Chapter 1

Introduction

Every information system has a *database* at its core. Therefore, it is essential for you to acquire an understanding of the area of *database systems*.

During the past 30 years, since its inception, the are of *Database Systems* has matured into a well established, conventional subject. However, new ideas are continuously developed and issues still exist, that have to be theoretically addressed. This subject can be perceived as consisting of a "classic" and a "young" component. The former is based almost entirely on the relational model and represents the, de facto, standard approach of most of the industrial applications. The latter proposes new models and architectures for database systems.

Aims

This course aims to provide you with an understanding of the main issues related to data storage and manipulation – the object of *Database Systems*. In particular, this course is aimed at the detailed presentation of the theory and practice of the relational model, on one side, and at the introduction of the emerging trends in database systems, on the other. You will also gain practical experience in a database language – SQL 92 – and in developing a relational database system, through the activities and coursework you will undertake.

Objectives

The objectives of this course are:

- to motivate the need for database systems;

- to provide a general description of a *database environment*;

- to thoroughly describe the *relational model* as the underlying framework of most of the industrial database management systems;

- to introduce a *relational database language*: SQL 92;

- to describe the issues pertaining to *database design*, focussing on *conceptual design* (through *E/R modelling*) and *logical design* (through *normalisation*);

- to develop *practical skills* in designing and implementing a relational database;

- to present issues pertaining to *data protection*: security, recovery and concurrency control;

- to introduce the emerging *distributed* architecture for database systems;

- to describe *new approaches* to database systems: deductive and object oriented.

Outline

Part *one* of the course is dedicated to the introduction of the basic concepts of database systems. If focuses on describing the *components* of a database environment, a prevalent *architecture* of a database environment and the concept of *data model*.

Part *two* presents the theory behind relational database systems – *the relational model*. It describes the relational data objects – *domains* and *relations* – relational operators – *relational algebra* – and issues about relational data integrity – *keys*. The theoretical

description of each of the components of the relational model are accompanied by a description of the operational issues – the way relational concepts are implemented in a database management system (DBMS).

Part *three* is dedicated to the introduction of a relational database language, namely *SQL-92*. This part is accompanied by activities and coursework aimed at developing *practical skills* in programming in SQL-92.

Part *four* describes the process of *database design*. It focuses on the *conceptual design* phase, by presenting the most popular conceptual model – the *entity-relationship (E/R) model* – and on the *logical design* phase, by presenting the *normal forms* associated with the relational model.

Part *five* is dedicated to the pragmatic issue of *data protection* in a DBMS. Even though this is more general issue, its presentation is made in the context of relational DBMSs. This part introduces the concepts of *transaction* and *transaction processing*. Then, it describes the way transaction processing can be employed in database *recovery* and *concurrency control*. As another facet of data protection, the issue of *security* in a database environment is also presented. Each of these subtopics is accompanied by a presentation of the way it is supported in SQL-92.

Part *six* describes the emerging *distributed architecture* for database systems. It focuses on presenting the *objectives* of distributed database systems development and the *problems* generated by this new architecture.

Part *seven* is dedicated to *new approaches* to database systems development (or *advanced database systems*). By outlining some *drawbacks of the relational model*, (drawn from its unsuitability to certain application areas of increasing importance, the need for new approaches is motivated. Then, this part focuses on the two main new approaches to database systems, namely *deductive databases* and *object oriented (OO) databases*.

Subject guide structure

This subject guide has two volumes and dedicates a chapter for each of the seven parts of the course.

The first four parts of the course are included in V*olume 1*. Accordingly, Volume 1 represents the description of the *fundamental issues of (relational) database systems*.

The last three parts of the course are included in V*olume 2*. Accordingly, Volume 2 represents the description some *advanced topics in database systems*.

Each chapter includes a set of *activities* that you are required to perform. They will help you in better understanding the theoretical issues and in achieving practical skills.

This course requires you to perform a set of assignments as *coursework*. They are not included in this guide. You will receive them separately.

Volume 2 contains three appendices. In *Appendix A* you are provided with a specimen examination paper. *Appendix B* contains detailed guidance to answering the examination paper; it also contains a marking scheme. *Appendix C* provides a further list of examination questions.

Reading

The following two books are the recommended text books for this course:

"An Introduction to Database Systems", sixth edition, by C. J. Date, published by Addison-Wesley, 1995.
[ISBN 0-201-82458-2]

and

> *"Database Systems: A Practical Approach to Design, Implementation and Management"*, second edition, by T. Connolly and C. E. Begg, 1999. [ISBN 0-201-34287-1]

You can choose to base your learning on only one of them, because either covers *almost entirely* the topics of this course. However, if possible, you should use both of them. In conjunction they provide a *perfect coverage* of all the topics of this course.

Each chapter starts with two sections – *Essential reading* and *Further reading*.

The *Essential reading* section directs you to the book sections that you have to read. For each chapter you are given an *alternative* between (Date 1995) and (Connolly 1999). In the light of the previous paragraph, you can choose to read only one of them. However, you are recommended to read both of the suggested alternatives.

The *Further reading* section provides you with pointers towards relevant literature for the presented topic. This is not compulsory reading. A short descriptive note is associated with each pointer; e.g., *for a more concise and basic reading, for a more comprehensive description.*

Further support

The same course – Database Systems – is also being taught internally, in the Department of Mathematical and Computing Sciences, Goldsmiths College, University of London. A site was developed in support of this teaching on the department's server. The address of this site may change. However, you will always find a pointer to it from either:

http://www.maths.gold.ac.uk

or

http://www.gold.ac.uk/~mas01mu

You may find some of the provided information helpful.

About this volume

This volume is *Volume 1* of the subject guide. It includes the first four parts of the *Database Systems* course, namely:

- basic concepts;
- the relational model;
- SQL-92;
- database design.

Notes

4