**ESPRIT BR Project RAND-REC**
**( EC-US Exploratory Collaborative Activity –**
**EC-US 030)**

# Annual Progress Report

# July 1, 1995 – June 30, 1996

## Contents

## 1 RAND-REC Research Sites

The research sites are:

- University of Bonn,
- University of Edinburgh,
- University of Lund,
- University of Oxford,
- University of Paris-Sud
- International Computer Science Institute, Berkeley

  and
- University of California, Berkeley

## 2 Overview of Research Activities

The research within the project RAND-REC has concentrated on the following research areas (see Section 3, Research Papers):

(1) Design of Efficient Randomized and Approximative Algorithms

(2) Efficient Parallel Algorithms

(3) VC Dimension of Sigmoidal and Pfaffian Neural Networks and Volume Approximation

(4) Derandomizing Algorithms and Probabilistic Methods

(5) Deterministic and Randomized PET (Priority Encoding Transmission) Systems

## 3 Research Papers (RAND-REC)

1. Andres Albanese, Johannes Blömer, Jeff Edmonds, Michael Luby, Madhu Sudan,
   *Priority Encoding Transmission,*
   $35^{th}$ **FOCS**, 1994, accepted March 1996 to special issue devoted to coding theory of *IEEE Transactions on Information Theory.*

2. Noga Alon, Michael Luby,
   *A Linear Time Erasure-Resilient Code With Nearly Optimal Recovery,*
   accepted March 1996 to special issue devoted to coding theory of *IEEE Transactions on Information Theory.*

3. A. Andersson, co-authors P.B. Miltersen, S, Riis and M. Thorup,
   em Static Dictionaries on /AC /RAMs: Query time $\Theta(\sqrt{\log n/\log\log n})$ is necessary and sufficient . In Proc. 37th Annual IEEE Symposium FOCS, 1994.

4. S. Arora, D. Karger, M. Karpinski,
   *Polynomial Time Approximation Schemes for Dense Instances of NP-Hard Problems,*
   submitted to J. Comput. Syst. Sciences, 1994 (preliminary version appeared in Proc. 27th ACM STOC (1995), pp. 284-293).

5. E. Bampis, M. El Haddad, Y. Manoussakis and M. Santha,
   *A parallel reduction of Hamiltonian cycle to Hamiltonian path in tournaments,*
   Journal of Algorithms (1995), preliminary version in LNCS 694.

6. P. Berman, M. Karpinski, L. Larmore, W. Plandowski, W. Rytter,
   *The Complexity of Two-Dimensional Compressed Pattern Matching,*
   submitted to the 8th ACM-SIAM SODA (1997).

7. S. Boucheron and D. Gardy,
*An Urn Model from Learning Theory,*
Random Structures and Algorithms (1996), accepted.

8. Gilles Brassard, C. Crepeau and M. Santha,
*Oblivious Transfers and Intersecting Codes,*
IEEE Transactions on Information Theory (1996), to appear in November.

9. E. Dahlhaus, M. Karpinski,
*On the Parallel Complextiy of Matching on Chordal and Strongly Chordal Graphs,*
Proc. 18th Australian Computer Science Conference ACSC (1995), pp. 108–112, submitted to Discrete Applied Mathematics, 1996.

10. C. Dorgerloh, J. Lüssem,
*A Simple Linear–Time Algorithm to Find the Contour in a Coloured Triangular Graph,*
Research Report No. 85146-CS, University of Bonn, (1996)

11. C. Dorgerloh,
*A Fast Randomized Parallel Algorithm for Finding Simpled Cycles in Planar Graphs,*
Research Report No. 85150-CS, University of Bonn, (1996)

12. C. Dürr, H. Lê Thanh and M. Santha,
*A decision procedure for well-formed quantum linear cellular automata,*
LNCS **1046** (1996), pp. 281-292, Springer.

13. W. Fernandez de la Vega,
*MAX-CUT has a Randomised Approximation Scheme in Dense Graphs,*
Random Structure and Algorithms **3** (1996), pp. 179-187.

14. A. Frieze and M. Jerrum,
*Improved approximation algorithms for MAX $k$-CUT and MAX BISECTION,*
Proceedings of the fourth Integer Programming and Combinatorial Optimization Conference (IPCO4), Springer-Verlag Lecture Notes in Computer Science **920** (1995), pp. 1–13.

15. A. Frieze, M. Jerrum and R. Kannan,
*Learning linear transformations,*
Proceedings of the 37th IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, 1996. [To appear soon.]

16. Oscar Garrido and Andrzej Lingas, co-authors S. Jarominek, W. Rytter,
*A simple randomized parallel algorithm for maximal $f$-matchings.* In Information Processing Letters **57** (1996) 83-87.

17. Oscar Garrido and Andrzej Lingas, co-author P. Kelsen,
*A simple NC-algorithm for a maximal independent set in a hypergraph*

*of poly-log arboricity.*
In Information Processing Letters **58** (1996) 55-58.

18. M. Goldmann, M. Karpinski,
*Simulating Threshold Circuits by Majority Circuits,*
to appear in SIAM J. Computing, 1996

19. Y. Grandvalet, S. Canu and S. Boucheron,
*Noise Injection: Theoretical Prospects,*
Neural Computation (1996), accepted.

20. D. Grigoriev, M. Karpinski, N. Vorobjov,
*Improved Lower Bound on Testing Membership to a Polyhedron by Algebraic Decision Trees,*
Proc. 36th IEEE FOCS (1995), pp. 258–265.

21. D. Grigoriev, M. Karpinski, F. Meyer auf der Heide, R. Smolensky,
*A Lower Bound for Randomized Algebraic Decision Trees,*
Proc. 28th ACM STOC (1996), pp. 612–619; submitted to J. Computational Complexity, 1996

22. D. Grigoriev, M. Karpinski, R. Smolensky,
*Randomization and the Computational Power of Analytic and Algebraic Decision Trees,*
submitted to J. Computational Complexity, 1996

23. D. Grigoriev, M. Karpinski, A. Yao,
*An Exponential Lower Bound on the Size of Algebraic Decision Trees for MAX,*
submitted to J. Computational Complexity, 1996

24. D. Grigoriev, M. Karpinski, A. Odlyzko,
*Short Proofs for Neudivisibility of Sparse polynomials under the Extended Riemann Hypothesis,*
to appear in Fundamenta Informaticae, 1996

25. D. Grigoriev, M. Karpinski,
*Computing Additive Complexity of Algebraic Circuits with Root Extracting,*
to appear in SIAM J. Computing, 1996.

26. M. Karpinski, L. Larmore, W. Rytter,
*Correctness of Constructing Optimal Alphabetic Trees Revisted,*
Research Report No. 85134-CS, University of Bonn, 1995

27. M. Karpinski, W. Rytter,
*On a Sublinear Time Parallel Construction of Optional Binary Search Trees,*
to appear in IPL, 1996.

28. M. Karpinski, A. Macintyre,
*Polynomial Bounds for VC Dimension of Sigmoidal and General Pfaffian Neural Networks,*
to appear in Special Volume on Neural Networks, J. Comput. Syst. Sciences, 1996.

