Linear, Semidefinite Programming and Randomization Methods for Combinatorial Optimization Problems

K. Jansen (Kiel)

J. Rolim (Geneva)

M. Sudan (MIT Cambridge)

1 Overview

During the week of January 23 - 28, 2000, the Seminar on Linear, Semidefinite Programming and Randomization Methods for Combinatorial Optimization Problems was organized by K. Jansen (Kiel), J. Rolim (Geneva) and Madhu Sudan (MIT Cambridge). 41 Participants came from universities or research institutes from Austria, Israel, Greece, France, Germany, Netherlands, Russia, Switzerland, United Kingdom and United States America.

The objective of the seminar was to focus on algorithmic and complexity aspects arising in the development of efficient solution techniques for computationally difficult optimization problems. The seminar promoted an exchange of different methods used in approximation algorithms. A general technique for efficient approximation algorithms is to formulate an optimization problem as an integer linear program and then to relax the integrality conditions. Recently, there has been striking success in obtaining also approximation algorithms on more general mathematical programming such as semidefinite programming. In this and other context, randomization has proved to be a powerful algorithmic method: it yields to simple and easy to analyse algorithms for many optimization problems, and it leads to a better performance guarantee.

Main topics of interests were:

- extending of randomization and semidefinite programming techniques to other optimization problems,

- improved approximation algorithms and structural insights by studying linear programming, semidefinite programming and randomization,

- development of approaches to solve (approximatively) large linear and semidefinite programs,

- complexity questions for randomization and semidefinite programming,

- practical implementation of the used techniques (randomization, semidefinite programming),

- exchange of informations on recent research and stimulation of further research.

The seminar was intended to bring together researchers from different areas in combinatorial optimization and from applications. Different algorithmic methods and techniques have been covered by 25 lectures. There was also an open problem session on Thursday evening and a lively discussion on problems from different fields

of application. Talks were given studying practical problems like facility location, scheduling, load balancing, data broadcasting, discrete tomography, and curve reconstruction and also theoretical oriented problems on satisfiability, polymatroids, lattices , semidefinite relaxations, metarounding, probabilistic methods, discrepancies, sorting by reversals and structed optimization.

Schloß Dagstuhl and its staff provided a very convenient and stimulating atmosphere. The organizers wish to thank the local organization and all those who helped to make the seminar a fruitful research experience.

On behalf of the participants,

Klaus Jansen, Jose Rolim, and Madhu Sudan

2   Program

Monday Morning Session Chair: Madhu Sudan

9:00 *Welcome*

9:15 - 9:45 *A constant approximation algorithm for the k-median problem*
David B. Shmoys, *Cornell University, USA*

10:00 - 10:30 *On the properties of basic solutions of linear relaxations for MAX DICUT with given sizes of parts*
Maxim I. Sviridenko, *BRICS, Danmark*

10:45 - 11:15 *Randomized metarounding*
Santosh S. Vempala, *MIT - Cambridge, USA*

Monday Afternoon Session Chair: Klaus Jansen

15:00 - 15:30 *Width-free bounds for structured optimization*
Michael D. Grigoriadis, *Rutgers University, USA*

16:15 - 16:45 *General multiprocessor scheduling: approximate solutions in linear time*
Lorant Porkolab, *Imperial College London, United Kingdom*

17:00 - 17:30 *Approximation algorithms for a weighted interval scheduling problem*
Frits C.R. Spieksma, *Maastricht University, Netherlands*

Tuesday Morning Session Chair: David B. Shmoys

9:00 - 9:45 *Approximation hardness of sorting by reversals & bounded dependency paradigm*
Marek Karpinski, *Universität Bonn, Germany*

10:00 - 10:40 *Approximating the MAX CUT problem on bounded degree graphs*
Michael Langberg, *Weizmann Institute Rehovot, Israel*

11:00 - 11:30 *Approximating MAX SAT with bounded occurrence of variables*

Johan Håstad, *KTH Stockholm, Sweden*

Tuesday Afternoon Session Chair: Peter Gritzmann

15:00 - 15:30 *Polynomial instances of the positive semidefinite matrix completion problem*

Monique Laurent, *CWI Amsterdam, Netherlands*

16:15 - 16:45 *SDP versus LP for graph problems: (a computational viewpoint)*

Franz Rendl, *Universität Klagenfurt, Austria*

17:00 - 17:30 *Computational methods for solving semidefinite relaxations*

Christoph Helmberg, *Konrad Zuse Zentrum Berlin, Germany*

Wednesday Morning Session Chair: Johan Hastad

9:00 - 9:45 *Computer assistent analysis of approximation algorithms for MAX 3-SAT and other problems*

