

Design Database

- **Step 1 – Specify requirements**
- Step 2 – Model data using Entity-Relationship (ER) model
- Step 3 – Derive ER model to a relation schema

Step 1: Specifying Requirements

- Defining data requirements need an understanding of how customers or users will use the data
- Things to consider
 - What data is available?
 - How the data is related?
- Exercise: What data is required in for a bank?

Step 1: Specifying Requirements

- Exercise: What data is required in for a bank
- Q1: What data is available?
 - **Bank accounts:** name, number, type (cheque or savings), opening date, opening branch, balance
 - **Bank branch:** name, address, phone
 - **Customer details:** Name (first, middle, last), ID, address, telephone number, email, tax number
 - **Account Transactions:** date, type (deposit, withdraw, transfer), amount
 - **Employee:** Name (first, middle, last), position, working branch, home address, telephone number, email, tax number
- Q2: How the data is related?
 - Can one customer has many bank accounts?
 - Will one bank branch has many customers & bank accounts?
 - Can customers has a joint bank account?
 - Can we store account transaction of all customers in one place?

Design Database

- Step 1 – Specify requirements
- **Step 2 – Model data using Entity-Relationship (ER) model**
- Step 3 – Derive ER model to a relation schema

Step 2: Modelling Data

- Entity-Relationship (ER) model is one model designed to define data requirements
- ER Diagram illustrates the relationship of the data
- Four main concepts:
 - Entity
 - A relevant thing or object being modelled
 - Attribute
 - Property of an entity
 - Relationship
 - An association between entities
 - Cardinality
 - The degree of the relationship

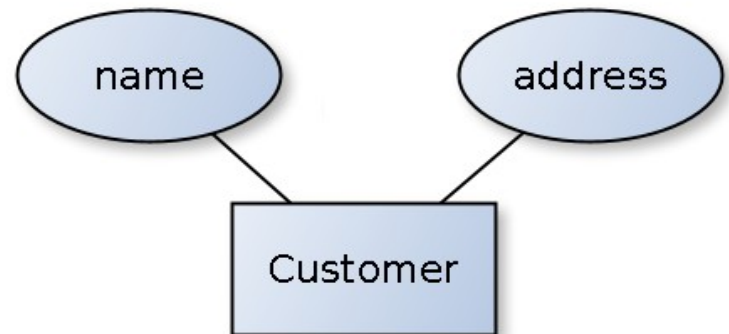
Step 2: Modelling Data - Entities

- Entity – An object that is distinguishable from other objects
 - John Smith with an ID a1234567
- Entity Set - A collection of entities that all share the same property
 - John Smith and Ashanta Pradeep are both **Customers**
- Entity Sets do not have to be disjoint
 - John and Ashanta may also be **Employees** as well as **Customers**
- ER Diagram: An Entity Set is represented by a rectangle
 - Labelled with name (a noun)



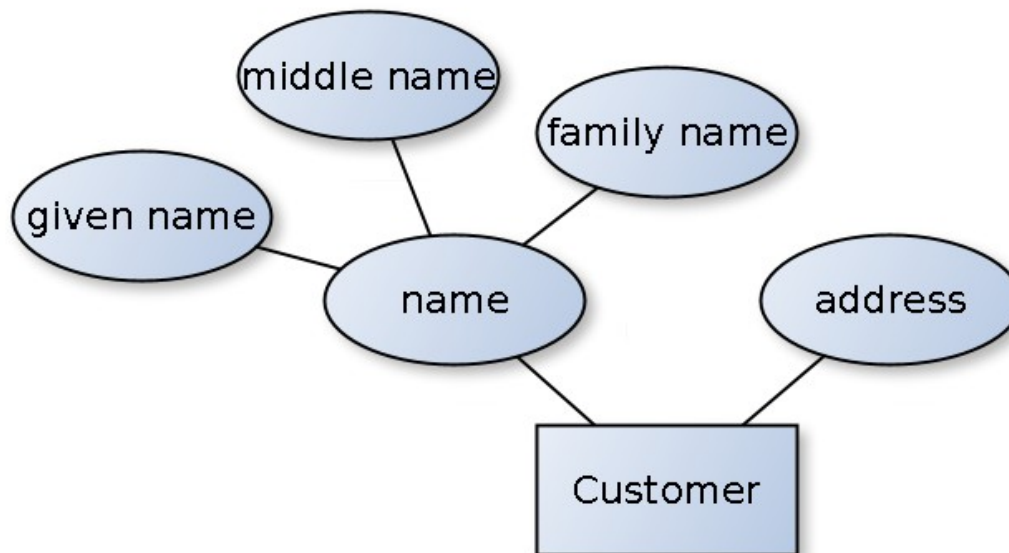
Step 2: Modelling Data - Entity Attributes

