INFO90002 Database Systems and Information Modelling

David Eccles

Lecture 23
Subject Review

Lecture 24
Exam Preparation

Assignment 2

- Inconsistencies in marking of the A2
- Students have the right to be accessed equitably
- Results will be finalised by end of SWOTVAC
- Assignment specifically stated
  - NO INLINE VIEWS

```
SELECT *
FROM (SELECT *
```

  - NO VIEWS

```
CREATE VIEW AS
(SELECT
```

Subject Revision

- High level overview of the semseter
- Focus on what is assessed

Week 1 The Development Lifecycle

- The Database Development Cycle

  - The database development life cycle
  - Three stages of database development (conceptual, logical, physical)
  - The ability to analyse a case study

Week 2 Intro & Conceptual Modelling

- L3. Introduction to Modelling
- Need to be able to draw conceptual diagrams on your own
  - Given a problem, *determine entities, attributes, relationships*
  - What is key constraint and participation constraint, weak entity?
  - Determine constraints for the given entities & their relationships
  - You must use CHEN notation for conceptual models
- L4. Conceptual Modelling
  - Be able to model a case study from conceptual to instance and all stages in between (conceptual, logical, physical, implementation and instance)
  - Translate conceptual (ER) into logical & physical design
  - Understand integrity constraints
  - Use DDL of SQL to create tables with constraints

Week 3: ER Models & Normalisation

- L5 Modelling with MySQL Workbench
  - Be able to use a modelling tool
  - Need to be able to draw conceptual, logical and physical diagrams
  - SQL DDL CREATE TABLE statements
- L6 Normalisation
  - Normalisation Process (1NF -> 2NF -> 3NF)
  - Anomalies
  - Armstrong's Axioms
  - Functional dependencies
  - Denormalisation

Week 4: Data Types & ER Modelling

- L7: Data Types and Data Modelling
  - Correct Datatype choice & Precision is required for Physical modelling
  - Writing SQL DDL CREATE statements
  - Normalise to 3NF (from Logical model)
  - Distinguish between & identify
    - Conceptual (chen notation)
    - Logical Relation model
    - Logical ER model (crows foot)
    - Physical ER model (crows foot)
    - Implementation (DDL)
    - Instance
- L8: Relational Algebra
  - IS examinable

Week 5 & Week 6

- SQL
  - SELECT
  - DML
  - VIEWS
  - DCL
  - JOINS
    - INNER JOIN
    - NATURAL JOIN
    - LEFT OUTER | RIGHT OUTER JOIN
    - CARTESIAN JOIN
  - RELATIONAL DIVIDES
  - DOMAIN CONSTRAINTS (Unique, Not Null, Data Type)
  - REFERENTIAL INTEGRITY CONSTRAINTS (Foreign Key, Primary Key)
  - CONCEPTS: Key, Superkey, Candidate Key, Primary Key

Advanced Database Concepts

- The remainder of the semester focussed on theoretical ideas
- It is important to demonstrate more than memorisation
  - Apply *concepts*
  - Demonstrate practical *application* of theoretical ideas
  - *Synthesize* knowledge in identifying and solving problems

Week 7 Transactions & Concurrency

- L13 Transactions and Concurrency
  - Why we need user-defined transactions
  - Properties of transactions
    - ACID
  - How to use transactions
    - BEGIN TRANSACTION; START; COMMIT; ROLLBACK;
  - Concurrent access to data
    - Concurrent access strategies
  - Locking and deadlocking
    - Types of Locks (Binary; Shared)
    - Deadlocks what they are; How they happen
  - Database recovery
    - Fundamentals of transaction recovery
    - Checkpoints

Week 7 Storage and Indexing

- L14 Storages and Indexes
  - Describe alternative file organizations
    - File hierarchy (record, page, file)
      - Double linked list (why what how)
    - Heap; Index; Clustered; Unclustered;
  - What is an index, when do we use them
  - Index classification
    - Primary, Secondary
    - Tree based
    - Heap based