29. M. Karpinski, A. Macintyre,
*Approximating the Volume of General Pfaffian Bodies,*
Research Report No. 85145-CS, Universtiy of Bonn, 1996

30. M. Karpinski, A. Zelikovsky,
*New Approximation Algorithms for the Steiner Tree Problems,*
to appear in J. of Combinatorial Optimization.

31. M. Karpinski, L. Larmore, W. Rytter,
*Sequential and Parallel Subquadratic Work Algorithms for Constructing Approximately Optimal Binary Search Trees,*
Proc. 28 ACM STOC (1996).

32. M. Karpinski, R. Verbeek,
*On Randomized versus Deterministic Computation,*
Theoretical Computer Science **154** (1996), pp. 23–39.

33. M. Karpinski, J. von zur Gathen, I. Shparlinski,
*Counting Curves and Their Projections,*
to appear in J. Computational Complexity.

34. M. Karpinski, W. Rytter, A. Shinohara,
*Pattern Matching for Strings with Short Descriptions,*
Proc. CPM '95.

35. M. Karpinski, L. Gasieniec, W. Plandowski, W. Rytter,
*Randomized Efficient Algorithms for Compressed Strings: the Finger–Print Approach,*
Proc. CPM '96.

36. M. Karpinski, I. Shparlinski,
*On Some Approximation Problems Concerning Sparse Polynomials over Finite Fields,*
Theoretical Computer Science **157** (1996), pp. 259–266

37. M. Karpinski,
*Lower Time Bounds for Randomized Computation,*
Proc. 22th ICALP'95, pp. 183–195.

38. M. Karpinski, F. Ablayev,
*On the Power of Randomized Branching Programs,*
Proc. 3rd ICALP'96, pp. 348–356.

39. Andrzej Lingas, co-author P. Berman,
*A Nearly Optimal Parallel Algorithm for the Voronoi Diagram of a Convex*

*Polygon*. In Proc. Scandinavian Workshop on Algorithm Theory, July 1994, Lecture Notes in Computer Science 824, Springer Verlag, pp. 73-82. Accepted for publication in Theoretical Computer Science in 1996.

40. Michael Luby and Avi Wigderson,
*Pairwise Independence and Derandomization,*
UC Berkeley Tech Report UCB/CSD-95-880,
ICSI Tech Report No. TR-95-035, July, 1995.

41. Michael Luby,
*Pseudorandomness and Cryptographic Applications,*
Princeton Computer Science Notes,
Editors David R. Hanson and Robert E. Tarjan,
Princeton University Press, January 1996.

42. Dominic Welsh,
*Randomised approximation of the number of bases: Contemporary Mathematics,*
(American Mathematical Society) to be published (1996), (with L. Chávez-Lomeli).

43. Dominic Welsh,
*Approximation algorithms: Surveys in Combinatorics,*
(ed. R.A. Bailey) London Mathematical Society Lecture Notes,
Cambridge University Press (1997) (to be published).

44. Dominic Welsh,
*The win polytope of a graph,*
(with J.E. Bartels and J. Mount) (to appear) (1996).

# 4 Conferences

C. Dorgerloh (Bonn)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

Chr. Günzel (Bonn)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

D. Wiggerich (Bonn)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

J. Wirtgen (Bonn)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

M. Karpinski (Bonn)

- 3rd ESA (1995), Corfu, Sept. 25-27, 1995.

- 36th IEEE FOCS (1995), Milwaukee, Oct. 23-25, 1995.

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995.

- Workshop on "Computation and Complexity", Dagstuhl, Nov. 6-10, 1995.

- Workshop on "Random Methods in Convex Geometry", MSRI, Berkeley, March 11-15, 1996.

- 28th ACM STOC (1996), Philadelphia, May 22-24, 1996.

C. Bazgan (Paris-Orsay)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

S. Boucheron (Paris-Orsay)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

C. Dürr (Paris-Orsay)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

M. Santha (Paris-Orsay)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

J. Stern (Paris-Orsay)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

F. De La Vega (Paris-Orsay)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

R. Bubley (Leeds)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

A. Lingas (Lund)

- "Randomness and Computation" RAC'95, RAND-REC-Workshop, Berkeley, Dec. 17-21, 1995

- 36th IEEE FOCS (1995), Milwaukee, Oct. 23-25, 1995.

- 28th ACM STOC (1996), Philadelphia, May 22-24, 1996.

D. Welsh (Oxford)

- Workshop on "Random Methods in Convex Geometry", MSRI, Berkeley, March 11-15, 1996.

## 5 Randomness and Computation Workshop, Berkeley, Dec. 17 - 21 1995 – Program and Abstracts

**SUNDAY, 12/17**

**Evening:**

| | | |
|---|---|---|
| 7:00 | - 11p.m. | Reception, SODA Hall Lounge, UC Berkeley |

**MONDAY, 12/18**

**Morning Session:**

| | | |
|---|---|---|
| 9:00 | - 9:30 | A. Lingas: Maximum tree packing is in RNC |
| 9:30 | - 10:00 | A. Fundia: Algorithmic matchings and coverings in nearly disjoint hypergraphs |
| 10:00 | - 10:30 | M. Henzinger: Randomization in dynamic graph algorithms |
| 10:30 | - 11:00 | ==Break== |
| 11:00 | - 11:35 | C. Scheideler [Cypher, Meyer auf der Heide, Vöcking]: Universal Algorithms for Store-and-Forward and Wormhole Routing |
| 11:35 | - 12:05 | C. Dürr [Thanh, Santha]: A decision procedure for well-formed linear quantum cellular automata |

**Afternoon Session:**

| | | |
|---|---|---|
| 3:00 | - 4:00 | A. Broder [Frieze, Suen, Upfal]: Surveying the Problem of Cosntructing Paths in Random Graphs |
| 4:00 | - 4:30 | A. Wigderson: Arithmetic Circuits |
| 4:30 | - 5:00 | ==Break== |
| 5:00 | - 5:30 | J. Kleinberg: Approximations for the Disjoint Paths Problem |
| 5:30 | - 6:00 | Y. Bartal [Leonardi, Fiat]: Lower Bounds to On-line Graph Problems with Applications to On-line Circuit and Optical Routing |
| 6:00 | - 6:30 | D. Karger [Benczur]: Nonuniform Sampling in Cut and Flow Problems |

## TUESDAY, 12/19

**Morning Session [learning]:**

| | | |
|---|---|---|
| 9:00 | - 9:30 | R. Kannan [Jerrum, Frieze]: Learning Product Distributions |
| 9:30 | - 9:55 | A. Blum [Frieze, Kannan, Vempala]: Learning linear threshold functions with noise |
| 9:55 | - 10:35 | M. Rabin [Micali]: An Efficient Zero-Knowledge Method for Answering 'Is He In Or Out?' Questions |
| 10:35 | - 11:00 | ==Break== |
| 11:00 | - 11:30 | D. Ron [Freund]: Learning to Model Sequences Generated by Switching Distributions |
| 11:30 | - 12:00 | R. Motwani: Randomized Robot Path Planning |

