Publications of Martin Schirneck

This document lists all peer-reviewed publications of Martin Schirneck, Chair for Algorithm Engineering, Hasso Plattner Institute, Potsdam, Germany. This listing was automatically generated on June 30, 2022. An up-to-date version is available online at hpi.de/friedrich/docs/publist/schirneck.pdf.

Journal articles

[1] Bläsius, T., Friedrich, T., Schirneck, M., The Complexity of Dependency Detection and Discovery in Relational Databases. In: *Theoretical Computer Science* 900, pp. 79–96, 2022.

Multi-column dependencies in relational databases come associated with two different computational tasks. The detection problem is to decide whether a dependency of a certain type and size holds in a given database, the discovery problem asks to enumerate all valid dependencies of that type. We settle the complexity of both of these problems for unique column combinations (UCCs), functional dependencies (FDs), and inclusion dependencies (INDs). We show that the detection of UCCs and FDs is W[2]-complete when parameterized by the solution size. The discovery of inclusion-wise minimal UCCs is proven to be equivalent under parsimonious reductions to the transversal hypergraph problem of enumerating the minimal hitting sets of a hypergraph. The discovery of FDs is equivalent to the simultaneous enumeration of the hitting sets of multiple input hypergraphs. We further identify the detection of INDs as one of the first natural W[3]-complete problems. The discovery of maximal INDs is shown to be equivalent to enumerating the maximal satisfying assignments of antimonotone, 3-normalized Boolean formulas.

[2] Bläsius, T., Friedrich, T., Lischeid, J., Meeks, K., Schirneck, M., Efficiently Enumerating Hitting Sets of Hypergraphs Arising in Data Profiling. In: *Journal of Computer and System Sciences* 124, pp. 192–213, 2022.

The transversal hypergraph problem asks to enumerate the minimal hitting sets of a hypergraph. If the solutions have bounded size, Eiter and Gottlob [SICOMP'95] gave an algorithm running in output-polynomial time, but whose space requirement also scales with the output. We improve this to polynomial delay and space. Central to our approach is the extension problem, deciding for a set X of vertices whether it is contained in any minimal hitting set. We show that this is one of the first natural problems to be W[3]-complete. We give an algorithm for the extension problem running in time O(m |X|+1 n) and prove a SETH-lower bound showing that this is close to optimal. We apply our enumeration method to the discovery problem of minimal unique column combinations from data profiling. Our empirical evaluation suggests that the algorithm outperforms its worst-case guarantees on hypergraphs stemming from real-world databases.

[3] Birnick, J., Bläsius, T., Friedrich, T., Naumann, F., Papenbrock, T., Schirneck, M., Hitting Set Enumeration with Partial Information for Unique Column Combination Discovery. In: *Proceedings of the VLDB Endowment* 13, pp. 2270–2283, 2020.

Unique column combinations (UCCs) are a fundamental concept in relational databases. They identify entities in the data and support various data management activities. Still, UCCs are usually not explicitly defined and need to be discovered. State-of-the-art data profiling algorithms are able to efficiently discover UCCs in moderately sized datasets, but they tend to fail on large and, in particular, on wide datasets due to run time and memory limitations. In this paper, we introduce HPIValid, a novel UCC discovery algorithm that implements a faster and more resource-saving search strategy. HPIValid models the metadata discovery as a hitting set enumeration problem in hypergraphs. In this way, it combines efficient discovery techniques from data profiling research with the most recent theoretical insights into enumeration algorithms. Our evaluation shows that HPIValid is not only orders of magnitude faster than related work, it also has a much smaller memory footprint.

[4] Shi, F., Schirneck, M., Friedrich, T., Kötzing, T., Neumann, F., Correction to: Reoptimization Time Analysis of Evolutionary Algorithms on Linear Functions Under Dynamic Uniform Constraints. In: *Algorithmica* 82, pp. 3117–3123, 2020.

In the article "Reoptimization Time Analysis of Evolutionary Algorithms on Linear Functions Under Dynamic Uniform Constraints", we claimed a worst-case runtime of $O(nD \log D)$ and $O(nD)$ for the Multi-Objective Evolutionary Algorithm and the Multi-Objective $(\mu + (\lambda, \lambda))$ Genetic Algorithm, respectively, on linear profit functions under dynamic uniform constraint, where $D = |B - B^*|$ denotes the difference between the original constraint bound B and the new one B^* . The technique used to prove these results contained an error. We correct this mistake and show a weaker bound of $O(nD^2)$ for both algorithms instead.

