Lecture 13 – Data Modelling

Karl R. Wilcox
K.R.Wilcox@reading.ac.uk

Objectives

- To examine the process of designing a database
- To discuss the scalability of databases

- Today's practical
  - Database Forms and Reports. There is a lot of material to go through!

Entities and Relationships

- Entities are things we want to know about
- Attributes are information about the entity
- Entities are stored as records in a table
- Tables have a primary key that uniquely identify each record
- There are relationships between entities

Relationships in Detail

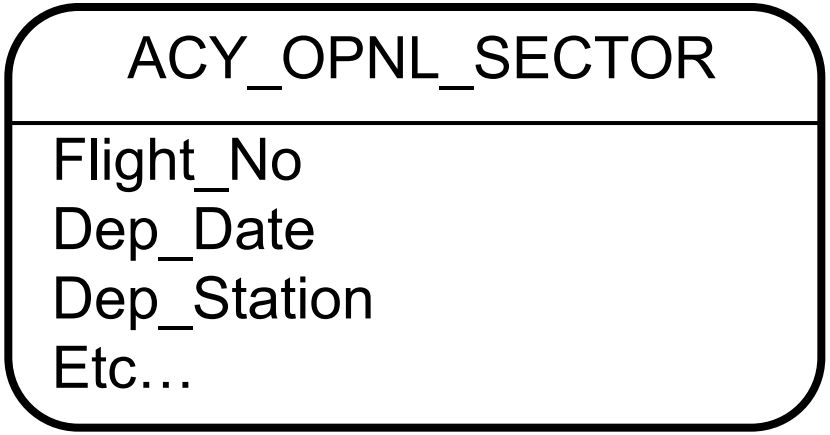
- More precisely, a relationship between entities exists where the value of an attribute *must* match one of the values in the primary key of another table
- This attribute is sometimes known as a Foreign Key
- The other table is sometimes know as a Lookup Table

Data Modelling

- The first stage in constructing a database is Data Modelling
- Step 1 – Identify entities
  - These must be uniquely identifiable
- Step 2 – Determine the attributes of the entity
  - Attributes may be either required or optional
- Step 3 – Identify the primary key attribute(s)
  - These ensure uniqueness
- Step 4 – Identify Relationships

