Introduction to Database Systems
CSE 414

Lecture 30: Final Review

The Final

- Monday, June 10, 2:30-4:20
- In class
- Open textbook, 1 sheet hand-written notes (+ sheet from midterm), no other notes/gadgets
- Review session: Sunday, 6/9, 2 pm, Loew 101

The Final

Entire class content is on the final!

But focus of questions on the final will be as follows:

1. SQL and Relational Query Languages
2. Relational algebra and query implementation
3. XML
4. Database design
5. Views
6. Transactions
7. Parallel Databases

How To Study

- Go over the lecture notes
- Read the book
- Go over the homeworks
- Practice
  - Practice web quiz (coverage is not completely balanced, but maybe useful for part of your review)
  - Finals from past 344s
  - Questions in the book
- Ask course staff questions! Trade questions/answers on discussion board!
- The goal of the final is to help you learn!

1. SQL including Views

SQL

- SELECT-FROM-WHERE
- DISTINCT, ORDER BY, renaming of attributes
- INSERT, DELETE, UPDATE
- GROUP-BY and HAVING: *different* from WHERE (why ?)
- NULLs, outer joins
- Nested queries (subqueries)

Know the syntax (but don't obsess over trivia)

Know the semantics (nested loops !)

1. SQL and Relational Query Languages

SQL

- CREATE TABLE, plus constraints
- INSERT/DELETE/UPDATE

2. Relational Algebra & Query Impl.

- Relational algebra
- Translation SQL $\leftrightarrow$ RA
- Pipelined execution vs. tuple materialization
- Cost measures: B(R), T(R), V(R,a), M
- Basic query algorithms / strategies
  - Hash join, nested loop join, sort-merge join

3. XML

- Basic syntax: elements, attributes; well-formed v.s. valid document
- XPath
- XQuery

4. Database Design

E/R diagrams:

- Entities, attributes
- Relationships:
  - Many-many, many-one, one-one, exactly one
  - Multi-way relationships
- Inheritance, weak entity sets, union types
- Constraints in E/R diagrams
- Translation to relations

4. Database Design

Constraints in SQL

- Keys and Foreign Keys
- Attribute level constraints
  - Predicates on values
  - NOT NULL

4. Database Design

Conceptual Design

- Data anomalies
- Functional dependencies
  - Definition
  - Make sure you can check if a table satisfies a set of FDs
- Attribute closure
- Keys and Super keys
- Definition of BCNF
- Decomposition to BCNF

5. Views

General ideas are the most important thing here

- Types of views: virtual vs. materialized views
- Definition and how to use them
- CREATE VIEW in SQL
- Query modification

6. Transactions

Transactions concepts

- Review ACID properties
- Definition of *serializability*
- Schedules, conflict-serializable and recoverable
- The four isolation levels in SQL – basic tradeoffs
- Concurrency control using locks
- Phantoms, dirty reads, and other problems
- Deadlocks
- Transactions in SQL

7. Parallel Data Processing

Parallel databases:

- Speedup/scaleup
- Shared memory, shared disk, shared nothing
- Horizontal data partition: block, hash, range
- How to implement simple algorithms: group-by, join
- How to execute a complete query in parallel

7. Parallel Data Processing

MapReduce

- Functions: map, (combine,) reduce
- Terminology: chunk, map job / reduce job; map task / reduce task; server (instance); failed server
- Basic implementation of MR
- Dealing with server failures and stragglers
- How to express simple computations in MapReduce

You will not be asked to write a Pig Latin query, but should have some basic understanding of how queries are implemented over MapReduce

Possible Questions about Last Lectures

- NoSQL: You can expect some sort of qualitative question that will ask you to
  - Explain the motivation for NoSQL systems
  - Discuss *some* similarities and differences between relational and NoSQL systems
  - Something else along those lines

- Data integration and data cleaning
  - You can similarly expect a qualitative question on these topics

And that's it…

Congratulations on a productive quarter!

Good luck on finals, last projects!!

Last piece: course evaluations

- CSE 414 was our first attempt at databases for non-majors. What worked well? Sortof? Not so well?
- What are the important things to keep? Drop? Fix?

Thank you for your time on this

See you Monday for the final exam!!