CMPT 354: Database System I

About Midterm

Midterm Coverage

1. Database History (7.5%)

2. Relational Model (7.5%)

3. SQL (40%)

4. Relational Algebra (25%)

5. Query Processing and Indexing (20%)

Database History (7.5%)

- You need to **understand**
  - The concepts of OLTP and OLAP
  - Why NoSQL
  - The contents of MapReduce
- Example

F (T) MongoDB is a typical system of NoSQL

Relational Model (7.5%)

- 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 (40%)

- 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 (25%)

- 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 (20%)

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

- Examples
  - What does SQL Parser do?
    - Convert the input SQL text to a logical plan
  - Similar to A3
  - Can the following index make this SQL query faster, slower or the same:

CREATE INDEX newI **ON** Student(name)

SELECT * FROM Student WHERE id = '301414'

The same!!

Query Processing and Indexing (20%): Additional Hint

- You need to know
  - If you create a relation with primary key, DBMS creates an index for this primary key automatically.
- Example

In [5]:

```
%%sql

CREATE TABLE students (
    id integer,
    name varchar(30) NOT NULL,
    gender char(30),
    age integer,
    PRIMARY KEY(id)
)
```

```
 * sqlite:///coursys.db
Done.
```

(1) Create relation students

'id' is the primary key

In [6]:

```
%%sql
EXPLAIN QUERY PLAN
SELECT * FROM students WHERE id=1
```

```
 * sqlite:///coursys.db
Done.
```

Out[6]:

id	parent	notused	detail
2	0	0	SEARCH students USING INTEGER PRIMARY KEY (rowid=?)

(2) Query relation students with clause 'id = 1'

the index of primary key 'id' is used

Notes

- Midterm
  - Wed. 10:30 – 11:20 am
  - EDB7618
  - Please come to the classroom at least 5 mins earlier

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