**Afternoon Session:**

| | | |
|---|---|---|
| 1:30 | - 2:30 | O. Goldreich [Bellare, Sudan]: Non-Approximability Results for MAX SNP – Towards Tight Results |
| 2:30 | - 3:00 | ==Break== |
| 3:00 | - 4:00 | N. Nisan: Extractors, Dispersers, and their Applications |
| 4:00 | - 4:30 | M. Ajtai: Generating Hard Instances of Lattice Problems |
| 4:30 | - 5:00 | ==Break== |
| 5:00 | - 5:30 | D. Zuckerman: Randomness-Optimal Sampling, Extractors, and Constructive Leader Election |
| 5:30 | - 6:00 | T. Rabin [Gennaro, Jarecki, Krawczyk]: Robust Threshold DSS Signatures |
| 7:00 | - 10:00 | BANQUET DINNER |

**WEDNESDAY, 12/20**

**Morning Session:**

| | | |
|---|---|---|
| 9:00 | - 9:30 | R. Rubinfeld [Goldreich, Sudan]: Learning Polynomials: The Highly Noisy Case |
| 9:30 | - 10:00 | M. Kiwi [Bellare, Coppersmith, Håstad, Sudan]: Linearity Testing |
| 10:00 | - 10:20 | F. Ergun [Ravi Kumar, Sivakumar]: Testing Multivariate Linear Functions: Overcoming the Generator Bottleneck |
| 10:20 | - 10:40 | ==Break== |
| 10:40 | - 11:10 | S. Ravi Kumar [Sivakumar]: Efficient Self-Testing of Linear Recurrences |
| 11:10 | - 11:40 | M. Naor: Evaluation may be easier than generation |
| 11:40 | - 12:10 | R. Ostrovsky [Kushilevitz, Rosen]: Characterizing Linear Size Circuits in Terms of Privacy |

**Afternoon Session:**

| | | |
|---|---|---|
| 3:00 | - 3:30 | D. Spielman: Disk Packings and Planar Separators |
| 3:30 | - 4:00 | C. Dwork [Lotspiech, M. Naor]: Digital Signets for Protection of Digital Information |
| 4:00 | - 4:20 | ==Break== |
| 4:20 | - 4:50 | A. Gal [Babai, Kollar, Ronyai, Szabo, Wigderson]: Extremal bipartite graphs and superpolynomial lower bounds for monotone span programs |
| 4:50 | - 5:20 | M.Karpinski [Grigoriev, Meyer auf der Heide, Smolensky]: Randomized $\Omega(n^2)$ Lower Bound for the Knapsack Problem (Randomized Lower Bounds for the Decision Trees Revisited) |
| 5:20 | - 5:50 | S. Arora: A New Rounding Procedure for Assignment-type Problems |

**THURSDAY, 12/21**

**Morning Session:**

| | | |
|---|---|---|
| 9:00 | - 9:30 | D. Aldous: An optimization problem on a random graph where the Metropolis algorithm is (almost) theoretically analyzable |
| 9:30 | - 10:00 | N. Kahale: A semidefinite bound for mixing rates of Markov chains |
| 10:00 | - 10:30 | S. Vempala [Kannan]: A simple random walk for counting degree sequences |
| 10:30 | - 11:00 | ==Break== |
| 11:00 | - 11:30 | R. Bubley [Dyer, Jerrum]: A new approach to polynomial time random walks for volume computation |
| 11:30 | - 12:00 | D. Randall: Sampling Domino Tilings on Regions With Free Boundary Conditions |

**Afternoon Session:**

| | | |
|---|---|---|
| 3:30 | - 4:00 | M. Bellare [Canetti, Krawczyk]: Cascading Pseudo-Random Functions– How to Key Merkle |
| 4:00 | - 4:30 | J. Stern [Fischer]: An efficient pseudo-random generator provably as secure as syndrome decoding |
| 4:30 | - 5:00 | ==Break== |
| 5:00 | - 5:30 | H. Krawczyk: Small Biased Toeplitz Hashing with Applications to Message Authentication |
| 5:30 | - 6:00 | R. Canetti [Feige, Goldreich, M. Naor]: Adaptively Secure Multiparty Computation |
| 6:00 | - 6:30 | Venkie: A Secure Pseudo-Random Generator from Theory to Practice |

# Abstracts

### Maximum tree packing is in RNC

Andrzej Lingas
Lund University

A randomized NC algorithm for determining the maximum number of node-disjoint subtrees of a tree isomorphic to a given tree is presented. The corresponding problems where topological embedding and subgraph homeomorphism are respectively substituted for subgraph isomorphism are also observed to admit randomized NC algorithms.

### Algorithmic Matchings and Coverings in Nearly Disjoint Hypergraphs

Andres Fundia

ITESM, Ciudad de México

We show a polynomial time algorithm that finds 'almost perfect' matchings and coverings in hypergraphs that for some constant $k$ are $k$-bounded (every edge contains at most k vertices), have minimum degree close enough to the maximum degree, and have codegrees (the numbers of edges containing a pair of vertices) negligible compared with the maximum degree.

The existence of such matchings and coverings was proved in [1] [1]. We apply a technique of derandomization based on the method of conditional expectations. It was first used in [2][2], and it is suitable to derandomize randomized

[1] Pippenger, N. and Spencer, J.(1989). "Asymptotic Behavior of the Chromatic Index for Hypergraphs," *Journal of Combinatorial Theory, Series A 51*, pp. 24-42.

[2] Fundia, A. (1995). "Derandomizing Chebyshev's Inequality to find Independent Sets in Uncrowded Hypergraphs," *Random Structures and Algorithms*, to appear.

algorithms whose good behavior depend on the fact that some random variables have sufficiently small variances. This technique is applied to a generalization of the existence result from [1] which is given in [3][3].

In [1] it is also proved the existence of partitions of the edges i nto 'almost perfect' matchings and coverings. We also show some progress toward an algorithmic version of this.

### Randomization in Dynamic Graph Algorithms

Monika Rauch Henzinger
Cornell University

A dynamic graph algorithm is a data structure that maintains a property of a graph during a sequence of edge insertions and deletions. The use of randomization has lead to an exponential improvement in the time per operation for various graph properties. For example, for connectivity, the previously best known running time was $O(\sqrt{n})$ per operation, we presented a randomized algorithm with $O(\log^3 n)$ amortized time per operation, where $n$ is the number of nodes in the graph (joint work with Valerie King).

I will discuss the above algorithm and an improved sampling routine that decreases the running time to $O(\log^2 n)$ (joint work with Mikkel Thorup). Additionally, I will discuss dynamic algorithms in directed graphs and pose open problems that arise in program verification, networking protocols, and relational data bases.

