Inf1B Data and Analysis
Tutorial 1 (week 3)

Byrne
Revised by Simpson and Hutchins-Korte (2008)

15 January 2008

- Please answer at least Question 1–5 of this worksheet in advance of the tutorial, and bring with you all work, including printouts of code and test results. Tutorials cannot function properly unless you do the work in advance.
- If possible, please also attempt the **(Drop-in Lab)** and **(Tutorial Discussion)** questions.
- Data & Analysis tutorial exercises are not assessed, but are a compulsory and important part of the course. If you do not do the exercises then you are unlikely to pass the exam.
- Attendance at tutorials is obligatory; please let your tutor know if you cannot attend.
- *Recommended Reading:* Chapter 2 of 'Database Management Systems' (Raghu Ramakrishnan and Johannes Gehrke, 2003). This chapter can be found online Google Books (`http://books.google.com/`).

Introduction

In this tutorial you are required to design an Entity Relationship Model for a database, based on a description of the scenario. You should read the description carefully, making a note of all the candidate entities you think are involved, and of their attributes. Think about the relationships between the entities as you do this.

An Inter-University Gliding Competition

The scenario is that you are organising an inter-university gliding competition, and you have decided to design a database to keep track of the administration of the competition.[1]

A number of universities have each entered a team in the competition (known as a Task Week), and one of the things you need to keep track of is whether or not they have paid the entry fee. Each university team consists of a variable number of people who will take part in the competition; everybody who competes must be a member of one of the teams.

The pilots will have different levels of experience. Some will be pre-solo, which means they can only fly as second pilot ("P2") in a two-seater glider. They can still compete for their team in this capacity, as long as there is an instructor flying with them as pilot-in-charge ("P1"). Pilots who are of cross-country standard can fly as P2 just like pre-solo people, but may also fly solo in any kind of aircraft. A pilot flying solo is always P1 of course. Instructors can fly solo in single- or two-seater gliders, or as P1 in a two-seater with a less experienced pilot. If two instructors are flying together they will simply decide between them who is P1 and who is P2.

There are a number of different types of glider involved in the competition. Some are two-seaters, such as K7, K13, K21 and DG505. The rest are single-seaters; their types include K8, Pirat, DG300, Discus and LS4. There may be more than one glider of a particular type, but every glider can be distinguished by its callsign — a short string which is used to identify it in radio communications. Typical callsigns include "MF", "P19", "FNS" and "CPG".

The competition is organised around "tasks", which are routes that each competing glider must attempt to fly around. On each competition day a task is set for the pilots to fly in their gliders. The task is defined by choosing a set of "turning points" taken from a list available from the BGA (British Gliding Association). There are almost a thousand such turning points defined for the UK, and each has a unique "trigraph" or three-letter acronym to identify it. For example, "STI" is for Stirling and "LOM" is the Lake of Menteith. The competition is being held at Portmoak Airfield (about 30 miles north of Edinburgh), which is "POR". The task-setter will decide on a suitable task for each competition day, which will involve trying to glide from the starting point at the airfield around one, two or more turning points. For example, a set like "POR, STI, MVN" would define a triangle of just over 100km, with the corners at Portmoak, Stirling and Methven (which is near Perth). For the purposes of this exercise we will assume that competitors are allowed to fly around the turning points in any order they choose. As well as specifying the trigraph, the BGA list of turning points gives the latitude and longitude of each turning point, so their positions can be precisely identified on a map.

Sometimes competitors can gain an unfair advantage by starting off much higher than other gliders, or by happening to pick a better time of day. The task-setter can therefore attach conditions to each task, specifying the maximum starting height allowed and the earliest time at which a glider can start.[2]

[1] Gliders are aeroplanes without engines, designed for maximal efficiency in extracting the energy freely available in the air, and (on a good day) using it to fly hundreds of kilometres.

[2] Incidentally, if you would like to try this kind of flying for real, why not have a look at `http://www.eusu.ed.ac.uk/clubs/gliding/`. Make sure to mention this Inf1B tutorial if you decide to become a member of the Edinburgh University Gliding Soctiety!

Question 1 - Determining of possible entity-sets

What are the candidate entity sets you would consider in order to build your ER model? Is it possible to map these entity sets directly and unambiguously from the information given, or is there more than one possibility for modelling these entity sets? Choose your final set of entity sets from the candidates, making sure to include both 'Person' and 'Glider', and justify your answer in as much detail as possible.

Question 2 - Defining attributes for a given entity-set

At this point you will have decided your sets of entities that can be used to design the ER diagram. For this question, focus on just the two entity sets below:

- Person
- Glider

What are the attributes you can assign to each of these given the available information? What are the attribute types?

The scenario told you that a pilot can fly a glider in crew capacity either P1 or P2. Is this an attribute? If not, what is it? If it is an attribute, does it belong to Person or to Glider or to neither of them? If it belongs to neither, where does it feature in the ER design? Explain your reasoning fully.

Note: In order to draw your diagram you will need to decide the attributes for *all* your ER objects. You do not need to list them all here, but they must appear on your final diagram.

Question 3 - Describing relationships

Describe the relationships between the entities you defined in the first question and give them names. What kinds of relationships are there? *one-to-one*? *one-to-many*? *many-to-many*? Are there any weak entities in your model? Are there any key constraints? Do any of the relationships involve full participation?

Question 4 - Defining primary keys

What are the primary keys for each entity in your model? Explain briefly why you chose them.

Question 5 - Drawing the diagram

Draw the diagram representing your ER model. Use the conventions given in the lecture slides.

(Drop-In Lab) - Using `dia`

Draw the ER diagram you created in Question 5 using the `dia` application. Brief instructions on how to start up and use `dia` can be found on the Inf1B website. You may choose to work on this task during one of the drop-in labs, where you will be able to ask a demonstrator for help. Drop-in labs are daily in the afternoons. The timetable is linked off the Inf1 webpage.

(Tutorial Discussion) - Thinking around the model

Do not alter your diagram for this question, but describe in outline changes that you might make (such as adding certain attributes to particular entities or relationships).

Imagine you really are running a gliding competition. Is there anything you think is missing from the scenario — any other information that you would like to collect and add to your model? Explain what you would add and why. Could your model be easily altered to include new information or would you need a major redesign?