DATABASE AND INFORMATION RETRIEVAL

DR. ELIEL KEELSON

LECTURE 02 – THE RELATIONAL DATA MODEL

SETS & RELATIONS

Sets and Relations

- The relational model is based on the mathematical theory of sets.

- A set can be viewed as a collection of zero or more items of similar type.

- Three most important characteristics of sets for our purposes are:
  - All members of the set are of the same type
  - Only one instance of any item is held in a set
  - The sequence of items in the set is not significant

Sets and Relations

- Given two sets X and Y, we can take any element x from X and y from Y to form an ordered pair (x,y).
- The set of all ordered pairs is called the product set and is denoted by X.Y
- A subset of X.Y is called a relation and can be denoted R(X,Y).
- A relation can be considered as a mapping from one set to another and given a functional name.

Sets and Relations

- The data of an application can be modeled as a two-dimensional table.

- Each relation defines and/or describes some area of the application and provides a mapping from an identifying value to other descriptive or qualifying attributes.

- The identifying value is sometimes called the ruling part while the rest of the attributes are collectively referred to as the dependent part.

Sets and Relations

- *Relation* is synonymous with *table* or *file*
- *Tuple* is synonymous with *row* or *record*
- *Attribute* is synonymous with *column* or *field*
- *Ruling part* is synonymous with *primary key*

- The domain of an attribute is the set of all possible values of that attribute.

Sets and Relations

A relation can be seen as having the following properties:

- Columns in the relation are all single values
- Entries in any column are all of the same datatype; e.g. integer, real number, character, etc
- No two rows of the relation are identical
- The order of the rows in the table are immaterial
- The order of the columns in the table are immaterial
- Each table contains an identifying column or columns

PRIMARY KEY

Primary Key

- The primary key of a table is a column (or a combination of two or more columns) that serves to identify the individual rows of the table.

Customer Number	Name	Address	Balance Owing
1234	Kwame	West Street	0.00
5676	Mary	North Lane	1.25
6565	Richard	River Lane	10.55

Primary Key

- The adoption of a set of unique codes is often used to simplify the definition of a primary key.

- Where there is no natural or existing code available for a relation, it is common practice simply to assign a sequential number to successive rows to serve as the key.

- Databases often provide a facility to generate these numbers automatically.

COMPOSITE PRIMARY KEY

Composite Primary Key

- The primary key may consist of more than one column.
- The key may be a concatenation of two or more columns.

Composite Primary Key

- In the table below, neither Project Number nor Engineer Surname can suffice as a primary key value on its own because the values in each column are not unique

- The combination of Project Number and Engineer Surname does produce a unique value suitable as a primary key.

Project Number	Engineer Name	Assignment Date
A2343	Mensah	25-Mar-09
Q9919	Daniels	09-Jun-09

FUNCTIONAL DEPENDENCY

Functional Dependency

Functional dependency can be defined as follows:

- If we say that one column B of a table is functionally dependent on another column A (or group of columns), it means that every value of A uniquely determines the value of B.

- This is often written using the notation A→B.

- If A→B, then it means that every time a particular value appears in the A column, then another particular value will appear in the B column.

Functional Dependency

Make	Model	Engine Size	Daily Rental GHC	Mileage Charge GHp
Ford	Escort	1400	30	20
Ford	Mondeo	1600	40	30
Nissan	Almera	1400	35	20
Renault	Megane	1400	37	20

Functional Dependency

- Functional dependency must be determined from knowledge of the application domain.

- You cannot determine whether a dependency exists simply by inspection of the table data

- In practice, such dependencies are often derived from ‘business rules’ of the application domain.

FOREIGN KEY

Foreign Key

- The primary key is used to refer to a specific row in a table.
- Primary key values can be included in a column of another table which is related in some way to the first table.
- Columns containing such values are called foreign keys.
- A foreign key is a column in one table that refers to the primary key of another table.

