

INFS7901

Database Principles

Course Summary and Exam Review

Rocky Chen

Course Summary

Final Exam Review

Future INFS Courses

Final Thoughts

Course-Level Learning Objectives

The course was designed to use tools and technologies that promote active learning and provide rich feedback to help you learn to

Course Objectives

1. Reason with the logical foundation of the relational data model and understand the fundamental principles of correct relational database design.
2. Express natural language queries using relational algebra and SQL.
3. Analyse the fundamental techniques and algorithms applied for sorting, tree manipulation, and hashing on structured data.
4. Reason with the logical foundation on how data is indexed and how a query is executed and optimised.
5. Design relational databases with considerations of data integrity and system performance.

Graduate Attributes

1. In-depth knowledge & skills in the field of study
2. Effective Communication
3. Independence and Creativity
4. Critical Judgement

Lectures

Ed

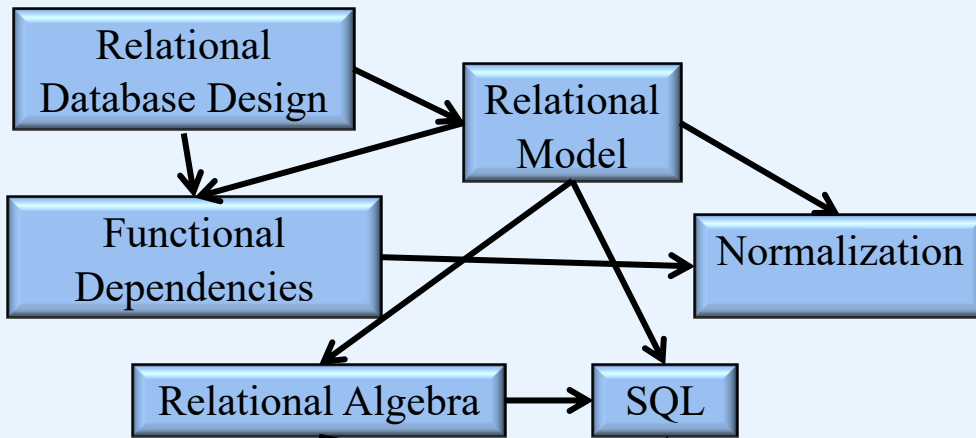
RiPPLE

Jupyter Notebooks

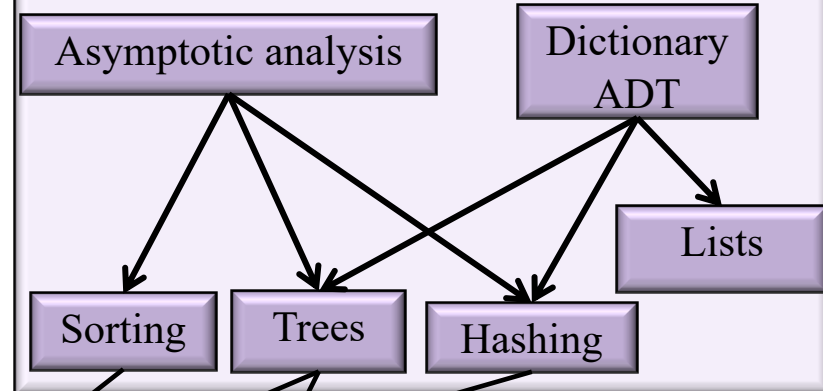
Project

INFS7901 Journey

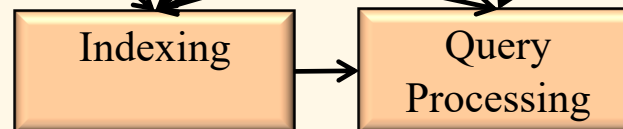
Module 1: Relational Databases



Module 2: Data Access Methods



Module 3 Indexing and Query Processing



Revision Checklist for ER Diagrams

- **Terminology**

1. Entity (Weak entity)
2. Relationship (Degree, Recursive)
3. Attribute (Key and Partial Key, Composite, Multivalued, Derived)
4. Constraints (Cardinality ratio, Participation constraints)
5. Extended ER (Specialization, Generalization)

- **Basic Concepts**

1. Variation in Notation
2. Subjectivity (Expressability, Design choices)
3. Mapping to Relational Model

Revision Checklist for the Relational Model

- **Terminology**

1. Relations (Is a Set, Table with Rows and Columns)
2. Domains (Atomicity, Data type)
3. Attributes (Degree of a relation, Prime or Key attribute)
4. Tuples
5. Key (Super key, Minimal key, Primary key, Candidate key, Foreign key)
6. Mapping