### Universal Algorithms for Store-and-Forward and Wormhole Routing

Authors: R. Cypher, F. Meyer auf der Heide, C. Scheideler, B. Vöcking

Christian Scheideler University of Paderborn

In this talk we present routing algorithms that are universal in the sense that they route messages along arbitrary (simple) paths in arbitrary networks. The algorithms are analyzed in terms of the number of messages, $n$, being routed, the maximum number of messages that must cross any edge in the network (edge congestion $C$), the maximum number of edges that a message must cross (dilation $D$), the buffer size, and the bandwidth of the links. We present two main results, both of which have applications to universal store-and-forward routing and universal wormhole routing. Our results yield significant performance improvements over all previously known universal routing algorithms for a wide range of parameters, and they even improve many time bounds for standard networks. In particular, we show that

[3] Kahn, J. (1991). "Recent Results on some not so recent Hypergraph Matching and Covering Problems," *Rutcor Research Report 4-91, Rutgers University.*

- Given any simple path collection in a network with bandwidth $\Theta(\frac{\log(C\cdot D)}{\log\log(C\cdot D)})$, there exists a packet routing algorithm that requires

$$O\left(D\log\log n + C + \frac{\log n \cdot \log\log n}{\log\log(C\cdot D)}\right)$$

  time, w.h.p., without buffering.

- Given any simple path collection in a network with bandwidth $B \leq \log n$, there exists a wormhole routing algorithm that requires

$$O\left(\frac{L\cdot C\cdot D^{1/B} + (D+L)\log n}{B}\right)$$

  time for worms of length $L$, w.h.p., without buffering.

In addition, we present adaptations of our main results for routing along shortest paths in arbitrary networks, and for routing in leveled networks, node-symmetric networks, edge-symmetric networks, expanders, butterflies, and meshes.

## A decision procedure for well-formed linear quantum cellular automata

Authors: C. Dürr, Thanh, Santha

Christophe Dürr
University of Orsay

In this paper we introduce a new quantum computation model, the linear quantum cellular automaton. Well-formedness is an essential property for any quantum computing device since it enables us to define the probability of a configuration in an observation as the squared magnitude of its amplitude. We give an efficient algorithm which decides if a linear quantum cellular automaton is well-formed. The complexity of the algorithm is $O(n^2)$ if the input automaton has continuous neighborhood.
**Classification of topics:** algorithms, automata and formal languages, computational complexity. (to be published in STACS'96)

## Good paths from random walks – a survey

Andrei Broder
Digital SRC

Consider the following class of problems: Given a graph $G = (V, E)$ and a set of $k$ pairs of vertices in $V$, we are interested in finding for each pair $(a_i, b_i)$ a path connecting $a_i$ to $b_i$, such that the set of paths has some disjointness properties. The related decision problems are typically $\mathcal{NP}$-complete; however in a series of papers Alan Frieze, Stephen Suen, Eli Upfal, and I have shown that for expanders the edge disjoint problem can be efficiently solved for $k$ within a polynomial-log factor of the optimum, and that in random graphs both the edge-disjoint problem and the vertex-disjoint problem can be efficiently solved, with high probability, for $k$ within a constant factor of the optimum.

The common basic paradigm used in these papers is to use random walks to produce (candidate) paths as follows: let $m_i$ be the endpoint of a random walk from $a_i$; construct a random walk from $b_i$ to $m_i$; use the catenation of the two walks as a (candidate) path from $a_i$ to $b_i$. Of course to make the analysis possible, numerous technical hurdles must be overcome, and the actual algorithms involve many details. However the fact remains that all paths are built from the catenation of random walks.

## Arithmetic Circuits

Avi Wigderson
The Hebrew University

I will discuss some recent results and old open problems on Arithmetic Circuits.

## Approximations for the Disjoint Paths Problem

Jon Kleinberg
MIT

This talk considers the problem of determining the maximum number of distinguished terminal pairs in a graph that can be simultaneously connected by disjoint paths. This is a classical NP-complete problem for which very little is known from the point of view of approximation algorithms. It has recently been brought into focus in papers concerned with routing in high-speed networks, where assigning paths to connection requests is a basic issue; in this setting, the current lack of understanding of the disjoint paths problem is an obstacle to the design of practical heuristics.

I will discuss recent work with Eva Tardos, in which we obtain a constant-factor approximation for the edge-disjoint paths problem in a class of locally planar graphs that includes the two-dimensional mesh. This is the first constant-factor approximation for this problem in any class of graphs other than trees. Our algorithm can also be adapted to work in the on-line setting, where we obtain an asymptotically optimal algorithm for the same class of graphs. This improves on a result of Awerbuch, Gawlick, Leighton, and Rabani for the special case of the mesh.

Also, I will discuss recent work with Alok Aggarwal and David Williamson on the node-disjoint paths problem on the mesh. This is a case of particular interest in the context of VLSI layout; from an algorithmic point of view, it also presents an appealing contrast to the edge-disjoint case, since the two problems are fundamentally different in the context of planar graphs. We obtain an improved trade-off between layout area and the number of layers required for routing in a model of Aggarwal, Klawe, Lichtenstein, Linial, and Wigderson. A special case of our main result is a significantly improved bound for the basic problem of routing a full permutation on the mesh using node-disjoint paths; our new bound is within polylogarithmic factors of the bisection bound.

## Lower Bounds to On-line Graph Problems with Applications to On-line Circuit and Optical Routing

Authors: Y. Bartal, S. Leonardi, A. Fiat

Yair Bartal
University of Rome

We present lower bounds on the competitive ratio of randomized algorithms for a wide class of on-line graph optimization problems and we apply such results to on-line virtual circuit and optical routing problems.

Lund and Yannakakis give unapproximability results for the problem of finding the largest induced subgraph satisfying any non-trivial, hereditary property. E.g., independent set, planar, acyclic, bipartite, etc. We consider the on-line version of this family of problems.

Furthermore, we study the on-line version of graph coloring whose off-line version has also been shown to be unapproximable by Lund and Yannakakis, on-line max edge-disjoint paths and on-line path coloring problems.

Irrespective of the time complexity, we show an $\Omega(n^\epsilon)$ lower bound of on the competitive ratio of randomized on-line algorithms for any of these problems.

As a consequence, we obtain an $\Omega(n^\epsilon)$ lower bound on the competitive ratio of randomized on-line algorithms for virtual circuit routing and optical routing on general networks, in contrast to the known results for some specific networks, and disproving a recently published claim of Kapoulas and Spirakis. Moreover, our lower bounds hold even if the use of preemption is allowed.

## Nonuniform Sampling for Cut and Flow Problems

David Karger
MIT

We improve on random sampling techniques for approximating problems that involve cuts in graphs. We give a linear-time construction that transforms any graph on $n$ vertices into an $O(n \log n)$-edge graph on the same vertices whose cuts have approximately the same value as the original graph's. In this new graph, for example, we can run the $\tilde{O}(mn)$-time maximum flow algorithm of Goldberg and Tarjan to find an $s$–$t$ minimum cut in $\tilde{O}(n^2)$ time. This corresponds to a $(1+\epsilon)$-times minimum $s$–$t$ cut in the original graph. In a similar way, we can approximate a minimum graph bisection in $\emptyset(n^2)$ time.

Joint work with Andras Benczur

## Learning Product Distributions

Authors: R. Kannan, M. Jerrum, A. Frieze

Ravi Kannan
Carnegie-Mellon University

We consider the following problem : suppose $x = (x_1, x_2, \ldots x_n)$ are $n$ independent real valued random variables with unknown distributions. Also $A$ is an unknown nonsingular matrix. We show that given samples of $y = Ax$, we can find approximately the columns of $A$ in poly time and also the distributions of the $x_i$.

