Ashish Gupta

agupta @ cs.stanford.edu

Education

Ph.D. in Computer Science, Stanford University (completed in 1994).
Advisors: Prof. Jeffrey D. Ullman and Prof. Jennifer Widom

Bachelor of Technology in Computer Science and Engineering. Indian Institute of Technology (IIT), Kanpur, India. (GPA 10.0/10.0)

Work Experience

8/98 – now *Director of Engineering, Amazon.com*

6/96 – now *Founder, VP Engineering, and Member Board of Directors Junglee Corp.*
Junglee developed technology to rapidly integrate multiple sources of data that could use different representations, storage mechanisms, access methods. For example, integrating different web sites that sell computer equipment. Junglee pioneered the web services business model using comparison shopping and employment classifieds as the vertical services it offered its customers like the Washington Post, Yahoo!, Compaq, LA Times and multiple other leading media companies.

Junglee was acquired in 1998 by Amazon.com. As a founder and spokesperson of Junglee I was involved in all aspects of growing the company from 5 to 70 persons - especially fund raising, setting business direction, recruiting.

8/95 – 6/96 *Senior Member of Technical Staff, Oracle Corporation.*
As a member of Distributed Database Group I implemented some of my past research into the replication server for Oracle, the leading replication product in the industry. I am worked with the decision support and warehousing groups.

8/94 – 8/95 *Research Staff Member, IBM Almaden Research Center.*
I was involved in the design and implementation of a Heterogeneous Multi-media Distributed Database System (Garlic). C++). I developed distributed query fragmentation and optimization algorithms for generating query execution plans that utilise the possibly different query capabilities of the sources. The algorithms are being implemented in the TSIMMIS (Stanford University) and Garlic systems (Publications 12, 2).

As a member of the Quest Datamining project I studied how to efficiently implement a multi-dimensional database system on a conventional RDBMS. We studied new indexing techniques and the relative performance of different aggregation strategies when computing multiple aggregates over a single relation (Publications 7).

1988 – 8/94 *Research Assistant, Stanford University.*

Co-designed and implemented a Distributed Constraint Management System in collaboration with the Center for Integrated Facilities Engineering (CIFE) at Stanford. The implementation was in C and C++ (using embedded SQL in Oracle-7) (Publications 5, 21).

Developed algorithms to efficiently verify integrity constraints and maintain materialized views. The algorithms reduce the amount of data accesed for constraint verification or view maintenance. Some of the algorithms are implemented in the CIFE project (Publications 5, 13, 14, 15, 25). We also developed algorithms for using materialized views to answer queries.

Developed optimization techniques for aggregate SQL queries. The optimization and view maintenance technologies both are very useful in building data warehousing and decision support systems and have elicited interest from several DB companies (Publication 10).

Summer, 1991 *Intern, IBM Almaden Research Center, San Jose*.
Developed efficient algorithms for 2-way parallel joins and compared their performance with existing parallel join algorithms over synthetic and benchmark data (Publications 6, 20).

Summer, 1989 *Intern, IBM Hawthorne, NY*.
Worked on the specification of reactive systems using object charts and implemented (in Prolog and C) a graphical interface for representing object charts.

Academic Services

- Teaching assistant for numerous courses at Stanford.
- Program Committee for VLDB98, organized workshops (SIGMOD, NSF), referee for conferences and journals.

Publications

Book

1. Materialized Views. Co-edited a collection of papers on the theory, implementation and applications of Materialized Views. Published by MIT Press (with I. S. Mumick).

Journal Articles

2. Capabilities-Based Query Rewriting in Mediator Systems. In *Distributed and Parallel Databases Journal Vol 6, No 1, Jan 1998*. Also appears in *PDIS* 1996 (with Y. Papakonstantinou, L. Haas)
3. Using Partial Information to Update Materialized Views. In *Information Systems* Vol 20, No. 8, 1995 (with J. Blakeley).
4. Maintenance of Materialized Views: Problems, Techniques, and Applications. in *IEEE Data Engineering Bulletin, Special Issue on Materialized Views and Data Warehousing*, Vol 18, No. 2, June 1995 (with I. S. Mumick).
5. Constraint Management on Distributed Design Configurations. In *Journal of Engineering with Computers*. Also an invited paper in *IEEE Data Engineering Bulletin, Special Issue on Database Constraint Management*, Vol 17, No. 2, 1994 (with S. Tiwari).
6. Algorithms for Handling Skew in Parallel Task Scheduling. In *Journal of Parallel and Distributed Computing*, 1992, Vol. 16. (with A. Swami and H. Young).

