Candidacy Exam in the Concentration of Data Management

Department of Computer Science, UC Irvine

Affiliated Faculty

- Michael Carey
- Chen Li
- Sharad Mehrotra

Requirements

- The student is expected to have taken UCI courses CS222 and CS223 (or equivalent courses elsewhere) before attempting the exam.
- The exam format will be a hybrid of written and oral testing in order to accommodate students who do better with one or the other of the two styles.
- The student is required to read and understand all of the papers on the reading list (see below). The 2-hour written portion of the exam will be first and will contain a set of questions pertaining to the material in these papers as well as the topics and material covered in CS222 and CS223.
- One week before the exam, the committee will assign a paper (not drawn from this list) to the student. The oral portion of the exam will be based on a 20 minute presentation of this paper prepared by the student. The student’s presentation will demonstrate his or her ability to read, understand, effectively summarize, and think critically about new research-related material. The exam may last up to 60 minutes including committee Q&A.

Reading List

Data Models


Historical Systems Projects

• Starburst Mid-Flight: As the Dust Clears; Haas et al., IEEE TKDE, 2(1), March 1990, pp. 143-160.

Indexing

• Volker and Gunther, Multidimensional access methods, ACM Computing Surveys, 1997.
• H. V. Jagadish: Linear Clustering of Objects with Multiple Attributes. SIGMOD Conference 1990: 332-342.

Query Processing


Transaction Processing

• C. Mohan et al., ARIES: A Transaction Recovery Method Supporting Fine-
  Granularity Locking and Partial Rollbacks Using Write-Ahead Logging;
  251-285.
• D. Lomet: MLR: A Recovery Method for Multi-level Systems. SIGMOD
  Acta Informatica 9, 1--21.
• D. Lomet, Key Range Locking Strategies for Improved Concurrency, VLDB
  1993.
• J. Gray et al., The Dangers of Replication and a Solution; Readings in Database

Data Integration


Semistructured and XML Data

• S. Pal, I. Cseri, G. Schaller, O. Seeliger, L. Giakoumakis, V. Zolotov: Indexing
  XML Data Stored in a Relational Database. VLDB 2004.

Data Warehousing and Mining

• R. Agrawal, R. Srikant. Fast Algorithms for Mining Association Rules in Large
• S. Chaudhuri, U. Dayal. An Overview of Data Warehousing and OLAP
• J. Gray, S. Chaudhuri, A. Bosworth, A. Layman, D. Reichart, M. Venkatrao, F.
  Pellow, H. Pirahesh. Data Cube: A Relational Aggregation Operator
  Generalizing Group-By, Cross-Tab, and Sub-Totals. Data Mining and
  Knowledge Discovery 1997.

Benchmarks

• D. DeWitt: The Wisconsin Benchmark: Past, Present, and Future. The