As a special case of our problem, we are able to learn a cube in $n$ space (with unknown axes) given uniform samples drawn from it. This generalizes to parallelopipeds as well. Curiously, we show that learning simplices is at least as hard as graph isomorphism.

The problem is also of interest in Factor Analysis in Statistics. It is simple to find $A$ upto rotations. The new contribution is to find the rotation which we do using nonlinear optimization.

## Learning linear threshold functions with noise

Avrim Blum
Carnegie-Mellon University

The problem of learning a linear threshold function (a halfspace in n dimensions, also called a "perceptron") is one of the oldest in machine learning. Methods for doing this generally fall into two categories. Greedy algorithms are simple and can be made noise tolerant; but, their running time depends on a separation parameter that may be exponentially small. On the other hand, linear programming algorithms such as the Ellipsoid Algorithm run in polynomial time, but seem to be intolerant of noise. We show how greedy methods can be used to find weak hypotheses (hypotheses that classify noticeably more than half of the examples) in polynomial time, without dependence on any separation parameter. This results in a polynomial-time algorithm for learning linear threshold functions in the PAC model in the presense of random classification noise.

This is joint work with Alan Frieze, Ravi Kannan, and Santosh Vempala.

## An Efficient Zero-Knowledge Method for Answering Is He In Or Out? Questions

Authors: S. Micali, M. Rabin

Michael O. Rabin
Harvard University

Assume a very large universe $U$ of elements, each having a distinct name. An example is the entire population of the United States, where the name of each person includes his or her social security number. In the course of some distributed activity, nodes in a network need to declare subsets $S$ of $U$, and later on respond to questions of the form: is element $N$ of $U$ a member of $S$? An example is an organization such as a credit card company or a corporation which has card holders or employees. A person presents himself at some network node $V$ and claims to be a member of organization $X$. Note that the person identifies himself by his universal name $N$, not an organization-$X$ specific name. The node $V$ wishes to verify whether $N$ is in or not in the set $S$ of all members of organization $X$. It just as important for $V$ to have proof that $N$ is not a member (if that is the case), as to have proof that $N$ is a member. Node $V$ obtains the answer to the membership question by interaction with node $X$. At the end of this interaction $V$ will have a verified yes or no answer, but no other information. Also, the verifier $V$ can initiate membership inquiries about $N$ only at the request of $N$. Thus $V$ cannot go on "fishing expeditions" on the membership of $S$. This last feature is optional and can be omitted.

Let the size of $S$ be $|S| = n$. We construct an algorithm for zero-knowledge proofs for membership/non-membership in $S$ which requires $O(n * logn)$ initial work for the prover $X$, and requires $O(logn)$ work for each proof. Furthermore, every node in the network (but not the members of $U$, or even not the members of $S$) initially receives from $X$ a small number of bits, and after that $X$ is committed to the set $S$. We use cryptography, GMR signatures, Goldreich-Levin bits, and a sophisticated randomized hash scheme in our construction.

## Learning to Model Sequences Generated by Switching Distributions

Dana Ron
MIT

In this work we study efficient algorithms for solving the following problem, which we refer to as the Switching Distributions learning problem. A sequence $S = s_1 s_2 ... s_n$, over a finite alphabet Sigma is generated in the following way. The sequence is a concatenation of K runs, each of which is a consecutive subsequence. Each run is generated by independent random draws from a distribution $p_i$ over Sigma, where $p_i$ is an element in a set of distributions $p_1, ..., p_N$. The learning algorithm is given the sequence $S$ and its goal is to find approximations of the distributions $p_1, ..., p_N$, and give an approximate segmentation of the sequence into its constituting runs. We give an efficient algorithm for solving this problem and show conditions under which the algorithm is guaranteed to work with high probability.

Our research was motivated by the problem of learning distributions generated by Hidden Markov Models (HMM's). In particular, we were interested in HMM's which have the property that the transition probability function assigns a relatively high value to the transition from each hidden state to itself. In other words, the model tends to stay at the same hidden state for long periods of time and switch from state to state only infrequently. Such an assumption is often made when using HMM's in the context of speech analysis. This assumption is justified by the fact that the time scale in which speech is sampled is usually an order of magnitude smaller than the time scale of changes in the vocal tract.

This is joint work with Yoav Freund from AT&T Bell Laboratories.

## Randomized Robot Path Planning

Rajeev Motwani
Stanford University

Path planning in high-dimensional configuration spaces is a challenging area of research in robotics, as existing planners are practically useless beyond 3-5 dimensions. High-dimensional spaces arise in planning for robot arms with many degrees of freedom or for multiple cooperating robots, and in applications such as manufacturing, medical surgery, and space exploration. Some non-robotic applications also lead to high-dimensional path planning problems, e.g., computer-assisted animation in movies and computer-aided drug design.

We will discuss recent work on applying randomization, in the form of random walks and random sampling, to these path planning problems. The attractiveness of such randomized planners stems from their applicability to virtually any type of robot, and their empirically observed success. Focusing primarily on the work done at the Stanford Robotics Lab, we will present some models for facilitating a theoretical analysis of randomized path planners that perform well in practice, describe some preliminary theoretical work, and suggest directions for future research.

[Based on the work of the following: Jerome Barraquand, Lydia Kavraki, Jean-Claude Latombe, Tsai-Yen Li, and Prabhakar Raghavan.]

## Non-Approximability Results for MAX SNP – Towards Tight Results

Oded Goldreich
Weizmann Institute of Science

The state-of-the art with respect to the non-approximability of MAX SNP hard problems has been steadily improving over the last few years. In combination with the development of new techniques for positive results on approximation, the lower bounds and upper bounds on the approximability of several benchmark problems are already within sight of each other and there's hope that they may get tighter. An interesting feature of the negative results is that while the analysis has gotten tighter, in many senses it has gotten simpler and more general. At this stage it is possible to see the tightness of almost all steps of the analysis. In this talk we will attempt to show all the ingredients leading to the current hardness results for MAX SNP hard problems.

This is joint work with Mihir Bellare (UCSD) and Madhu Sudan (IBM). This talk covers a portion of the work titled "Free bits, PCP and Non-approximability: Towards tight results."
Previous talks on this work were concentrated on the Max Clique results.

### Extractors, Dispersers, and their Applications

Noam Nisan
Hebrew University

Extractors are boolean functions that allow, in some precise sense, extraction of randomness from "somewhat random" distributions. Extractors, and the closely related Dispersers, exhibit some of the most "random-like" properties of explicitly constructed combinatorial structures. In turn, extractors and dispersers have many applications in "removing randomness" in various settings, and in making randomized constructions explicit.

This talk surveys extractors and dispersers: what they are, sketch how they can be designed, and some of their applications. The work described is due to of a long list of research papers by various authors – most notably by David Zuckerman.

