

Database Management Systems

- Exam Review

Doug Shook

Logistics

- Thursday in class
- Open book/notes
- No electronic devices

Topic Outline

- General DB Structure
 - Tables, rows, columns
 - Data types
 - Metadata
 - Relationships
- Basic SQL
 - SELECT, FROM, WHERE, GROUP BY, ORDER BY
- Design
 - ER Diagrams
 - Normalization
 - CREATE TABLE, INSERT, UPDATE, DELETE

Topic Outline

- Physical Layer
 - Heap file structure
 - Insertion and deletion of data
 - Retrieval of data
- Relations and Constraints
 - Types of relationships
 - Primary keys / foreign keys
 - Triggers
- Relational Operations
 - Types of operations
 - Converting queries to relational operations

Topic Outline

- Indexing
 - Types of indexes
 - Properties of indexes
 - B+ tree operation
- Query optimization
 - Query tree creation
 - Optimization heuristic

Exercise

- We wish to create a database that tracks what movies users have watched.
- Construct an ER diagram for this database including all entities, relationships, and at least a few attributes per entity

Exercise

- Use SQL to create the database you designed in the previous exercise. Include all primary and foreign keys in your implementation.
- Insert at least 3 rows into each table of your database
- Explain the datatypes and constraints for each of the columns in your tables.

Exercise

- Given the following tables, write queries to answer the following questions:
- What is the name of each user?
- How many movies have been watched by all users?
- How many movies has each user watched?
- How many times has each movie been watched?

uid	fname	lname
0	Doug	Shook
1	Jenny	Weber

mid	name
0	The Big Lebowski
1	The Matrix
2	The Usual Suspects

uid	mid
0	0
0	1
1	0

Exercise

- For each of the queries you wrote in the previous problem:
 - Construct a query tree
 - Optimize the query using the heuristic

Exercise

- Are the tables on slide 8 properly normalized?
Explain how you know.

Exercise

- For each column in the users and movies table on slide 8, explain which of the four types of indexes (primary, clustered, secondary sorted, secondary unsorted) apply, and explain why.

Exercise

- For the users page on slide 8:
 - What is the size of one tuple, using the DB from hw1?
 - How many tuples could fit on one page?
 - How many bytes would be in the header of a page?
- Explain the difference between fixed size and variable sized tuples. How would our heap file have to change if the tuple sizes were allowed to vary?

Exercise

- Construct a B+ tree with $b = 2$ by inserting the following values:
 - 42, 15, 7, 53, 54, 9, 12, 75, 1, 24, 61
- Repeat this process with $b = 3$

Exercise

- Show the result of deleting the following values from the B+ tree you created in the previous exercise
 - 9, 61, 1, 7, 53, 75, 12, 24
- Repeat this process with $b = 3$