Report for workshop Approximation Algorithms and Parameterized Complexity (15w5118)

The workshop "Approximation Algorithms and Parameterized Complexity" was organized in cooperation by Michael Fellows, Klaus Jansen, Hadas Shachnai, and Roberto Solis-Oba. Within the 5-day-workshop from the 29th of November until the 4th of December 2015 18 talks were given in Banff center concentrating on the following subjects.

NP-hard problems are optimization problems that are notoriously difficult to solve. They are in fact so difficult that the only known solutions for them require impractically large amounts of time even on the fastest existing computers. There are thousands of these problems arising from a very rich and varied number of applications.

Two approaches that have been independently, but very successfully used to deal with NP-hard problems are approximation algorithms and parameterized algorithms. Approximation algorithms are efficient methods which do not necessarily find optimum solutions; yet, they do guarantee that the output solution achieves a bounded ratio to the optimum. Parameterized algorithms identify and exploit properties of a problem that make it hard to solve to produce efficient solutions for many instances. This workshop explored relationships between techniques used in the design of approximation and parameterized algorithms to gain a better understanding of what makes a problem difficult to solve, with the aim of developing better tools for tackling NP-hard problems.

There are a number of ways in which parameterization and approximation interact mathematically, and ways in which the two approaches could be fruitfully combined to obtain better tools for handling and understanding the computational complexity of hard optimization problems. Some of these have barely been explored; thus, a main goal of the workshop was to provide a venue for experts in parameterized and approximation algorithms to come together and discuss them, furthering our understanding of the complexity of NP-hard optimization problems. Some of the possible ways in which both fields can collaborate and were confirmed by this workshop are listed below.

- EPTAS versus PTAS
- Lower bounds via ETH
- Kernels for approximation algorithms
- Specifically parameterized forms of approximation
- Aggregate parameterizations including approximation
- Parameterized approximations for weighted problems

We had a collection of survey talks, e.g. by Mohammad Taghi Hajiaghayi, by Stefan Kratsch, Mike Fellows, and Klaus Jansen.

- MohammadTaghi Hajiaghayi (University of Maryland at College Park)
  Fixed-Parameter Tractability and Approximability: A Survey of Connections
  In his talk he discussed briefly classes of fixed-parameter tractability as well as approximation algorithms and he surveyed several connections between the two areas in terms of both results and approaches.

- Stefan Kratsch (University of Bonn)
  A brief introduction to kernelization
  Kernelization is a notion from parameterized complexity that captures the concept of efficient preprocessing for NP-hard problems. A kernelization is a polynomial-time algorithm that given an instance (x, k) with parameter k will return an equivalent instance of size bounded only in terms of k. In particular, we were interested in polynomial kernels where the bound depends polynomially on k.

  The talk gave an introduction to core concepts from kernelization. Relations to approximability of the considered problems were discussed.

- Michael Fellows (Charles Darwin University)
  Using Parameterization to Move Approximation into Problem Legislation
  The talk gave a few examples of moving approximation concerns into the definition of parameterized problems — into the modeling of the problem! Which is where, considering the nature of worst-case asymptotic complexity analysis, approximation often realistically belongs. The talk pointed to some large horizons for this approach.

- Klaus Jansen (University of Kiel)
  Lower bounds on the running time for packing and scheduling problems
  Klaus presented lower bounds on the running time for both exact and approximation algorithms based on the exponential time hypothesis (ETH). Then we discussed lower and upper bounds on the running time for exact algorithms for subset sum, partition, knapsack, bin packing, and scheduling on identical machines. Next he gave lower bounds on the running time of approximation schemes for the multiple knapsack, multi-dimensional knapsack and scheduling problem on identical, uniform, and unrelated machines.

The other contributed talks and discussions can be divided in the following research areas.

Packing and Scheduling

Sebastian Berndt (University of Lübeck) studied together with K. Jansen and K.-M. Klein the fully dynamic bin packing problem, where items arrive and depart in an online fashion and repacking of previously packed items is allowed. The goal is to minimize both the number of bins used as well as the amount of repacking. They measured the repacking by the migration factor, defined as the total size of repacked items divided by the size of an arriving or departing item. If one wishes to achieve an asymptotic competitive ration of $1 + \epsilon$ for the number of bins, a relatively simple argument proves a lower bound of $\Omega(1/\epsilon)$ for the migration factor. They established a nearly matching upper bound of $O((1/\epsilon)^4 \log(1/\epsilon))$ using a new dynamic rounding technique and new ideas to handle small items in a dynamic setting such that no amortization is needed. The running time of their algorithm is polynomial in the number of items n and in $1/\epsilon$. The previous best trade-off was for an asymptotic competitive ratio of 5/4 for the bins (rather than $1+\epsilon$) and needed an amortized number of $O(\log n)$ repackings (while in the new scheme the number of repackings is independent of n and non-amortized).