Conference Articles

7. Modeling Multidimensional Databases. In *ICDE* 1997. (with Rakesh Agrawal and Sunita Sarawagi)

8. On the computation of multidimensional aggregates. In *VLDB*1996. (with Sameet Agarwal, Rakesh Agrawal, Prasad Deshpande, Jeffrey F. Naughton, Raghu Ramakrishnan, and Sunita Sarawagi)
9. Making Views self-Maintainable for Data Warehousing. In *PDIS* 1996. (with Dallan Quass, I. S. Mumick, J. Widom).
10. Aggregate-Query Processing in Data Warehousing Environments. In *VLDB* 1995 (with V. Harinarayan and D. Quass).
11. Adapting Materialized Views after Redefinitions. In *SIGMOD* 1995 (with I. S. Mumick and K. Ross).
12. A Query Translation Scheme for Rapid Implementation of Wrappers. In *Deductive and Object-Oriented Databases 1995*. (with Y. Papakonstantinou, H. Garcia-Molina, and J. Ullman)
13. Constraint Checking with Partial Information. In PODS 1994 (with Y. Sagiv, J. D.|Ullman, and J. Widom).
14. Local Checking of Global Integrity Constraints. In *SIGMOD 1993* (with J. Widom).
15. Maintaining Views Incrementally. In *SIGMOD 1993* (with I. S. Mumick and V. S. Subrahmanian).
16. Magic-sets Transformation in Non-recursive Systems. In *PODS* 1992 (with I. S. Mumick).
17. Data Integration using Self-Maintainable Views. In *EDBT* 1996 (with I. S. Mumick and H. V. Jagadish).
18. Optimization Using Tuple Subsumption. In *ICDT* 1995 (with V. Harinarayan).
19. Validating Constraints with Partial Information: Research Overview. Invited paper in *DAISD* 1994 (with Y. Sagiv, J. D. Ullman, and J. Widom).
20. Centralized, Decentralized and Pipelined Algorithms to Parallelize Joins. In *Proceedings of the 1993 International Conference on Parallel and Distributed Systems* (with R. Morris, A. Swami, and H. Young).
21. Distributed Constraint Management for Collaborative Engineering Databases. In *Proceedings of the Second International Conference on Information and Knowledge Management (CIKM)* 1993 (with S. Tiwari).
22. DScheduler: A Deductive Database for Scheduling Building Construction Tasks. In 5^{th} *International Conference on Computing in Civil and Building Engineering (ASCE)* 1993 (with S. Tiwari).

Others

23. Efficient and complete tests for database integrity constraint checking. In *Lecture Notes in Computer Science*, Number 874 (with Y. Sagiv, J. D. Ullman, and J. Widom).
24. Integrated Data Exchange and Concurrent Design for Engineered Facilities. In *Proceedings of NSF Scientific Database Projects 1991-1993, AAAS Workshop on Data Management for the Scientest and Engineer* 1993 (with H. C. Howard, J. D. Ullman, K. Law, A. Keller, S. Tiwari, and K. Krishnamurthy).
25. Generalizing Conjunctive Query Containment for View Maintenance and Integrity Constraint Checking. In *Workshop on Deductive Databases, JICSLP* 1992 (with J. D. Ullman). Institute of

Patents

- US5963949: Method for data gathering around forms and search barriers (with Anand Rajaraman, Peter Norvig).
- US5832475: Database system and method employing data cube operator for group-by operations (with Rakesh Agrawal and Sunita Sarawagi)
- Encoded bit-vector Indexes, IBM Santa Teresa Laboratories (pending).

Honors

- President of India Gold Medal, 1988, IIT Kanpur (Awarded to the valedictorian of the graduating class).
- National Talent Scholar, 1982-1988 (Awarded to approx. 250 students on the basis of a national examination).