### Generating Hard Instances of Lattice Problems

Miklos Ajtai
IBM Almaden

We give a random class of lattices in $Z^n$ so that, if there is a probabilistic polynomial time algorithm which finds a short vector in a random lattice with a probability of at least 1/2 then there is also a probabilistic polynomial time algorithm which solves the following three lattice problems in every lattice in $Z^n$ with a probability exponentially close to one. (1) Find the length of a shortest nonzero vector in an $n$-dimensional lattice, approximately, up to a polynomial factor. (2) Find the shortest nonzerovector in an $n$-dimensional lattice $L$ where the shortest vector $v$ is unique in the sense that any other vector whose length is at most $n^c|v|$ is parallel to $v$, where $c$ is a sufficiently large absolute constant. (3) Find a basis $b_1, ..., b_n$ in the $n$-dimensional lattice $L$ whose length, defined as $max_{i=1}^n |b_i|$, is the smallest possible up to a polynomial factor.

### Randomness-Optimal Sampling, Extractors, and Constructive Leader Election

David Zuckerman
University of Texas

We present the first universal oblivious sampler that uses an optimal number of random bits, up to an arbitrary constant factor bigger than 1. In particular, using slightly more random bits than required to get one sample, we can approximate the average value of an arbitrary function $f : \{0,1\}^n \to [0,1]$ to within a polynomially small additive factor, with an exponentially small error probability.

Our proof is based on an improved extractor construction. An extractor is a procedure which takes as input the output of a defective random source and a small number of truly random bits, and outputs a nearly-random string. We present the first optimal extractor, up to constant factors, for defective random sources with constant entropy rate.

We give two applications of these tools. First, we exhibit a constructive $O(\log n)$ round protocol for leader election in the full information model that is resilient against any coalition of size $\beta n$ for any constant $\beta < 1/2$. Each player sends only $\log n$ bits per round. Second, given a $2g(n)$ round AM proof for $L$ in which Arthur sends $l(n)$ random bits per round and Merlin responds with a $q(n)$ bit string, we construct a $g(n)$ round AM proof for a language $L$ in which Arthur sends $O(l(n)+q(n))$ random bits per round and Merlin's response remains of polynomial length.

## Robust Threshold DSS Signatures

Tal Rabin
MIT

We present threshold DSS (Digital Signature Standard) signatures where the power to sign is shared by $n$ parties such that for a given parameter $t < n/2$ any subset of $2t+1$ signers can collaborate to produce a valid DSS signature on any given message, but no subset of $t$ corrupted parties can forge a signature (in particular, cannot learn the signature key). In addition, we present a robust threshold DSS scheme that can also stand the participation of $t$ malicious parties that attack the system by refusing to participate in the signature protocol ($t < n/3$ in this case) or generating incorrect partial signatures at time of signature computation ($t < n/4$ in this case). This results in a highly secure and resilient DSS signature system applicable to the protection of the secret signature key, the prevention of forgery, and increased system availability.

Joint work with : R. Gennaro, S. Jarecki and H. Krawczyk

## Learning Polynomials: The Highly Noisy Case

Authors: R. Rubinfeld, O. Goldreich, M. Sudan

Ronitt Rubinfeld
MIT

Given a function $f$ mapping $n$-variate inputs from a finite field $F$ into $F$, we consider the task of reconstructing a list of all $n$-variate degree $d$ polynomials which agree with $f$ on a tiny but non-negligible fraction, $\delta$, of the input space. We give a randomized algorithm for solving this task which accesses $f$ as a black box and runs in time polynomial in $\frac{1}{\delta}, n$ and exponential in $d$, provided $\delta$ is $\Omega(\sqrt{d/|F|})$. For the special case when $d = 1$, we solve this problem for all $\epsilon \equiv \delta - \frac{1}{|F|} > 0$. In this case the running time of our algorithm is bounded by a polynomial in $\frac{1}{\epsilon}, n$ and exponential in $d$. Our algorithm generalizes a previously known algorithm, due to Goldreich and Levin, that solves this task for the case when $F = GF(2)$ (and $d = 1$).

Joint work with Oded Goldreich and Madhu Sudan.

## Linearity Testing

Marcos Kiwi
MIT

Given a function $f$ mapping from a group $G$ to a group $H$, the probability that the BLR (Blum-Luby-Rubinfeld) test rejects $f$, denoted $\mathsf{Rej}(f)$, is the probability that $f(u) + f(v) \neq f(u + v)$ when $u$ and $v$ are randomly chosen in $G$. Linearity testing is the study of the relationship between $\mathsf{Rej}(f)$ and how far away $f$ is from the space of linear functions. Several analyses of the BLR test are known, but none is tight.

The problem arises when we have access to a function and want to estimate its closeness to a linear function as accurately as possible, but using few queries.

The case of interest in the construction of PCPs and the derivation of non-approximability results is when the underlying groups are $G = GF(2)^n$ and $H = GF(2)$. We focus on this case and present an analysis of the Linearity Test which is nearly complete in all its aspects. The lower bound shown here has been used in recent works to present the best known hardness results for Max3SAT and other MaxSNP problems.

Part of our results are obtained by showing a new connection between the linearity testing problem and Fourier analysis. We will discuss this connection at length and show that it is of independent interest.

Part of the talk will discuss joint work with Mihir Bellare, Don Coppersmith, Johan Håstad and Madhu Sudan.

## Testing Multivariate Linear Functions: Overcoming the Generator Bottleneck

Funda Ergun
Cornell University

Self-testing programs provide an approach to the problem of program correctness that has the advantage of being program independent. One can construct self-testers by exploiting the set of properties that uniquely define the function that the program purportedly computes and testing that they hold at random inputs. The previous testers introduced very small overhead while testing univariate functions, but became significantly more costly in the case of multivariate functions, since the number of properties that define the function grows infeasibly large in those instances. In this paper we develop techniques for finding a much smaller set of such properties, which lead to more efficient testers for multivariate linear functions. We present efficient self-testers for the following functions that did not have self-testers before: the Discrete Fourier Transform, evaluation of polynomials, dot product (and therefore vector 2-norm), and pointwise evaluation of linear functions on vectors. We present a tester for polynomial multiplication that makes $O(1)$ calls to the program, in contrast to $O(\log n)$ of the best previously known tester, and present a new tester for matrix multiplication. All of the testers presented make $O(1)$ calls to the program that is being tested, therefore the asymptotic complexity of the whole operation remains unchanged after the addition of the testing overhead. We then generalize these results and place them in one common framework to present a general method for dealing with multivariate linear functions.

This is joint work with S. Ravikumar and D. Sivakumar.

## Efficient Self-Testing of Linear Recurrences

S. Ravi Kumar
Cornell University

We consider the problem of designing efficient self-testers for linear recurrences, and present a complete package of self-testers for this class of functions. The results are proved by demonstrating an efficient reduction from this problem to the problem of testing linear functions over certain matrix groups. Our tools include spectral analysis of matrices over finite fields, and various counting arguments that extend known techniques. The matrix twist yields completely new degree tests over the finite field Z/p. The efficiency of our polynomial self-testers is better than all previously known testers, and in the univariate case, we are able to match the efficiency of the Blum-Luby-Rubinfeld linearity tester. We also present self-testers for convolution identities over groups, and improved self-testers for polynomials over rational domains.

