DBS211
Week 9 - Entity Relationship Diagrams ERDs

Agenda

- Welcome
- Conceptual Data Modelling
- 8 Steps of Data Modelling
- Entity Relationship Diagrams
- Symbology and Formatting
- Summary

What is an ERD?

- This conceptual data model represents the data used in an organization and the relationships between the data
- It is a graphical representation of the proposed database

Why ERDs?

- Documentation used to represent the database in an abstract way
- The data model can be reviewed by the end user and the person responsible for the physical database design
- Useful tool for the person creating the data model.

Entities and Events

- Entities - People, places, things or concepts about which information must be recorded
- Entities as Events – Placing an order or approving a loan
- Attributes for entities, events and relationships would be things like customer names or dates on which orders were placed
- Attribute – one single valued fact about an entity that we may want to record

E (Relationship)Ds

- Relationships are found between entities
- An employee is in a department
- A department has many employees
- Business rules must be taken into account
- Every employee must be in a single department

8 Steps of Data Modelling

Step 1 - Identify Entities
Identify the major objects of which we need to store. These will likely be **people, places, events** about which the client needs to store data. It is easy to lose focus on this first step and start thinking attributes (or characteristics) are entities. You must be clear in determining what is an entity vs an attribute.

For Example: For a used car lot, some entities might be: Vehicles, Staff, Locations, Services offered. Be careful not to chose something **like type of car as that is an attribute and not necessarily an entity.** It is almost always more beneficial to start with less entities than more.

8 Steps of Data Modelling (contd…)

Step 2 - Identify Attributes
Identify the characteristics, or properties, of the entities that need to be stored.

Examples: Vehicle Colour, Staff First Name, Staff Last Name, Vehicle Make, Vehicle Model, etc
In addition to the attribute names, determine if the attribute is mandatory or optional data. Mandatory data are included in bold format.

8 Steps of Data Modelling (contd…)

Step 3 - Specify Unique Identifiers (PKs)
Define attributes within each entity that could be used as a unique identifier for individual instances of the entity. This will ultimately become the primary key, and therefore must meet the criteria for a primary key. If a suitable one could not be identified, then a new surrogate key can be added to auto-generate a key field. Primary Keys are indicated in the diagram by being underlined. They will always be bold as well as the unique identifiers will always be required.

Example: Social Insurance Number (SIN), Vehicle Identification Number (VIN), productSerialNumber, etc.

8 Steps of Data Modelling (contd…)

Step 4 - Define Attributes Which Form Relationships
Identify entities that have relationships with each other. At this point it does not have to be tied to a specific attribute, although it ultimately will be.

Examples:
- a sales rep may be related to customers
- products may be related to product categories
- vehicles related to manufacturers
- vehicle models associated with manufacturers

8 Steps of Data Modelling (contd…)

Step 5 - Define Relationship Cardinality / Optionality

Determine the number of occurrences of one entity for a single occurrence of the related entity. This typically will only have three options - "none", "one" or "more than one". More than one is known as "many". Each relationship cardinality is defined by a range.
i.e. none or one, one or more, none or many, one and only one.

These are the possible results and are determined in BOTH directions of the relationship. (i.e. # entity 1 for each entity 2 and # entity 2 for each entity 1)

Be careful not to write these based on the existing data, but based on data that may occur in the future as well.

Lastly in this step, the Key attributes are identified as Foreign Keys that relate to their related entities Primary Key.

8 Steps of Data Modelling (contd…)

Step 6 - Eliminate Many to Many Relationships

If the cardinalities for any one relationship contains a many in both directions, then it must be eliminated. This is because it is physically impossible to create a many-to-many relationship in a standard relational database.

To eliminate the many-to-many relationship, erase the relationship completely and create a new entity between the two previous entities (called a junction or bridge entity). The recreate 2 new relationships between the new entity and one each to the two original entities. In most, if not all, cases, these new relationships will be opposite 1-to-many relationships.

8 Steps of Data Modelling (contd…)

Step 7 - Normalize the Database

The normalization process is the largest step, by far, and involves determining dependencies between attributes, and their unique identifiers. The next 3 weeks of this course covers this large topic and then you will understand that the data modelling process via ERDs overlaps with normalization here and you need to know both in order to perfect your database designs.

