Advanced Database Systems

Autumn 2020-21

Week 1

Lecture 1 (02/09/2020)
- Introduction to the course

Lecture 2 (03/09/2020)
- Recap of database indexing, B+ Tree
- Introduction to multi-dimensional indexing, high dimensional indexing
- Structure of R Tree

Lecture 3+4 (04/09/2020)
- R Tree
- Insert, delete and search algorithms
- Nearest neighbor search on R Trees
- Introduction to pruning strategies

Week 2

Lecture 5 (09/09/2020)
- Use of MINDIST and MINMAXDIST for NN search

Lecture 6 (10/09/2020)
- Other types of trees
- RStar, RPlus and SS trees
- High dimensional image database search

Lecture 7+8 (11/09/2020)
- Properties of high dimensional data
- Different distance metrics
- Approximate nearest neighbor search in high dimensions
- Document database search, inverted index
- Vector space model
- TF.IDF for document vector representation
- Document similarity and searching of document databases

Week 3

Lecture 9 (16/09/2020)
- Introduction to temporal databases
- Basic examples requiring support for temporal data elements
- Valid time and transaction time
Lecture 10 (17/09/2020)
- SQL-2011 support for temporal elements
- Time period concept
- Application-time period table and System-time period tables (System versioned tables)
- Creating Application-time period tables
- Insert, update, delete in application-time period tables
- Primary key and foreign key in application-time period tables

Lecture 11+12 (18/09/2020)
- Querying application-time period tables
- Creating system versioned tables
- Insert, update, delete in system-versioned tables
- Primary key and foreign in system-versioned tables
- Querying system-versioned tables
- Support for application-time period tables in Oracle - demonstration
- Support for system versioning in MS SQL Server – demonstration
- Summary and discussion

Week 4

Lecture 13 (23/09/2020)
- Deductive databases
- Basic Datalog concepts
- Facts and rules
- Head, goal and body of rules
- Conjunction and disjunction
- Simple Datalog program

Lecture 14 (24/09/2020)
- Datalog with negation
- RA, TRC, DRC, DL equivalence
- Safe datalog
- Converting safe non-recursive datalog with and without negation to RA

Lecture 15+16 (25/09/2020)
- Class test 1 (topic temporal databases) held
- Recursive Datalog rules
- XSB for writing and executing Datalog programs

Week 5

Lecture 17 (30/09/2020)
- Advanced topics in transaction management
- Review of basic techniques
• ACID properties
• Lock based protocol
• Timestamp-based protocol
• Validation based protocol (Optimistic concurrency control)

Lecture 18 (01/10/2020)
• Multi-level schemes
• Multi-version schemes
• Weak levels of consistency
• Snapshot isolation

Week 6

Lecture 19 (07/10/2020)
• Weak levels of consistency revisited
• Phantom phenomenon
• Serializability considerations for insert and delete operations
• Index locking

Lecture 20 (08/10/2020)
• Long lived transactions and Sagas
• Compensatory transactions
• Recovery mechanisms
• Notion of stable storage
• Log based recovery
• Undo and redo operations

Lecture 21+22 (09/10/2020)
• Distributed databases
• Homogeneous and heterogeneous distributed databases
• Replication and fragmentation
• Data transparency
• Two phase commit
• Failure modes and their handling in 2PC

Week 7

Lecture 23 (14/10/2020)
• Alternative models of transaction processing
• Persistent Messaging
• Concurrency control in distributed transaction processing
• Single and distributed lock manager approaches
• Majority protocol, biased protocol, quorum consensus protocol

Lecture 24 (15/10/2020)
• Timestamping
• Replication with weak concurrency
• Deadlock handling
• Centralized approach
• Global wait for graphs
• False cycles
• Availability in distributed databases

Lecture 25+26 (16/10/2020)
• Class Test 2 held
• Consistency vs. Availability
• CAP Theorem
• BASE property
• Availability vs Latency

Week 8

Lecture 27 (21/10/2020)
• In-memory databases introduction
• Characteristics of in-memory databases
• Design and architectural considerations
• Garcia-Molina’s paper

Lecture 28 (22/10/2020)
• Design considerations continued
• H-store – In-memory database (row store)
• C-store – a column data store (not in-memory)

Week 9

Lecture 29 (28/10/2020)
• Motivation behind re-architecting database design
• C-Store continued

Lecture 30 (29/10/2020)
• Join index
• Plattner’s paper on both OLTP and OLAP workloads on In-memory databases
• Oracle Times Ten
• Overview of VoltDB, Vertica, MariaDB, MonetDB