[5] Friedrich, T., Kötzing, T., Lagodzinski, J. A. G., Neumann, F., Schirneck, M., Analysis of the (1+1) EA on Subclasses of Linear Functions under Uniform and Linear Constraints. In: *Theoretical Computer Science* 832, pp. 3–19, 2020.

Linear functions have gained great attention in the run time analysis of evolutionary computation methods. The corresponding investigations have provided many effective tools for analyzing more complex problems. So far, the runtime analysis of evolutionary algorithms has mainly focused on unconstrained problems, but problems occurring in applications frequently involve constraints. Therefore, there is a strong need to extend the methods for analyzing unconstrained problems to a setting involving constraints. In this paper, we consider the behavior of the classical (1+1) evolutionary algorithm on linear functions under linear constraint. We show tight bounds in the case where the constraint is given by the OneMax function and the objective function is given by either the OneMax or the BinVal function. For the general case we present upper and lower bounds.

[6] Shi, F., Schirneck, M., Friedrich, T., Kötzing, T., Neumann, F., Reoptimization Time Analysis of Evolutionary Algorithms on Linear Functions Under Dynamic Uniform Constraints. In: *Algorithmica* 81, pp. 828–857, 2019.

Rigorous runtime analysis is a major approach towards understanding evolutionary computing techniques, and in this area linear pseudo-Boolean objective functions play a central role. Having an additional linear constraint is then equivalent to the NP-hard Knapsack problem, certain classes thereof have been studied in recent works. In this article, we present a dynamic model of optimizing linear functions under uniform constraints. Starting from an optimal solution with respect to a given constraint bound, we investigate the runtimes that different evolutionary algorithms need to recompute an optimal solution when the constraint bound changes by a certain amount. The classical $(1+1)$ EA and several population-based algorithms are designed for that purpose, and are shown to recompute efficiently. Furthermore, a variant of the $(1+(\lambda,\lambda))$ GA for the dynamic optimization problem is studied, whose performance is better when the change of the constraint bound is small.

[7] Doerr, B., Fischbeck, P., Frahnow, C., Friedrich, T., Kötzing, T., Schirneck, M., Island Models Meet Rumor Spreading. In: *Algorithmica* 81, pp. 886–915, 2019.

Island models in evolutionary computation solve problems by a careful interplay of independently running evolutionary algorithms on the island and an exchange of good solutions between the islands. In this work, we conduct rigorous run time analyses for such island models trying to simultaneously obtain good run times and low communication effort. We improve the existing upper bounds for both measures (i) by improving the run time bounds via a careful analysis, (ii) by balancing individual computation and communication in a more appropriate manner, and (iii) by replacing the usual communicate-with-all approach with randomized rumor spreading. In the latter, each island contacts a randomly chosen neighbor. This epidemic communication paradigm is known to lead to very fast and robust information dissemination in many applications. Our results concern island models running simple (1+1) evolutionary algorithms to optimize the classic test functions OneMax and LeadingOnes. We investigate binary trees, d-dimensional tori, and complete graphs as communication topologies.

Conference papers

[8] Friedrich, T., Kötzing, T., Radhakrishnan, A., Schiller, L., Schirneck, M., Tennigkeit, G., Wietheger, S., Crossover for Cardinality Constrained Optimization. In: *Genetic and Evolutionary Computation Conference (GECCO)*, 2022. Best Paper Award (Theory Track) Nomination.

In order to understand better how and why crossover can benefit optimization, we consider pseudo-Boolean functions with an upper bound B on the number of 1s allowed in the bit string (cardinality constraint). We consider the natural translation of the OneMax test function, a linear function where B bits have a weight of $1+\varepsilon$ and the remaining bits have a weight of 1. The literature gives a bound of $\Theta(n^2)$ for the (1+1) EA on this function. Part of the difficulty when optimizing this problem lies in having to improve individuals meeting the cardinality constraint by flipping both a 1 and a 0. The experimental literature proposes balanced operators, preserving the number of 1s, as a remedy. We show that a balanced mutation operator optimizes the problem in $O(n \log n)$ if $n - B = O(1)$. However, if $n - B = \Theta(n)$, we show a bound of $\Omega(n^2)$, just as classic bit flip mutation. Crossover and a simple island model gives $O(n^2/\log n)$ (uniform crossover) and $O(n\sqrt{n})$ (3-ary majority vote crossover). For balanced uniform crossover with Hamming distance maximization for diversity we show a bound of $O(n \log n)$. As an additional contribution we analyze and discuss different balanced crossover operators from the literature.