Uri Zwick, *Tel Aviv University, Israel*

10:00 - 10:45 *On the complexity of lattice problems*

Shmuel Safra, *Tel Aviv University, Israel*

11:00 - 11:30 *Approximating minimum spanning sets in polymatroids*

Stefan Hougardy, *HU Berlin, Germany*

Thursday Morning Session Chair: Hans - Jürgen Prömel

9:00 - 9:40 *Circle method and discrepancies*

Anand Srivastav, *Universität Kiel, Germany*

10:00 - 10:30 *Probabilistic & deterministic methods for counting sum-free sets*

Tomasz Schoen, *Universität Kiel, Germany*

10:45 - 11:30 *Randomized approximate counting and bounding mixing times*

    Leslie A. Goldberg, *University of Warwick, United Kingdom*

Thursday Afternoon Session Chair: Ingo Schiermeyer

15:00 - 15:45 *Cooperative facility location games*

    Martin Skutella, *TU Berlin, Germany*

16:15 - 16:45 *Approximation algorithms for packing and covering problems in discrete tomography*

    Peter Gritzmann, *TU München, Germany*

17:00 - 17:30 *An approximation algorithm for the precedence constrained scheduling problem with hierarchical communications*

    Rodolphe Giroudeau, *Universite Evry, France*

17:30 - 18:00 *TSP - based curve reconstruction*

    Kurt Mehlhorn, *Max-Planck-Institut Saarbrücken, Germany*

20:00 Open Problems Chair: Marek Karpinski

Friday Morning Session Chair: Frits Spieksma

 9:00 -   9:30 *Balanced allocation: the heavily loaded case*

    Angelika Steger, *TU München, Germany*

 9:45 - 10:15 *A PTAS for data broadcast*

    Nicolas Schabanel, *Universite Lyon, France*

10:45 - 11:30 *Semidefinite programming for semi-random graphs*

    Uri Feige, *Weizmann Institute Rehovot, Israel*

3 Abstracts

Semidefinite Programming for Semirandom Graphs Problems

Uri Feige

feige@narkis.wisdom.weizmann.ac

We consider semirandom graph models for finding large independent sets, colorings and bisections in graphs. These models generate problem instances by blending random and adversarial decisions. To generate semirandom independent set problems, an independent set S of αn vertices is randomly chosen. Each edge connecting S with $\bar{S}$ is chosen with probability p, and an adversary is then allowed to add new edges arbitrarily, provided that S remains an independent set. We show that an algorithm based on semidefinite programming recovers with high probability an independent set of size αn whenever $p > (1 + \epsilon) \ln n / \alpha n$, for any constant $\epsilon > 0$. This is essentially best possible, because the problem becomes NP-hard once $p < (1 - \epsilon) \ln n / \alpha n$. Our algorithm for finding independent set can be used to color k-colorable graphs in the semirandom model of Blum and Spencer, greatly extending the range of paramaters for which a k-coloring is found. We also remark that a variant of Boppana's algorithm for graph bisection works not only for random models, but also for semirandom models. Again, this algorithm involves semidefinite programming. (Joint work with Joe Kilian from NEC Research Institute. Extended abstract appeared in proceedings of FOCS98.)

An approximation algorithm for the precedence constrained scheduling problem with hierarchical communications

Rodolphe Giroudeau

giroudea@buxtehude.lami.univ-evry.fr

We study the problem of minimizing the makespan for the precedence multiprocessor constrained scheduling problem with hierarchical communications. We propose an $\frac{8}{5}$-approximation algorithm for the UET-UCT (Unit Execution Time Unit Communication Time) hie-rarchical problem with an unbounded number of biprocessor machines. Moreover, we extend this result in the case where each cluster has m processors (where m is a fixed constant) by presenting an ρ-approximation algorithm where $\rho = (2 - \frac{2}{2m + 1})$.

Randomised Approximate Counting, Randomised Sampling, and Bounding Mixing Times

Leslie Ann Goldberg

leslie@dcs.warwick.ac.uk

This talk is concerned with approximate counting problems such as "Estimate the number of (proper) k-colourings of an n-vertex graph." The basic approach is the Markov chain Monte Carlo approach:

1. Define a Markov chain whose stationary distribution is uniform on k-colourings of G. 2. Use the Markov chain to sample k-colourings of G uniformly at random. 3. Use the samples to estimate the number of k-colourings, following the approach of Jerrum, Valiant and Vazirani.

