

## Research Statement

R. Teal Witter

Algorithms form the backbone of modern computing. They appear throughout our lives, from everyday tasks like searching the internet to medical tasks like diagnosing illnesses. For each application, algorithms must satisfy problem-specific criteria such as speed, efficiency, and accuracy. For example, when we search the internet, we evaluate an algorithm's performance primarily based on speed but, when we diagnose illnesses, we evaluate an algorithm's performance primarily based on accuracy. Generally, we want algorithms that both have demonstrated effective performance on reasonable inputs and have theoretical performance guarantees on the worst inputs that we may encounter. Designing algorithms that are both practically effective and theoretically robust requires a deep understanding of problem-specific criteria and creative applications of mathematical techniques.

My research focuses on designing and analyzing algorithms that satisfy problem-specific criteria like speed, efficiency, and accuracy. My work has focused on applications in 1) sequential decision-making, 2) social dynamics, and 3) quantum computing.

Even though the areas I work in are seemingly unrelated, I apply the same algorithmic design method. I first identify problems that are impactful from discussions with domain experts and consider previously studied algorithms for these problems. Based on the problem-specific criteria, I pinpoint what properties of the existing algorithms can be improved and then develop algorithms that satisfy these properties. The development process consists of iteratively testing on real data, analyzing the algorithm's properties mathematically, and modifying the algorithm to improve its practical performance and mathematical properties.

I apply the algorithmic design method in collaboration with fellow researchers and domain experts. I am grateful for the excellent collaborators I have worked with across seven institutions and the many nonprofits I have spoken with to identify impactful problems. I will continue applying the algorithmic design method in the future and look forward to working with new colleagues and students.

### Sequential Decision-Making

Deciding what to do next is a fundamental algorithmic problem. For example, in healthcare, doctors must decide what tests to run on a patient in order to diagnose them. In this context, we want an algorithm that can both diagnose patients with only a few tests and identify what test to run next with only a few doctor consultations. Unfortunately, the two goals are in tension. An *adaptive* algorithm that seeks doctor advice after each test will minimize the number of tests but require many consultations. In contrast, a *non-adaptive* algorithm that decides all tests in advance will minimize the number of consultations but will require many unnecessary tests. Determining whether to use an adaptive or non-adaptive algorithm requires carefully analyzing the gap in performance. Prior work has considered this adaptivity gap for simple Boolean functions where all tests are interchangeable. In my work, I analyzed the adaptivity gap for more general classes of Boolean functions that are common in practice.

Using the algorithmic design method, I showed that the gap in performance between the best adaptive and non-adaptive algorithms can be large for general Boolean functions. In practical terms, my findings suggest that adaptive algorithms should generally be preferred unless adaptivity is prohibitively expensive or the Boolean function is simple to evaluate. One of the Boolean functions I analyzed is a DNF formula (OR of ANDs). For this function, I used insights about the optimal adaptive algorithm to prove a matching upper and lower logarithmic bound on the adaptivity gap. Another one of the Boolean functions I analyzed is a read-once formula (AND-OR trees where each test appears at most once). For this function, I used branching processes from probability theory to construct a lower bound where the best adaptive algorithm

can be almost linearly better than the best non-adaptive algorithm. A simple, greedy algorithm gives a basically matching upper bound. Along with results for several other classes of functions, my work was published in the proceedings of the Workshop on Approximation and Online Algorithms [3].

When adaptively making decisions is not possible, the problem becomes determining in which order to perform a set of actions. For example, in healthcare, an expanding hospital must decide in which order to open clinics so that the average time until each patient can access a clinic nearby them is minimized. For high-stakes applications, we want an algorithm that offers both theoretical guarantees and diverse solutions so that the best solution can be chosen with stakeholder input. For the setting where the reward of actions (e.g. number of patients nearby a clinic) satisfies a diminishing returns property, prior work showed a greedy algorithm gives a constant approximation to the problem. However, the greedy algorithm outputs only one solution that may not satisfy stakeholder constraints. I used the algorithmic design method to develop a local search algorithm that exhaustively explores the solution-space through iterative improvements. By testing it on real data, I showed the algorithm outputs a set of solutions that is diverse. In addition, I used linear programming and duality to prove the algorithm gives nearly the best constant approximation factor to the problem unless  $P=NP$ . My work was published in the proceedings of the International Symposium on Algorithms and Computation [4].