Felix Land (University of Kiel) presented (joint work with Klaus Jansen) a fully polynomial $(\frac{3}{2} + \epsilon)$-approximation for scheduling monotone moldable jobs. A moldable job is a job that can be executed on an arbitrary number of processors, and whose processing time depends on the number of processors allotted to it. He considered the problem of scheduling monotone moldable jobs to minimize the makespan. Most existing approximation algorithms have running time polynomial in the number n of jobs and the number m of processors. He argued that for compact input encodings, such running times are actually exponential in the input size, and that a fully polynomial algorithm has a running time polynomial in n and $\log m$. The best known approximation algorithm with such a running time was due to Mouni, Rapine, and Trystram and achieved approximation ratio $\sqrt{3} + \epsilon \approx 1.73$. Another algorithm, also due to Mouni, Rapine, and Trystram, had approximation ratio $(\frac{3}{2} + \epsilon)$, but had running time $\mathrm{O}(nm)$. He described different techniques to improve the running time of the latter to polynomial in n and $\log m$. In particular, they showed how to solve a knapsack problem with n items and capacity m in time $\mathrm{O}(\frac{n^2}{\epsilon} \log \epsilon m)$ when items larger than $b = \Theta(\frac{1}{\epsilon})$ can be compressed by a factor $1 - \Theta(\epsilon)$. For their scheduling problem, the compression increases the makespan by a factor of $1 + \epsilon$, and they expect a wide applicability of their techniques.
Furthermore, they proved that scheduling monotone moldable jobs to minimize the makespan is strongly NP-hard, which was previously known only for the variant without monotony.

Kati Land (University of Kiel) estimated the makespan of the two-valued Restricted Assignment Problem. Together with Klaus Jansen and Marten Maack she considered a special case of the scheduling problem on unrelated machines, namely the Restricted Assignment Problem with two different processing times.

They showed that the configuration LP has an integrality gap of at most $5/3 \approx 1.667$ for this problem. This allowed to estimate the optimal makespan within a factor of $5/3$, improving upon the previously best known estimation algorithm with ratio $11/6 \approx 1.833$ due to Chakrabarty, Khanna, and Li.

Monaldo Mastrolilli (IDSIA Instituto Dalle Molle di Studi sull'Intelligenza Artificiale) described research on a Lasserre Lower Bound for the Min-Sum Single Machine Scheduling Problem in joint work with Adam Kurpisz and Samuli Leppanen. The Min-sum single machine scheduling problem (denoted $1||\sum f_j$) generalizes a large number of sequencing problems. The first constant approximation guarantees have been obtained only recently and are based on natural time-indexed LP relaxations strengthened with the so called *Knapsack-Cover* inequalities (see Bansal and Pruhs, Cheung and Shmoys and the recent $(4+\epsilon)$-approximation by Mestre and Verschae). These relaxations have an integrality gap of 2, since the Min-knapsack problem is a special case. No APX-hardness result is known and it is still conceivable that there exists a PTAS. Interestingly, the Lasserre hierarchy relaxation, when the objective function is incorporated as a constraint, reduces the integrality gap for the Min-knapsack problem to $1+\epsilon$. In their paper they studied the complexity of the Min-sum single machine scheduling problem under algorithms from the Lasserre hierarchy. They proved the first lower bound for this model by showing that the integrality gap is unbounded at level $\Omega(\sqrt{n})$ even for a variant of the problem that is solvable in $O(n \log n)$ time by the Moore-Hodgson algorithm, namely Min-number of tardy jobs. They considered a natural formulation that incorporates the objective function as a constraint and proved the result by partially diagonalizing the matrix associated with the relaxation and exploiting this characterization.