The main difficulty is usually determining the mixing time of the Markov chain (that is, how long one must simulate the Markov chain in order to obtain an almost-uniform sample). In the talk, I will survey coupling-based methods for bounding mixing time, including the path coupling method of Bubley and Dyer and the delayed path coupling method of Czumaj, Kanarek, Kutylowski and Lory's. Then I will describe a new extension of path coupling which is joint work with Martin Dyer, Catherine Greenhill, Mark Jerrum, and Michael Mitzenmacher. We apply the new method to analyse the mixing time of the Glauber dynamics Markov chain for sampling graph colourings. We show that the mixing time is O(n log n) for triangle-free Delta-regular n-vertex graphs if at least (2 - eta) Delta colours are used, for a small positive constant eta. (This gives an optimal upper bound for the mixing time in this case.)

Width-free iteration bounds for structured optimization

Michael D. Grigoriadis

grigoria@aramis.rutgers.edu

For a given convex compact set B and convex vector function $f : B \rightarrow R_+^M$, $M \geq 1$, the *min-max* model is: $\lambda^* = \min\{\lambda \mid f(x) \leq \lambda e, x \in B\}$, where e is the M-vector of all ones. The *width* of B is the largest component of f over all $x \in B$. Iteration complexities of exponential potential Lagrangian decomposition approximation schemes for this general model depend on the width of B, requiring elaborate "width-reduction" procedures and solution of subproblems over further-restricted B. Here we discuss the first width-free approximation algorithm for such problems. Our method is a logarithmic potential-reduction, Lagrangian decomposition algorithm

that computes an $x \in B$: $f(x) \leq (1 - \varepsilon)\lambda^* e$, in $O(M(\varepsilon^{-2}\ln(1/\varepsilon) + \ln M))$ iterations, where $\varepsilon \in (0, 1]$ is a fixed relative tolerance, and each subproblem solved per iteration of this method is on the *original B*. We also give another logarithmic potential reduction approximation algorithm, requiring a separate analysis, for the *max-min* model, i.e., compute an $x \in B$: $f(x) \geq (1 - \varepsilon)\Lambda^* e$, where Λ^* is the unknown exact max-min value, for concave $f : B \rightarrow R_+^M$. This method runs in $O(M(\varepsilon^{-2} + \ln M))$ iterations, again with subproblems on the original B. Such methods are most advantageous when optimization over B is simpler than for the overall problem, such as when B is given as the product of many convex compact "blocks" and when f is block separable. Computational results for large concurrent multicommodity flow instances are presented. (Parts in collaboration with L. Khachiyan, L. Porkolab and J. Villavicencio.)

Approximation algorithms for packing and covering problems in discrete tomography

Peter Gritzmann

`gritzman@sungritzmann1.mathematik.tu-muenchen.de`

The talk deals with the problem of approximating binary images that are only accessible through few evaluations of their discrete X-ray transform, i.e., through their projections counted with multiplicity along some lines. This inverse discrete problem is motivated by demands from material sciences for the (approximate) reconstruction of crystalline structures from images produced by quantitative high resolution transmission electron microscopy. It belongs to a class of generalized set partitioning problems and allows natural packing and covering relaxations. For these (NP-hard) optimization problems we present various approximation algorithms, and provide estimates for their worst case performance. Further, we report on computational results for various variants of these algorithms. In particular, the corresponding integer programs are solved with only small absolute error for instances up to 250000 binary variables.

On approximating CSP-B

Johan Håstad

`johanh@nada.kth.se`

We prove that any constraint satisfaction problem where each variable appears a bounded number of times admits a nontrivial polynomial time approximation algorithm.

Computational Methods for Solving Semidefinite Relaxations

Christoph Helmberg
helmberg@zib.de

We survey various algorithmic approaches for solving semidefinite programs and investigate possibilities offered by these algorithms for exploiting structural properties present in many semidefinite relaxations of combinatorial optimization problems. In praticular, we discuss primal-dual interior point methods, a pure dual interior point approach suggested by Benson, Ye, and Zhang, and the spectral bundle method of Helmberg and Rendl. We point out some new approaches in nonlinear programming that model semidefinite variables as quadratic forms of factorizations.

Approximating minimum spanning sets in polymatroids

Stefan Hougardy
hougardy@sunbay.informatik.hu-berlin

We present a polynomial time approximation algorithm for finding a minimum spanning subset in a k-polymatroid. Our algorithm has a performance ratio of approximately $\ln k$ which is best possible for large k. It implies an $\ln k$ approximation algorithm for the minimum spanning subhypergraph problem. This problem includes as special cases the k-Steiner tree problem and the k-set cover problem.