- An entity is represented by a set of attributes
 - An entity is often called a *row* of a table (or a tuple)
- Domain of an attribute is the set of permitted values for that attribute
 - Numbers, strings, dates, etc.
- ER Diagram: An entity attribute is represented by an ellipse
 - Connected to an entity set (rectangle)
 - Also named as nouns
- Attribute types:
 - Simple vs Composite attributes
 - Single-value vs Multi-value attribute
 - Derived attributes



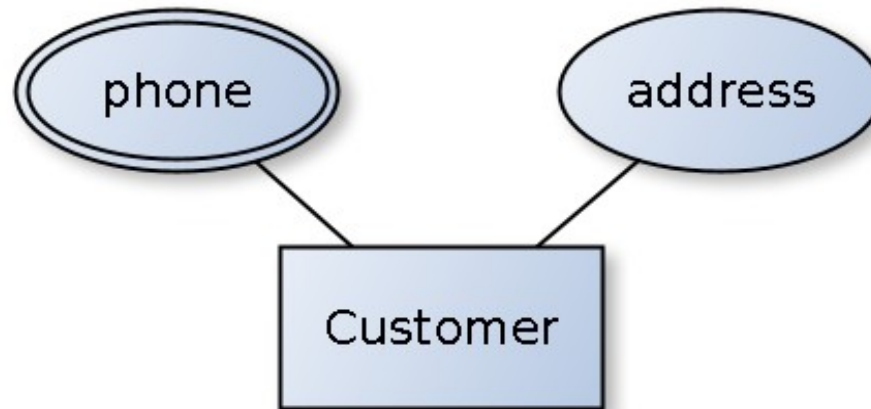
Step 2: Modelling Data - Entity Attributes

- **Composite Attributes** are attributes that are made up of multiple parts
 - For example a name: first name, middle name, last name
 - Drawn as attributes of attribute



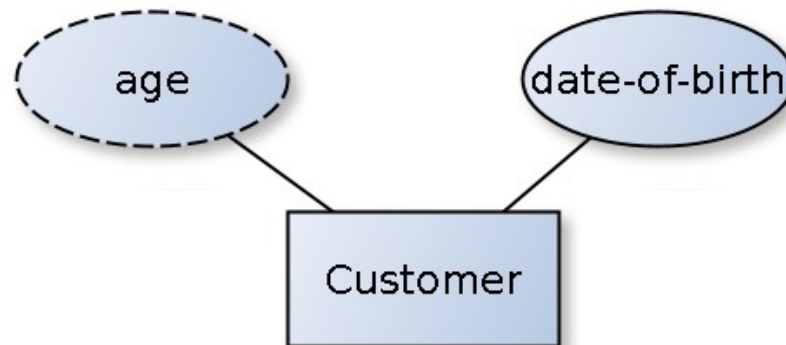
Step 2: Modelling Data - Entity Attributes

- **Multi-value Attributes** may have zero, one or more values
 - For example, a phone number: home, mobile, home + mobile, or even no phone.
 - Drawn as double ellipses