Week 8 Architecture & Administration

- What a DBA and Data Administrator do
  - And the difference in each role
- Database Architecture
  - Label all memory structures & know their role
  - Buffer types (current, active, stale aged)
- What affects database performance (nb: not *how)*
  - Caching; Datafile placement; Fast Storage; Indexes; Data types; Query Execution plans; Efficient code;
- When to create an index

Week 8 DDBMS

- Distributed Database
  - Advantages & Disadvantages
- Replicated Databases
  - Advantages & Disadvantages
- Synchronous v Asyncrhonous
  - Difference between
  - Advantages & Disadvantages
- Partitioning Options
  - Vertical, Physical, Vertical and Physical
- The five configurations
  - Advantages & Disadvantages
- Given a scenario choose the best fit for that scenario justifying *strengths* and *weaknesses*

Week 9 Applications

- L17 Database Applications
  - Identify the limitations of SQL
  - Advantages and Disadvantages of Stored Procedures
  - Distribution of Processing Logic (thin, fat, distributed)
  - Database Architectures (1-n tier architecture)
- L18
  - Web languages
    - HTML CSS SQL JAVASCRIPT
  - Web architecture
  - HTML elements
    - Be able to interpret a basic static HTML page
      - <body> <p> <h2><head><style>
  - How static and dynamic web pages work (high level)

Week 10

- L19 Data Warehouse
  - Differences between transactional and informational databases
  - Designing a star schema
  - Defining facts and dimension tables

- L20 Security and Ethics
  - Five Component Framework of Security
    - Technical Human and Data safeguards
  - Rights and Responsibilities
    - Code of Ethics
    - Kant's Categorical Imperative
  - Backup types and differences
    - Online v Offline; Physical v Logical; Incremental v Full
  - Web Security
    - Including vulnerabilities and preventions

Week 11

- NoSQL
- NoSQL Database types
  - Key column, document, column type, node and tie
- ACID v BASE
- CAP Theorem

INFO90002 Database Systems and Information Modelling

Lecture 24: Exam Preparation

Topics Covered

- Handbook Intended Learning Outcomes
- Expected minimum take-away from INFO90002
- Exam
  - Structure
  - Process
- Hurdles

ILO (Intended Learning Outcomes

- Understand the different technologies available to manage structured data, and the evolutionary process that led to them
- Be familiar with how databases work within a larger application architecture
- Understand the relationship of database systems to a variety of fields such as data warehousing, health informatics and Geospatial applications
- Through the combination of seminars, labs and assignments, students gain expertise and confidence to make informed decisions about database systems and appropriate modelling techniques for the structured informational needs of modern organisations. They will gain considerable hands-on experience in modelling a number of diverse informational situations, drawing upon the first principles and techniques taught, useful to both organisations and individuals
- Be able to construct data models at the conceptual, logical and physical level from real-world, natural language requirements documents and apply data normalisation to these models
- Be able to competently use a CASE tool (computer-aided software engineering)
- Be competent in basic SQL and familiar with the usage of advanced SQL commands
- Understand the need and mechanism for database transactions, including the so-called ACID properties

See https://handbook.unimelb.edu.au/2020/subjects/info90002

INFO90002 Take away skills

- School faculty members expect students to be able to do 3 things when they leave INFO90002
  - Model
    - Case Study → Conceptual → Logical → Physical → Implement → Instance
  - Fluent in basic SQL
    - Competent in moderately complex SQL
  - Normalise data

EXAM

Tuesday 24th November 2020 9.00 a.m.

