG14
David
Ford
Supervisor
M
24-Mar-58
18000
B003
SA9
Mary
Howe
Assistant
F
19-Feb-70
9000
B007
SG5
Susan
Brand
Manager
F
3-Jun-40
24000
B003
SL41
Julie
Lee
Assistant
F
13-Jun-65
9000
B005
Relation

Student-project database

STUDENT

Rollno	Sname	Sem	Branch	Marks	Pno
1	RAM	3	CSE	40	121
2	SHYAM	5	CSE	50	122
3	MOHAN	7	CSE	55	123
4	GOPAL	5	IT	65	121
5	RINKI	3	MECH	40	122
6	PINKI	3	ETC	90	123

PROJECT

Pno	Pname	Duration
121	P1	10
122	P2	20
123	P3	30

Key

- A value used to identify a record in a database is called as the key.
- In other words a subset of the fields within a table for which data must be entered and validated before a new record may be added to the table is called as the key field or key attribute.

STUDENT

Rollno	Sname	Sem	Branch	Marks	Pno
1	RAM	3	CSE	40	121
2	SHYAM	5	CSE	50	122
3	MOHAN	7	CSE	55	123
4	GOPAL	5	IT	65	121
5	RINKI	3	MECH	40	122
6	PINKI	3	ETC	90	123

Types of Keys in DBMS

1. Super Key
2. Candidate Key
3. Primary Key
4. Alternate Key
5. Secondary Key
6. Composite Key
7. Foreign Key
8. Semantic or natural Key
9. Technical or Surrogate or Artificial Key

STUDENT

Sno	Rollno	Enrollno	Fname	Lname	Sem	City
1	3012205001	AA-12345	Ram	Sharma	7	Bhilai
2	3012205002	AA-30122	Ram	Sharma	7	Raipur
3	3012206001	AB-20333	Shyam	Verma	5	Bhilai
4	3012206002	AB-33339	Mohan	Verma	5	Raipur
5	3012207001	AC-44456	Gopal	Gupta	3	Durg

Super Key

- A Super key is a set of one or more attributes which taken collectively, allow us to identify uniquely an entity instance (tuple) in the entity set (relation).
- For e.g. From the student table the super keys are {sno}, {rollno}, {enrollno}, and {fname, lname, sem, city}.
- Once specific characteristic with super key is that, as per its definition any combination of attributes with the super key is also a super key.
- For e.g. From the student table the combination of {sno} and {rollno} i.e. {sno, rollno} is also a super key.

Composite Key

- A composite key is a set of two or more attributes which taken collectively, allow us to identify uniquely an entity instance (tuple) in the entity set (relation).
- For e.g. From the student table the composite keys is {fname, lname, sem, city}.

Candidate Key

- A super key for which no subset is a super key is called a candidate key,
- or the minimal super key is the candidate key.
- It means that there are two conditions for the candidate key,
  1. It identifies the entity instances uniquely, as is required in case of super key,
  2. It should be minimum, that is, no proper subset of candidate key is a key.
- So if we have a simple super key, that is, that consists of single attribute, it is definitely a candidate key, 100%.
- However, if we have a composite super key and if we take any attribute out of it and remaining part is not a super key anymore then that composite super key is also a candidate key since it is minimal super key.
- For e.g., From the student table the {sno}, {rollno}, and {enrollno} are the candidate keys.

Primary Key

- A candidate key chosen by database designer as a principle means of identifying an entity (tuple) in an entity set (relation) is called as primary key.
- An entity type (relation) may have more than one candidate keys; in that case the database designer has to designate one of them as primary key, since there is always only a single primary key in an entity type (reation).
- If there is just one candidate key then obviously the same will be declared as primary key.
- For e.g., From the student table any one of the candidate key from {sno}, {rollno}, and {enrollno} is chosen to serve as primary key.
- A certain value that may be associated with any attribute is NULL, that means “not given” or “not defined”.
- A major characteristic of the PK is that it cannot have the NULL value.
- If PK is a composite, then none of the attributes included in the PK can have the NULL.

Alternate Key

- Candidate keys which are not chosen as the primary key are known as alternate keys.
- For e.g., we have three candidate keys of student, {sno}, {rollno} and {enrollno}, if we select {rollno} as PK then the {sno} and {enrollno} will be alternate keys.

Secondary Key

- Many times we need to access certain instances of an entity type using the value(s) of one or more attributes other than the PK.
- The difference in accessing instances using the value of a key or non-key attribute is that the search on the value of PK will always return a single instance (if it exists), where as uniqueness is not guaranteed in case of non-key attribute.
- Such attributes on which we need to access the instances of an entity type that may not necessarily return unique instance is called the secondary key.
- For e.g., we want to see how many of our students belong to Bhilai, in that case we will access those instances of the student entity type that contain “Bhilai” in their address. In this case city will be called secondary key, since we are accessing instances on the basis of its value, and there is no compulsion that we will get a single instance.