[9] Bilò, D., Casel, K., Choudhary, K., Cohen, S., Friedrich, T., Lagodzinski, J. G., Schirneck, M., Wietheger, S., Fixed-Parameter Sensitivity Oracles. In: *Innovations in Theoretical Computer Science (ITCS)*, pp. 23:1–23:18, 2022.

We combine ideas from distance sensitivity oracles (DSOs) and fixed-parameter tractability (FPT) to design sensitivity oracles for FPT graph problems. An oracle with sensitivity f for an FPT problem Π on a graph G with parameter k preprocesses G in time $O(g(f,k)poly(n))$. When queried with a set F of at most f edges of G, the oracle reports the answer to the Π-with the same parameter k-on the graph $G-F$, i.e., G deprived of F. The oracle should answer queries in a time that is significantly faster than merely running the best-known FPT algorithm on $G-F$ from scratch. We design sensitivity oracles for the k-Path and the k-Vertex Cover problem. Our first oracle for k-Path has size $O(k^{f+1})$ size and query time $O(f \min\{f, \log(f)+k\})$. We use a technique inspired by the work of Weimann and Yuster [FOCS 2010, TALG 2013] on distance sensitivity problems to reduce the space to $O\big(\big(\frac{f+k}{f}\big)^f \big(\frac{f+k}{k}\big)^k fk \cdot \log n\big)$ at the expense of increasing the query time to $O\big(\big(\frac{f+k}{f}\big)^f \big(\frac{f+k}{k}\big)^k f \min\{f,k\} \cdot \log n\big)$. Both oracles can be modified to handle vertex-failures, but we need to replace k with $2k$ in all the claimed bounds. Regarding k-Vertex Cover, we design three oracles offering different trade-offs between the size and the query time. The first oracle takes $O(3^{f+k})$ space and has $O(2^f)$ query time, the second one has a size of $O(2^{f+k^2+k})$ and a query time of $O(f+k^2)$; finally, the third one takes $O(fk+k^2)$ space and can be queried in time $O(1.2738^k + fk^2)$. All our oracles are computable in time (at most) proportional to their size and the time needed to detect a k-path or k-vertex cover, respectively. We also provide an interesting connection between k-Vertex Cover and the fault-tolerant shortest path problem, by giving a DSO of size $O(poly(f,k) \cdot n)$ with query time in $O(poly(f,k))$, where k is the size of a vertex cover. Following our line of research connecting fault-tolerant FPT and shortest paths problems, we introduce parameterization to the computation of distance preservers. Given a graph with a fixed source s and parameters f,k, we study the problem of constructing polynomial-sized oracles that reports efficiently, for any target vertex v and set F of at most f edge failures, whether the distance from s to v increases at most by an additive term of k in $G-F$. The oracle size is $O(2^k k^2 \cdot n)$, while the time needed to answer a query is $O(2^k f^\omega k^\omega)$, where $\omega < 2.373$ is the matrix multiplication exponent. The second problem we study is about the construction of bounded-stretch fault-tolerant preservers. We construct a subgraph with $O(2^{fk+f+k} k \cdot n)$ edges that preserves those s-v-distances that do not increase by more than k upon failure of F. This improves significantly over the $\tilde{O}(fn^{2-\frac{1}{2^f}})$ bound in the unparameterized case by Bodwin et al. [ICALP 2017].

[10] Bilò, D., Choudhary, K., Cohen, S., Friedrich, T., Schirneck, M., Deterministic Sensitivity Oracles for Diameter, Eccentricities and All Pairs Distances. In: *International Colloquium on Automata, Languages and Programming (ICALP)*, pp. 68:1–68:19, 2022.

