Final Lecture

SWEN 304

Trimester 1, 2021

Lecturer: Dr Hui Ma

Engineering and Computer Science

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

What We Learned (1)

- Database Basics
  - What is a database
  - What is a DBMS and what are its tasks
  - Program - Data Independence
  - Three Schema Architecture
- Relational Data Model
  - Tuple
  - Relation
  - Relational Schema
  - Relational Database
  - Relational Database Schema
  - Relational Schema Key
  - Attribute Constraint
  - Referential Integrity Constraint

What We Learned (2)

- Entity-Relationship (ER) Data Model
  - Regular entity type
  - Relationship type
  - Cluster

- Mapping EER to the Relational Data Model
  - Mapping an entity type
  - Mapping a relationship type
  - Mapping with clusters

What We Learned (2)

- Structured Query Language
  - DDL, DML
  - VDL
- Relational Algebra
  - Basic Operations (select, project, join)
  - Set Theoretic Operations (union, intersect, difference)
- Query optimization
  - Heuristic Optimization
    - Reordering of algebraic operations
    - First execute unary than binary operations
  - Cost Based Optimization
    - Start from heuristic optimization tree
    - Calculate cost of each successive operation

What We Learned (3)

- Update Anomalies
  - Insertion, Deletion, and Modification Anomaly
- Lossless Join
  - Two Relations – A Key is in the Intersection
  - More Than Two Relations – A Relation Schema Contains A Key of the Universal Relation Schema
- Functional Dependencies
  - Inference Rules
  - Closure of a Set of Attributes
  - Finding a Minimal Cover of a Set of Functional Dependencies
  - A Relation Schema Set of Keys Is A Consequence of Functional Dependencies
    - A Key Finding Algorithm
    - Additional key Finding Algorithm

What We Learned (4)

- Normal Forms Based On Functional Dependencies
  - Definitions: 2NF, 3NF, BCNF

- Normalization Algorithms
  - Synthesis Algorithm: 3NF
  - Decomposition Algorithm: BCNF
  - Dependency preservation

What We Learned (5)

- Transaction Processing
  - Dirty Read,
  - Unrepeatable Read,
  - Lost update
  - Phantom Record
- Concurrency Control
  - Shareable and Exclusive Locks
  - Basic 2-phase locking (acquire all locks before releasing any)
  - Strict 2-phase locking and Conservative 2-phase lock
  - Transaction Isolation Levels
  - Deadlock
  - Deadlock Prevention Protocols

SWEN304 Assessment

- Assessment:

  Assignment 1: (marked) 15%
  Assignment 2: (marked) 10%
  Assignment 3: (being marked) 10%
  Project 1: (marked) 20%
  Project 2: (to be marked) 10%

  } **65%**

  Final test 35%

- Test will be at **9:30am Monday, 14 June**
- SWEN304 - HMLT104, HMLT205
- Remote students will be informed about test via Zoom

SWEN439 Assessment

- Assessment:
  - Assignment 1: (marked) 15%
  - Assignment 2: (marked) 5%
  - Assignment 3: (being marked) 5%
  - Project 1: (marked) 20%
  - Project 2: (to be marked) 10%

  } **55%**

  - Essay 15%
  - Final test 30%
- Test will be at **9:30am Monday, 14 June**
- SWEN439 - HMLT002, HMLT205;
- Remote students will be informed about test via Zoom

This Year's Test

- Testing you understanding of what we:
  - discussed during the lectures and tutorials,
  - used to solve assignment and project questions
- 100 marks, 2 hours
- Test questions will cover all the important topics we learned
- Help Desk?

Summary or Relational Operations

- SELECT $\sigma_c(r(N))$: choose rows
- PROJECT $\pi_{A1,\ldots,Ak}(r(N))$: choose columns
- RENAME $\delta_{A1\rightarrow B1,\ldots,Ak\rightarrow Bk}(r(N))$: rename attributes
- JOIN: combine tables
  - Natural Join $r(N_1) * r(N_2)$ or
  - Equi-Join $r(N_1) \bowtie_{A1=B1,\ldots,Ak=Bk} r(N_2))$
- CARTESIAN PRODUCT (x): combine tables
- Set operations
  - UNION ($\cup$),
  - INTERSECTION ($\cap$),
  - DIFFERENCE (or MINUS, –)
- Additional Relational Operations
  - OUTER JOINS

Queries

- **More than two** and **exactly one** questions.
- Which students have enrolled more than one paper (two or more)

ENROLMENT	
ID	paper
101	157221
101	157331
102	157223
103	157331

- we first rename paper to paper' and join the resulting relation with Enrolment, hence

$$X := \text{ENROLMENT} * \delta_{\text{paper}\mapsto\text{paper'}}(\text{ENROLMENT})$$

ID	paper	paper'
101	157221	157221
101	157221	157331
101	157331	157221
101	157331	157331
102	157223	157223
103	157331	157331

Queries

- We're not interested in pairs which have same entries; these are

$$Y := \sigma_{\text{paper=paper'}}(\text{ENROLMENT} * \delta_{\text{paper}\mapsto\text{paper'}}(\text{ENROLMENT}))$$

ID	paper	paper'
101	157221	157221
101	157331	157331
102	157223	157223
103	157331	157331

- So we take these away, got tuples of students enrolled more than one course:

$$Z := X - \sigma_{\text{paper=paper'}}(X)$$

ID	paper	paper'
101	157221	157331
101	157331	157221

Queries

- We get the student ID that have enrolled **two or more** courses

$$\pi_{id}(Z)$$

ID
101

- Finally we remove those students who have enrolled in more than one paper to get those students who have enrolled in **exactly one** paper, this gives:

$$\pi_{id}(\text{ENROLMENT}) - \pi_{id}(Z)$$

ID
102
103

- In one query:

$$\pi_{id}(\text{ENROLMENT}) - \pi_{id}(\text{ENROLMENT} * \delta_{\text{paper}\mapsto\text{paper'}}(\text{ENROLMENT}) - \sigma_{\text{paper}=\text{paper'}}(\text{ENROLMENT} * \delta_{\text{paper}\mapsto\text{paper'}}(\text{ENROLMENT}))))$$

SWEN304/ SWEN439 Test

Good Luck!