**Theory Seminar - Fall 2015**

**Venue: 4403 SC, Monday 10 A.M. - 11 A.M.**

**Organizer - Vivek Madan, vmadan2 [at] illinois [dot] edu**

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker</th>
<th>Title</th>
<th>Slides</th>
<th>Additional Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>08/31/2015</td>
<td>Vivek Madan</td>
<td>Simple and Fast Rounding Algorithms for Directed and Node-weighted Multiway Cut</td>
<td>Multi-way-Cut.pptx</td>
<td>Multi-way Cut</td>
</tr>
</tbody>
</table>

In Multiway Cut problem, input is an edge/node weighted graph and a set of k terminals; the goal is to remove min-weight subset of edges/nodes such that there is no path between any two terminals. Directed Multiway cut (Dir-MC) admits 2-approximation and Node-weighted Multiway cut (Node-MC) admits $2(1-1/k)$ approximation. Previous rounding algorithms for these problems are based on careful rounding of an "optimum" solution to an LP relaxation. We describe extremely simple and near linear-time rounding algorithms with same approximation ratio via a natural distance based LP relaxation. Additionally, we also prove that integrality gap of the LP relaxation for Dir-MC is 2 even in directed planar graphs with two terminals ($k=2$).

If you are around and would like to speak in the seminar, please drop me an email. To get emails about weekly seminar, please subscribe to Theory-Seminar Mailing List.

For registered Students: >75% attendance to get pass grade.
Submodular function minimization is a fundamental optimization problem that arises in several applications in machine learning and computer vision. The problem is known to be solvable in polynomial time, but general purpose algorithms have high running times and are unsuitable for large-scale problems. Recent work have used convex optimization techniques to obtain very practical algorithms for minimizing functions that are sums of "simple" functions. In this talk, we give two algorithms based on random coordinate descent with faster linear convergence rates and cheaper iteration costs.
We prove the first nontrivial worst-case lower bounds for two closely related problems. First, $\Omega(n^{3/2})$ degree-1 reductions, series-parallel reductions, and Y transformations are required in the worst case to reduce an $n$-vertex plane graph to a single vertex or edge. The lower bound is achieved by any planar graph with treewidth $\Theta(\sqrt{n})$. Second, $\Omega(n^{3/2})$ homotopy moves are required in the worst case to reduce a closed curve in the plane with self-intersection points to a simple closed curve. For both problems, the best upper bound known is $O(n^2)$, and the only lower bound previously known was the trivial $\Omega(n)$.
The first lower bound follows from the second using medial graph techniques ultimately due to Steinitz, together with more recent arguments of Noble and Welsh [J. Graph Theory 2000]. The lower bound on homotopy moves follows from an observation by Haiyashi et al. [J. Knot Theory Ramif. 2012] that the standard projections of certain torus knots have large defect, a curve invariant introduced by Aicardi and Arnold. Finally, we prove that every closed curve in the plane with $n$ crossings has defect $O(n^{3/2})$, which implies that better lower bounds for our algorithmic problems will require different techniques.
Given a point set in high dimensional feature space, it is natural to define a similarity measure between points that allow ignoring a certain number of coordinates when computing the distance between two points. Thus, two entities are close if the points encoding them "agree" on almost all features, except for some "few" noisy coordinates.

We show how to build a data-structure that can answer quickly such approximate nearest-neighbor queries. The main technical difficulty here is that the coordinates being ignored, when computing the similarity between points, are a function of the two points being compared.
<table>
<thead>
<tr>
<th>Date</th>
<th>Presenter</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/05/2015</td>
<td>Chandra Chekuri</td>
<td>The Nagamochi-Ibaraki Algorithm for Global Mincut in Graphs and Beyond</td>
</tr>
</tbody>
</table>

Nagamochi and Ibaraki in the early 90's came up with a beautiful algorithm for finding the global mincut of an "undirected" graph. Queyranne extended the algorithmic idea to find the global mincut of a symmetric submodular function $f$. We will describe this extremely simple algorithm and its analysis and finish with some open problems on finding the global mincut of a hypergraph that the speaker is excited about recently.