This is joint work with D. Sivakumar (SUNY, Buffalo).

## Evaluation may be easier than generation

Moni Naor
The Weizmann Institute of Science

Kearns et al. (STOC 94) defined two notions for learning a distribution $D$. The first is with a generator, where the learner presents a generator that outputs a distribution identical or close to $D$. The other is with an evaluator, where the learner presents a procedure that on input $x$ evaluates correctly (or approximates) the probability that $x$ is generated by $D$. They showed an example where efficient learning by a generator is possible, but learning by an evaluator is computationally infeasible.

Though it may seem that generation is, in general, easier than evaluation, in this talk we show that the converse may be true: we provide a class of distributions where efficient learning with an evaluator is possible, but coming up with a generator that approximates the given distribution is infeasible. We also show that some distributions may be learned (with either a generator or an evaluator) to within any $\epsilon > 0$, but the learned hypothesis must be of size proportional to $\epsilon$ (and not $\log \epsilon$ which is always the case in the distribution-free PAC model).

## Characterizing Linear Size Circuits in Terms of Privacy

Rafail Ostrovsky
Bellcore

In this paper we prove an unexpected relationship between the complexity class of linear size circuits, and n-party private protocols. Specifically, let $f$ : $0-, 1^n \rightarrow 0, 1$ be a boolean function. We show that $f$ has a linear size circuit if and only if $f$ has 1-private, $n$-party protocol in which the total number of random bits used by all players is constant.

¿From the point of view of complexity theory, our result gives a characterization of the class of linear size circuits in terms of another class of a very different nature. From the point of view of privacy, this result provides 1-private $O(1)$-random protocols for many important functions for which no such protocol was known. On the other hand, it suggests that proving, for any NP function, that it has no 1-private $O(1)$-random protocol might be quite difficult.

Joint-work with Eyal Kushilevitz and Adi Rosen.

## Disk Packings and Planar Separators

Dan Spielman
UC Berkeley

We demonstrate that the geometric separator algorithm of Miller, Teng, Thurston, and Vavasis finds a 3/4-separator of size $1.84\sqrt{n}$ in every $n$ node planar graph.

This is joint work with Shang-Hua Teng.

## Digital Signets for Protection of Digital Information

Cynthia Dwork
Almaden IBM

The problem of protecting digital content – software, video, documents, music, etc. – from illegal redistribution by an authorized user, is the focus of considerable industrial and academic effort. In the absence of special-purpose tamper-proof hardware, the problem has no cryptographically secure solution: once a legitimate user has purchased the content, the user, by definition, has access to the material and can therefore capture it and redistribute it. A number of techniques have been suggested or are currently employed to make redistribution either inconvenient or traceable. The problem with traceability is that it requires a "digital content police" that has no automatic means of generating suspicion.

In this work we introduce *digital signets*, a new technique for protecting digital content from illegal redistribution. In broad terms, signets work as follows. There is some common public data, some of which is encrypted *content*, an *authorization center*, and any number of potential *users*. Each user has access to the common public data. To gain access to the content, user $U$ interacts with an authorization center to obtain a short *digital signet*, which is a function of information private to $U$. Using only the public data, its own private information, and the signet, $U$ can decrypt the content. Signets are designed in such a way that for $U$ to help any other $U'$ to access the content, $U$ would have to either transmit something very long (such as the content itself), or reveal $U$'s private information. Thus, users have incentive to police *themselves.*

The work motivates the study of the previously unexamined class of *incompressible* functions, analysis of which adds a cryptographic twist to communication complexity.

This is joint work with Jeffrey Lotspiech and Moni Naor.

pagebreak

## Extremal bipartite graphs and superpolynomial lower bounds for monotone span programs

Anna Gal
Institute for Advanced Studies

Introduced by Karchmer and Wigderson in 1993, "span programs" provide a linear algebraic model of computation with applications to lower bounds in other models (contact schemes, symmetric branching programs, formula size), and to cryptography (secret sharing).

We consider monotone span programs. Our first result is that this monotone model can be more powerful than monotone circuits. We exhibit a function that is computable by monotone span programs in linear size but requires superpolynomial size monotone circuits.

We then present the first superpolynomial lower bounds for the size of monotone span programs computing explicit functions. The best previous lower bound was $\Omega(n^{5/2})$ by Beimel, Gal, Paterson (FOCS'95). Our analysis exploits a criterion from that paper, that allows to prove lower bounds for monotone span programs by considering a problem in extremal set theory.

Our lower bounds are based on explicit constructions of bipartite graphs which do not contain certain complete bipartite graphs. We give two such constructions, both based on Paley-type bipartite graphs. One of the constructions is of independent interest, it gives maximal density, up to a constant factor, under this constraint. The construction beats the previously known probabilistic lower bound on density. This problem has been studied in combinatorics as the Zarankiewicz problem, and has applications to other questions in monotone circuit complexity.

Joint work with Laszlo Babai, Janos Kollar, Lajos Ronyai, Tibor Szabo and Avi Wigderson.

## Randomized $\Omega(n^2)$ Lower Bound for the Knapsack Problem (Randomized Lower Bounds for the Decision Trees Revisited)

Marek Karpinski
University of Bonn

We establish superlinear lower bounds on the depth of *randomized* algebraic decision trees computing finite unions of hyperplanes and the intersections of halfspaces, solving a long standing open problem. As an application, among other things, we derive for the first time an $\Omega(n^2)$ *randomized* lower bound for the *Knapsack Problem.*

Joint work with D. Grigoriev, F. Meyer auf der Heide, and R. Smolensky.

## A New Rounding Procedure for Assignment-type Problems

Sanjeev Arora
Princeton University

Many NP-hard problems can be phrased as the problem of finding a matching in a bipartite graph subject to linear or higher degree constraints. We show how to find "almost perfect" matchings that "almost" satisfy the constraints. As a result we get approximation schemes for many such problems (or some of their important subcases).

## An optimization problem on a random graph where the Metropolis algorithm is (almost) theoretically analyzable.

David Aldous
UC Berkeley

The (artificial) problem involves a certain random $n$-vertex 3-regular graph with a certain objective function on the vertices; we seek algorithms to find near-optimal vertices in poly-$logn$ steps. The problem was designed so that the analysis of the Metropolis algorithm is conceptually similar to topics in the theoretical probability-statistical physics field. The model involves a parameter $p$ which has a critical value. On one side of the critical value there is no poly-$logn$ algorithm. On the other side a simple greedy algorithm (which however requires remembering the entire past) works in $O(logn)$ time, and (setting aside technical difficulties) so does the memoryless Metropolis algorithm.

## A semidefinite bound for mixing rates of Markov chains

Nabil Kahale
UC Berkeley

We study the method of bounding the spectral gap of a reversible Markov chain by establishing canonical paths between the states. We provide natural examples where improved bounds can be obtained by allowing variable length functions on the edges. We give a simple heuristic for computing good length functions. Further generalization using multicommodity flow yields a bound which is an invariant of the Markov chain, and which can be computed at an arbitrary precision in polynomial time via semidefinite programming. We show that, for any reversible Markov chain on $n$ states, this bound is off by a factor of at most $O(\log^2 n)$, and that this can be tight.

