CMPT 354 D1: Course Review

(2004-2 semester)

Chapter 1: introduction
- DBMS vs. file system, DBMS structure
- Data model
- Levels of abstraction

Chapter 2: database design with ER model
- Entity, entity set, relationship, relationship set
- Key constraint, participation constraint, weak entity, class hierarchy, aggregation
- Representations

Chapter 3: relational model
- SQL language, integrity constraints (PK, FK)
- ER model to relational
- Views

Chapter 4: formal relational query languages
- Relational algebra (selection, projection, set operations, renaming, join / cross-product, division)
- Relational calculus
- Expressive power of relational algebra and calculus

Chapter 5: SQL language
- Basic query and conceptual evaluation strategy
- Set operations
- Aggregate operations (COUNT, SUM, AVERAGE, MAX, MIN)
- Group aggregation (GROUP BY, HAVING) and conceptual evaluation strategy
- General constraints
- Triggers

Chapter 8: overview of storage and indexing
- File organizations, I/O cost comparison
- Index: alternatives (1), (2), (3), clustered vs. unclustered
- Hash index, B+ tree index
- Choice of index, index-only plan, composite search key

Chapter 9: disks and database files
- Disk and I/O operation cost
- Disk space management
- Buffer management, buffer pool, buffer replacement policy
- Record format (fixed-length, variable-length), page format (fixed / variable-length records), format of file of records (heap file)

Chapter 10: tree-structured indexing
- ISAM
- B+ tree: order, insertion, deletion, prefix key compression, bulk loading

Chapter 11: hash-based indexing

- Static hashing
- Extendible hashing: global depth of directory, local depth of bucket, searching, insertion, deletion
- Linear hashing: a family of hash functions, searching, insertion, deletion
- Similarity between extendible hashing and linear hashing

Chapter 13: external sorting

- Internal vs. external sorting
- 2-way external merge sort, I/O cost
- General external merge sort, I/O cost
- Replacement sort
- Blocked I/O, double buffering
- B+ tree for sorting

Chapter 14: relational operators evaluation

- Selection: with / without index, refinement for unclustered index, general selection conditions
- Projection: eliminate duplicates based on sorting / hashing
- Set and aggregate operations
- Nested loops join, sort-merge join, hash join, I/O costs

Chapter 15: relational query optimization

- Query evaluation plan
- Query block
- Relational algebra equivalences
- I/O cost estimation, result size estimation
- Enumeration of alternative plans (single-relation queries, multi-relation queries)

** This document is intended as an overview of the major contents in each chapter. It gives NO indication on what may or may not be included in the final exam.
** The final exam will cover the chapters 1~5, 8~15, excluding those parts not discussed in the lectures.
** The final exam will take place at: 12: 00 ~ 15: 00, Saturday, 2004-Aug-07, WMC 3260.