Data Analysis 2

Unit 2.2

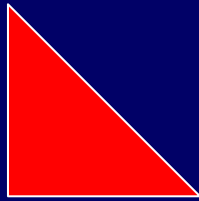# Entity Relationship Modelling - 2

Overview

- construct an ER model
- understand the problems associated with ER models
- understand the modelling concepts of Enhanced ER modelling

Country Bus Company

A Country Bus Company owns a number of busses. Each bus is allocated to a particular route, although some routes may have several busses. Each route passes through a number of towns. One or more drivers are allocated to each stage of a route, which corresponds to a journey through some or all of the towns on a route. Some of the towns have a garage where busses are kept and each of the busses are identified by the registration number and can carry different numbers of passengers, since the vehicles vary in size and can be single or double-decked. Each route is identified by a route number and information is available on the average number of passengers carried per day for each route. Drivers have an employee number, name, address, and sometimes a telephone number.

Entities

- Bus - Company owns busses and will hold information about them.

- Route - Buses travel on routes and will need described.

- Town - Buses pass through towns and need to know about them

- Driver - Company employs drivers, personnel will hold their data.

- Stage - Routes are made up of stages

- Garage - Garage houses buses, and need to know where they are.
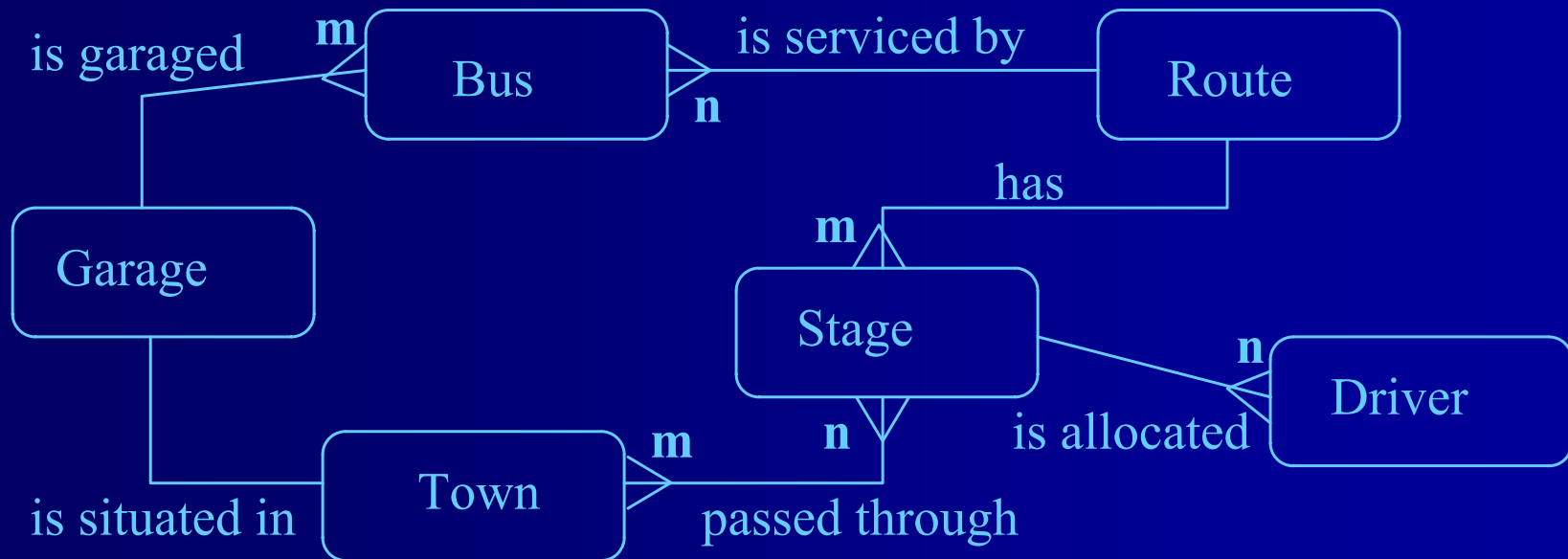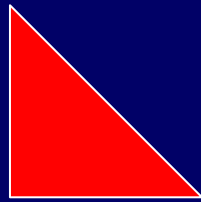
Relationships
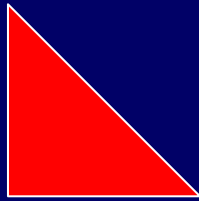
- A bus is allocated to a route and a route may have several buses.
    - Bus-route (m:1) is serviced by
- A route comprises of one or more stages.
    - route-stage (1:m) comprises
- One or more drivers are allocated to each stage.
    - driver-stage (m:1) is allocated
- A stage passes through some or all of the towns on a route.
    - stage-town (m:n) passes-through
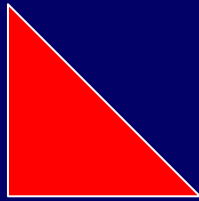
Relationships cont...
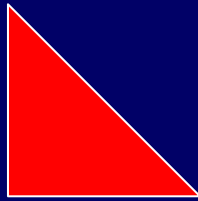
- A route passes through some or all of the towns
  - route-town (m:n) passes-through
- Some of the towns have a garage
  - garage-town (1:1) is situated
- A garage keeps buses and each bus has one 'home' garage
  - garage-bus (m:1) is garaged