- **Basic Concepts**

1. Constraints (Domain, Key, Entity, Referential)
2. Constraint violations, constraints and operations, inconsistent database state
3. Step by step process for mapping

Revision Checklist for FDs and Normalisation

- **Terminology**

1. Anomalies
2. Functional Dependencies
3. Normal Forms

- **Basic Concepts**

1. How to determine and how to infer FDs
2. Closure computation
3. Definitions of Normal Forms
4. Normalization is a Process

Revision Checklist for SQL

- **Terminology**

1. Declarative vs. Procedural
2. DDL statements (CREATE TABLE, ALTER TABLE, DROP TABLE)
3. DML statements (INSERT, DELETE, UPDATE, **SELECT**)

- **Basic Concepts**

1. Selection, Projection, Sorting (WHERE/HAVING, SELECT, ORDER BY)
2. Aggregation (COUNT, SUM, AVG, MIN, MAX) and Grouping (GROUP BY)
3. Conditions on groups and aggregates (HAVING)
4. Multiple relation queries (Joins, Nesting)
5. When to join and when to nest?
6. Correlated and non-correlated sub-queries
7. Sub-query operators (IN, comparison with (or without) ANY/ALL, EXISTS)

Revision Checklist Asymptotic Analysis

- **Terminology**

1. Input size
2. Big-O
3. Tractable and intractable algorithms
4. Best-, worst- and average case analysis

- **Basic Concepts**

1. Categorise an algorithm into a common complexity class
2. Given code, measure number of steps as function of the input size

Revision Checklist for Sorting

- **Terminology**

1. Time complexity
2. Space complexity
3. Stability

- **Basic Concepts**

1. Understanding how different learning algorithms work
2. Analyse the complexity of different sorting algorithms
3. Compare and contrast different sorting algorithms

Revision Checklist for Binary Trees

- **Terminology**

1. Dictionary data type
2. Tree terminology
3. Binary search trees

- **Basic Concepts**

1. Demonstrate how arrays, linked lists and trees can be used to implement dictionary ADT
2. Compare and contrast algorithms for storing and retrieving values from a dictionary ADT

Revision Checklist for Hashing

- **Terminology**

1. Hash functions
2. The pigeonhole principle
3. collision in hashing
4. Chaining and open addressing
5. Primary and secondary clustering
6. Linear probing, quadratic probing and double hashing

- **Basic Concepts**

1. Have the ability to create good hash functions
2. Demonstrate how hashing can be used to implement dictionary ADT
3. Compare and contrast hashing algorithms for storing and retrieving values from a dictionary ADT

Revision Checklist for Indexing

- **Terminology**

1. Primary, secondary and clustering indexing
2. Sparse and dense indexes
3. Multi-level indexes
4. Tree-based and hash-based indexes

- **Basic Concepts**

1. Explain the benefits of using indexes
2. Explain how different indexing methods work
3. Given information about the most common queries in a database, recommend which attributes using which methods should be indexed.

Revision Checklist for Relational Algebra and Query Optimization

- **Terminology**

1. Relational algebra and its supported operations
2. Initial query tree and final query tree

- **Basic Concepts**

1. Write queries in relational algebra
2. Construct query trees
3. Understand heuristic rules to transform an initial query tree into a final query tree that is efficient to execute.

Course Summary

Final Exam Review

Future INFS Courses

Final Thoughts

Final Exam

- When: **Confirm online**
- Where:
 - **External: Blackboard invigilated (proctorU)**
 - **Internal: UQ (paper-based)**
- Open book: **any printed/written (not digital) materials allowed – good indexing may help**
- Limited time: **10 min reading + 120 min exam.**
 - Feel free to come to the consultation hour next week.

Final Exam

- The final exam will address (theoretical) material from the entire semester
- Double-pass: Students are required to pass the final exam i.e. to obtain **at least 50% in the Final Exam** to pass the subject.
- In ECP: *Students must receive a passing grade on the final exam in order to pass this course (i.e., achieve at least 50%). If you fail the exam, your final mark will be capped at 49 and your final grade will be capped at 3. Fractional marks achieved during assessments will be rounded up before calculation of final results.*

Final Exam Questions

5 MCQs (approx. 5 min each)

9 Short Answers:

1. ER Diagrams, relational model and mapping
2. Functional dependency and normalisation
3. SQL
4. Asymptotic analysis
5. Sorting
6. Trees
7. Hashing
8. Indexing
9. Relational algebra and optimization

