CS346-15 Advanced Databases

21/22

Department
Computer Science
Level
Undergraduate Level 3
Module leader
Hakan Ferhatosmanoglu
Credit value
15
Module duration
10 weeks
Assessment
Multiple
Study location
University of Warwick main campus, Coventry

Description

Introductory description

The aim of the module is to introduce students to current techniques, methods and results from the active field of database systems and data management. Typical topics include query planning and optimisation; transaction processing and concurrency control; big management; data warehousing and OLAP; theory of databases.

Module aims

The aim of the module is to introduce students to current techniques, methods and results from the active field of database systems and data management. Typical topics include query planning and optimisation; transaction processing and concurrency control; big management; data warehousing and OLAP; theory of databases.

Outline syllabus

This is an indicative module outline only to give an indication of the sort of topics that may be covered. Actual sessions held may differ.

The topics will be drawn from core conceptual topics in advanced databases, and current ideas in database systems. This will be drawn from:

- Refresher on databases and modelling
- Relational algebra, tuple relational calculus, SQL, and equivalences between them
- Query planning, evaluation and optimisation
- Transaction processing, concurrency, ACID rules, OLTP
- Online analytical processing (OLAP), data warehouses
- Data storage and indexing, B-trees and hashing
- NoSQL to relax ACID rules; consistency, availability, partition tolerance
- Parallel databases, hardware and software
- Big data, MapReduce, PigLatin
- Special purpose databases, e.g. temporal, spatial, or multimedia databases

Learning outcomes

By the end of the module, students should be able to:

- - Demonstrate understanding of issues surrounding concurrency control and parallelism in data management.
- - Express queries in different forms (relational algebra, SQL, etc).
- - Devise appropriate ways to store and index data.
- - Show understanding of modern data processing paradigm such as NoSQL and MapReduce/PigLatin.
- - Explain methods suitable for particular types of data such as temporal, multimedia or spatial data.

Indicative reading list

Please see Talis Aspire link for most up to date list.

[View reading list on Talis Aspire](#)

Research element

The students do some critical thinking and research on how to compare different big data systems in terms of efficiency, scalability, and usability. This enhances the depth of their understanding of the latest technologies.

Subject specific skills

Concurrency control and parallelism in data management. Data management design and modelling.
Devising appropriate ways to store and index data.
Show understanding of modern data processing paradigms such as NoSQL and MapReduce.
Explain methods suitable for particular types of data such as temporal, multimedia or spatial data.
Analytical and computational thinking.
Methods and tools for the above

Transferable skills

Teamwork
Time management
Organisation
Problem solving

Study

Study time

Type	Required
Lectures	30 sessions of 1 hour (20%)
Seminars	8 sessions of 1 hour (5%)
Private study	112 hours (75%)
Total	150 hours

Private study description

Background reading
Systems and programming practices
Solving problem sheets
Revision

Costs

No further costs have been identified for this module.

Assessment

You do not need to pass all assessment components to pass the module.

Students can register for this module without taking any assessment.

Assessment group D4

	Weighting	Study time	Eligible for self-certification
Project Report	15%		No

Project report - Group work. This assignment is group work and is not, therefore, eligible for self-certification.

	Weighting	Study time	Eligible for self-certification
Project Software output	15%		No

Project Software output - Group work. This assignment is group work and is not, therefore, eligible for self-certification.

	Weighting	Study time	Eligible for self-certification
In-person Examination	70%		No

CS346 Exam

- Students may use a calculator
- Answerbook Pink (12 page)

Assessment group R1

	Weighting	Study time	Eligible for self-certification
In-person Examination - Resit	100%		No

CS346 resit exam

- Answerbook Pink (12 page)

Feedback on assessment

Written feedback on each assignment.

[Past exam papers for CS346](#)

Availability

Pre-requisites

Students must have studied CS258 Database Systems or be able to show that they have studied equivalent relevant content.

Courses

This module is Optional for:

- Year 3 of UCSA-G4G1 Undergraduate Discrete Mathematics
- UCSA-G4G3 Undergraduate Discrete Mathematics
  - Year 3 of G4G1 Discrete Mathematics
  - Year 3 of G4G3 Discrete Mathematics

- Year 4 of UCSA-G4G2 Undergraduate Discrete Mathematics with Intercalated Year

This module is Option list A for:

- Year 4 of UCSA-G504 MEng Computer Science (with intercalated year)
- Year 3 of UCSA-G500 Undergraduate Computer Science
- Year 4 of UCSA-G502 Undergraduate Computer Science (with Intercalated Year)
- UCSA-G503 Undergraduate Computer Science MEng
  - Year 3 of G500 Computer Science
  - Year 3 of G503 Computer Science MEng
- Year 3 of USTA-G302 Undergraduate Data Science
- Year 3 of USTA-G304 Undergraduate Data Science (MSci)
- Year 4 of USTA-G303 Undergraduate Data Science (with Intercalated Year)

This module is Option list B for:

- Year 3 of UCSA-G406 Undergraduate Computer Systems Engineering
- Year 3 of UCSA-G408 Undergraduate Computer Systems Engineering
- Year 4 of UCSA-G407 Undergraduate Computer Systems Engineering (with Intercalated Year)
- Year 4 of UCSA-G409 Undergraduate Computer Systems Engineering (with Intercalated Year)