

Fundamentals of Databases notes

Todd Davies

December 22, 2014

Introduction

Databases are core, if largely invisible, components of modern computing architectures in both commercial and scientific contexts. The management of data has evolved from application-specific management of myriad files to organisation-wide approaches that see data as one of the most important assets of modern organisations and, as such, a key factor in their ability to compete and thrive. At this organisation-wide scale, database management systems (DBMSs) are the crucial piece of software infrastructure needed to achieve the desired results with consistent quality and robust efficiency. A modern DBMS is a thing of wonder and embodies in its internal construction and in its wide usability many advances in algorithms and data structures, programming language theory, conceptual modelling, concurrency theory, and distributed computing. This makes the study of databases a data-centric traversal of many of the most exciting topics in modern computing.

Aims

The aim of this course unit is to introduce the students to the fundamental concepts and techniques that underlie modern database management systems (DBMSs).

The course unit studies the motivation for managing data as an asset and introduces the basic architectural principles underlying modern DBMSs. Different architectures are considered and the application environments they give rise to.

The course unit then devotes time to describing and motivating the relational model of data, the relational database languages, and SQL, including views, triggers, embedded SQL and procedural approaches (e.g., PL/SQL).

The students learn how to derive a conceptual data model (using the Extended Entity Relationship paradigm), how to map such a model to target implementation model (for which the relational model is used), how to assess the quality of the latter using normalisation, and how to write SQL queries against the improved implementation model to validate the resulting design against the data requirements originally posed. For practical work, the Oracle DBMS is used.

The course unit also introduces the fundamentals of transaction management including concurrency (e.g., locking, 2-phase locking, serialisability) and recovery (rollback and commit, 2-phase commit) and of file organisation (e.g., clustering) and the use of indexes for performance.

Finally, the course unit addresses the topic of database security by a study of threats and countermeasures available.

In the case of the former, these include potential theft and fraud as well as loss of confidentiality, privacy, integrity and availability. In the case of the latter these primarily include mechanisms for authorization and access control, including the use of views for that purpose. The course unit also addresses the topic of legal frameworks that give rise to obligations on the part of database professionals by introducing exemplars such as the 1995 EU Directive on Data Protection and the 1998 UK Data Protection Act.

Additional reading

Contents