Abstracts for the Talks

A randomized approximation scheme for metric MAX-CUT

W. Fernandez De La Vega

Metric MAX-CUT is the problem of dividing a set of points in metric space into two parts so as to maximize the sum of the distances between points belonging to distinct parts. We show that metric MAX-CUT has a polynomial time approximation scheme.

(Joint work with Claire Kenyon)

On counting independent sets in sparse graphs

Martin Dyer, Leeds

I will discuss two results concerning approximate counting of independent sets in graphs with constant maximum degree Δ. The first implies that the Monte Carlo Markov chain technique is likely to fail if $\Delta \geq 6$. The second shows that no fully polynomial randomized approximation scheme can exist if $\Delta \geq 25$, unless P=NP under randomized reductions. (Joint work with Alan Frieze and Mark Jerrum.)

Inferring evolutionary trees

Leslie Goldberg

The General Markov Model of evolution (due to Steel) is a stochastic model concerned with the evolution of strings over a fixed-size alphabet. The Two-State General Markov Model (in which the alphabet is taken to be binary) generalises the well-known Cavender-Farris-Neyman model of evolution. In the last RAND meeting, Mary Cryan described preliminary work in which we discovered a PAC (Probably Approximately Correct) learning algorithm which recovers a Markov Evolutionary Tree, given the topology of the tree, and polynomially-many samples from the output distribution of the tree. Since then, we have discovered how to learn the topology (with sufficient accuracy). Thus, we obtain a polynomial-time PAC-learning algorithm (in the sense of Kearns et al.) for the two-state General Markov Model. This generalises work of Farach and Kannan, who showed how to PAC-learn Markov Evolutionary Trees in the Cavender-Farris-Neyman model provided that the target tree satisfies the additional restriction that all pairs of leaves have a sufficiently high probability of being the same.

(Joint work with Mary Cryan and Paul Goldberg)

On Markov chains for the Widom-Rowlinson model

Catherine Greenhill, Leeds

The Widom-Rowlinson model is a model of a gas with two or more different types of particles. I will present some new results from work in progress involving rapidly mixing Markov chains for the two-particle Widom-Rowlinson model.

[Joint work with Artur Czumaj and Martin Dyer]

The "Burnside process" converges slowly

Mark Jerrum, Edinburgh

We consider the problem of sampling "unlabelled structures", i.e., sampling combinatorial structures modulo a group of symmetries. The main tool which has been used for this sampling problem is Burnside's lemma. In situations where a significant proportion of the structures have no non-trivial symmetries, it is already fairly well understood how to apply this tool. More generally, it is possible to obtain nearly uniform samples by simulating a Markov chain that we call the Burnside process; this is a random walk on a bipartite graph which essentially implements Burnside's lemma. For this approach to be feasible, the Markov chain ought to be "rapidly mixing", i.e., converge rapidly to equilibrium. The Burnside process was known to be rapidly mixing for some special groups, and it has even been implemented in some computational group theory algorithms. In this paper, we show that the Burnside process is not rapidly mixing in general. In particular, we construct an infinite family of permutation groups for which we show that the mixing time is exponential in the degree of the group.

[Joint work with Leslie Goldberg]

Efficient approximation algorithms for the Hamming center problem

Thore Husfeldt

The Hamming center problem for a set S of k binary strings, each of length n, is to find a binary string beta of length n that minimizes the maximum Hamming distance between beta and any string in S. Its decision version is known to be NP-complete. We provide several approximation algorithms for the Hamming center problem. Our main result is a randomized (4/3+epsilon)-approximation algorithm running in polynomial time if the Hamming radius of S is at least superlogarithmic in k. Furthermore, we show how to find in polynomial time a set B of O(log k) strings of length n such that for each string in S there is at least one string in B within Hamming distance not exceeding the radius of S.

Joint work with Leszek Gasieniec, Jesper Jansson, Andrzej Lingas

A method for experimentally exploring search spaces: the case of graphs with small planted bisections

Russell Impagliazzo, UC San Diego

There is a large gap between the experimentally observed success of many optimization heuristics (e.g., local optimization, Metropolis, simulated annealing, WalkSat, and genetic algorithms) and a theoretical understanding of their performance. As well as being frustrating to theorists, this gap limits the utility of these heuristics: it is currently a matter of guess-work to determine which approach is suitable for a specific problem domain, or how to set the various parameters for best performance.

One reason for this gap is the fact that even for simple problems, the associated search space, which determines the heuristics' performance, can be complicated and hard to understand. This makes an experimental approach to understanding connections between the combinatorial structure of search graphs and the performance of search algorithms also seem intractable than theoretical study. Each heuristic method has an enormous variety of variations and parameters, and most are not analyzed in terms of the properties of the search graph. So merely implementing various heuristics and comparing their performance tells us very little about the structure of the search graph or about the relative merits of the heuristic approaches. Also, since the search graphs are exponentially large, it is infeasible to directly compute their combinatorial structure for any meaningful instance.

However, we propose a method for feasibly performing a meaningful experimental study of search spaces to identify combinatorial features relevant to the performance of optimization heuristics. Our method uses the Go-With-the-Winners optimization algorithm ([AV,DI1]) as a tool for making random samples from relevant regions of the search graph underlying a combinatorial optimization problem.

We apply this method to the minimum bisection problem on low density graphs with a "planted" bisection. Using our experimental data, we show how to select a "cooling schedule" for a search algorithm similar to Simulated Annealing that is tailored to the distribution studied.