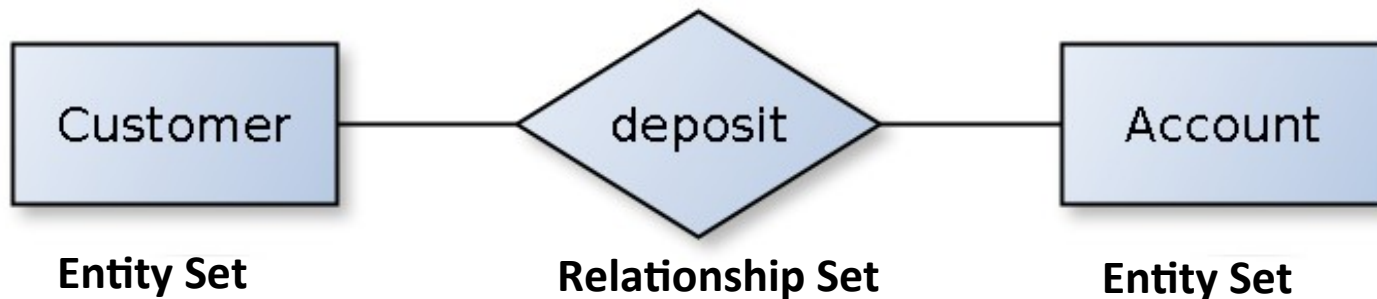
Step 2: Modelling Data - Entity Attributes

- **Derived Attributes** can be determined by existing attributes
 - For example, age can be derived from the date-of-birth attribute and current date
 - Drawn as dashed ellipses



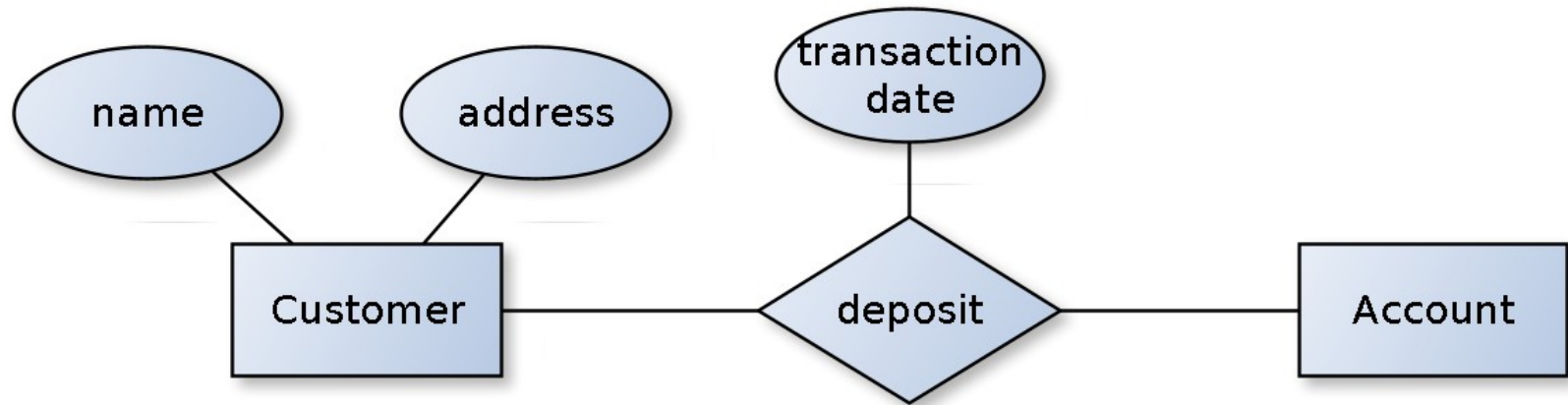
Step 2: Modelling Data - Relationships

- A relationship is an association among entities
 - John Deposit A123
(Customer entity) (a relationship) (Account entity)
- Relationship set is a set of relationships of similar type
 - $\{(John, A123), (Marry, A124), \dots\} \in \text{Deposit}$
- ER diagram: a relationship set is represented by diamond
 - Connected to rectangles
 - Named with verbs (usually)



Step 2: Modelling Data - Relationship Sets

- An attribute can also be a property of a relationship set not just entity set
 - (Sometimes called descriptive attributes)
- Still represented by ellipse but connected to relationship set diamond

