

Database Query Languages: The Evolution of SQL (Ramakrishnan)

- Aho and Ullman, "The Universality of Data Retrieval Languages", POPL 1979. (1 lec)
- Datalog and Recursive Query Processing (1-2 lecs)
- Beyond Sets: Multisets, Sequences and Streams (1 lec)
- OLAP: The CUBE Operator; materialized views; data warehousing (1 lec)
- Data Mining: Clustering/Classification/Frequent Itemsets; Scalability; Exploratory Mining; SQL support (2 lecs)

Databases and the Web (Ramakrishnan)

- Asynchronous Replication: Queries; Concurrency Control (1-2 lecs)
- Parallel and Distributed Databases (1-2 lecs)
- Web-Scale Infrastructure (1-2 lecs)
- Querying Text: Inverted Indexes; Vector Space Model; Ranking Web Search Results (1 lec)
- XML and XQuery (1 lec)
- Structure and Information Extraction (1 lec)

Data Reduction & Approximate Query Answering (Garofalakis)

- Acharya et al. "Join synopses for approximate query answering", ACM SIGMOD'99
- Poosala et al. "Improved Histograms for Selectivity Estimation of Range Predicates", ACM SIGMOD'96
- Poosala and Ioannidis "Selectivity Estimation Without the Attribute Value Independence Assumption", VLDB'97
- Vitter and Wang "Approximate Computation of Multidimensional Aggregates of Sparse Data Using Wavelets", ACM SIGMOD'99
- Garofalakis and Kumar "Deterministic wavelet thresholding for maximum error metrics", ACM PODS'04
- Polyzotis, Garofalakis "XSketch Synopses for XML databases", ACM SIGMOD'02
- Polyzotis et al. "Approximate XML query answers", ACM SIGMOD'04

- Approximate query processing tutorial in VLDB'01 -- slides at:
<http://www2.berkeley.intel-research.net/~minos/Talks/vldb01-tutorial.ppt>

Data Stream Management (Garofalakis)

- Alon, Matias, Szegedy "The space complexity of approximating the frequency moments", ACM STOC'96
- Alon, Gibbons, Matias, Szegedy "Tracking Join and Self-Join Sizes in Limited Storage", ACM PODS'99
- Gibbons "Distinct Sampling for Highly-Accurate Answers to Distinct Values Queries and Event Reports" VLDB'01
- Garofalakis et al. "Processing Set Expressions over Continuous Update Streams", ACM SIGMOD'03
- Cormode, Garofalakis "Distributed approximate query tracking", VLDB'05

- Selected (preliminary) chapters from an upcoming Springer-Verlag volume on Data-Stream Processing.

- Data stream management tutorial in VLDB'02 -- slides at:

<http://www2.berkeley.intel-research.net/~minos/Talks/streams-tutorial02.ppt>

- Distributed data stream management tutorial in VLDB'06 -- slides at:

<http://www2.berkeley.intel-research.net/~minos/Talks/vldb06tut-slides.ppt>

Probabilistic Data Management (Garofalakis)

- Dalvi, Suciu "Efficient Query Evaluation on Probabilistic Databases", VLDB'04

- Benjelloun et al. "ULDBs: Databases with uncertainty and lineage", VLDB'06

- Das Sarma et al. "Working models for uncertain data", IEEE ICDE'06

- Sen, Deshpande "Efficient Query Evaluation on Probabilistic Databases", IEEE ICDE'07

- Burdick et al. "OLAP Over Uncertain and Imprecise Data", VLDB'05

- Getoor et al. "Probabilistic Models of Relational Structure"

- Probabilistic databases tutorial in SIGMOD'05 -- slides at:

<http://www.cs.washington.edu/homes/suciu/tutorial-sigmod2005.ppt>