### A simple random walk for counting degree sequences

Authors: S. Vempala, R. Kannan

Santosh Vempala
Carnegie-Mellon University

We consider the problem of randomly generating bipartite graphs with a given degree sequence. We analyze a Markov Chain for the problem. We cannot prove that this chain is rapidly mixing in general, but in the regular case when all the degrees are equal, we give a proof of rapid mixing. Jerrum and Sinclair solved the corresponding problem for general (nonbipartite) regular graphs. Our chain has one state for every bipartite graph with the degree sequence. (There are no auxiliary states as in the chain used by jerrum and Sinclair.)

The corresponding problem for bipartite multi graphs (where we are allowed multiple edges) is the problem of random generation of matrices with nonnegative integer entries and given row and column sums. This problem is of importance in Statistics where such matrices arise as "contingency tables". A Markov Chain similar to the one we analyze has been proposed for this problem by Diaconis and it is a long-standing open problem as to whether this mixes rapidly. We hope our techniques shed some light on this problem.

### A new approach to polynomial time random walks for volume computation

Authors: R Bubley, M. Dyer, M. Jerrum

Russ Bubley
University of Leeds

We present a fully-polynomial randomized almost uniform generator for a class of log-concave functions; this extends naturally into a fully polynomial randomized approximation scheme for the volume of a convex body in Euclidean $n$-space. Prior algorithms for computing volumes have relied on complicated conductance arguments and isoperimetric inequalities to show that their Markov processes mix rapidly; the much simpler classical coupling method is used here to prove that a certain Markov process is rapidly mixing.

## Sampling Domino Tilings on Regions With Free Boundary Conditions

Dana Randall
Princeton University

We present an efficient algorithm for sampling domino tilings on a finite region $R_n$ with *free* boundary conditions. Here we allow dominoes to cross the boundary of the region (as though part of a larger tiling), and we sample from the set of distinct configurations we can see within a window of shape $R_n$. This builds on techniques from Luby, Randall and Sinclair [LRS] for sampling domino tilings on regions with *fixed (or Dirichlet)* boundary conditions, where domino are restricted to stay within the boundary of the region.

The motivation for both of these problems comes from statistical mechanics, where tilings on successively larger regions are studied to determine properties of dimer systems, represented by domino tilings on the infinite lattice. For example, consider a nested set of regions $\{R_1 \subset R_2 \subset ...\}$ and let $c_n$ be the number of domino configurations on $R_n$ (with either type of boundary condition). If $s_n = \frac{\ln c_n}{\text{area}(R_n)}$, then the entropy is defined as $\lim_{n\to\infty} s_n$. It turns out that when the boundary conditions are fixed, the entropy depends crucially on the shape of the regions (where the square regions have maximal entropy). In contrast, when the boundary conditions are free, there is a large class of nested regions which all have maximal entropy. Furthermore, there is compelling empirical evidence that certain other statistical properties of tilings on the infinite lattice are bounded above and below by the corresponding statistics of finite regions with free and fixed boundary conditions, respectively. This suggests that our sampling schemes can be used together to derive explicit bounds on estimates of these statistics, not only for finite regions, but also for the infinite lattice.

## Greedy Approximate Counting

Lars Rasmussen
UC Berkeley

Suppose we want to approximate the size of some finite set $S$, and have available a probabilistic algorithm capable of generating every element $x$ of $S$ with non-zero probability $P(x)$. Then the random variable $1/P(x)$ has expected value exactly $|S|$, and, if the variance of the generator is sufficiently low, we may build a "fully polynomial randomized approximation scheme" for $|S|$ from it.

We present a survey of the use of the technique and a new application to the problem of counting cliques in random graphs.

## Cascading Pseudo-Random Functions– How to Key Merkle

Mihir Bellare
UC San Diego

Suppose we are given a finite pseudo-random function $F_a$ mapping $l$ bit strings to $k$ bit strings. (Eg. $l = 512$ and $k = 128$.) How can we extend this to a function taking inputs of arbitrary length?

We will describe a natural construction which we call "cascading." It can be viewed as a keyed version of the well known Merkle method of extending the domain of a hash function.

We will look at a few variants of this construction and analyze their security. The emphasis will be on "exact security." We will see how, for an adversary who sees a given number of examples and runs for a given amount of time, the success probability can be exactly analyzed, and differs according to the scheme.

The problem is motivated by the need to find provably good ways of constructing message authentication codes out of a certain common cryptographic primitive called a "compression function."

We hope to illustrate by this construction and approach a more general theme: the usage of finite pseudo-random functions as a tool to analyze basic cryptographic constructions, and the importance and intrinsic technical interest of exact security analysis.

Joint work with: Ran Canetti (MIT/Weizmann) and Hugo Krawczyk (IBM)

## An efficient pseudo-random generator provably as secure as syndrome decoding

Authors: J.B. Fischer, J. Stern

Jacques Stern
ENS, France

We show a simple and efficient construction of a pseudo-random generator based on the intractability of an NP-complete problem from the area of error-correcting codes. The generator is proved as secure as a hard instance of the syndrome decoding problem. Each application of the scheme generates a linear amount of bits in only quadratic computing time.

## Small Biased Toeplitz Hashing with Applications to Message Authentication

Hugo Krawczyk
IBM - T.J. Watson Research Center

We show how to construct almost universal hashing using Toeplitz matrices generated out of small biased sequences. This leads to simple and efficient hashing schemes with essentially the same hashing strength of a completely random matrix but at a substantially lower cost in randomness, description size, and implementation complexity. The schemes are especially advantageous to hash large amounts of data.

We show cryptographic applications of these results to message authentication, where information is authenticated by encrypting the hash value under a (random or pseudorandom) one-time pad. We present specific efficient and practical constructions (e.g., based on linear feedback shift registers) that require short keys and short authentication tags, yet providing provable security. These results include a full characterization of hash families that are secure for message authentication in the one-time pad model.

## Adaptively Secure Multiparty Computation

Ran Canetti
MIT

A fundamental problem in the area of secure multiparty computation is how to deal with adaptive adversaries in the computational setting. (Adaptive adversaries are adversaries that may choose the corrupted parties during the course of the computation, based on the information gathered so far.)

The power of an adaptive adversary is greatly affected by the extent to which honest parties carry out instructions that cannot be externally verified, such as erasing all records of the history of the execution. It has been shown that if the parties are trusted to erase all history records, then adaptively secure computation can be carried out using known primitives. However, this total trust may be unrealistic in many scenarios. An important question, open since 1986, is whether adaptively secure multiparty computation can be carried out in the computational setting, even if all parties keep all history records of the execution.

We answer this question in the affirmative, by introducing and using a novel property of probabilistic encryption protocols. We show that if encryption enjoying this property is used, instead of standard encryption, then known constructions become adaptively secure. Next we construct, based on the RSA assumption, an encryption protocol that enjoys this property.

□