We construct data structures for extremal and pairwise distances in directed graphs in the presence of transient edge failures. Henzinger et al. (ITCS 2017) initiated the study of fault-tolerant (sensitivity) oracles for the diameter and vertex eccentricities. We extend this with a special focus on space efficiency. We present several new data structures, among them the first fault-tolerant eccentricity oracle for dual failures in subcubic space. We further prove lower bounds that show limits to approximation vs. space and diameter vs. space trade-offs for fault-tolerant oracles. They highlight key differences between data structures for undirected and directed graphs. Initially, our oracles are randomized leaning on a sampling technique frequently used in sensitivity analysis. Building on the work of Alon, Chechik, and Cohen (ICALP 2019) as well as Karthik and Parter (SODA 2021), we develop a hierarchical framework to derandomize fault-tolerant data structures. We first apply it to our own diameter/eccentricity oracles and then show its usefulness by derandomizing algorithms from the literature: the distance sensitivity oracle of Ren (JCSS 2022) and the Single-Source Replacement Path algorithm of Chechik and Magen (ICALP 2020).

[11] Bilò, D., Cohen, S., Friedrich, T., Schirneck, M., Near-Optimal Deterministic Single-Source Distance Sensitivity Oracles. In: *European Symposium on Algorithms (ESA)*, pp. 18:1–18:17, 2021.

Given a graph with a distinguished source vertex s, the Single Source Replacement Paths (SSRP) problem is to compute and output, for any target vertex t and edge e, the length $d(s,t,e)$ of a shortest path from s to t that avoids a failing edge e. A Single-Source Distance Sensitivity Oracle (Single-Source DSO) is a compact data structure that answers queries of the form (t,e) by returning the distance $d(s,t,e)$. We show how to compress the output of the SSRP problem on n-vertex, m-edge graphs with integer edge weights in the range $[1,M]$ into a deterministic Single-Source DSO that has size $O(M^{1/2}n^{3/2})$ and query time $\widetilde{O}(1)$. We prove that the space requirement is optimal (up to the word size). Our techniques can also handle vertex failures within the same bounds. Chechik and Cohen [SODA 2019] presented a combinatorial randomized $\widetilde{O}(m\sqrt{n}+n^2)$ time SSRP algorithm for undirected and unweighted graphs. We derandomize their algorithm with the same asymptotic running time and apply our compression to obtain a deterministic Single-Source DSO with $\widetilde{O}(m\sqrt{n}+n^2)$ preprocessing time, $O(n^{3/2})$ space, and $\widetilde{O}(1)$ query time. Our combinatorial Single-Source DSO has near-optimal space, preprocessing and query time for dense unweighted graphs, improving the preprocessing time by a $\sqrt{n}$-factor compared to previous results. Grandoni and Vassilevska Williams [FOCS 2012, TALG 2020] gave an algebraic randomized $\widetilde{O}(Mn^{\omega})$ time SSRP algorithm for (undirected and directed) graphs with integer edge weights in the range $[1,M]$, where $\omega < 2.373$ is the matrix multiplication exponent. We derandomize their algorithm for undirected graphs and apply our compression to obtain an algebraic Single-Source DSO with $\widetilde{O}(Mn^{\omega})$ preprocessing time, $O(M^{1/2}n^{3/2})$ space, and $\widetilde{O}(1)$ query time. This improves the preprocessing time of algebraic Single-Source DSOs by polynomial factors compared to previous results. We also present further improvements of our Single-Source DSOs. We show that the query time can be reduced to a constant at the cost of increasing the size of the oracle to $O(M^{1/3}n^{5/3})$ and that all our oracles can be made path-reporting. On sparse graphs with $m = O(\frac{n^{5/4-\varepsilon}}{M^{7/4}})$ edges, for any constant $\varepsilon > 0$, we reduce the preprocessing to randomized $\widetilde{O}(M^{7/8}m^{1/2}n^{11/8}) = O(n^{2-\varepsilon/2})$ time. To the best of our knowledge, this is the first truly subquadratic time algorithm for building Single-Source DSOs on sparse graphs.

[12] Bilò, D., Cohen, S., Friedrich, T., Schirneck, M., Space-Efficient Fault-Tolerant Diameter Oracles. In: *Mathematical Foundations of Computer Science (MFCS)*, pp. 18:1–18:16, 2021.

