COS 425:
Database and Information
Management Systems

Final Remarks

Where we have been

- MODELS and QUERIES
  - Relational
  - XML
  - Information retrieval
- STORING and ANSWERING
  - File organization and indexing
  - Relational query evaluation and optimization
  - Indexing for search
  - Indexing the XML tree model
- TRANSACTIONS: Correctness and Durability
  - Managing concurrency
  - Managing failure

A "look under the hood" course

- Understand what going on
    => Make better choices in design and use
- Major CS methods applied
    - Specification/modeling/correctness
    - Algorithms and cost analysis
    - Concurrency control and reliability

What we have missed

- In "classic" data base studies:
    - DB optimization and tuning
        - Table design
        - Normal forms
    - Security
        - Access, privacy, views
    - Even more DB models
    - Distributed System Aspects
    - Applications Programming

What we have missed

- In studies of information search:
  - LOTS! Including:
    - All but small sample of search techniques
      - Machine learning techniques
    - Representation and search of non-text media
    - Semantic web
      - W3C Initiative
      - For specification beyond XML
      - Goals
        - » Allow automated understanding and use of info.
        - » Specify wide variety of information types
        - » Provide common frameworks
        - » Allows sophisticated functionality

What we have missed

- Data mining
  - Discovery of information from data
  - Some of techniques
    - Clustering
    - Pattern analysis
    - Sampling
  - Wide variety of data types
    - Text
    - Biological
    - Images
    - …
  - COS 424: Interacting with Data
    - Prof.s Blei and Schapire

Where Info Management going?

- Energetic research and development
  - New models and functionality
  - Search on collections of non-text objects
  - Sophisticated Web search
    - Example: Semantic Web
  - DB as partner in much larger endeavors
  - Data mining