Approximation hardness of sorting by reversals and the bounded depency paradigm

Marek Karpinski
marek@cs.uni-bonn.de

We study approximation hardness of the problem MIN-SBR of sorting a permutation by the minimum number of reversals. We prove, somewhat surprisingly, that this problem is hard to approximate (NP-hard by randomized reductions) within any constant factor less than some explicit threshold. This excludes an existence of a PTAS for this problem under usual assumptions, thus settling a question which was open for some time now. The proof method uses certain new explicit approximation hardness techniques for bounded dependency, and bounded degree optimization

problems. We discuss recent improvements on approximation ratios for MIN-SBR and the bounded dependency problems based on semidefinite and linear programming.

The MIN-SBR problem has been motivated and extensively studied in computational molecular biology. This problem connects also to the well known problem of sorting by prefix reversals (or, so called, pancake sorting). Our inapproximability result on MIN-SBR is in sharp contrast to its signed version for which efficient exact algorithms have been designed recently. (Joint work with P. Berman.)

Approximating Max-Cut on graphs of bounded degree

Michael Langberg
mikel@narkis.wisdom.weizmann.ac

We analyze the addition of a simple *placement* step to various known randomized approximation algorithms. Adding such a step to the semidefinite based approximation algorithm for the Max-Cut problem ([GW95]), we achieve an improved approximation ratio on graphs with bounded maximal degree. In particular, for graphs of maximal degree 3, we obtain an approximation ratio of at least 0.919.

Polynomial instances of the positive semidefinite matrix completion problem

Monique Laurent
monique@cwi.nl

The positive semidefinite matrix completion problem asks whether a partially specified matrix can be completed to a positive semidefinite matrix. This problem is an instance of the problem of testing feasibility of a semidefinite program and its complexity is not known.

When the graph corresponding to the specified entries is chordal, one can construct a positive semidefinite matrix completion or decide that none exists in polynomial time for rational data; the algorithm is combinatorial and uses only simple matrix computations. More generally, the problem is polynomially sovable for the graphs having a fixed minimum fill-in; the minimum fill-in being the minimum number of edges needed to be added to the graph in order to make it chordal. The complexity is not known for circuits although it can be shown to be polynomial in the real number model for series-parallel graphs.

The above results extend to the closely related distance matrix completion problem, where a distance matrix is a matrix whose entries can be realized as the pairwise Euclidean distances among a set of points in some space. This problem has applications to multidimensional scaling in statistics or the molecule conformation problem in chemistry.

TSP - based curve reconstruction

Kurt Mehlhorn
mehlhorn@mpi-sb.mpg.de

The input of a curve reconstruction problem is a finite sample V of an unknown curve γ and the task is to connect the points in V in the order in which they lie on γ. The aim is to find algorithms which come with a reconstruction and a performance guarantee: For curves from a certain class of curves and sample set V satisfying a certain sampling condition, the algorithm is guaranteed to find the reconstruction. The algorithm should also be efficient. We report on recent advances on the problem. The papers combine methods from computational geometry and combinatorial optimization.

General multiprocessor task scheduling: Approximate solutions in linear time

Lorant Porkolab
porkolab@doc.ic.ac.uk

We consider the problem of scheduling a set of n independent tasks on a fixed number of parallel processors, where the execution time of a task is a function of the subset of processors assigned to the task. We propose a fully polynomial approximation scheme that for any fixed $\epsilon > 0$ finds a preemptive schedule of length at most $(1+\epsilon)$ times the optimum in O(n) time. We also discuss the non-preemptive variant of the problem and present a linear programming based polynomial approximation scheme that computes an approximate solution of any fixed fixed accuracy in linear time. (Joint work with Klaus Jansen, University of Kiel, Germany.)

Semidefinite versus Linear Relaxations for graph problems: a computational view.

Franz Rendl
franz.rendl@uni-klu.ac.at

We consider the MAX-CUT, the k-EQUIPARTTION and the Traveling Salesman Problem. For all these problems both linear and semidefinite relaxations are known. We consider these relaxations with respect to quality and computational effort to obtain them, and come to the following observations.

(a) The semidefinite relaxation, perhaps combined with (some) triangle inequalities, is an extremly useful tool to get tight approximations for MAX-CUT.

(b) In case of partitioning the vertices of a graph into k (equal-sized) sets, so as to minimize the total weight of all edges cut by the partition, it depends on the relative size of k versus n, the number of vertices, whether LP is preferable over SDP. For large values of k, a simple LP produces tighter bounds much faster than SDP, while SDP is better for small values of k. A theoretical explanation for this is provided.