We design f-edge fault-tolerant diameter oracles (f-FDO, or simply FDO if $f=1$). For a given directed or undirected and possibly edge-weighted graph G with n vertices and m edges and a positive integer f, we preprocess the graph and construct a data structure that, when queried with a set F of edges, where $|F| \leq f$, returns the diameter of $G-F$. An f-FDO has stretch $\sigma \geq 1$ if the returned value $\widehat{D}$ satisfies $\operatorname{diam}(G-F) \leq \widehat{D} \leq \sigma \operatorname{diam}(G-F)$. For the case of a single edge failure ($f=1$) in an unweighted directed graph, there exists an approximate FDO by Henzinger et al. [ITCS 2017] with stretch $(1+\varepsilon)$, constant query time, space $O(m)$, and a combinatorial preprocessing time of $\widetilde{O}(mn + n^{1.5}\sqrt{Dm/\varepsilon})$, where D is the diameter. We present a near-optimal FDO with the same stretch, query time, and space. It has a preprocessing time of $\widetilde{O}(mn + n^2/\varepsilon)$, which is better for any constant $\varepsilon > 0$. The preprocessing time nearly matches a conditional lower bound for combinatorial algorithms, also by Henzinger et al. When using fast matrix multiplication instead, we achieve a preprocessing time of $\widetilde{O}(n^{2.5794} + n^2/\varepsilon)$. We further prove an information-theoretic lower bound showing that any FDO with stretch better than $3/2$ requires $\Omega(m)$ bits of space. Thus, for constant $0 < \varepsilon < 3/2$, our combinatorial $(1+\varepsilon)$-approximate FDO is near-optimal in all the parameters. In the case of multiple edge failures ($f > 1$) in undirected graphs with non-negative edge weights, we give an f-FDO with stretch $(f+2)$, query time $O(f^2 \log^2 n)$, $\widetilde{O}(fn)$ space, and preprocessing time $\widetilde{O}(fm)$. We complement this with a lower bound excluding any finite stretch in $o(fn)$ space. Many real-world networks have polylogarithmic diameter. We show that for those graphs and up to $f = o(\log n/\log\log n)$ failures one can swap approximation for query time and space. We present an exact combinatorial f-FDO with preprocessing time $mn^{1+o(1)}$, query time $n^{o(1)}$, and space $n^{2+o(1)}$. With fast matrix multiplication, the preprocessing time can be improved to $n^{\omega+o(1)}$, where $\omega < 2.373$ is the matrix multiplication exponent.

[13] Bläsius, T., Friedrich, T., Schirneck, M., The Minimization of Random Hypergraphs. In: *European Symposium on Algorithms (ESA)*, pp. 21:1–21:15, 2020.

We investigate the maximum-entropy model $\mathcal{B}_{n,m,p}$ for random n-vertex, m-edge multi-hypergraphs with expected edge size pn. We show that the expected size of the minimization $\min(\mathcal{B}_{n,m,p})$, i.e., the number of inclusion-wise minimal edges of $\mathcal{B}_{n,m,p}$, undergoes a phase transition with respect to m. If m is at most $1/(1-p)^{(1-p)n}$, then $\mathrm{E}[|\min(\mathcal{B}_{n,m,p})|]$ is of order $\Theta(m)$, while for $m \geq 1/(1-p)^{(1-p+\varepsilon)n}$ for any $\varepsilon > 0$, it is $\Theta(2^{(\mathrm{H}(\alpha)+(1-\alpha)\log_2 p)n}/\sqrt{n})$. Here, H denotes the binary entropy function and $\alpha = -(\log_{1-p} m)/n$. The result implies that the maximum expected number of minimal edges over all m is $\Theta((1+p)^n/\sqrt{n})$. Our structural findings have algorithmic implications for minimizing an input hypergraph. This has applications in the profiling of relational databases as well as for the Orthogonal Vectors problem studied in fine-grained complexity. We make several technical contributions that are of independent interest in probability. First, we improve the Chernoff–Hoeffding theorem on the tail of the binomial distribution. In detail, we show that for a binomial variable $Y \sim \mathrm{Bin}(n,p)$ and any $0 < x < p$, it holds that $\mathrm{P}[Y \leq xn] = \Theta(2^{-\mathrm{D}(x\,\|\,p)n}/\sqrt{n})$, where D is the binary Kullback–Leibler divergence between Bernoulli distributions. We give explicit upper and lower bounds on the constants hidden in the big-O notation that hold for all n. Secondly, we establish the fact that the probability of a set of cardinality i being minimal after m i.i.d. maximum-entropy trials exhibits a sharp threshold behavior at $i^* = n + \log_{1-p} m$.

[14] Bläsius, T., Friedrich, T., Lischeid, J., Meeks, K., Schirneck, M., Efficiently Enumerating Hitting Sets of Hypergraphs Arising in Data Profiling. In: *Algorithm Engineering and Experiments (ALENEX)*, pp. 130–143, 2019.