## **Social Modeling**

Modeling human behavior and concepts is a challenging yet important problem. Because behavior is complicated, research tends to focus on similarly complicated algorithms built in ad hoc ways to mimic the observed behavior. In my work, I used mathematical ideas to propose approaches which capture the complexity of behavior and concepts in elegant and simple ways.

Studying how people form political opinions is important for building functional discourse. Prior work has generally focused on modeling opinion dynamics using complicated algorithms to emulate the political polarization we observe in the real world. In contrast, I used mathematical insights to show how a simple averaging algorithm, when viewed through the right lens, can produce realistic polarization. Instead of viewing opinions as absolute, my work formulated opinions as relative to the mean opinion of the social network. Through a connection to spectral graph theory, the averaging opinion algorithm can then be seen as converging to the second eigenvector of a matrix related to the social network. A corollary of the result is that relative opinions are formed solely based on the connections in the social network. The findings suggest that augmenting social networks to make them more connected would reduce opinion polarization. My work was published in the proceedings of the International Workshop on Mining and Learning with Graphs [6].

Defining and designing algorithms to achieve fairness is important for building a more equitable world. A popular but complicated notion of fairness is based on counterfactuals; an algorithm is counterfactually fair if each person is treated the same as the version of themselves where only their protected attribute (e.g. race, gender, religion) is changed. Unfortunately, measuring counterfactual fairness is challenging and designing counterfactually fair algorithms is even more so. In my work, I used probability techniques to show that counterfactual fairness is basically equivalent to another fairness definition called demographic parity. Demographic parity is easier to measure and design algorithms to achieve. My work was published in the proceedings of the AAAI Conference on Artificial Intelligence [7].

## **Quantum Computing**

Quantum computers offer a new paradigm for computing that promises algorithms which can run much faster than their classical counterparts. Because building quantum computers at scale is so challenging, it is important to theoretically analyze how useful quantum computers will be. My research has focused on designing and analyzing quantum algorithms for graph theory problems.

Using an elegant formulation from prior work, I showed how quantum computers can be used to solve graph theory problems related to  $st$ -connectivity like cycle and bipartite detection. Using the algorithmic design method, I developed simple algorithms with provable guarantees. For many of the problems I considered, the guarantees generally matched known lower bounds so the algorithms were optimal. My work was published in the proceedings of the Conference on the Theory of Quantum Computation, Communication, and Cryptography [2]. Another important graph theory problem is maximum matching. Maximum matching cannot be solved using an  $st$ -connectivity subroutine so I turned to a recent method for designing quantum algorithms from decision trees. By combining the decision tree method with an efficient classical algorithm for maximum matching, I designed a quantum algorithm with the best known query complexity for maximum matching. My work was published in the proceedings of the Algorithms and Data Structures Symposium [5].

The two methods I previously used for designing quantum algorithms originate from a semi-definite program (SDP). An optimal solution to the SDP gives a query-optimal quantum algorithm. However, finding an optimal solution is difficult because the SDP is brittle to small errors and the number of variables grows exponentially with the size of the input. I showed that the SDP can still give a query-optimal quantum algorithm even when the SDP is not solved exactly but instead is solved within some tolerance. Further, I showed that the number of variables can be reduced by using high-dimensional compression techniques like the Johnson-Lindenstrauss projection. My work was published in the proceedings of the European Symposium on Algorithms [1].

In addition to the research described above, I have also developed algorithms for the board game Ticket to Ride [9] and analyzed the computational hardness of the board game backgammon [8].