The main goals for normalization is to:

- eliminate redundant data
- eliminate repeating data
- remove irrelevant associated data
- ensure full and direct dependency between all attributes and their entities unique identifier.

8 Steps of Data Modelling (contd…)

Step 8 - Define Attribute Types and Sizes

This final step is simply to define the attribute data types and sizes. These values are not required for any of the previous steps and therefore it is recommended to do this last to avoid having to redo them several times. Some things to note:

- chose data types to be optimal for the data type while still allowing future potential data to be inserted.
- for referential integrity, FKs and PKs
- use numeric fields over string fields wherever possible

Entity Relationship Diagrams

ERDs are a graphical representation of the proposed database showing the entities, attributes or characteristics of the entities, and the relationships between the entities (people, places, concepts, and events).

There are 4 Goals for the ERDs:

- capture all the required data
- ensure no data repetition
- do not include any data that is derived from other data
- arrange the data in a logical manner

Symbology and Formatting

CROW's Foot style of UML ERD. The standards for a UML Crow's Foot ERD are:

•Entities are contained in boxes with the entity name in a top box and written in ALL CAPS

•Entity and Attribute names must be **singular** in nature. Singular means written as a single word, no spaces.

•Attribute names are written in one of, lower case, pascal case, or camel case.

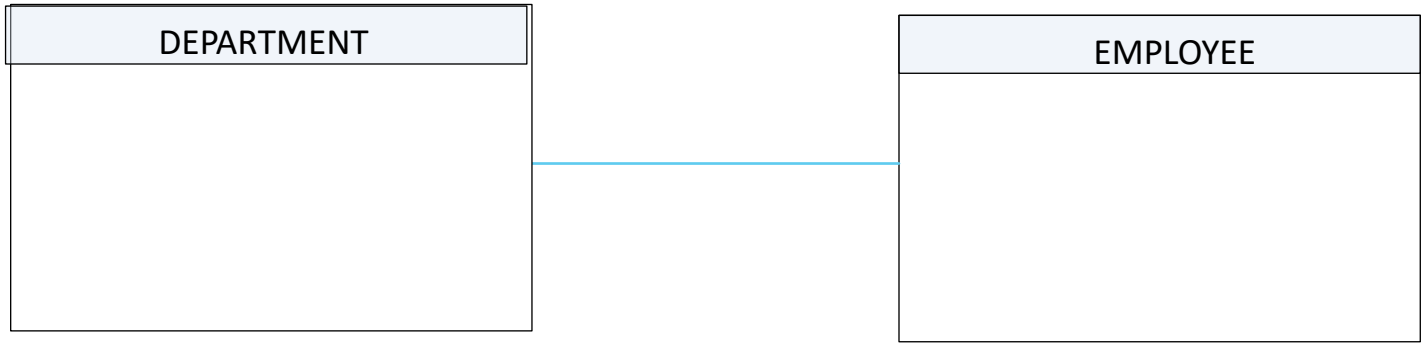
•Unique identifiers are underlined or included in a separate section of the box.

•required fields are included using **bold** format.

•foreign keys are identified using FK

Determining relationships

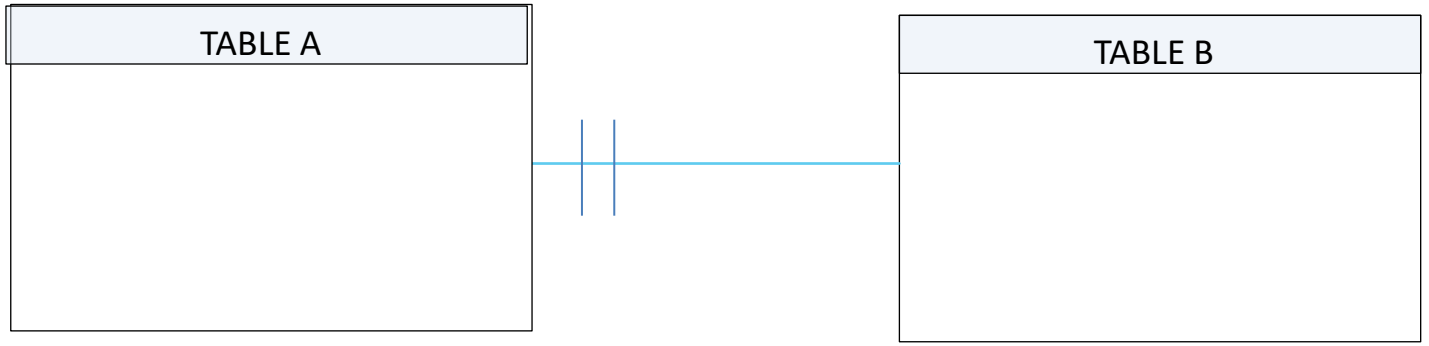
A relationship is like a verb that shows some dependency or natural association between two entities