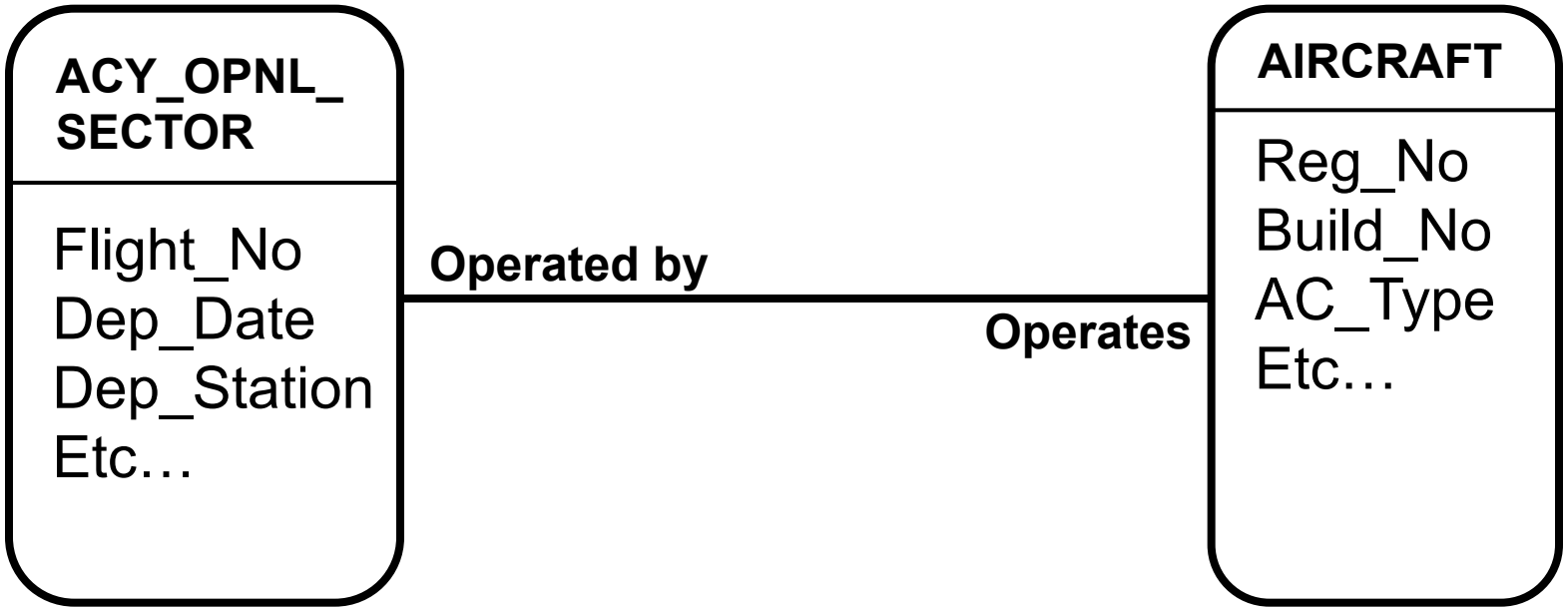
Entity Relationship Diagrams

- **Data Modelling is usually done using Entity Relationship Diagrams (ERDs)**
- **There is a widely accepted convention for these diagrams**
- **Entities are drawn in a box**
  - **Attributes may be listed inside the box…**



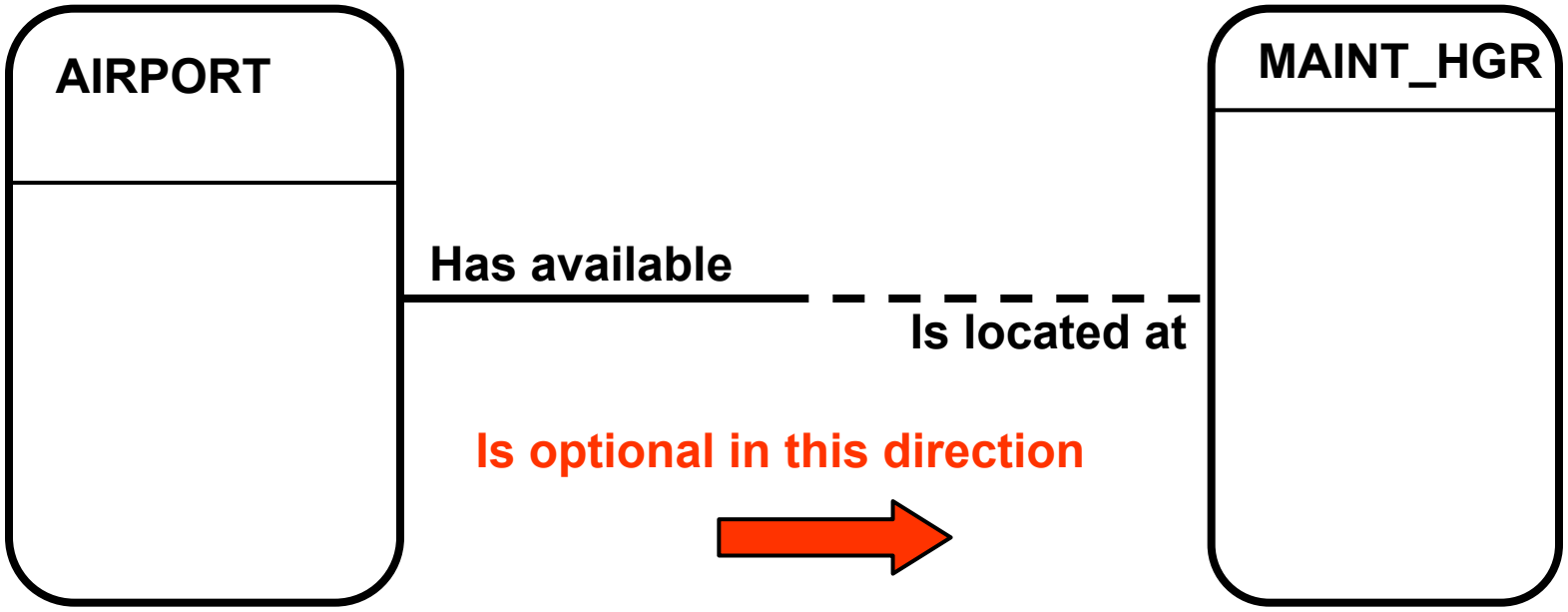
Relationship Modelling

- **Relationships are shown by lines between entities**
  - **Lines are labelled with the relationship**
  - **Normally at both ends…**



