Review

DB Design

- Entities, Relationships, and Attributes
  - (Primary) Keys of entities
  - Datalog-like notation & EGDs
  - Cardinality Constraints on Relationships
  - (Primary) Keys of relationships
  - Participation Constraints & TGDs
  - Weak and strong entities
  - ISA hierarchies – generalization & specialization (disjoint vs overlapping; partial vs total)
  - Aggregation
- Designing a conceptual DB schema involves picking features and constructs from the above to specify the DB schema
- Several constraints of application not captured in a ER diagram

Relational DB design

- Basic concepts
  - Creating & altering tables in SQL
  - Specifying constraints
  - Primary vs candidate keys
  - Foreign keys and referential integrity constraints (RICs)
  - Converting an ER diagram to a relational db schema in SQL
  - Specifying constraints in SQL and in Datalog
  - Use of views

Schema Refinement (Normalization)

- Pitfalls in DB scheme design – redundancy & anomalies
  - FDs
  - Lossless-join split (decomposition)
  - Armstrong's Axioms, FD Inference, super keys and (candidate) keys
  - Proofs, Closure, Finding all candidate keys
  - Normal Forms
    - 3NF
    - BCNF
  - Lossless-join decomposition into BCNF
  - … into 3NF
  - 3NF design via synthesis (need Minimal Cover)
  - FD-preserving designs

RA & Datalog

- RA: basic ops and additional/derived ops.
  - Diff. b/w − and ≠.
  - Various joins, division, ∩, and assignment.
  - Beyond division.
  - Max/min queries.
- Datalog Lite and Datalog Full.
  - Basic (SPJ) queries.
  - ¬ vs ≠.
  - Queries with multiple rules.
  - Max/min queries.

SQL

- Full-fledged data manipulation language
- Queries
  - Duplicate elimination using DISTINCT
  - String manipulation, Ordering output
  - Set ops and their bag variants
  - IN, NOT IN, op ANY, op ALL, EXISTS, NOT EXISTS, UNIQUE, NOT UNIQUE
  - Aggregation, Group By, Having
  - Null Values
  - Views, WITH clause, special Joins
  - Integrity Constraints
- Insert, Delete, Update & Transactions

Concurrency Control

- ACID properties of transactions
- Schedules and Conflicts
  - RW, WR, and WW
  - (un)recoverable schedules and cascaded aborts
  - Conflict serializability
  - View serializability
- protocols for ensuring CC
  - Regular, strict, conservative 2PL
    - Deadlocks – detection, resolution, and prevention
    - Phantom problem
  - Timestamp based

Crash Recovery

- Buffer management
  - (no) steal, (no) force
- Key steps of ARIES algorithm
- Logging and Checkpointing
- Undo/Redo Logging
- Recovering from a crash
- Interaction w/ CC mechanism