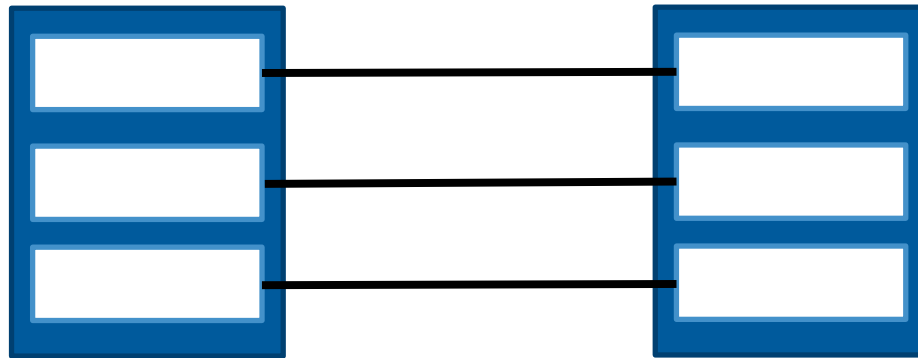


Step 2: Modelling Data : Cardinalities

- Cardinalities express the number of entities to which another entity can be associated via a relationship set
- 4 types of cardinalities
 - One to one
 - One to many
 - Many to one
 - Many to many

Step 2: Modelling Data : Cardinalities

- One to one relationship
 - One entity in an entity set is associated with one and only one entity in another entity set

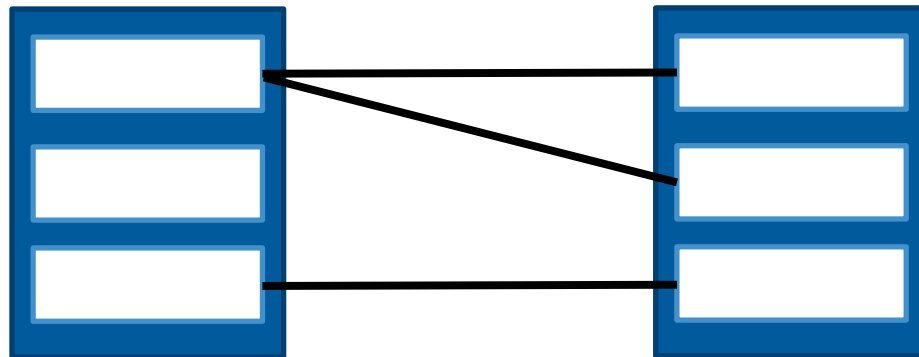


- ER diagram: 'One' is represented by a stroke or a 1 on the line between the relationship set and the entity set

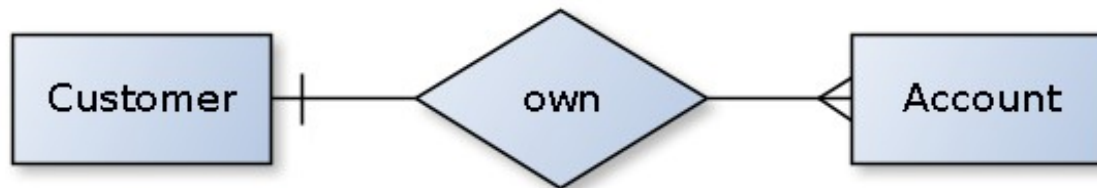


Step 2: Modelling Data : Cardinalities

- One to Many relationship
 - One entity in an entity set is associated with many entities in another entity set
 - Example: One customer can have several accounts

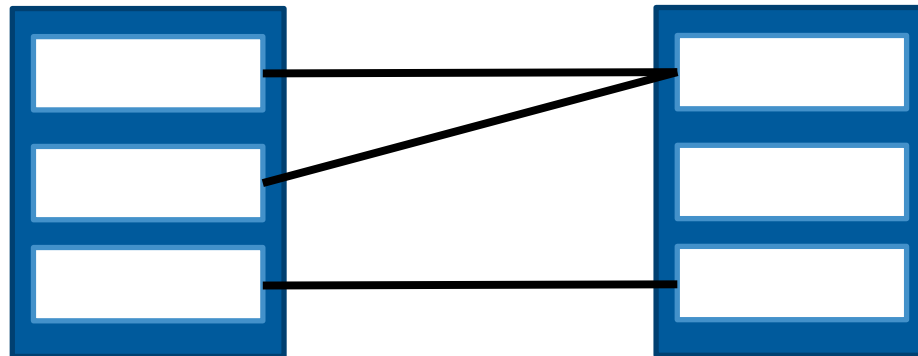


- ER diagram: 'Many' is represented by an forked line or a N between the relationship set and the entity set