<table>
<thead>
<tr>
<th>Date</th>
<th>Presenter</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/09/2015</td>
<td>Venkatesan Guruswami</td>
<td>$(2+\epsilon)$-SAT is NP-hard, and related results</td>
</tr>
</tbody>
</table>

Given a $k$-SAT instance with the promise that there is an assignment satisfying at least $t$ out of $k$ literals in each clause, can one efficiently find a satisfying assignment (setting at least one literal to true in every clause)?

Extensions of some 2-SAT algorithms solve this problem when $t \geq k/2$. We prove that for $t < k/2$, the problem is NP-hard. Thus, SAT becomes hard when the promised density of true literals falls below 1/2. One might thus say that the transition from easy to hard in 2-SAT vs. 3-SAT takes place just after two and not just before three.
A strengthening of this result shows that given a \((2k+1)\)-uniform hypergraph that can be 2-colored such that each hyperedge has near-perfect balance, it is NP-hard to even find a 2-coloring that avoids a monochromatic edge. This shows extreme hardness of discrepancy minimization for systems of bounded-size sets.

The talk will sketch the proof of the SAT result, which is based on the fact that the only functions passing a natural "dictatorship test" are "juntas" depending on few variables. We will also elucidate the general principle, based on the paucity of "weak polymorphisms," that underlies tractability when the notion of approximation is in the kind of predicates satisfied in the Yes and No instances.

Based on joint work with Per Austrin and Johan Håstad. We might also briefly mention some ongoing work (with Joshua Brakensiek) on using the weak polymorphism principle to prove hardness of coloring graphs with very small chromatic number.
<table>
<thead>
<tr>
<th>Date</th>
<th>Author</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/12/2015</td>
<td>Patrick Lin</td>
<td>Discrete Stochastic Submodular Maximization</td>
</tr>
<tr>
<td></td>
<td></td>
<td>We consider the problem of stochastic monotone submodular function maximization, subject to state-independent constraints. This problem is motivated by problems in application areas such as machine learning, social networks, and recommendation systems. Algorithms for such problems work in an on-line setting, and can be adaptive. We give a result on the adaptivity gap, which is the ratio between the optimal adaptive and non-adaptive solutions, by presenting a procedure that transforms a decision tree (adaptive algorithm) into a non-adaptive chain.</td>
</tr>
<tr>
<td>10/19/2015</td>
<td>Kent Quannud</td>
<td>Approximation Algorithm for Polynomial Expansion and Low-Density Graphs</td>
</tr>
</tbody>
</table>
We investigate the family of intersection graphs of low density objects in low dimensional Euclidean space. This family is quite general, includes planar graphs, and in particular is a subset of the family of graphs that have polynomial expansion.

We present efficient $(1+\eps)$-approximation algorithms for polynomial expansion graphs, for Independent Set, Set Cover, and Dominating Set problems, among others, and these results seem to be new. Naturally, PTAS's for these problems are known for subclasses of this graph family. These results have immediate interesting applications in the geometric domain. For example, the new algorithms yield the only PTAS known for covering points by fat triangles (that are shallow).

We also prove corresponding hardness of approximation for some of these optimization problems, characterizing their intractability with respect to density. For example, we show that there is no PTAS for covering points by fat triangles if they are not shallow, thus matching our PTAS for this problem with respect to depth.
In this talk, we discuss the strong exponential time hypothesis (SETH), an unproven assumption about the hardness of \( k \)-SAT. SETH has become popular recently as a way of proving conditional lower bounds for problems like the edit distance, Frechet distance, Hopcroft's problem, and independent set. This talk will have two parts. In the first part, we prove some lower bound results relying on SETH by describing reductions from \( k \)-SAT to various problems; many of these lower bounds match the best-known upper bounds. In the second part of the talk, we describe one possible limitation of this approach by introducing the nondeterministic strong exponential time hypothesis (NSETH). If NSETH turns out to be true, then for many problems, SETH is insufficient for proving lower bounds.
Denoising and inference algorithm usually rely on a generative noise model or a number of parameters. In this talk, I will introduce sampling conditions that account for outliers and are satisfied with high probability if the sampling is obtained by classical generative models. I will present an algorithm based on the distance to a measure which denoises datasets requiring only one parameter, which is the minimum number of required parameter if no further assumptions are made. This algorithm serves as a building block for a parameter free denoising algorithm of uniform noisy samples with outliers.