(c) It has recently been proposed to model the Traveling Salesman Problem (TSP), using the concept of algebraic connectivity. This leads to an SDP relaxation of TSP. It is shown, that this model is dominated by the classical linear model, based on the subtour elimination constraints.

In summary, the interplay between linear and semidefiniteness constraints should be analysed carefully to come to the right algorithmic machinery.

On the complexity of lattice problems

Shmuel Safra
safra@math.tau.ac.il

Various computational problems regarding Lattices have been investigated through the last 3(!) centauries. Dirichlet and Minkowsky implicitly sought an algorithm, in particular, for the shortest-vector problem (SVP) and the closest-vector problem (CVP). Still up to today the best known algorithm, of Lenstra,Lenstra,Lovasz [L cubed], approximates these problems only to within exponential factors.

The talk will survey recent developments regarding the computational complexity of approximating these Lattice problems, the relation between their complexity in the worst case and the average case (by Ajtai), and applications of their seeming hardness to cryptographic systems.

A State of the Art in Data Dissemination

Nicolas Schabanel
nschaban@ens-lyon.fr

Data dissemination (also named data broadcast, pushed-based system,...) is a client-server protocol that consists in broadcasting over an infinite time horizon, a set of information pieces (the messages) on some reserved channels so as to minimize the expected waiting time of a random client who is interested in one of the messages, connects at some random time (according to some Poisson process) and monitors the channels until the message he is interested in is broadcast. The input of the problem is the distribution of clients requests over the messages, from which the schedule is computed. This protocol was introduced in middle 80's (see seminal work of Ammar and Wong) to reduce the load of information servers and networks by allowing to get ride off the requests and then by handling with no extra cost an arbitrary number of simultaneous clients. This protocol is now used by some information servers on the Internet.

This problem is a quadratic optimization problem that is known to be NP-hard when broadcasting messages has a cost, or when messages have different lengths (with or without preemption). It is still unknown whether the uniform length case without broadcast cost is NP-hard or even belongs to NP (or P) (only exponential bounds are known on the period of an optimal schedule). In joint work with C. Kenyon and N. Young, we have recently exhibited a PTAS for this case (the best former ratio was 9/8 by Bar-Noy, Bhatia, Noar and Schieber, 1998). In joint work with C. Kenyon, we have also designed a first constant factor approximation for the non-uniform length case without preemption. We have also proposed a preemptive model for which we have a 2-approximation that improves by an arbitrary factor the expected wait of the clients when the messages have very different lengths.

Many facets of this problem are still widely unexplored: Is the non-uniform length case approximable to $1 + \epsilon$? How to design efficient client caching strategies in this context? How to index the broadcast to reduce client's tuning time?

Probabilistic & deterministic methods for counting sum-free sets

Tomasz Schoen
tos@numerik.uni-kiel.de

A subset $A \subseteq G$ of an (additively written) group G is said to be *sum-free* if no $a_1, a_2, a_3 \in A$ satisfy the equation $a_1 + a_2 = a_3$. Alon proved that the number of sum-free subsets of any group G of cardinality $n = |G|$ is at most $2^{(1/2+o(1))n}$ (as $n \to \infty$) and asked about the sharp form of this result. Lev, Łuczak and the author showed the following thoerem, which answers Alon's question for G abelian.

Theorem 1 *There is an absolute constant $\delta > 0$ such that the number of sum-free subsets of any abelian group G of cardinality $n = |G|$ is*

$$(2^{\nu(G)} - 1)\, 2^{n/2} + O(2^{(1/2-\delta)n}),$$

where $\nu(G)$ is the number of even order components in the canonical decomposition of G into a direct sum of its cyclic subgroups, and the implicit constant in the O-sign is absolute.

A constant approximation algorithm for the k-median problem

David Shmoys
shmoys@CS.Cornell.EDU

In the k-median problem, we are given a collection of n points in a metric space and a parameter k; we wish to select k of the points (as medians) and assign each of the n points to its closest median, so as to minimize the total distance of these assignments. We give the first approximation algorithm for this problem with a constant performance guarantee (which is, in fact, equal to 20/3). The algorithm is based on solving a linear programming relaxation of the problem, and then rounding the optimal fractional solution to an integral one while increasing the cost by only this constant factor. (Joint work with Moses Charikar, Sudipto Guha, and Eva Tardos.)