We devise an enumeration method for inclusion-wise minimal hitting sets in hypergraphs. It has delay $O(m^{k^*+1} \cdot n^2)$ and uses linear space. Hereby, n is the number of vertices, m the number of hyperedges, and k^* the rank of the transversal hypergraph. In particular, on classes of hypergraphs for which the cardinality k^* of the largest minimal hitting set is bounded, the delay is polynomial. The algorithm solves the extension problem for minimal hitting sets as a subroutine. We show that the extension problem is W[3]-complete when parameterised by the cardinality of the set which is to be extended. For the subroutine, we give an algorithm that is optimal under the exponential time hypothesis. Despite these lower bounds, we provide empirical evidence showing that the enumeration outperforms the theoretical worst-case guarantee on hypergraphs arising in the profiling of relational databases, namely, in the detection of unique column combinations.

[15] Bläsius, T., Fischbeck, P., Friedrich, T., Schirneck, M., Understanding the Effectiveness of Data Reduction in Public Transportation Networks. In: *Workshop on Algorithms and Models for the Web Graph (WAW)*, pp. 87–101, 2019.

Given a public transportation network of stations and connections, we want to find a minimum subset of stations such that each connection runs through a selected station. Although this problem is NP-hard in general, real-world instances are regularly solved almost completely by a set of simple reduction rules. To explain this behavior, we view transportation networks as hitting set instances and identify two characteristic properties, locality and heterogeneity. We then devise a randomized model to generate hitting set instances with adjustable properties. While the heterogeneity does influence the effectiveness of the reduction rules, the generated instances show that locality is the significant factor. Beyond that, we prove that the effectiveness of the reduction rules is independent of the underlying graph structure. Finally, we show that high locality is also prevalent in instances from other domains, facilitating a fast computation of minimum hitting sets.

[16] Kötzing, T., Schirneck, M., Seidel, K., Normal Forms in Semantic Language Identification. In: *Algorithmic Learning Theory (ALT)*, pp. 493–516, 2017.

We consider language learning in the limit from text where all learning restrictions are semantic, that is, where any conjecture may be replaced by a semantically equivalent conjecture. For different such learning criteria, starting with the well-known TxtGBc-learning, we consider three different normal forms: strongly locking learning, consistent learning and (partially) set-driven learning. These normal forms support and simplify proofs and give insight into what behaviors are necessary for successful learning (for example when consistency in conservative learning implies cautiousness and strong decisiveness). We show that strongly locking learning can be assumed for partially set-driven learners, even when learning restrictions apply. We give a very general proof relying only on a natural property of the learning restriction, namely, allowing for simulation on equivalent text. Furthermore, when no restrictions apply, also the converse is true: every strongly locking learner can be made partially set-driven. For several semantic learning criteria we show that learning can be done consistently. Finally, we deduce for which learning restrictions partial set-drivenness and set-drivenness coincide, including a general statement about classes of infinite languages. The latter again relies on a simulation argument.

[17] Friedrich, T., Kötzing, T., Lagodzinski, J. A. G., Neumann, F., Schirneck, M., Analysis of the (1+1) EA on Subclasses of Linear Functions under Uniform and Linear Constraints. In: *Foundations of Genetic Algorithms (FOGA)*, pp. 45–54, 2017.

Linear functions have gained a lot of attention in the area of run time analysis of evolutionary computation methods and the corresponding analyses have provided many effective tools for analyzing more complex problems. In this paper, we consider the behavior of the classical (1+1) Evolutionary Algorithm for linear functions under linear constraint. We show tight bounds in the case where both the objective function and the constraint is given by the OneMax function and present upper bounds as well as lower bounds for the general case. Furthermore, we also consider the LeadingOnes fitness function.

[18] Shi, F., Schirneck, M., Friedrich, T., Kötzing, T., Neumann, F., Reoptimization Times of Evolutionary Algorithms on Linear Functions Under Dynamic Uniform Constraints. In: *Genetic and Evolutionary Computation Conference (GECCO)*, pp. 1407–1414, 2017.

Thee investigations of linear pseudo-Boolean functions play a central role in the area of runtime analysis of evolutionary computing techniques. Having an additional linear constraint on a linear function is equivalent to the NP-hard knapsack problem and special problem classes thereof have been investigated in recent works. In this paper, we extend these studies to problems with dynamic constraints and investigate the runtime of different evolutionary algorithms to recompute an optimal solution when the constraint bound changes by a certain amount. We study the classical $(1+1)$ EA and population-based algorithms and show that they recompute an optimal solution very efficiently. Furthermore, we show that a variant of the $(1 + (\lambda, \lambda))$ GA can recompute the optimal solution more efficiently in some cases.