Draw E-R Diagram

is garaged **m** Bus is serviced by Route

Garage

n

has

m

Stage

n is allocated Driver

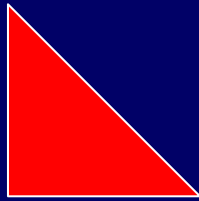is situated in Town **m** **n** passed through

Attributes

- Bus (reg-no,make,size,deck,no-pass)
- Route (route-no,avg-pass)
- Driver (emp-no,name,address,tel-no)
- Town (name)
- Stage (stage-no)
- Garage (name,address)

Problems with ER Models

There are several problems that may arise when designing a conceptual data model. These are known as connection traps.

There are two main types of connection traps:

1. fan traps
2. chasm traps

Fan traps

A fan trap occurs when a model represents a relationship between entity types, but the pathway between certain entity occurrences is ambiguous. It occurs when 1:m relationships fan out from a single entity.

Department >— is on ———— Site ———— employ —< Staff
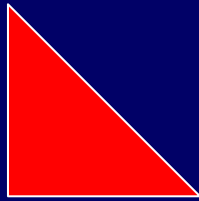n   m

Fan traps cont...

A single site contains many departments and employs many staff. However, which staff work in a particular department?

The fan trap is resolved by restructuring the original ER model to represent the correct association.
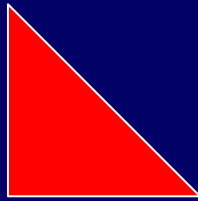
Chasm traps

A chasm trap occurs when a model suggests the existence of a relationship between entity types, but the pathway does not exist between certain entity occurrences.

It occurs where there is a relationship with partial participation, which forms part of the pathway between entities that are related.

Chasm traps cont...

- A single branch is allocated many staff who oversee the management of properties for rent. Not all staff oversee property and not all property is managed by a member of staff.

- What properties are available at a branch?

- The partial participation of Staff and Property in the oversees relation means that some properties cannot be associated with a branch office through a member of staff.
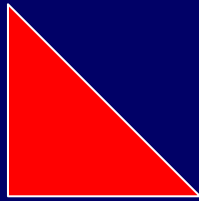
Chasm Traps cont...
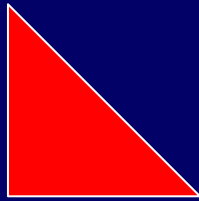
- We need to add the missing relationship which is called 'has' between the Branch and the Property entities.
- You need to therefore be careful when you remove relationships which you consider to be redundant.

is_allocated **n** Staff **0** oversees
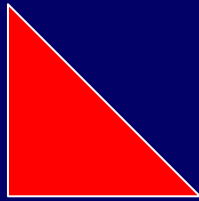
0

Branch has Property
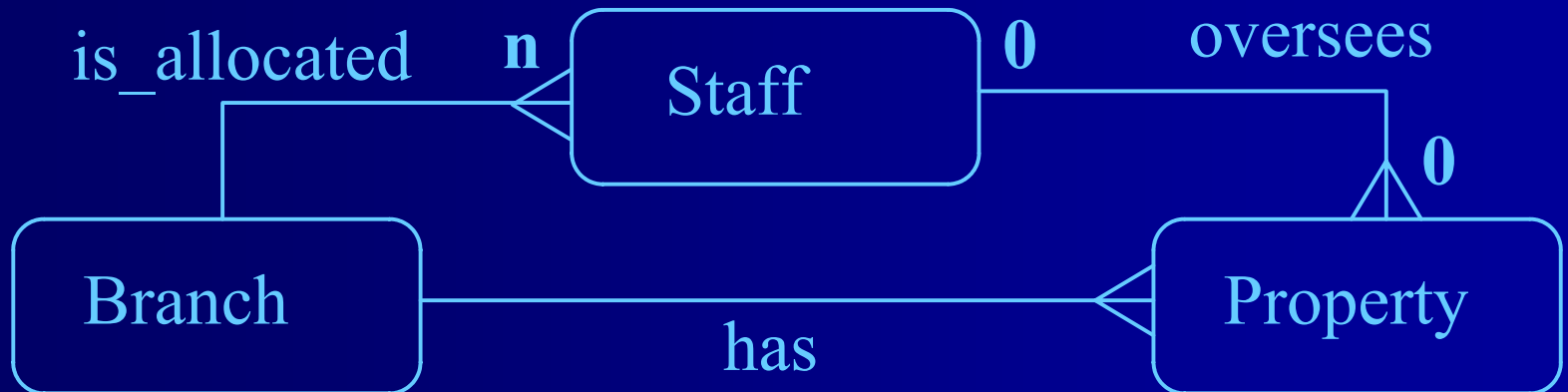
Enhanced ER Models (EER)