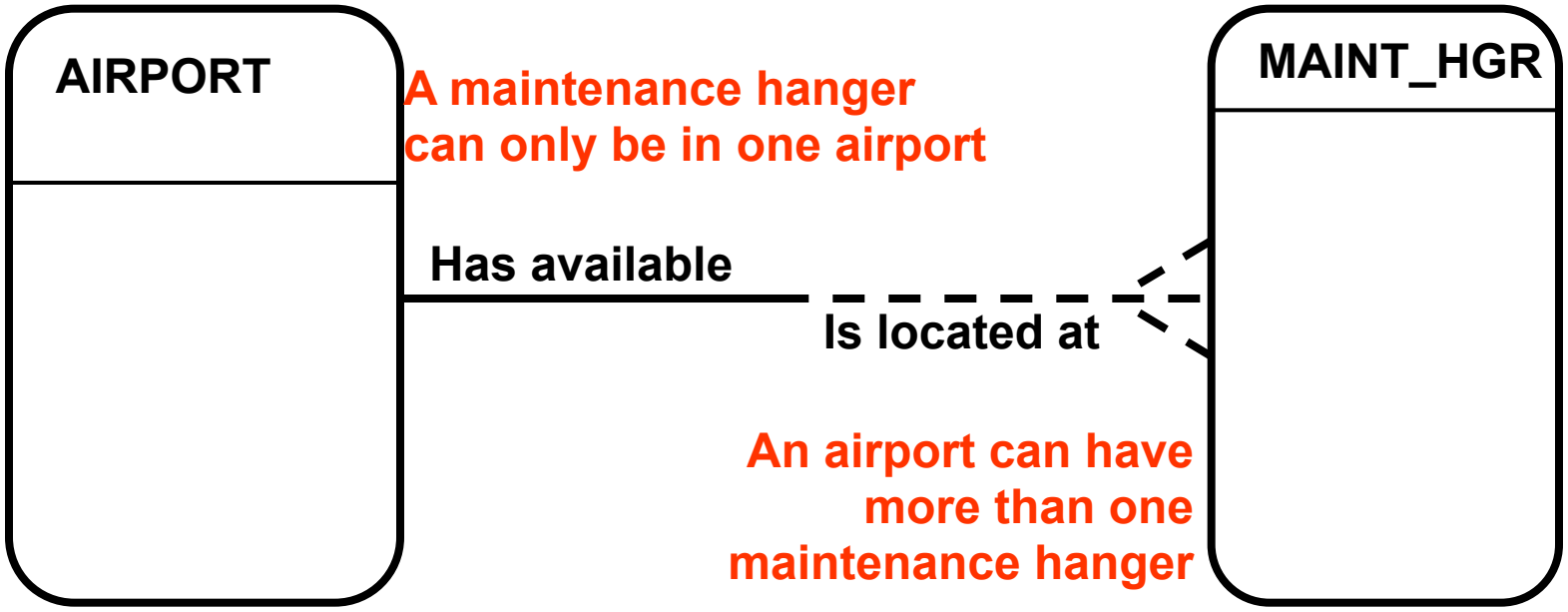
Relationship Optionality

- **Relationships are shown by lines between entities**
  - **An optional relationship is shown by a dotted line**