A department contains employees
An employee is assigned to a department

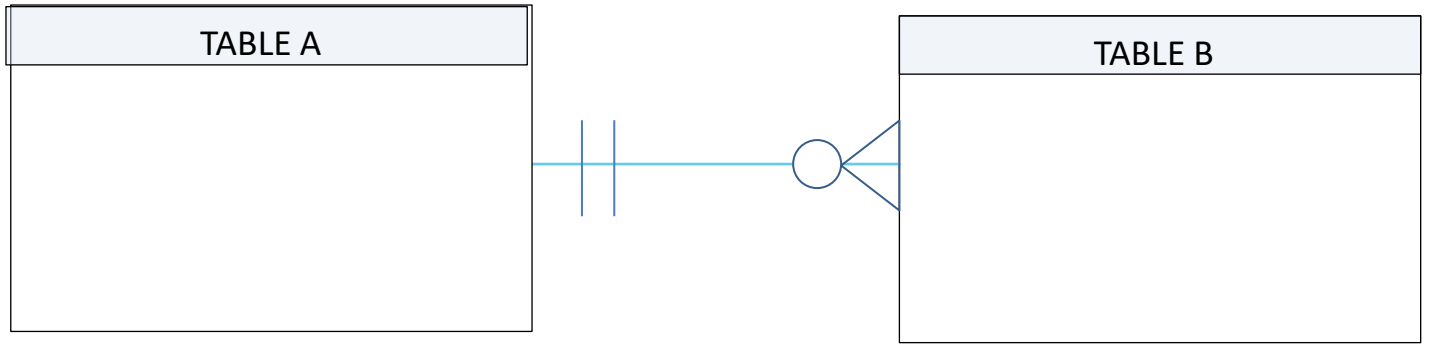
Determining optionality and cardinality

Cardinality. Determines the number of entities on one side of the relationship that can be joined to a single entity on the other side.
Optionality. Specifies if entities on one side must be joined to an entity on the other side.
Each instance of Table B is related to a maximum of one and a minimum of one instance of Table A

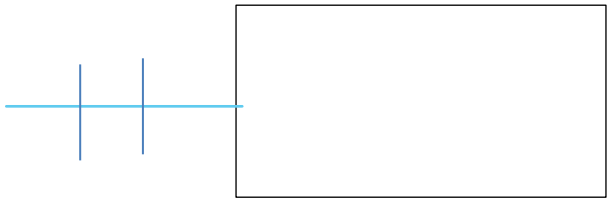


Optionality and Cardinality

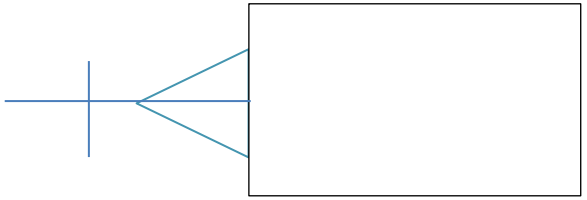
- Each instance of Table A is related to zero, one or more instances of Table B



