CMPT 354: Database System I

About Midterm

Midterm Coverage

1. Database History (10%)

2. Relational Model (10%)

3. SQL (45%)

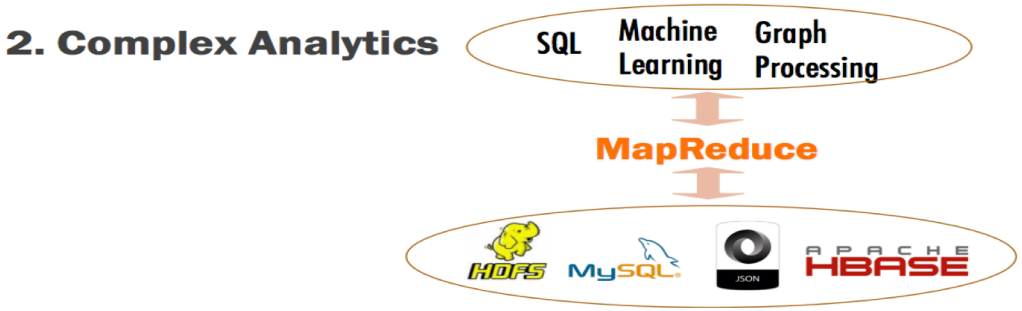
4. Relational Algebra (25%)

5. Query Processing and Indexing (10%)

Database History (10%)

- You need to **understand**
  - Why Relational Model
  - Why MapReduce
  - Why NoSQL
- Example

  F (T) It is possible to implement SQL using MapReduce



Relational Model (10%)

- You need to **understand**
  - Basics of Data Models
  - Terminologies
  - Keys

- Example

F T A primary key is a single column that uniquely identifies a record in a relation

SQL (45%)

- You need to be **familiar** with
  - SQL DDL
    - Create/Insert/Alter
    - Constraints
  - SQL DML
    - Selection, Projection
    - Set Operators (UNION, INTERSECT, EXCEPT)
    - Joins (INNER, OUTER)
    - Aggregation, Group By, Having
    - Order By, Distinct, NULL
    - Subqueries
- Example
  - Similar to A1 and A2

Relational Algebra

- You need to **know**
  - How to write an RA query
  - How to optimize an RA query
  - How to convert an SQL query to an RA query
- Examples
  1. All students whose birth is larger than 1995
     - $\sigma_{birth > 1995}$(Student)
  2. Optimize this query $\sigma_{cNum=354}$ (R $\bowtie$ S)
     - $\sigma_{cNum=354}$ (R) $\bowtie$ S
  3. Convert "SELECT name FROM student" to an RA query
     - π_{name} (Student)

Query Processing and Indexing

- You need to know
  - Query Processing Steps
  - Which index is better/useful?

- Examples
  - What does SQL Parser do?
    - Convert the input SQL text to a logical plan
  - Can the following index speed up this SQL query:
    SELECT * FROM Student WHERE id = ?
    - Index on Student(id) YES NO

Notes

- Midterm
  - Thursday 2:30 – 3:20 pm
  - AQ3149
  - Don’t be late

- **Bring your Student IDs**
- Please **budget your time** so you get to all questions
- **Relax**. You are here to learn