- Section 1: Modelling, SQL, RA, & Normalisation (70 marks)
  - "Take Home exam"
  - No invigilation
  - Submitted via the LMS within 48 hours from release date
    - Release date: Tuesday November 24th 2020 AEDT
    - Submitted by Thurdsay November 26th 9am AEDT – No exceptions

Monday 23rd November 2020 4.00pm

- Section 2: Multiple Choice and Short Answer (70 marks)
  - Conducted via Canvas Quiz
  - Timed
    - once you start you have 2 hours to complete*
  - 25 multiple choice questions from a variety of topics (2pts each)
  - Must select the BEST answer (more than 1 answer may be correct)
    - More than 1 answer may be correct
  - 5 Short Answer Questions (4 marks each)
    - Bullet points are fine

Exam Breakdown

- Part 1
  - Modelling 30 marks
  - SQL & Relational Algebra 25 marks
  - Normalisation 15 marks
- Part 2
  - All database theory (weeks 1-12)
  - MCQ 50 marks
  - Short Answer 20 marks
    - PLAN YOUR TIME Accordingly!

Hurdle Requirements

- To pass INFO90002 you must pass two hurdles
- Hurdle 1:
  - 15/30 for Assignment 1 and 2
- Hurdle 2:
  - 35/70 for the Exam

Hurdle Examples

- Alice
  - Assignment 1: 11.8/20
  - Assignment 2: 9.4/10
  - 21.2/30 (hurdle 15)
  - Exam
  - 40/70 (hurdle 35)
  - 61.2 → 62% P
- Bob
  - Assignment 1: 19.3/20
  - Assignment 2: 9.7/20
  - 29/30 (hurdle 15)
  - Exam:
  - 33/70 (hurdle 35)
  - 62/100 NH 49 Hurdle fail

Exam Preparation

- Practice!
- Practice the exam question answers
- Engage!
  - Ask questions on the LMS
  - Try the www.w3schools.com SQL tutorials
- Check your understanding

Remote Exams

- Prepare for the exam anxiety / stress
- PLAN THE EXAM
  - Reading time in the first 3 hours of the first day of the Assignment component
- Understand the Instructions
  - IF NOT ASK THE Examiner

- Plan how and when you'll tackle both sections
- 2 days for Part 1 then 2 hours for Part 2
  - Pace Yourself
    - Have something in the tank for Part 2
  - Rest breaks (consider Pomodoro for Part 1)
  - Make summary notes
- Part 2 will feel like usual "exam stress"
  - 2 hours to complete 25 questions and 5 short answer questions
  - Majority of the time is MCQ (50 marks) Short Answer (20 marks)
- Part 1 is different but the same

David's Advice for a career in I.T.

1. Get a blue chip on your CV
   - Amazon AWS, Google, Microsoft, Oracle, IBM, Big4 Accounting, Major Banks and Insurance companies
   - It shortlists your CV
   - Tech companies are misogynist and have a bullying culture. Don't put up with that crap.
2. Change your job every 3-5 years (especially if an intern!). Don't rust on!
   - Best way for salary growth and job variety
   - OK to step down in salary and role if it is a new growing market
3. Small companies/Contract roles = great job variety / good networking (long hours)
4. Large companies = pigeonholed (well resourced & good pay + conditions)
5. Seek Job Variety (tech is everywhere!). Volunteer for new projects.
6. It's okay to say "I don't know". It's okay to ask for help.
   - Better than costing the company $$$$$ for guessing and getting it wrong!
7. When in doubt ESCALATE and ask your peers for advice.
   - Escalate within your chain of command or **informally** approach H.R team.
8. Treat those underneath you in the org chart as you wish to be treated – fairly and honestly
   - They may be your manager one day!
9. Government & Statutory Authorities are bureaucratic. it's what they do. Slow change slow innovation. It suits some personality types more than others
10. You'll know when you have a great job – it won't feel like a job!

THANK YOU

On behalf of the tutors and staff of INFO90002 and the School of Computing and Information Systems

THANK YOU!

We wish you well with your upcoming assessments and for your future studies here at the University of Melbourne.

David, Veronica, Fraser, Ibrahim, Neven and Nick.

THE UNIVERSITY OF
MELBOURNE