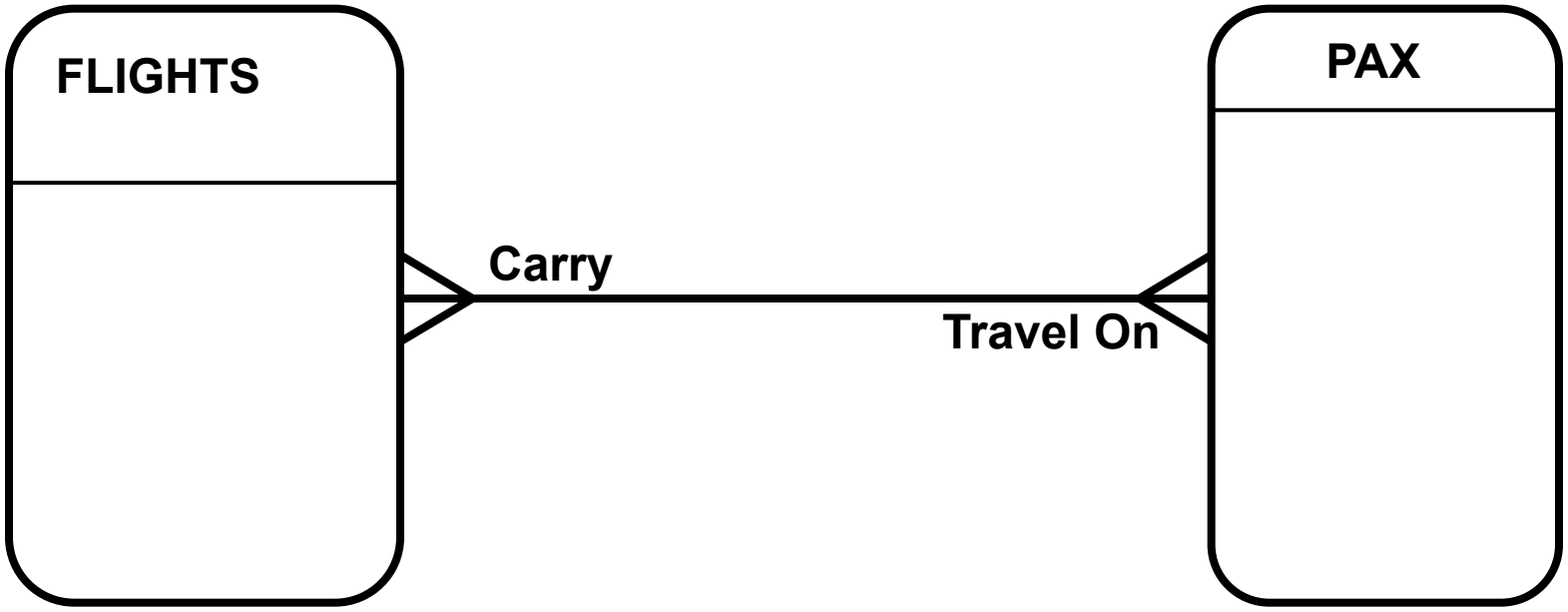
Relationship Ordinality

- **Relationships are shown by lines between entities**
  - **Relationships can be one to one or one to many**



