Research in Data Structures

Mihai Pătraşcu

Lower Bounds in the Cell-Probe Model. My most influential research contribution has been a new lower bound technique for dynamic data structures. In most areas of complexity theory, lower bounds in general models of computation are renowned to be very difficult to prove, and the current barriers in the magnitude of the lower bounds are severe. For dynamic data structures, the best known lower bound was $\Omega(\lg n / \lg \lg n)$, by the chronogram method of [Fredman and Saks, 1989]. Surpassing this barrier has repeatedly been recognized as a fundamental problem throughout one and a half decades. In two papers, one from SODA'04 and another one from STOC'04 (both published with my advisor, Erik Demaine), I developed a new technique which could prove a lower bound of $\Omega(\lg n)$. A referee commented: "this is an outstanding article, breaking through a long existing barrier in natural cell-probe lower bounds". I applied this technique to obtain tight or almost tight bounds for two very important and intensively-studied problems: dynamic connectivity and maintaining partial sums. My lower bound technique also has the advantage of being conceptually simpler and cleaner that the chronogram method; most previous results find simpler and more intuitive proofs in this new framework. At present, I am continuing my work on better lower bound techniques, and hope to break the $\Omega(\lg n)$ barrier.

Competitiveness of Binary Search Tree Algorithms. Binary search trees (BSTs) are the prototypical data structure for comparison-based searching. The BST model has been studied extensively, and a myriad of BST algorithms have been proposed. This raises the question whether there exists a "best binary search tree", i.e. one which is (almost) as good as any other tree on *any* sequence of accesses. Formally, we say that an algorithm is c-competitive if the cost of any sequence of accesses is bounded by c times the best possible cost for that particular sequence. Research in this direction has been fueled by the famous splay conjecture of [Sleator and Tarjan 1983], which states that splay trees are $O(1)$-competitive. Competitiveness in the BST model is now recognized as one of the most important open problems in the theory of data structures. Despite this, no progress had been made in two decades, and the best known competitive ratio remained the trivial $O(\lg n)$. Together with Erik Demaine, Dion Harmon and John Iacono, we described an $O(\lg \lg n)$-competitive BST algorithm, which is an exponential improvement over the old bound. The paper was accepted at FOCS 2004; the referee called this "a major progress on an old problem" and "an elegant result". This came as a result of over half a year of team effort; at present, we are enthusiastically continuing research in this direction.

Integer Search Problems. Beginning with the 1990s, integer search problems have been one of the most active areas of research in data structures. A central problem in this field is predecessor search. For w-bit integers, this problem can be solved in $O(\lg w)$ time per operation using the classic van Emde Boas data structure. This data structure is based on an ingenious recursion on the height of a trie representing the integers. This recursion idea has been extremely influential and has found numerous applications in distant areas of data structures. I considered another equally natural and fundamental integer search problem, namely dynamic range reporting in one dimension. The problem can easily be solved using predecessor search. However, I found a solution with update time $O(\lg w)$, and query time $O(\lg \lg w)$, which is an exponential improvement over the van Emde Boas structure. This solution is based on a fundamentally new recursion idea, going significantly beyond the van Emde Boas approach. Together with Christian Mortensen and Prof. Rasmus Pagh (both from IT Univ. of Copenhagen), we found a way to use my idea to improve the bounds for approximate range counting in one dimension. We also found a way to reduce the space requirement of my data structure to linear; this required developing an optimal solution to dynamic Bloomier filters (a previously-stated open problem). All these results are described in a joint paper which we have submitted for publication.

Recently, I also considered the static predecessor-search problem. Together with Mikkel Thorup (AT&T Research), we extended the well-known bounds of [Beame and Fich 1999] to obtain optimal tradeoffs between space and query time. We will continue to work on the dynamic case and/or publish our current result for the static case.

Mihai Pătraşcu
Curriculum Vitae

540 Memorial Dr, #907
Cambridge, MA 02139
Tel: (617) 577-5576

mip@mit.edu
http://www.mit.edu/~mip/

Education

2002–2006 Massachusetts Institute of Technology. Anticipated B.S. degree in Mathematics with Computer Science. My academic adviser is Michael Sipser. Current GPA: 5.0/5.0

Graduate coursework: Advanced Data Structures, Advanced Algorithms, Geometric Computing, Sublinear Time Algorithms, Theory of Computation, Pseudorandomness, The Probabilistic Method, Theory of Parallel Hardware, Computer System Architecture, Database Systems

2001–2002 University of Craiova, Romania, Dept. of Automatic Control, Computers and Electronics. University exceptional award for academic achievement. GPA: 10/10

1997–2001 (high school) C.N. Carol I, Craiova, Romania. School merit scholarship for outstanding academic achievement, 1998–2001. GPA: 9.6/10; Baccalaureate: 9.55/10

Positions

02/2003–present Undergraduate Research Opportunity in Theory of Computation Group, Prof. Erik Demaine, MIT CSAIL

09/2002–12/2002 Undergraduate Research Opportunity in Program Compilation and Verification Group, MIT LCS (logic, theorem provers, functional programming)