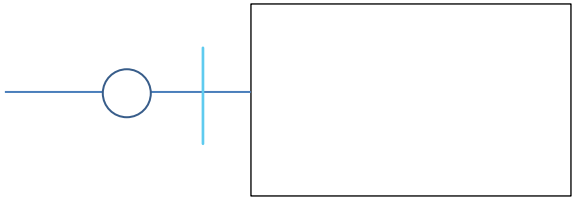
Optionality and Cardinality



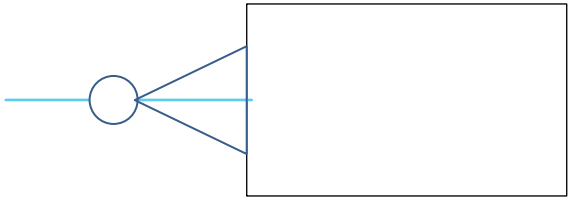
One and only one



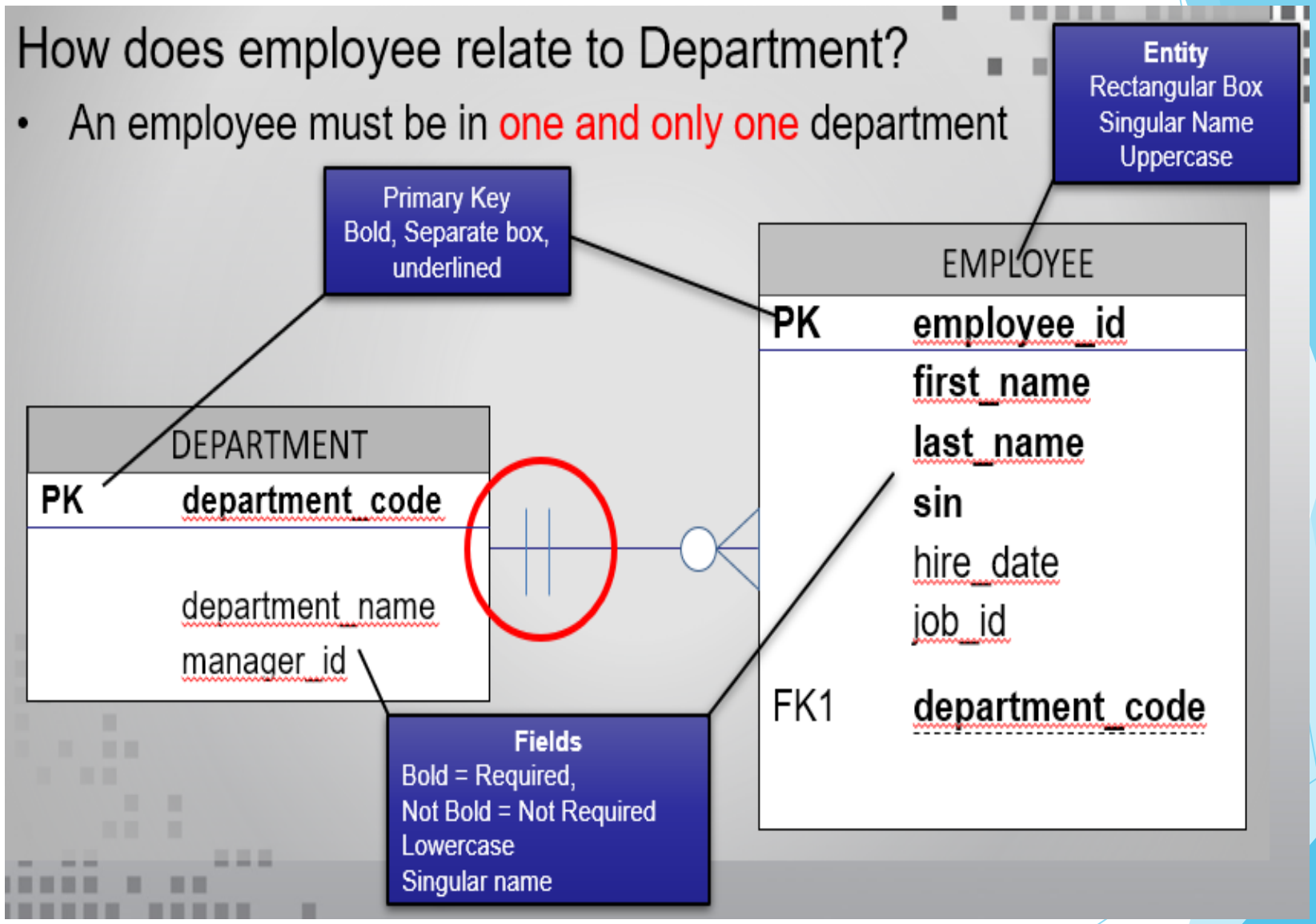
One or more



Zero or one



Zero or many


How does employee relate to Department?
An employee must be in one and only one department
Entity
Rectangular Box
Singular Name
Uppercase
Primary Key
Bold, Separate box,
underlined
EMPLOYEE
PK
employee_id
first_name
last_name
sin
hire_date
job_id
FK1
department_code
DEPARTMENT
PK
department_code
department_name
manager_id
Fields
Bold = Required,
Not Bold = Not Required
Lowercase
Singular name