Cooperative facility location games

Martin Skutella
skutella@math.TU-Berlin.DE

The location of facilities in order to provide service for customers is a well-studied problem in the operations research literature. In the basic model, there is a predefined cost for opening a facility and also for connecting a customer to a facility, the goal being to minimize the total cost. Often, both in the case of public facilities (such as libraries, municipal swimming pools, fire stations, ...) and private facilities (such as distribution centers, switching stations, ...), we may want to find a 'fair' allocation of the total cost to the customers — this is known as the cost allocation problem. A central question in cooperative game theory is whether the total cost can be allocated to the customers such that no coalition of customers has any incentive to build their own facility or to ask a competitor to service them.

We establish strong connections between fair cost allocations and linear programming relaxations for several variants of the facility location problem. In particular, we show that a fair cost allocation exists if and only if there is no integrality gap for a corresponding linear programming relaxation. We use this insight in order to give proofs for the existence of fair cost allocations for several classes of instances based on a subtle variant of randomized rounding. This also leads to polynomial-time algorithms to solve the facility location problem in these cases. We also prove that it is in general NP-complete to decide whether a fair cost allocation exists and whether a given allocation is fair. (This is joint work with Michel X. Goemans.)

Approximating a weighted interval scheduling problem

Frits Spieksma
Spieksma@MATH.UNIMAAS.NL

We study the following problem. The input consists of jobs, each of which is given as a set of intervals on the real line, and a number m of machines. Each interval i has a positive weight w(i). A feasible solution is a subset of the given intervals such that (1) at most 1 interval is selected from each job, and (2) for any point p on the real line, at most m intervals overlapping p are selected. The goal is to find a solution that maximizes the sum of the weights of the selected intervals. We consider three variants of this problem: i) w(i) = 1 for all intervals i, ii) equal weights per job, iii) arbitrary weights. We present a greedy algorithm that achieves, in case m=1, performance ratios of 1/2, 3-2sqrt(2), 1/8 for variants i), ii) and iii) respectively. We also show that within a special class of algorithms (to which this greedy algorithm belongs) no algorithm can do better than 1/2, 3-2sqrt(2), 0.141 for variants i), ii) and iii) respectively. (Joint work with Thomas Erlebach.)

Circle Method and Discrepancies

Anand Srivastav
asr@numerik.uni-kiel.de

In this article the combinatorial discrepancy of the hypergraph $\mathcal{H}$ of cartesian products of d arithmetic progressions in the $[N]^d$–lattice is studied ($[N] = \{0, 1, \ldots, N-1\}$). For simplicity, we call such a hyperedge a d–dimensional arithmetic progression. Such higher dimensional arithmetic progressions are motivated by a high-dimensional version of van der Waerden's theorem, namely the Gallai-Witt-theorem (1990).

We resolve the discrepancy problem for d–dimensional arithmetic progressions by proving $(\mathcal{H}) = \Theta(N^{\frac{d}{4}})$ for every fixed integer $d \geq 1$. This extends the famous lower bound of $\Omega(N^{1/4})$ of Roth (1964) and the optimal upper bound $O(N^{1/4})$ of Matoušek and Spencer (1994) from $d = 1$ to arbitrary, fixed d. To establish the lower bound we use harmonic analysis on locally compact abelian groups. For the upper bound a product coloring arising from the theorem of Matoušek and Spencer is sufficient. Finally, some special cases, e.g. symmetric arithmetic progressions and infinite arithmetic progressions, are in vestigated. Extensions to multi-color discrepancy are discussed as well. (Joint work with Benjamin Doerr (UNiversity of Kiel) and Petra Wehr (SAP AG Duesseldorf).)

Balanced Allocations: The Heavily Loaded Case

Angelika Steger

`steger@informatik.tu-muenchen.de`

Balls and bins algorithms using multiple choices for each ball have caused furor in recent years. In contrast to the classical balls and bins game in which each ball is placed into a randomly selected bin, these algorithms consider several random locations for each ball and place the ball into the bin with minimal load.

Almost all previous analyses focus on the lightly loaded case, i.e., $m = O(n)$. In this case, the multiple-choice algorithms improve greatly upon the classical one-choice algorithm. The known results for the heavily loaded case, i.e., $m \gg n$, however, were not impressive at all. Indeed, here the best known result was that the fullest bin contains $\frac{m}{n} + O(\sqrt{m \cdot \ln n / n})$ balls, which just corresponds to the maximum load produced by the classical one-choice algorithm.

In this talk, we show that the actual distribution produced by the multiple-choice algorithms is much better than this prediction. We explicitely calculate the distributions produced by different multiple-choice algorithms. For example we show that for the sequential two-choice algorithm the deviation of the fullest bin from the average is at most $(\ln \ln n)/\ln 2 + \Theta(1)$, independently of the number of balls. (Joint work with Petra Berenbrink, Artur Czumaj and Berthold Vöcking.)