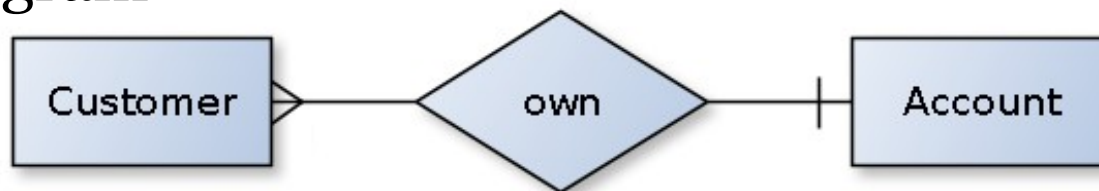


Step 2: Modelling Data : Cardinalities

- Many to one relationship
 - Many entities in an entity set is associated with one entity in another entity set
 - Example: Many customers can own a single account

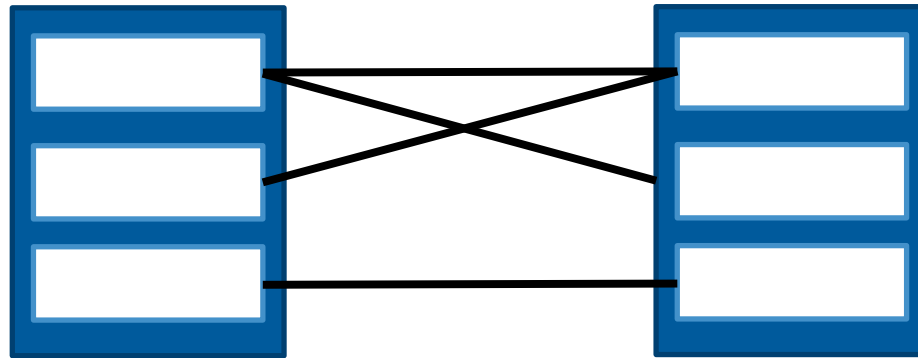


- ER Diagram

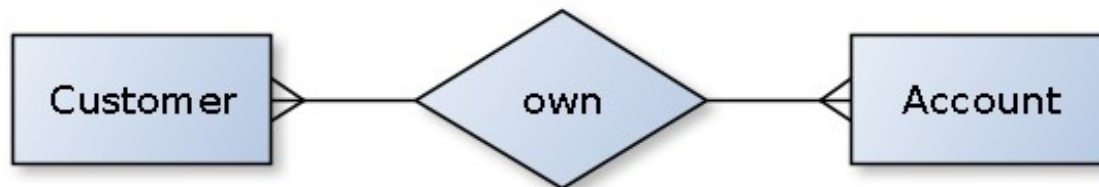


Step 2: Modelling Data : Cardinalities

- Many to many relationship
 - Many entities in an entity set is associated with many entities in another entity set
 - Example: Customers can have many accounts & accounts can be owned by many customers