Joint work with Ted Carson

Boolean functions: influence, noise sensitivity, measure concentration, Fourier expansions, applications and speculations

Gil Kalai, Hebrew University

I will review my recent work with Itai Benjamini and Oded Schramm on matters described in the title. Most applications are to percolation. Many speculations are to complexity theory.

On some tighter inapproximability results

Marek Karpinski, University of Bonn

We prove a number of new inapproximability results including the first explicit approximation thresholds for bounded dependency satisfiability problems, like MAX-2SAT and E2-LIN-2, and the bounded degree problems, like MIS, Node-Cover, and MAX-CUT. We prove also for the first time inapproximability of the problem of Sorting by Reversals and display an explicit approximation threshold for that problem.

(Joint work with P. Berman.)

Fully dynamic algorithms for maintaining the transitive closure

Valerie King

The fully dynamic transitive closure problem is to maintain the transitive closure of a directed graph which is changing, that is, as edges are inserted and deleted. Clearly, this can be done by recomputing the transitive closure after each update to the graph. But until now, it was not known if there was any more efficient way to update the transitive closure.

I'll describe a method for updating the transitive closure that is more efficient than matrix multiplication, $O(n^2)$ per update for acyclic graphs, and a bit more for general graphs. The algorithm is Monte Carlo with one-sided error.

This is joint work with Garry Sagert.

Concentration for random minimum spanning trees

Colin McDiarmid, Oxford

Suppose that the complete graph K_n has independent random edge lengths, each uniformly distributed on the interval $(0,1)$. Let X_n be the minimum length of spanning tree. It is well known that $E(L_n) \sim \zeta(3)$, and that the random variable L_n concentrates around this value. We discuss how to establish strong concentration using either the martingale or the Talagrand approach.

Searching for a near neighbor in high dimensional spaces

Yuval Rabani, Technion

We address the following search problem: Given a database consisting of a set of vectors in high dimensional space, we want to construct a small data structure that would allow us to search quickly, given a query vector, for a vector in the database which is close to the query. The problem has many flavors, depending on the vector space, distance function, and required proximity. Finding the database vector closest to the query is expected to suffer from the "curse of dimensionality" in most cases (i.e., it is conjectured that high dimensional problems are intractable).

For near-optimal search we present randomized constructions of data structures whose size is polynomial in the size of the database, and randomized search algorithms that run in time nearly linear in the dimension. In contrast, we show that algorithms with similar features do not exist for optimal search (when the closest vector is desired). In the talk, we emphasize the case of the Hamming cube. Our results extend to Euclidean and other spaces.

The talk combines joint work with Allan Borodin, Eyal Kushilevitz, and Rafail Ostrovsky.

On extractors and pseudo-randomness

Omer Reingold, Weizmann Institute

In this talk we will describe the recent construction of extractors due to Luca Trevisan and our improvements upon this construction. Informally, an extractor is a function that extracts almost uniformly distributed bits from a weakly random source using a few additional uniformly distributed bits as a catalyst. Recently Trevisan presented a new direction to the construction of extractors based on the Nisan-Wigderson pseudo-random generators. In some scenarios these extractors are optimal in the number of truly random bits they use.

Based on Trevisan's extractors, we give explicit constructions of extractors which work for a source of any min-entropy k on strings of length n. These construction reduce the number of truly random bits used by Trevisan's extractors in two cases:

1. When trying to extract all the min-entropy of the weak source (or a constant fraction of it).
2. When the output of the extractor should be very close to uniform.

In particular to output $m = \Omega(k)$ bits that are ϵ-close to uniform the number of truly random bits used by our extractors is $O(\min(\log(n/epsilon))^2, (\log n)^2\epsilon)$.

Joint work with Ran Raz and Salil Vadhan.

Approximate testing with relative error

Miklos Santha

We formalize the notion and initiate the investigation of approximate testing for arbitrary forms of the error term. Until now only the case of absolute error had been addressed ignoring the fact that often only the most significant figures of a numerical calculation are valid. This work considers approximation errors whose magnitude grows with the size of the input to the program. We demonstrate the viability of this new concept by addressing the basic and benchmark problem of self–testing for the class of linear and polynomial functions. We obtain stronger versions of results of Ergun, Ravi Kumar, and Rubinfeld exploiting elegant techniques from Hyers–Ulam stability theory.

This is a joint work with Marcos Kiwi and Frederic Magniez.

Quadratic Euclidean clustering

Leonard J. Schulman, Georgia Tech

We consider the problem of partitioning a set of n points in rset^d into clusters, so as to minimize the sum over clusters of sum of the squares of the Euclidean distances between points. We obtain a sequence of results culminating in a highly efficient approximation algorithm for this minimization problem.

The results are: (a) A characterization of "locally optimal" clusterings. (b) A deterministic algorithm identifying optimal clusterings, running in time polynomial in n for fixed dimension and number of clusters. (c) A randomized, error-free approximation algorithm in which the runtime depends only weakly on the dimension. (d) A randomized approximation algorithm running in *linear* time even when the dimension and number of clusters grow slowly as a function of the number of points.

Approximation algorithms for MAX CUT and MAX SAT

Uri Zwick, Tel-Aviv University

We present a tool, outward rotations, for enhancing the performance of several semidefinite programming based approximation algorithms. Using outward rotations, we obtain an approximation algorithm for MAX CUT that, in some interesting cases, performs better than the algorithm of Goemans and Williamson. We also present some improved results for MAX 4-SAT and for MAX SAT,

Some of the results are joint with Eran Halperin