On the properties of basic solutions of linear relaxation for MAX DI-CUT with given sizes of parts

M. Sviridenko

`sviri@brics.dk`

In this talk we show that the straightforward LP relaxation of the MAX DICUT with given sizes of parts satisfies a very strong structural property. We prove that all variables corresponding to vertices in an input graph in any basic optimal solution of relaxation can have only five possible values 0, e, 1/2, 1-e, 1 where $0 < e < 1/2$ is some number depending on an input graph and a basic solution. This theorem provides a theoretical ground for constructing an 1/2-approximation algorithm for MAX DICUT with given sizes of parts. (Joint work with A. Ageev and R. Hassin.)

Randomized Metarounding

Santosh Vempala
vempala@math.mit.edu

Let P be a linear relaxation of an integer polytope Z such that the integrality gap of P w.r.t. Z is at most r, as verified by a poly-time heuristic A, which on any cost function c returns an integer solution (extreme point of Z) whose cost is at most r times the optimal cost over P. Then for any point x^* in P (fractional solution), rx^* dominates some convex combination of extreme points of Z. A constructive version of this theorem is presented with applications to approximation algorithms.

Computer assisted analysis of approximation algorithms for MAX 3-SAT and other problems

Uri Zwick
zwick@math.tau.ac.il

We present an exact and rigorous analysis of several semidefinite programming based approximation algorithms for problems such as MAX 3-SAT. One of the main tools used in the computerized analysis is interval arithmetic. We also mention several other problems that were recently resolved using similar means, such as the Kepler conjecture and the double bubble conjecture.

4 Open problems

Some open problems (by Marek Karpinski)

1. Define MIN-2SAT as the search problem for an assignment satisfying the minimum number of clauses of a given 2SAT instance. It is known that the bounded occurrence instances of MIN-2SAT, like the 3-OCC-MIN-2SAT, are APX-complete. Their first explicit inapproximability ratios were computed recently. On the other hand, define PAIRED MIN-BISECTION problem for a given graph, and a list of pairs of its vertices, as a problem of computing a minimum bisection of a graph which separates all the paires of vertices in a list. It is known that PAIRED MIN-BISECTION is approximation hard for (3-OCC) MIN-2SAT, and hence it does not have a PTAS. The question of an existence of a PTAS for MIN-BISECTION is still wide open. Is there any argument for showing that MIN-BISECTION is also approximation hard (say, under randomized GP reductions?) for 3-OCC-MIN-2SAT, thus proving approximation hardness of MIN-BISECTION to within a certain constant, and excluding an existence of a PTAS for this problem under usual assumptions ?

2. It is well known that planar MAX-CUT can be solved exactly in polynomial time (this can also be achieved by the SDP method by looking at its cuts-polyhedron). It is also well known that planar MIN-CUT can be solved exactly in polynomial time for arbitrary graphs. Is MIN-BISECTION problem for planar graphs also solvable exactly in polynomial time? What is its approximation complexity? Is it possible to design a PTAS for that problem? How about MAX-BISECTION for planar graphs ? Can the SDP and a careful analysis of its bisections-polyhedron help?

3. k-MIS is the problem of computing a maximum independent set in a graph of degree bounded by k. The only known method of proving inapproximability of the problem of Sorting by Reversals uses an approximation preserving reduction from 4-MIS. The best known up to date approximation lower bound for 4-MIS is 1.0136 and its approximation upper bound is 1.40 (the corresponding bounds for 3-MIS are: 1.0071 and 1.20). An interesting open question is to improve both bounds and to close, respectively, huge gaps between them. Any improvement in approximation ratios for 4-MIS will shed some light also on possible improvements towards the problem of Sorting by Reversals. Can SD programs be used for k-MIS problems to achieve better approximation ratios than the best known to date approximation algorithms? In particular, can the approximation ratio of 1.40 be improved for the 4-MIS problem? How about corresponding Coloring problems for graphs of degree bounded by k?

Touching numbers and equidistant sets (by Monique Laurent)

Given a k-dimensional convex body K, its touching number $t(K)$ is defined as the maximum number of translates of K that are pairwise touching; two convex bodies being touching if they meet but their relative interiors are disjoint. The convex body K can be assumed to be centrally symmetric, since its touching number is equal to the touching number of its symmetrization $K - K$. Then, $t(K)$ is equal to the maximum number of equidistant points in the associated normed space with K as unit ball.