Foreign Key

- Foreign key is an attribute (or set of attributes) that serves to link one table to another table in the database, thereby establishing a relationship.
- Foreign key often (but not always) link to primary key of the other table.
- It is recommended that where a foreign key is required the same name as that of the associated key on the foreign table be used.
- It is requirements of a relational join that two relations can only be join when they share at least one common attribute, and this should be taken to mean the attribute name(s) as well as value(s).
- For e.g., consider the STUDENT and PROJECT tables, in which the Pno attribute of PROJECT table becomes the foreign key for STUDENT table.

Semantic or Natural Key

- A semantic or Natural Key is a key for which the possible values have an obvious meaning to the user or the data.
- In relational model database design, a natural key is a candidate key that has a logical relationship to the attributes within that row.
- The main advantage of a natural key over a surrogate key, which has no such logical relationship, is that it already exists; there is no need to add a new, artificial column to the schema.

Semantic or Natural Key

- The main disadvantage of choosing a natural key is that its value may change and the relational database engine may not be able to propagate that change across the related foreign keys.
- For example, if person_name is used as the primary key for the person table, and a person gets married and changes name, then all of the one-to-many related tables need to be updated also.
- The secondary disadvantage of choosing a natural key is identifying uniqueness. The primary key must consist of the attributes that uniquely identify a row. However, it may be difficult (or it may add constraints) to create a natural key on a table.
- For example, if person_name is used as a primary key for the person table, many persons may share the same name and all but the first entry will be rejected as a duplication.
- The uniqueness constraint may be overcome by adding an additional column to the primary key, like street_address, to increase the likelihood of uniqueness.

Artificial Key or Technical Key or Surrogate Key

- A artificial key is a key for which the possible values have no obvious meaning to the user or the data. They are used instead of semantic key for any of the following reason:
  - When the value in a semantic key is likely to be changed by the user or can have duplicates.
- When none of the existing attributes can be used to guarantee uniqueness. In this case adding an attribute whose value is generated by the system.
- For e.g., Sno attribute of STUDENT relation is the surrogate key as whose contain a sequence of numbers which is the only way to provide a unique value.

GATE Questions

- Consider a relational table with a single record for each registered student with the following attributes:
  I. Registration_Number: Unique registration number for each registered student.
  II. UID: Unique Identity number, unique at the national level for each citizen.
  III. BankAccount_Number: Unique account number at the bank. A student can have multiple accounts or joint accounts. This attribute stores the primary account number.
  IV. Name: Name of the Student.
  V. Hostel_Room: Room number of the hostel.

Which of the following options is INCORRECT?

(a) BankAccount_Number is a candidate key.

(b) Registration_Number can be a primary key.

(c) UID is a candidate key if all students are from the same country.

(d) If S is a superkey such that S ∩ UID is NULL then S ∪ UID is also a superkey.

(GATE 2011: 1 Mark)

GATE Questions

Given the basic ER and relational models, which of the following is INCORRECT?

a) An attribute of an entity can have more than one value.

b) An attribute of an entity can be composite.

c) In a row of a relational table, an attribute can have more than one value.

d) In a row of a relational table, an attribute can have exactly one value or a NULL value.

(GATE 2012: 1 Mark)

GATE Questions

Given an instance of the STUDENTS relation as shown below :

StudentD	StudentName	StudentEmail	StudentAge	CPI
2345	Shankar	shankar@math	X	9.4
1287	Swati	swati@ee	19	9.5
7853	Shankar	shankar@cse	19	9.4
9876	Swati	swati@mech	18	9.3
8765	Ganesh	ganesh@civil	19	8.7

For (StudentName, StudentAge) to be a key for this instance, the value X should NOT be equal to_____.

(GATE 2014: 1 Mark)

GATE Questions

The maximum number of super keys for the relation schema R (E, F, G, H) with E as the key is ___.

(GATE 2014: 1 Mark)

GATE Questions

A prime attribute of a relation scheme R is an attribute that appears

a) in all candidate keys of R.
b) in some candidate key of R.
c) in a foreign keys of R.
d) only in the primary key of R.

(GATE 2014: 1 Mark)

GATE Questions

Which of the following is NOT a super key in a relational schema with attributes V, W, X, Y, Z and primary key VY?

(A) VXYZ

(B) VWXZ

(C) VWXY

(D) VWXYZ

(GATE 2014: 1 Mark)

For Video lecture on this topic please subscribe to my youtube channel.

The link for my youtube channel is

https://www.youtube.com/channel/UCRWGtE76JlTp1iim6aOTRuw?sub_confirmation=1