Nicole Megow (TU Berlin) talked about an $O(\log m)$-Competitive Algorithm for Online Machine Minimization which was her joint work with Lin Chen and Kevin Schewior. They considered the online machine minimization problem in which jobs with hard deadlines arrive online over time at their release dates. The task was to determine a feasible preemptive schedule on a minimum number of machines. Their main result was an $O(\log m)$-competitive algorithm, with m being the optimal number of machines used in an optimal offline solution. This was the first improvement on an intriguing problem in nearly two decades. To date, the best known result is a $O(\log p_{\min}/p_{\max})$-competitive algorithm by Phillips et al. (STOC 1997) that depends on the ratio of maximum and minimum job sizes, $p_{\max}$ and $p_{\min}$. Even for $m = 2$ no better algorithm was known. Their algorithm is in this case constant-competitive. When applied to laminar or agreeable instances, their algorithm achieves a competitive ratio of $O(1)$ even independently of m. The following two key components lead to their new result. Firstly, they derived a new lower bound on the optimum value that relates the laxity and the number of jobs with intersecting time windows. Then, they designed a new algorithm that was tailored to this lower bound and balanced the delay of jobs by taking the number of currently running jobs into account.

Frits Spieksma (K.U. Leuven) presented his joint work with Annette Ficker and Gerhard Woeginger on the topic of the so-called balanced optimization with vector costs. They proposed a framework containing such problems; this framework allowed them to investigate the complexity and approximability of these problems in a general setting. More concretely, each problem in the framework admitted a 2-approximation, and for many problems within the framework this result was best-possible, in the sense that having a polynomial-time algorithm with a performance ratio better than 2 would imply P=NP. Special attention was paid to the balanced assignment problem with vector costs: they showed that the problem remains NP-hard even in case of sum costs.

Andreas Wiese (MPI für Informatik) found with Anna Adamaszek and Giorgi Nadiradze better approximation guarantees for geometric packing problems. A common setting in geometric packing problems is that we are given a set of two-dimensional items, e.g., rectangles, and a rectangular container and the goal is to pack these items or a subset of them items into the container to optimize objective functions like the total profit of the packed items or the necessary height of the container. A typical obstacle in these problem settings is that in the input there are different types of items, i.e., items that are wide and thin, that are high and narrow, or items that are large in both dimensions. Andreas Wiese presented a method to handle this obstacle. In a nutshell, the key was to prove that there are near-optimal solutions in which the given container can be partitioned into few rectangular boxes such that in each box there are only items of one of the mentioned types. This leads to better approximation guarantees for two specific problems: a $(1+\epsilon)$-approximation algorithm in quasi-polynomial time for the two-dimensional knapsack problem and a $(1.4+\epsilon)$-approximation algorithm in pseudo-polynomial time for the strip-packing problem. Note that the latter bound is strictly smaller than the lower bound of 3/2 that holds for (non-pseudo-)polynomial time algorithms for the problem.

Guochuan Zhang (Zhejiang University) and Lin Chen studied packing group items. They considered a natural generalization of the classical multiple knapsack problem where instead of packing single items they were packing groups of items. In this setting, they had multiple knapsacks of unit capacity, and a set of items, each of a size within (0,1). These items appeared in groups, where each group was associated with a profit. The profit could be attained if and only if every item of this group was packed into the knapsacks. Such a general model finds applications in delivering bundles of goods. Apart from that, the theoretical issue is of particular interests. It is obvious that no finite bounds are possible, unless P=NP, if a group size (the total size of items in the group) can be arbitrarily large. They thus paid attention to the parameterized version while every group size was bounded by a factor of the total capacity of knapsacks. Along this line, they provided deep insights into the approximability with respect to the factor and derive, respectively, approximation algorithms and inapproximability results.