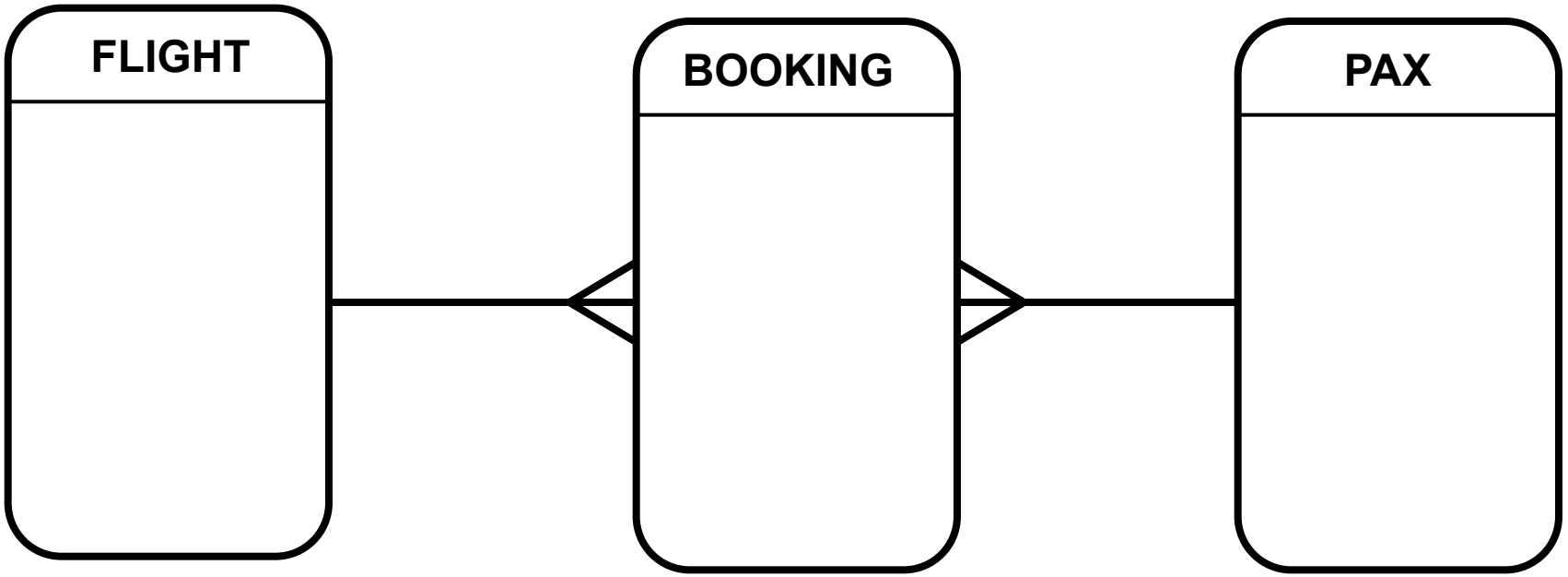
Many To Many Relationships

- **Many-to-Many Relationships cannot be handled directly by relational database systems**