Foreign Key

Batch Number	Tutor Code	Date Sent	Date Returned
23	JS	26-Apr-07	07-May-07
24	GH	28-Apr-07	06-May-07
25	GH	1-May-07	06-May-07

Tutor Code	Assigns Marked
JS	87
GH	91

CANDIDATE KEYS

Candidate Keys

- In some tables it is possible to find that more than one column, or combination of columns, could serve as a primary key.
- Such alternative primary keys are called candidate keys.
- One of the possible candidate keys is chosen to be the primary key.

Candidate Keys

- Consider a lecturer table (from a university database) below:

Lecturer Id	Name	Department	Room No	Course
123	Daniels	Computer Engineering	B705	Database Systems
145	Mensah	Electrical Engineering	A111	Applied Electricity

NULLS

Nulls

- It often happens when inserting data into a database table that some of the attribute values cannot be entered for a variety of reasons.

- Possible reasons are:
  - The data is not available
  - The data is not applicable to this entity

- To provide a standard means of filling in columns of a table that would otherwise be empty, the null concept was devised.

- Although sometimes referred to as a ‘null value’, a null is not a value.

DB INTEGRITY

Other Relational Concepts & Terminology

Entity Integrity

- It is a defining principle of relational tables that each row of a table uniquely represents one entity in the application domain.

- It is necessary that no two rows of a table are the same; if this were allowed it would mean that the same application domain entity was represented by two rows of the table.

- Preservation of this principle is referred to as **Entity Integrity**.

Entity Integrity

- The use of unique primary key values guarantees that is principle is complied with.

- Entity integrity is the principle that no part of a primary key is null

Referential Integrity

- Referential integrity is concerned with the linkages between tables defined by the foreign and primary key fields.
- A foreign key is an attribute in one table that refers to the primary key in another table.
- For a set of database tables, all foreign key values in all tables must be matched by a row in another table.
- A database for which this is true is said to conform to referential integrity.

Referential Integrity : Anomalous Example

CourseCode	CourseTitle	CourseDept
COE251	C Programming	Computer Engineering
EE151	Applied Electricity	Electrical Engineering
MATH151	Mathematics I	Mathematics

IndexNum	StudentName	CourseCode
12345	Johnny	EE151
54321	Samuel	COE454
12444	Ben	COE251

Referential Integrity

- A database that does not exhibit referential integrity is in an anomalous and impractical condition and will likely produce serious run-time failures.

- It is important that referential integrity be maintained throughout the database and most database systems now provide facilities to assist in complying with this.

RELATIONAL ALGEBRA

Relational Algebra

- Relational database systems are expected to be equipped with a query language that can assist its users to query the database instances.

- There are two kinds of query languages: **Relational Algebra** and **Relational Calculus**.

Relational Algebra

- Relational databases are based on the mathematical notion of a relation; i.e. a set of mappings from independent values (keys) to dependent values.

- The mathematical theory defines a number of algebraic operations on relations that produce new relations from one or more originals.

Relational Algebra

- Relational algebra is a procedural query language, which takes instances of relations as input and yields instances of relations as output.

- It uses unary or binary operators to perform queries.

- They accept relations as their input and yield relations as their output.

Relational Algebra

The fundamental operations of relational algebra are as follows:

- SELECTION: Form new relation from selected rows of input relation.

- PROJECTION: Form new relation from selected columns of input relation.

- JOIN: Form new relation by ‘joining’ rows of two or more input relations.

Relational Algebra

- UNION: Form new relation by combining rows from two input tables.
- PRODUCT: Form new relation by joining every row in one table with every row of a second
- DIFFERENCE: Form the difference of two relations, a third relation containing rows that occur in the first relation but not in the second
- INTERSECTION: Form the intersection of two relations, a third relation containing rows that appear in both the first and second relations.

Relational Algebra

- UNION: Form new relation by combining rows from two input tables.
- PRODUCT: Form new relation by joining every row in one table with every row of a second
- DIFFERENCE: Form the difference of two relations, a third relation containing rows that occur in the first relation but not in the second
- INTERSECTION: Form the intersection of two relations, a third relation containing rows that appear in both the first and second relations.

THANKS!

Any questions?

You can find me at elielkeelson@gmail.com &
ekeelson@knust.edu.gh