09/2001–03/2002 Part-time software developer at Softwin Romania (signal processing, AI with neural networks, analyzing a mechanical system)

07/2001–08/2001 Internship at SyncRo Soft, Romania (signal processing, voice recognition)

05/1999–06/1999 Internship at Idaco Systems, Romania (assembly language, micro-controllers, interfacing with mechanical and electrical components in real-time)

Research Publications

10. **On Dynamic Range Reporting in One Dimension**, with Christian Worm Mortensen and Rasmus Pagh. Submitted for publication.
9. **Logarithmic Lower Bounds in the Cell-Probe Model**, with Erik Demaine. Journal version of 6 and 3, invited to special issue of SIAM Journal on Computing (*SICOMP*) dedicated to STOC'04 (to appear).
8. **Subquadratic Algorithms for 3SUM**, with Erik Demaine. Submitted for publication.
7. **Dynamic Optimality – Almost**, with Erik Demaine, Dion Harmon and John Iacono. Proc. 45th IEEE Symposium on Foundations of Computer Science (*FOCS'04*), to appear.
6. **Lower Bounds for Dynamic Connectivity**, with Erik Demaine. Proc. 36th ACM Symposium on Theory of Computing (*STOC'04*), pp 546-553.
5. **Computing Order Statistics in the Farey Sequence**, with Corina Pătraşcu. Proc. 6th Algorithmic Number Theory Symposium (*ANTS 2004*) – LNCS vol. 3076, pp 358-366.
4. **Finding a Divisible Pair and a Good Wooden Fence**, with Stelian Ciurea, Erik Demaine and Corina Pătraşcu. Proc. 3rd International Conference on Fun with Algorithms (*FUN 2004*), pp 206-219. A poster on the divisible pair problem was displayed at the 6th Algorithmic Number Theory Symposium (ANTS 2004). An invited extended abstract of the divisible-pair material will appear in the ACM *SIGSAM Bulletin*.

3. **Tight Bounds for the Partial-Sums Problem**, with Erik Demaine. Proc. 15th ACM-SIAM Symposium on Discrete Algorithms (*SODA'04*), pp 20-29. Invited to special issue of ACM Transactions on Algorithms. Declined, and merged with 6 for final journal version 9.
2. **Interpolation Search for Non-Independent Data**, with Erik Demaine and Thouis Jones. Proc. 15th ACM-SIAM Symposium on Discrete Algorithms (*SODA'04*), pp 522-523.
1. **On Two Problems from the National Olympiad in Informatics 2002** (in Romanian). Gazeta Informatică (*GInfo*), February 2002, pp 13-14.

Coauthors

Stelian Ciurea (Univ. Lucian Blaga, Sibiu, Romania), Erik Demaine (MIT), Dion Harmon (MIT), John Iacono (Polytechnic University, NY), Thouis Jones (MIT), Christian Worm Mortensen (IT Univ. of Copenhagen), Rasmus Pagh (IT Univ. of Copenhagen), Corina Pătraşcu (Harvard).

Research Talks

09/2004 "Logarithmic Lower Bounds in the Cell-Probe Model", invited talk at The IT University of Copenhagen.
06/2004 "Lower Bounds for Dynamic Connectivity", conference talk at STOC'04.
05/2004 "Finding a Divisible Pair and a Good Wooden Fence", conference talk at FUN'04.
01/2004 "Tight Bounds for the Partial-Sums Problem", conference talk at SODA'04.

Awards

- first prize for age group, Romanian National Olympiad in Informatics, 1993–2001
- President of Romania's Award for Excellence, 2000 and 2001
- gold medal, International Olympiad in Informatics, Tampere, Finland, 2001
- gold medal, International Olympiad in Informatics, Bei Jing, China, 2000
- silver medal, Central European Olympiad in Informatics, Cluj, Romania, 2000
- silver medal, Balkan Olympiad in Informatics, Ohrid, Macedonia, 2000
- silver medal, International Olympiad in Informatics, Antalya, Turkey, 1999
- gold medal, Central European Olympiad in Informatics, Brno, Czech Rep., 1999
- first prize in informatics, Tuymaada Olympiad, Yakutsk, Russia, 1998
- first prize in individual and team contests, Applied Mathematics Competition, Chisinau, Rep. of Moldavia, 1996
- Romanian National Olympiad in Physics, first (1996) and second (1997) prize
- various prizes in regional Romanian competitions (Computer Science and Physics)

Contributions to Olympiads

After I graduated from high school, I got involved in organizing the high-school olympiads in informatics (as volunteer work). My involvement was as follows:

- Scientific Committee, Balkan Olympiad in Informatics, 2003. Author of 3 contest problems.
- Scientific Committee, Romanian National Olympiad in Informatics, 2002–2004. Author of 9 problems used in the senior division and olympic team selection contests.
- Taught 2 short courses during the national team preparation in 2003.
- Developed the grading software used in recent National Olympiads and the Balkan Olympiad 2003.