[19] Doerr, B., Fischbeck, P., Frahnow, C., Friedrich, T., Kötzing, T., Schirneck, M., Island Models Meet Rumor Spreading. In: *Genetic and Evolutionary Computation Conference (GECCO)*, pp. 1359–1366, 2017.

Island models in evolutionary computation solve problems by a careful interplay of independently running evolutionary algorithms on the island and an exchange of good solutions between the islands. In this work, we conduct rigorous run time analyses for such island models trying to simultaneously obtain good run times and low communication effort. We improve the existing upper bounds for the communication effort (i) by improving the run time bounds via a careful analysis, (ii) by setting the balance between individual computation and communication in a more appropriate manner, and (iii) by replacing the usual communicate-with-all-neighbors approach with randomized rumor spreading, where each island contacts a randomly chosen neighbor. This epidemic communication paradigm is known to lead to very fast and robust information dissemination in many applications. Our results concern islands running simple (1+1) evolutionary algorithms, we regard d-dimensional tori and complete graphs as communication topologies, and optimize the classic test functions OneMax and LeadingOnes.

[20] Friedrich, T., Kötzing, T., Krejca, M. S., Nallaperuma, S., Neumann, F., Schirneck, M., Fast Building Block Assembly by Majority Vote Crossover. In: *Genetic and Evolutionary Computation Conference (GECCO)*, pp. 661–668, 2016.

Different works have shown how crossover can help with building block assembly. Typically, crossover might get lucky to select good building blocks from each parent, but these lucky choices are usually rare. In this work we consider a crossover operator which works on three parent individuals. In each component, the offspring inherits the value present in the majority of the parents; thus, we call this crossover operator majority vote. We show that, if good components are sufficiently prevalent in the individuals, majority vote creates an optimal individual with high probability. Furthermore, we show that this process can be amplified: as long as components are good independently and with probability at least $1/2 + \delta$, we require only $O(\log 1/\delta + \log\log n)$ successive stages of majority vote to create an optimal individual with high probability! We show how this applies in two scenarios. The first scenario is the Jump test function. With sufficient diversity, we get an optimization time of $O(n \log n)$ even for jump sizes as large as $O(n^{(1/2-\epsilon)})$. Our second scenario is a family of vertex cover instances. Majority vote optimizes this family efficiently, while local searches fail and only highly specialized two-parent crossovers are successful.

[21] Bläsius, T., Friedrich, T., Schirneck, M., The Parameterized Complexity of Dependency Detection in Relational Databases. In: *International Symposium on Parameterized and Exact Computation (IPEC)*, pp. 6:1–6:13, 2016.

We study the parameterized complexity of classical problems that arise in the profiling of relational data. Namely, we characterize the complexity of detecting unique column combinations (candidate keys), functional dependencies, and inclusion dependencies with the solution size as parameter. While the discovery of uniques and functional dependencies, respectively, turns out to be W[2]-complete, the detection of inclusion dependencies is one of the first natural problems proven to be complete for the class W[3]. As a side effect, our reductions give insights into the complexity of enumerating all minimal unique column combinations or functional dependencies.

[22] Kötzing, T., Schirneck, M., Towards an Atlas of Computational Learning Theory. In: *Symposium on Theoretical Aspects of Computer Science (STACS)*, pp. 47:1–47:13, 2016.

A major part of our knowledge about Computational Learning stems from comparisons of the learning power of different learning criteria. These comparisons inform about trade-offs between learning restrictions and, more generally, learning settings; furthermore, they inform about what restrictions can be observed without losing learning power. With this paper we propose that one main focus of future research in Computational Learning should be on a structured approach to determine the relations of different learning criteria. In particular, we propose that, for small sets of learning criteria, all pairwise relations should be determined; these relations can then be easily depicted as a map, a diagram detailing the relations. Once we have maps for many relevant sets of learning criteria, the collection of these maps is an Atlas of Computational Learning Theory, informing at a glance about the landscape of computational learning just as a geographical atlas informs about the earth. In this paper we work toward this goal by providing three example maps, one pertaining to partially set-driven learning, and two pertaining to strongly monotone learning. These maps can serve as blueprints for future maps of similar base structure.