- ER Diagram

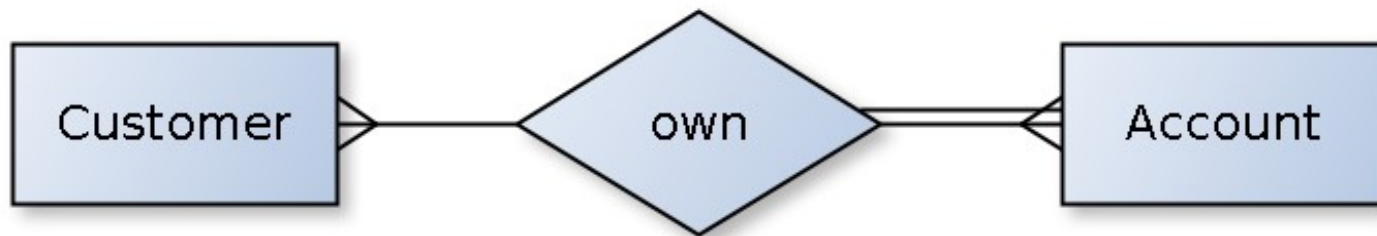


Primary Keys

- A primary key is a set of attributes that can be used to distinguish entities in an entity set
 - Each entity has unique values of key attributes
- Keys are used to link various entity sets (tables) in a database to each other in an efficient way
- Sometimes a suitable set of attributes will already be present in data model
 - e.g. Bank account {Branch number, Account number}
- Sometimes they will not ...
 - e.g. Books with *{title, year}*
 - In such cases, need to invent one or more artificial attributes which are designed to be unique
 - Examples are: passport number, driving licence number, Medicare number, bar code, ISBN, etc ...
- ER Diagram: The attribute(s) that are underlined will be the primary key

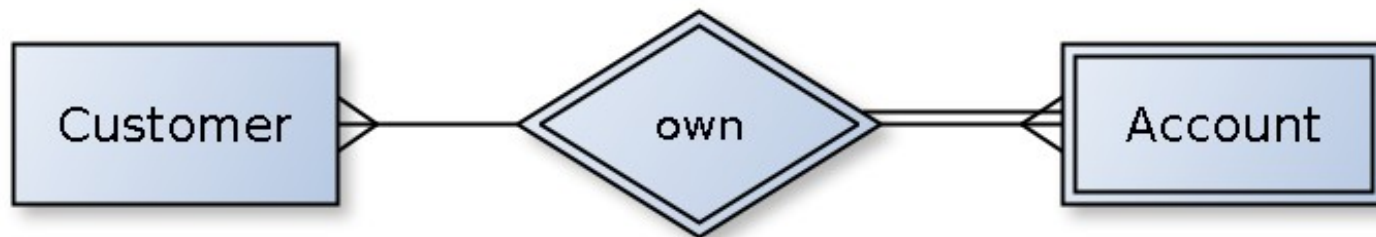
Participation

- Total participation: All entities in an entity set has to be involved in the relationship set.
 - Example: all accounts must be owned by customers
- Partial participation: Not all entities are involved in the relationship
- ER Diagram:
 - Total participation is represented by double lines.
 - Partial participation is represented by single lines.

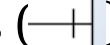



Weak Entity Sets

- Weak entity is existence-dependent on another entity
 - The entity can exist only when it has a relationship with another entity
 - Example: An Account can only exist when owned by a Customer
- Weak entity always has total participation of the relationship
- ER Diagram: Use double rectangle for weak entity with double diamond for relationship connecting to its associated strong entity



ER Diagrams - summary

- **Rectangles** represent entity sets
 - **Double rectangles** indicate weak entity sets
 - **Ellipses** represent attributes
 - Attributes may be simple or composite
 - **Double ellipses** indicate multi-valued attributes
 - **Dashed ellipses** indicate derived attributes
 - **Diamonds** represent relationship sets
 - **Double diamonds** indicate the the relationship of weak entity sets
 - **Lines** link attributes to entity sets and entity sets to relationship sets
 - **Single lines** indicate partial participation of the relationship
 - **Double lines** indicate total participation of the relationship
 - **Underline** indicates primary key attributes
 - Cardinality
 - **Stroked lines** () , or lines marked with 1, signify “one”
 - **Forked lines** () , or lines marked with N, signify “many”
-

ER Modelling Exercise: Phone Company

- A customer has a name (first and last), customer number (unique) and address.
- Each customer hires one or more phone lines.
- Each line has a unique phone number (the start date for the hire should be recorded).
- A phone line is of a certain line type. There are two line types: the primary line has a higher monthly hire than the additional line type.
- Local calls are free and are included in the hire fee.
- History of domestic and international calls should be stored in DB.
- For each call made from a certain phone line, the database should store its timing (the date and start time), end time and the length of the call.
- A call is classified as international call or domestic.
- The cost per minute for international & domestic calls should be stored in DB.
- The cost per minute for domestic calls depends on the area code.
- The cost for international calls depends on the country code.
- The total cost of a call should also be stored, and is the product of the length of the call and the price per minute for the corresponding type of call.