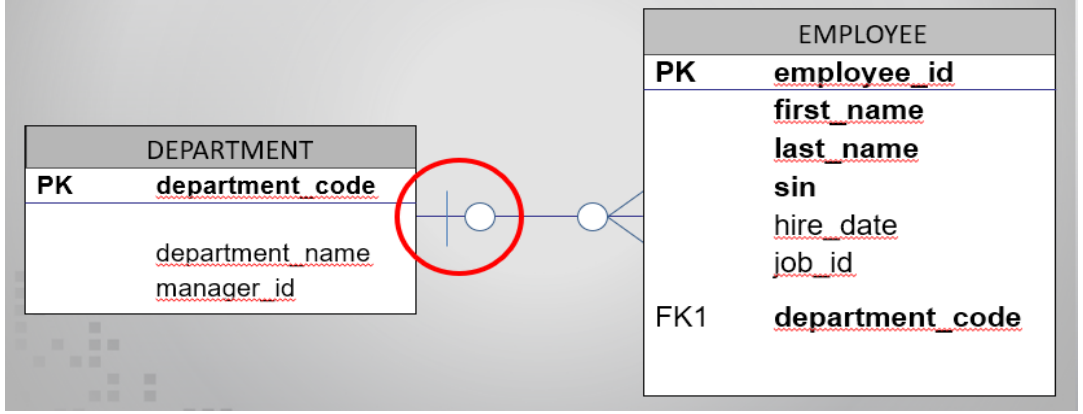
The cardinalities of the diagram are displayed using only 3 symbols.
•a circle - representing 0
•a line - representing 1
•a crow's foot - representing more than one. Sometimes called a fan.

EVERY relationship line will have TWO symbols on each side representing the range of values.
There are only 4 possible combinations that are used.
•Circle and Line - 0 to 1 entities per each related entity
•2 lines - 1 and only 1 entity per each related entity
•a Circle and a Foot - 0 to Many entities per each related entity
•a line and a foot - 1 or Many entities per each related entity

In the above example:
•each department can have 0 or many employees
•each employee must be in one department and only one department

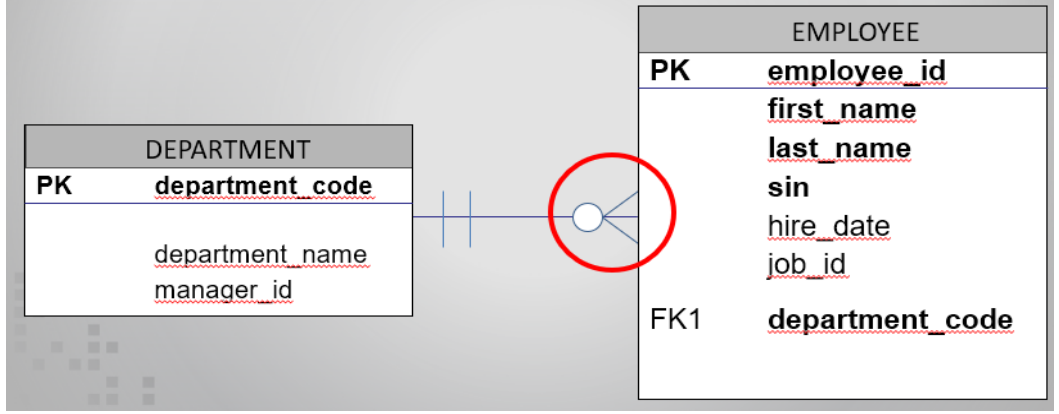
How does employee relate to Department?

- An employee must be in zero or one department



How does employee relate to Department?

- An department may have zero, one, or many employees



Lab8 :Using Diagrams.net (Draw.io) to Create a Crow's Foot ERD

The link to the referred to website is https://app.diagrams.net/
There are other online tools or you can use MicroSoft Visio to complete the ERDs.
This site will demo draw.io as it is fairly simple, free and produced clean ERDs.

Creating a New ERD

When you first visit the site, you are prompted to either create a new diagram or open a new one. Your ERDs can be saved to your local harddrive and reopened later for further editing.

Summary

- Entity or Event
- Relationship
- Optionality
- Cardinality
- Attributes
- UID
- ERD Diagrams