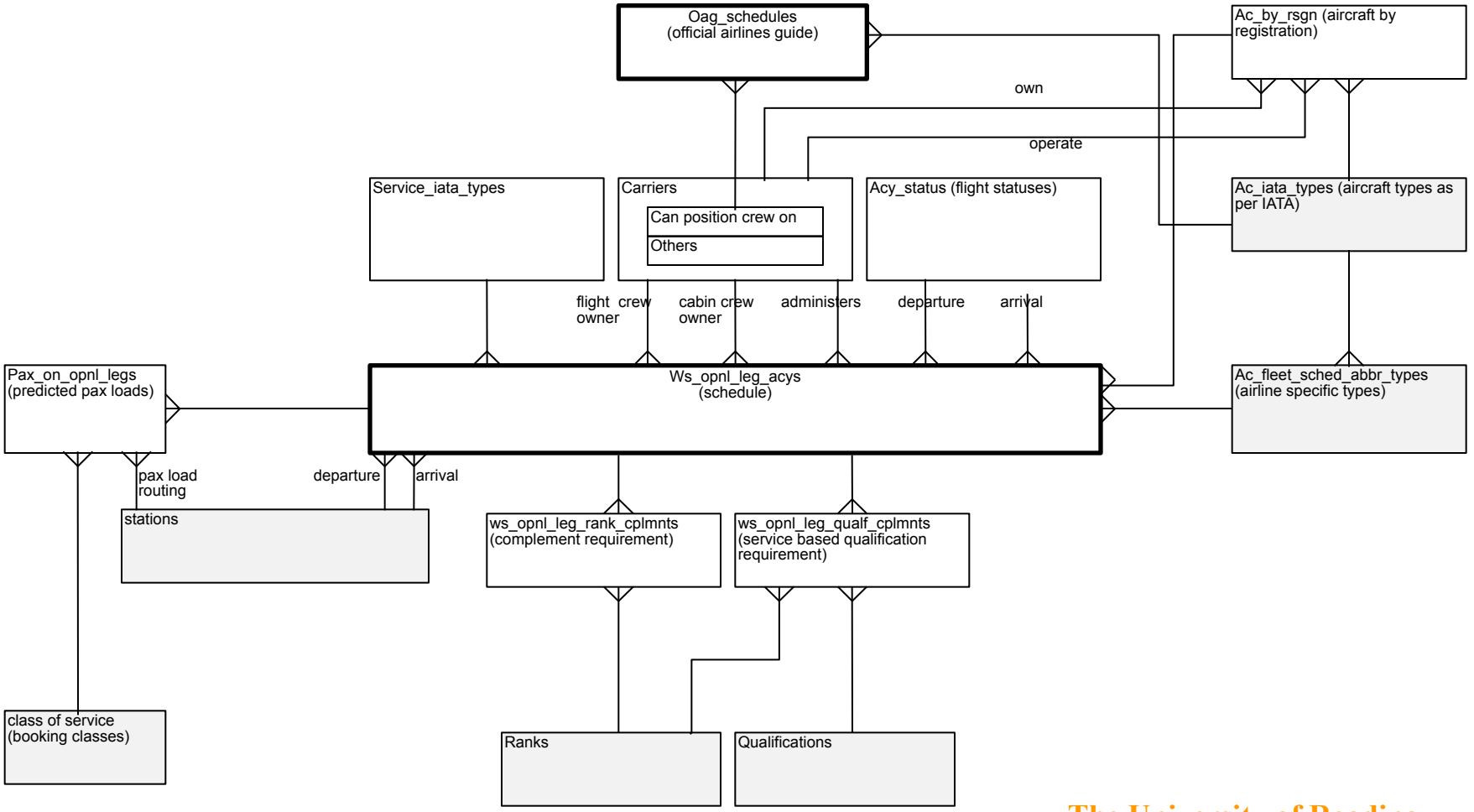


Intersection Entity

- An Intersection Entity must be created to link many-to-many entities



Example Data Model



Data Modelling Problems

- Data modelling appears straightforward
- There are often problems however
- Consider an entity of an “operational aircraft movement”
  - I.e. a Flight operated by a commercial aircraft as part of a scheduled service
- We wish to model all the activities around this aircraft movement, for example
  - Aircraft loading
  - Aircraft departure
  - Aircraft arrival

Our First Attributes

- We need an entity for an aircraft activity
- We believe we need the following attributes:
  - The Carrier Code (e.g. AA)
  - The Flight Number (e.g. 123)
  - The Date of departure (e.g. 23/5/02)
  - The Origin Airport (e.g. DFW)
  - The Destination Airport (e.g. PHX)

  - + others not shown

Potential Problems

- We wish to have a record for each activity, uniquely identifiable, however:
  - What happens if the aircraft has to return to origin airport with a technical problem, and then sets off again?
    - We need a new attribute for “departure number”
  - What if the aircraft is diverted to a different destination due to weather etc?
    - We need a new attribute for “diversion airport”
    - OR
    - We need a new attribute for “original destination”
  - What happens if the aircraft is diverted again?
    - ???

Business Rules

- Remember also that entity relationship diagrams only model entities and their relationships(!)
- Referential integrity is maintained by the database but this is not usually sufficient
- Individual attributes may need further validation
  - (can departure numbers be non-contiguous?)
- There may also be arbitrarily complex “business rules” that must also be enforced
  - Recall triggers and stored procedures

Example Business Rules

PLAN_DEP_DTM_UTC	Latest timing as notified through a SSIM Chapter 7 publication or an ASM/SSM change message.		N	Y	Must be >= FLT_DT_UTC
DEP_STS	This indicates the status of the departure timing, e.g. estimate, actual, etc. In the case where DEP_STS is null, the BEST_DEP_DTM_UTC represents a timing as notified through a SSIM Chapter 7 publication or an ASM/SSM change message.		Y	Y	Valid ACY_STATUS.ACY_STS_CD. Must be set if ARR_STS is set where ARR_DEP_BOTH_FLAG is either D or B. Valid combinations of arrival and departure status codes will not be enforced as MVT messages to get delayed or lost.

Data Modelling Tools

- Some database systems provide graphical tools for producing ERDs
  - e.g. Oracle CASE*Designer
- These can be used to generate the tables described by the diagrams
- They may also be able to “reverse engineer” existing database designs into ERDs

- Access allows some graphical linking of tables
  - This is not the same as a full blown ERD tool

Database Scalability

- Can a database developed as on Microsoft Access scale up to support an entire large enterprise?
  - Probably not
- Although you will learn a lot from the Access version
  - Should be regarded as a prototype
- Access is best for small workgroup requirements
  - Needs VB programming to implement business rules
- More next year when we cover Oracle & Access

Today's Practical

- Database Forms & Reports
- There is a lot of material to cover – you may NOT be familiar with all of it

- **REMEMBER TO SIGN OFF ON THE REGISTRATION SHEET!**