Preparing for The Final

- Lecture Notes (incl. clicker questions)
- Past final exams
- Tutorials
- RiPPLE Questions
- Textbook

Course Summary

Final Exam Review

Future INFS Courses

Final Thoughts

Core Courses

- **INFS1200 Information Systems** will give you foundations of what a database system does and how to use it
- **INFS2200 Relational Database Systems** on what the database management software does and how to administer it
- What effect multiple computers and huge amounts of data have in **INFS3200 Advanced Database Systems**

Advanced Courses

- **INFS3202 Web Information Systems** (how the Web affects information systems)
- **INFS3204 Service-Oriented Architectures** (how organizations can be tied together using information systems technology).
- **INFS4205 High Dimensional Data** covers techniques on managing spatial and multimedia data
- **INFS4203 Data Mining** covers discovery of patterns and anomalies in large volumes of data, including the Web
- **INFS7410 Information Retrieval** where you can learn how search engines work and are built

Course Summary

Final Exam Review

Future INFS Courses

Final Thoughts

Have you rated INFS7901?



External:

 [SECaT: INFS7901-25037 \(STLUC-EX-S1-2023\) - Database Principles](#)

Internal:

 [SECaT: INFS7901-24459 \(STLUC-IN-S1-2023\) - Database Principles](#)

Technology has revolutionised the way we live



Upskilling

Skill type 1: Technical skills

- **Data analysis methods:** statistical methods, machine learning methods (e.g., deep learning), principles and pipelines of data analytics
- **New coding languages:** SQL, Python, R, Java, MatLab
- The ability to **determine the right technique for the right task**

Skill type 2: Soft skills

- **Communication:** Present data-driven insights to different audiences
- **Critical thinking and creativity:** Critically justify the pros and cons of solutions
- **Leadership, vision, and more:** The ability to develop business acumen and leadership to derive values from data science applications

Where I Might See You In The Future

ORACLE

AI Scientist - Machine Learning in NLP

21000W5F

AU-AU,Australia-Melbourne; AU-AU,Australia-Sydney

AI Scientist - Machine Learning in NLP

21000W5E

AU-AU,Australia-Melbourne; AU-AU,Australia-Sydney

Analyst - Data Scientist

21000V1M

AU-AU,Australia-Melbourne; AU-AU,Australia-Sydney

amazonjobs

Machine Learning Engineer - AWS

Posted September 3, 2021

(Updated about 1 month ago)

AU, NSW, Sydney | Job ID: 1713564

Amazon Web Services (AWS) is the world leader in providing a highly reliable, scalable, low-cost infrastructure platform in the cloud that powers hundreds of thousands of businesses in 190 countries around...[Read more](#)

Machine Learning Applied Scientist - Anton van den Hengel

Posted October 18, 2021

(Updated 10 days ago)

AU, Sydney | Job ID: 1778641

There's never been a more exciting time to join Amazon Australia! Who are we and what do we do? We are a world class ML team based in Adelaide and created in April 2020 with the hire of the Director...[Read more](#)

Machine Learning Applied Scientist - Anton van den Hengel

Posted October 17, 2021

(Updated 14 days ago)

AU, QLD, Brisbane | Job ID: 1776834

There's never been a more exciting time to join Amazon Australia! Who are we and what do we do? We are a world class ML team based in Adelaide and created in April 2020 with the hire of the Director...[Read more](#)

Where I Might See You In The Future



Consultant/Senior Consultant - Technology Strategy Consulting

Date: 06-Oct-2021

Location: Canberra, ACT, AU

Consultant / Senior Consultant - Digital Finance Technology - Anaplan (Perth)

Date: 17-Oct-2021

Location: Perth, WA, AU

Consultant/Senior Consultant - Technology Strategy Consulting

Date: 06-Oct-2021

Location: Canberra, ACT, AU



Commonwealth
Bank

Technical Business Analyst

REQ148535 | Sydney CBD Area | Posted Today

Senior Insights Analyst

REQ137433 | Redfern, NSW - 166 Redfern Street | Posted 2 Days Ago

Business Analyst

REQ144552 | Sydney CBD Area | Posted Today

From my experience

- **Feel free to communicate!** We are here to help, not to discourage.
- Life is all about experiencing things, including a major, a job, or a sport...
- Starting early is a good habit
- Don't lock yourself in a single long-term vision
- Only the 1% luckiest can combine hobby with career, but almost 99% can contribute to a hobby with a career

Good Luck!

Thanks for the semester!

Good luck to you all!