These theoretically guaranteed algorithms can then be used for topological inference for example.
A faster Pseudo polynomial time algorithm for Subset Sum

We introduce a faster pseudopolynomial time algorithm for the classical subset sum problem: Given a set of $n$ positive integers $S$ and an integer target value $t$, decide if there is a subset of $S$ that sums to $t$, in $\tilde{O}(n^{1/2}t)$ time, where $n$ is the size of $S$. In fact, we answer a more general question than that, we compute this for all target numbers $t$ in $\tilde{O}(\min(n^{1/2}t, u^{4/3}, u))$, where $u$ is the sum of all elements in $S$. Our algorithm improves on the textbook $O(nu)$ dynamic programming algorithm, and as far as we know, is the fastest general algorithm for the problem when $n$ is $(\log^c u)$ for some constant $c$. It use a simple idea to speeding up the naive divide-and-conquer algorithm.
In this work, we consider the question of improving the rate of split-state non-malleable codes. We focus on the standard computational setting since it is not possible to go beyond rate 1/2 otherwise. In this setting, each tampering function must be efficiently computable, and the message in the tampered codeword is required to be either the original message $m$ or a "computationally independent" one. Assuming only the existence of one-way functions, we present a compiler which converts any poor rate, two-state, (sufficiently strong) non-malleable code into a rate 1, two-state, computational non-malleable code. Furthermore, for the qualitative optimality of our result, we show that the existence of one-way functions is necessary to achieve rate $> 1/2$ for such codes (generalizing a result of Cheraghchi and Guruswami in ITCS 2014).
Approximation Algorithms for the Bregman k-Median Problem

In this talk, we discuss the k-median problem with respect to a dissimilarity measure $D$ from the family of Bregman divergences: Given a finite set $P$ of size $n$ from $\mathbb{R}^d$, the goal is to find a set of centers $C$ of size $k$ such that the sum of distances (bregman) to the closest center is minimized. This problem plays an important role in applications from many different areas of computer science, such as information theory, statistics, data mining, and speech processing.

We will see a fast, practical, randomized approximation algorithm that computes an $O(\log k)$-approximate solution for arbitrary input instances. In the second half of the talk, we will focus on coreset construction for a large subclass of Bregman divergences. In a nutshell, a coreset is a small (weighted) set that has the same clustering behavior as the original input set. Using these coresets, we will show an $(1 + \epsilon)$-approximation algorithm for the Bregman k-median problem.
Following the latest hype, in this talk we present the Graph Isomorphism (GI) problem, its importance in both theory and practice, and the "practical" difficulty of designing general algorithms for the problem. We go over milestone past results and briefly touch upon the latest claimed result by Babai. The main part of the talk is demonstrating Spielman's result on isomorphism of strongly regular graphs, a class of graphs once considered to be the "bane" class of any general algorithm. His approach is to analyze the individualization and refinement algorithm in light of Neumaier's claw bound, which implies that low degree strongly regular graphs have a small second-largest eigenvalue, unless they are Steiner or Latin square graphs.

(2+\eps)-SAT is NP-hard, and related results

Recent space activity

- Kolliaris, Konstantinos
- Theory Seminar - Fall 2016

Space contributors

- Madan, Vivek (1142 days ago)
- Julieanne Chapman (1247 days ago)