Graph Problems and Algorithms

In this area of research **Liming Cai** (University of Georgia) talked about Maximum Spanning Backbone k-Tree: Tractability and Approximability. The Maximum Spanning Backbone k-Tree (BkT) problem, for $k \geq 2$, is to find a maximum weight spanning k-tree from the input edge-weighted graph with a designated Hamiltonian path to be desired in the output spanning graph. Originally motivated by research in bio-molecular 3D structure prediction, BkT turns out a typical problem in a new class of languages logic-definable beyond MSOL. They showed that, unlike the Maximum Spanning k-Tree problem that is NP-hard for even $k = 2$, the BkT problem is solvable in time $O(n^{k+1})$, for every fixed $k \geq 2$. While there is evidence of difficulty to improve the polynomial degree $k + 1$ to any number lower, they showed that there are $O(n^3)$-time algorithms to approximate the BkT problem to the ratio $k(k - 1)$, for every fixed $k \geq 3$. The tractability results also hold with the constraint of a designated spanning tree instead of a designated Hamiltonian path, a scenario that often arises in learning of Markov networks of bounded tree width.
Thomas Erlebach (University of Leicester) pesented his research results on Temporal Graph Exploration in joint work with Michael Hoffmann and Frank Kammer. A temporal graph is a graph in which the edge set can change from step to step. The temporal graph exploration problem TEMPEX is the problem of computing a foremost exploration schedule for a temporal graph, i.e., a temporal walk that starts at a given start node, visits all nodes of the graph, and has the smallest arrival time. They considered only temporal graphs that are connected at each step. For such temporal graphs with n nodes, they showed that it is NP-hard to approximate TEMPEX with ratio $O(n^{1-\epsilon})$ for any $\epsilon > 0$. Thomas also provided an explicit construction of temporal graphs that require $\Theta(n^2)$ steps to be explored. He then considered TEMPEX under the assumption that the underlying graph (i.e. the graph that contains all edges that are present in the temporal graph in at least one step) belongs to a specific class of graphs. Among other results, they showed that temporal graphs can be explored in $O(n^{1.5}k^2 \log n)$ steps if the underlying graph has treewidth k and in $O(n \log^3 n)$ steps if the underlying graph is a $2 \times n$ grid. They also showed that sparse temporal graphs with regularly present edges can always be explored in $O(n)$ steps.
Martin Fürer (The Pennsylvania State University) studied Multi-Clique-Width and defined it as a powerful new width parameter. Multi-clique-width is obtained by a simple modification in the definition of clique-width. It has the advantage of providing a natural extension of tree-width. Unlike clique-width, it does not explode exponentially compared to tree-width. Efficient algorithms based on multi-clique-width are still possible for interesting tasks like computing the independent set polynomial or testing c-colorability. In particular, c-colorability can be tested in time linear in n and singly exponential in c and the width k of a given multi-k-expression. For these tasks, the running time as a function of the multi-clique-width is the same as the running time of the

fastest known algorithm as a function of the clique-width. This results in an exponential speed-up for some graphs, if the corresponding graph generating expressions are given. The reason is that the multi-clique-width is never bigger, but is exponentially smaller than the clique-width for many graphs.
In addition **Matthias Mnich** (University of Bonn) showed interesting improved Approximation Algorithm for Minimum Feedback Vertex Sets in Tournaments. His research together with Lászlo A. Végh (London School of Economics and Political Science) considered the minimum feedback vertex set problem in tournaments, which finds applications in ranking scenarios. This problem is NP-hard to solve exactly, and Unique Games-hard to approximate by a factor better than two. They presented an approximation algorithm for this problem that improves on the previously best known ratio 5/2, given by Cai et al. in FOCS 1998 / SICOMP 2001.

Constraint Satisfaction and Voting

Valia Mitsou (Hungarian Academy of Sciences) started with thoughts on Complexity and Approximability of Parameterized CSP which were written down together with Holger Dell, Eunjung Kim, Michael Lampis, and Tobias Mömke. The complexity of various Constraint Satisfaction Problems (CSP) when parameterized by structural measures (such as treewidth or clique-width) is a well-investigated area. They took a fresh look at such questions from the point of view of approximation, considering four standard CSP predicates: AND, OR, PARITY, and MAJORITY. By providing new or tighter hardness results for the satisfiability versions, as well as approximation algorithms for the corresponding maximization problems, they showed that already these basic predicates display a diverse set of behaviors, ranging from being FPT to optimize exactly for quite general parameters (PARITY), to being W-hard but well-approximable (OR, MAJORITY) to being W-hard and inapproximable (AND). Their results indicate the interest in posing the question of approximability during the usual investigation of CSP complexity with regards to the landscape of structural parameters.
She was followed by **André Nichterlein** (Technische Universität Berlin) with his talk *FPT approximation schemes for Shift Bribery* which is a joint work with Robert Bredereck, Jiehua Chen, Piotr Faliszewski, and Rolf Niedermeier. In the Shift Bribery problem, they were given an election (based on preference orders), a preferred candidate p, and a budget. The goal was to ensure that p wins by shifting p higher in some voters preference orders. However, each such shift request came at a price (depending on the voter and on the extent of the shift) and they had to minimize the overall costs. They showed FPT approximation schemes for the Copeland voting rule (the winner is the candidate winning the most head-to-head competitions) with respect to each of the parameters number of voters and number of candidates.