It is known that $t(K) \leq 2^k$, with equality when K is the cube. Moreover, $t(K) = k + 1$ when K is the Euclidean ball. It is an open question to determine the touching number $t(\beta_k)$ of the cross-polytope β_k (i.e., the unit ball for the ℓ_1-norm). It is conjectured that $2k$ is the right value; it is obviously a lower bound (since the ± 1 unit vectors are at pairwise ℓ_1-distance 2). The conjecture has been shown to hold when $k \leq 3$ (by Bandelt, Chepoi and Laurent in Discrete and Computational Geometry (19) 1998) and recently for $k = 4$ (by Koolen, Laurent, and Schrijver, article to appear).

This conjecture is also related to the following problems: Determine the maximum number $h(k)$ of equidistant points lying in the hyperplane $\sum_{i=1}^k x_i = 0$; determine the touching number $t(\alpha_k)$ of the k-dimensional simplex; determine the maximum cardinality $a(n)$ of an antichain in a design on n points. Namely, the following holds (cf. [KLS]): $h(k) = t(\alpha_{k-1})$; $h(k) \geq n \iff a(n) \leq k$; $a(n) \geq 2k \implies t(\beta_k) \leq n$; $t(k) = k$ for $k \leq 3$ and $t(k) \geq k + 1$ for $k \geq 4$; $a(n) = n$ for $n \leq 4$ and $a(n) \leq n - 1$ for $n \geq 5$; $a(5) = 4$; $a(6) = 5$; $h(4) = t(\alpha_3) = 5$.

Some open questions in Data Dissemination (by Nicolas Schabanel)

Data dissemination consists in: given a distribution (p_i) over a finite set of messages $M_1, \ldots, M_m$, design a infinite time horizon schedule of the broadcasts of the messages on C channels so as to minimize the expected waiting time of a random client who connects at some random time (according to some Poisson process), requests a random message according to the distribution (p_i), and monitors the channels until the messages he is interested in is broadcast. Many facets of this problem are widely unexplored, here is a list of some interesting open questions.

Problem 1. The basic case consists of unit length messages to be broadcast over discrete time synchronized channels. Which is the exact complexity of this case? Is it NP-hard? Is it in NP? in P? Only exponential upper bounds are known on the period of an optimal schedule (Ammar and Wong, 1987; Anily, Glass et Hassin, 1995). Recently a PTAS has been designed for this case (Kenyon, Schabanel and Young, 2000), can one design an exact solution?

Problem 2. Constant factor approximation have been designed for the non-uniform length case (Kenyon and Schabanel, 1999; Schabanel 2000). Is this case approximable to $(1 + \epsilon)$?

Problem 3. Imagine that clients now request not only one message, but some subsets of the set of messages. Can we design constant factor approximation? Is it still tractable?

Problem 4. An interesting question consists in studying client caching strategies. This problem is really different from classic memory caching strategies where the cost of a cache miss is constant and only depends on the state of the cache. Here, the cost of a cache miss not only depends on the cache state but also on its occurrence date and on the broadcast schedule: it is the length of the period of time until the next broadcast of the missing message. Only very particular cases of this variant have been studied yet (Khanna and Liberatore, 1998). The tractability of this variant is still an open question.

The 3-colourability problem (by Ingo Schiermeyer)

The 3-colourability problem is a well-known NP-complete problem. Let 3-col(F) denote the 3-colourability problem restricted to the class of all graphs containing no induced copy of a prescribed graph F, and let P_n denote the induced path on n vertices. Recently, B. Randerath, M. Tewes and I. Schiermeyer obtained a polynomial time algorithm to decide 3-colourability for 3-col(P_5).

Question 1: Can 3-col(P_6) be (also) decided in polynomial time?

Question 2: Does there exist an integer $k \geq 6$ such that 3-col(P_k) remains NP-complete?

Job shop problem (by Maxim Sviridenko)

Job shop problem is one of the most known scheduling problems. In this problem there are few important parameters: m – a number of machines, s – a maximum number of operations per job and n – a number of jobs. It is known that the makespan minimization problem has a PTAS if both m and s are bounded by a constant. If both parameters are part of the input then best known approximation algorithm has a polylogarithmic (in m and s) performance guarantee, it is also known that there is no approximation algorithm with performance guarantee better than 5/4 unless P=NP. First question is to close the gap between positive and negative results for this problem, the second question is to study problem with one fixed

parameter (m or s), the best known algorithms have performance guarantees m and $s^{0.5}$, respectively. From negative viewpoint we only know that these problems are both strongly NP-hard even in very restricted settings.