

General Information

Course Organization

- Lecturer:** Andreas Pieris (apieris@inf.ed.ac.uk)
- Course Page:** <http://homepages.inf.ed.ac.uk/apieris/courses/atfd2020/>
- Lecture:** Friday, 16:10 – 18:00
50 George Square, room G02
- Assessment:** Essay 1 (15%), due 7 February, ITO before 4pm
Essay 2 (15%), due 3 March, ITO before 4pm
Essay 3 (15%), due 3 April, ITO before 4pm
Final Project (40%), due 20 April, ITO before 4pm
Project Presentation (15%), in class, to be scheduled

The Main Goal

- Explain some of the underlying **theoretical principles** and characteristics of Database Management Systems (DBMSs)
- Explain how **real-life concepts** (such as a database and a query) and **phenomena** (such as incompleteness and inconsistency of data), can be abstracted from their physical implementation and formalised using tools coming from other areas
- Study the main **database tasks** following a mathematically rigorous approach.

Prerequisites

While there are no formal prerequisites, it is recommended that students:

1. Have passed an introductory course in Databases such as the undergraduate course Introduction to Databases; in particular, some familiarity with the relational model, the main relational query languages (calculus and algebra), and integrity constraints is welcome.
2. Have some basic familiarity with complexity theory (standard complexity classes such as PTIME and NP, and the notion of completeness).

In any case, this course is self-contained, and all the necessary tools will be properly introduced and explained during the lectures.

General Plan - Theme 1

Relational Model and Foundations of Relational Query Languages

1. **Relational Model:** data model, relational algebra, relational calculus, query evaluation, static analysis
2. **Conjunctive Queries (CQ):** syntax and semantics, CQ evaluation, static analysis of CQs, minimization of CQs
3. **Fast CQ Evaluation:** acyclic CQs, evaluation of acyclic CQs (Yannakaki's algorithm), semantic acyclicity, evaluation of semantically acyclic CQs

General Plan - Theme 2

Uncertainty - Reasoning over Possible Worlds

1. **Inconsistent Databases:** integrity constraints, primary keys, database repairs, consistent query answering
2. **Incomplete Databases:** the model, querying incomplete databases
3. **Probabilistic Databases:** the model, querying probabilistic databases
4. **Knowledge-enriched Databases:** ontologies, querying knowledge bases

General Plan - Theme 3

Semi-structured Data

1. **Tree-structured Data:** data model, tree pattern queries (syntax and semantics), tree pattern query evaluation and minimization
2. **Graph-structured Data:** data model, basic query languages (regular path queries and extensions thereof), query evaluation

Essays

- Choose a research paper (from a given list) and present
 1. a summary of the paper, and
 2. analysis and critical thoughts (criticism of the paper, and discussion on follow-up works that show how the ideas of the paper have influenced the field)
- Essay 1 is based on theme 1, essay 2 on theme 2, and essay 3 on theme 3
- The length of each essay must be 5 to 7 pages (including references). Each essay should be clearly written in sentences with appropriate punctuation, display of formulae, appropriate use of “Definition”, “Lemma”, “Theorem”, “Proof”, etc. The work should be properly and adequately referenced in the text, with the full list of references at the end of the essay, following any of the standard labelling conventions as technical papers.
- The essay will be marked on its clarity and technical accuracy

Final Project

- Same as for the essays, i.e., choose a research paper (from a given list) and present
 1. a summary of the paper,
 2. analysis and critical thoughts (criticism of the paper, and discussion on follow-up works that show how the ideas of the paper have influenced the field), and
 3. a new contribution:
 - Extend some of the results
 - Close an open problem
 - Implement an algorithm and analyze its performance
 - Isolate special cases that improve existing solutions
 - Etc. (the list is not exhaustive)
- The project is based on **all the themes** covered during the course
- The length of the project must be **7 to 9 pages** (including references)