Informatics 1B
Data and Analysis: Assessed Coursework ASN-02
Structured Data

Kate Byrne

26 January 2007

Your solution to this assignment will consist of *two* separate files. One of them is the answer sheet you should fill in with your answers (including your name and matriculation number), that was posted beside this handout on the website. (It is at

`http://www.inf.ed.ac.uk/teaching/courses/inf1/da/assignments/a02/ASN-02_ans.txt`.)

The second file is the ER diagram you will design for this assignment. We encourage the use of the `dia` tool that was introduced in the lab session. Create your ER model in a `dia` document called `ASN-02.dia`. If you decide to use another drawing tool, this is fine, however only documents called ASN-02 of type `PDF` or `EPS` will be accepted in this case.

Once you have your two files ready for submission, create a new directory called `ASN-02` and put your two files in this directory (ask a lab demonstrator if you are not sure about how to do this). When your two documents are in the directory, submit your answers by running:

```
submit inf1 inf1b da2 ASN-02
```

Note that here `ASN-02` is the name of your directory, by doing this you are making sure that your *two* files (answers and diagram) are submitted together.

Alternatively, you may submit in hardcopy to the Informatics Teaching Office, 5th floor, Appleton Tower. Print your two files and make sure that your name and matriculation number is clearly stated on *both* of them.

Keep a backup copy of work that you submit. The deadline for this coursework is **Friday 9th February, 12:00 noon**. Please make sure you submit either in hardcopy to the ITO or electronically before this deadline, as late work will not be marked unless you have been granted an extension by the course organiser.

Introduction

In this assignment 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.

The first part of the assignment consists of answering a series of questions about your design and producing a final ER diagram for it. The `dia` application (as used in the Lab exercise) is the recommended tool to use. In the second part of the assignment you are asked to map your ER model to a relational schema, using the techniques described in the lectures.

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

[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.

(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.

From this information you now need to come up with an ER design.[2]

Question 1

Use the `dia` application introduced in the lab session to design your ER diagram. (You can use other drawing tools if you prefer, as long as you use the correct representation for ER items and submit in the required format, as described in the rubric at the start of this handout.)

Question 1.1 - Determining of possible entity-sets [*10% of marks*]

What are the candidate entities you would consider in order to build your ER model? Is it possible to map these entities directly and unambiguously from that information, or is there more than one possibility for modelling these entities? Explain how you chose your final set from the candidates. Justify your answers in as much detail as possible.

Question 1.2 - Defining attributes for a given entity-set [*10% of marks*]

At this point you will have decided your set of entities that can be used to design the ER diagram. Whatever your answer was for the previous question, focus now on just these two entities (which may or may not be amongst your set):

- Person
- Glider

[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/`. Alas, it is not possible to award extra marks for membership of EUGC.

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 1.3 - Describing relationships [*10% of marks*]

Describe the relationships between the entities you defined in the first question. What kinds of relationships are there? *one-to-one*? *one-to-many*? *many-to-many*? Name each relationship as we did in the first tutorial and explain how each relationship affects the model. Are there any weak entities in your model? Are there any key constraints? Do any of the relationships involve full participation?

Question 1.4 - Defining primary keys [*5% of marks*]

What are the primary keys for each entity in your model? Explain why you chose them and whether there are alternatives.

Question 1.5 - Drawing the diagram [*15% of marks*]

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

Question 1.6 - Thinking around the model [*10% of marks*]

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?

Question 2 - Mapping ER to relational schemata [*40% of marks*]

With the ER diagram you built in question 1.5 you will now produce the relational schema for your database.

How is each entity and relationship mapped? State the SQL `create table` statements to define *all* the relations required to create this database. Note that in these SQL statements you will also need to define any *key* and *foreign key* constraints.

Some Comments on Plagiarism

1. Submitting another student's code (even if modified) as if it is your original work is plagiarism.
2. Assisting another student to plagiarise (eg. by sharing code without the borrowing student acknowledging the use) is also penalised.
3. Discussing the assignment in broad terms is OK, but not at the level of detailed answers or code.
4. Even partial, non-working submissions get partial credit.
5. A failed assignment is not a ruined career. Getting caught at plagiarism could ruin it.
6. We use various techniques to detect plagiarism, including automated tools.
7. If you have difficulty doing the assignment, use the Inf1B newsgroup or ask for help from your demonstrators or tutor. If you are still having problems, contact your TA, Kate Byrne (k.byrne@ed.ac.uk). If some other external reason prevents you completing the assignment on time, please contact your Director of Studies and/or the Course Organiser.
8. Change the protections on your homework files and directories so potential plagiarists cannot access your solution. If your file is called XXX, then use the UNIX command: `chmod go-rwx XXX`.