COURSE DESCRIPTION

Department and Course Number: CSC 244 **Coordinator:** Ying Jin

Course Title: Database System Design **Units:** 3

Catalog Description

Topics in the design and implementation of database management systems. Database system concepts and architectures; query compiler, query processing algorithms, logical and physical query plans, query optimization; recovery, concurrency control; transaction management in centralized database management systems and distributed database management systems. Exploration of current research directions, issues, and results related to databases and data management. **Prerequisite**: CSC 174 or CSC 204.

Textbook

H. Garcia-Molina, J. D. Ullman, and J. Widom, Database System Implementation, 1st Edition, Prentice-Hall Publishing, 2000.

References

R. Elmaseri and S. Navathe, Fundamentals of Database Systems, 4th Edition, Addison-Wesley Publishing, 2004.

Course Goals

Understanding of the design and implementation of database management systems, including:

1. Database management system architecture.
2. Query processing issues and query optimization.
3. Database recovery techniques.
4. Concurrency control techniques.
5. Transaction management in centralized and distributed database management systems.
6. Recent research directions and issues related to databases and data management.

Prerequisites by Topic

Thorough understanding of:

- File structures and organization.
- Hashing and indexing structures for files.
- Relational data model.
- SQL syntax and usage.
- Relational algebra.

Basic Understanding of:
- Transaction processing and database recovery concepts.
- Fundamental concepts of concurrency control.

Major Topics Covered in the Course

1. Database system concepts and architectures (1.5 hours).
2. Query execution algebra and operators (3 hours).
3. Query processing algorithms: one-pass algorithms, nested-looping joins, two-pass algorithms based on sorting, two-pass algorithms based on hashing, index-based algorithms (10 hours).
4. Query compiler (2 hours).
5. Algebraic Laws for improving query plans (3 hours).
6. Logical and physical query plans (3 hours).
7. Query optimization and cost-based plan selection (5 hours).
8. Database recovery techniques (3 hours).
9. Concurrency control techniques (3 hours).
10. Transaction management in centralized and distributed database management systems (4 hours).
11. Special topics on recent database research (6 hours).
12. Exam (1.5 hours).

Outcomes

Thorough understanding of:
- Query execution algebra and algebraic laws for improving query plan.
- One-pass and two-pass query processing algorithms.
- Cost-based plan selection.
- Database recovery techniques.
- Concurrency control techniques.

*Basic understanding of***:**
- Concepts of distributed databases.
- Distributed transaction management issues.

Exposure to:
- Recent research issues related to databases and data management.

Laboratory Projects

1. Design and implement an iterator.
2. Generate parse tree; translate a parse tree to an initial logical query plan; optimize an initial logical query plan to final logical query plan.
3. Design recovery solutions based on different recovery techniques.

Estimated Curriculum Category Content (Semester hours)

Area	*Core*	*Advanced*	*Area*	*Core*	*Advanced*
Algorithms		1.5	Data Structures		.25
Software Design		.25	Prog. Languages		
Comp. Arch.					

Oral and Written Communications

Students will be required to write a report on the current research issues related to databases and data management. The topics are given by the instructor. Each student will be required to give an oral presentation for approximately 10 minutes.

Social and Ethical Issues

No significant component.

Theoretical Content

- Query execution algebra and algebraic laws.
- Query processing algorithms.

Analysis

Query plans are selected based on cost-based analysis.

Design

Design is part of each laboratory assignment.

YJ
11/9/05