The basic concepts of ER modelling is not powerful enough for some complex applications... We require some additional semantic modelling concepts:
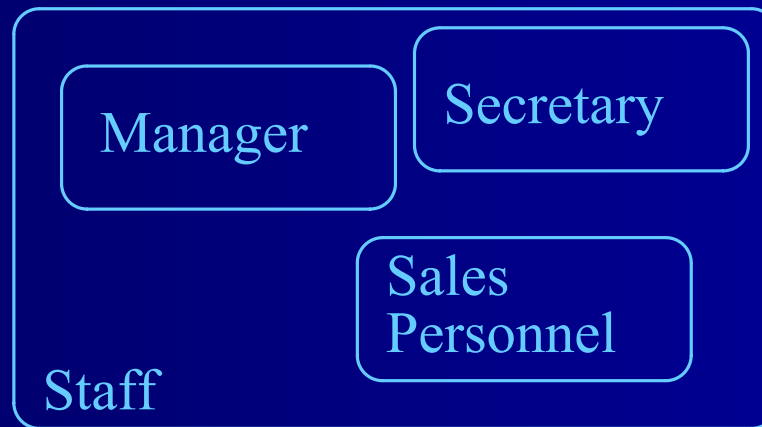
- Specialisation
- Generalisation
- Categorisation
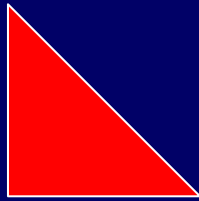- Aggregation

EER Models cont...

First we need some new entity constructs.

- Superclass - an entity type that includes distinct subclasses that require to be represented in a data model.
- Subclass - an entity type that has a distinct role and is also a member of a superclass.

Manager

Secretary

Sales
Personnel

Staff

EER Models cont...

Subclasses need not be mutually exclusive; a member of staff may be a manager and a sales person.

The purpose of introducing superclasses and subclasses is to avoid describing types of staff with possibly different attributes within a single entity. This could waste space and you might want to make some attributes mandatory for some types of staff but other staff would not need these attributes at all.
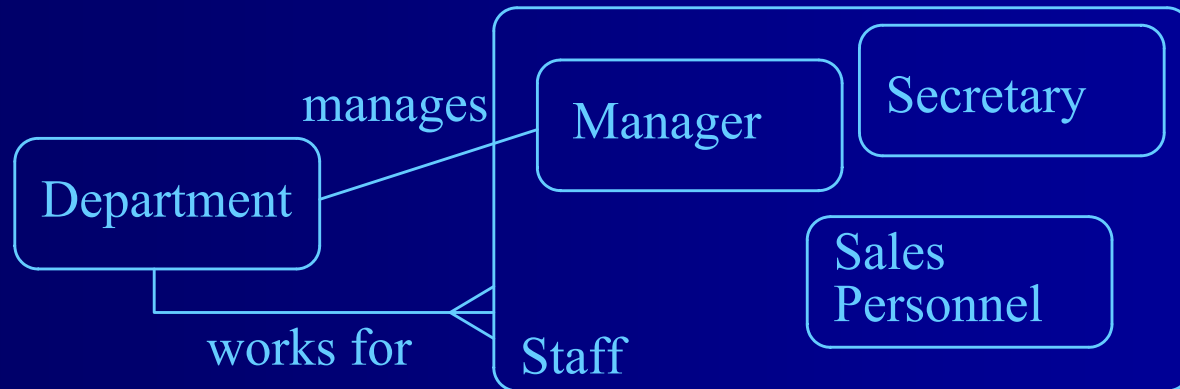
Specialisation

This is the process of maximising the differences between members of an entity by identifying their distinguishing characteristics.
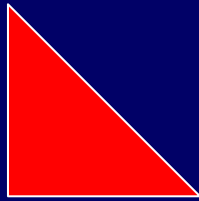
- Staff(<u>staff_no</u>,name,address,dob)
- Manager(bonus)
- Secretary(wp_skills)
- Sales_personnel(sales_area, car_allowance)

Specialisation cont...

- Here we have shown that the manages relationship is only applicable to the Manager subclass, whereas the works_for relationship is applicable to all staff.
- It is possible to have subclasses of subclasses.

Generalisation

Generalisation is the process of minimising the differences between entities byidentifying common features.

This is the identification of a generalised superclass from the original subclasses. This is the process of identifying the common attributes and relationships.