## Future Work

My goal is to continue applying the algorithmic design method to achieve problem-specific criteria in impactful domains. I'm particularly excited to collaborate on projects with new colleagues and students.

**Restless Multi-Armed Bandits** In my prior work, I studied decision-making optimization at a single time step. However, decisions tend to have long-term consequences that are reflected in future time steps. The restless multi-armed bandit formulation is a way to integrate the complicated temporal dynamics of sequential decision-making. The restless bandit formulation is well-studied particularly in the context of problems with social impact. I plan to extend the formulation in two key ways that broaden the applicability of the approach. The first way I plan to extend the formulation is to introduce flexibility by considering a constrained *minimization* version of the problem. The minimization version allows flexibility in the standard budget constraint which is useful for domains like healthcare where the primary goal is positive outcomes. Based on initial results, the minimization formulation is actually quite different from the standard formulation and requires its own algorithms. The second way I plan to extend the restless bandit formulation is to introduce stochastic and supermodular costs. The extension more accurately reflects the real world where costs are unpredictable and accumulate in interesting ways. Based on initial results, the extension requires its own algorithms and mathematical tools.

**Principled Data Driven Decision-Making** There are many high stakes domains with big data available but yet decisions are still made in a heuristic way. I am interested in applying data driven tools like deep learning and reinforcement learning to make decisions in principled ways. One problem in the infrastructure domain is choosing streets to “open” (open to people, close to cars) for social enrichment. Using publicly available data, I plan to design a reinforcement learning algorithm to select streets which, when “opened,”

reduce traffic congestion and collisions. Another problem in the environmental domain is choosing how to deploy fire fighting resources in large wildfires. Using publicly available data, I plan to design a deep learning algorithm to predict what the wildfire will do and a reinforcement learning algorithm to suggest where resources should be directed to minimize negative impacts.

**Treatment Effect Estimation for Natural Experiments** Estimating treatment effect is an important problem for understanding the impact of nonprofit work. Unfortunately, the standard approach is to use randomized trials which are expensive and often infeasible especially for under-resourced non-profits. Instead, I plan to design algorithms that estimate treatment effect from natural experiments. The technical challenge is that the treatment assignment is not random and the treatment and control groups may be different. The problem has been studied by the statistics community but the algorithms are optimized only for the asymptotic regime rather than the finite sample regime. I will design estimators that have provably low variance and benefit from learning the structure of the outcomes. I will then compare the estimators against prior work on real data-sets.

## Conclusion

Even though the applications are varied, my research applies the algorithmic design method to achieve problem-specific criteria. I view my research as both a social good and a social pursuit. Designing algorithms requires making connections to mathematical ideas and learning from my collaborators. As I transition to a new stage in my academic career, I look forward to collaborating with my new colleagues and students.

## References

- [1] M. Czekanski, S. Kimmel, and R. T. Witter. [Robust and Space-Efficient Dual Adversary Quantum Query Algorithms](#). In *European Symposium on Algorithms*, 2023.
- [2] K. DeLorenzo, S. Kimmel, and R. T. Witter. Applications of the quantum algorithm for st-connectivity. In *Conference on the Theory of Quantum Computation, Communication and Cryptography*, 2019.
- [3] L. Hellerstein, D. Kletenik, N. Liu, and R. T. Witter. Adaptivity gaps for the stochastic boolean function evaluation problem. In *Workshop on Approximation and Online Algorithms*, 2022.
- [4] L. Hellerstein, T. Lidbetter, and R. T. Witter. A local search algorithm for the min-sum submodular cover problem. In *International Symposium on Algorithms and Computation*, 2022.
- [5] S. Kimmel and R. T. Witter. A query-efficient quantum algorithm for maximum matching on general graphs. In *Algorithms and Data Structures Symposium*, pages 543–555, 2021.
- [6] C. Musco, I. Ramesh, J. Ugander, and R. T. Witter. How to quantify polarization in models of opinion dynamics. In *International Workshop on Mining and Learning with Graphs*, 2022.
- [7] L. Rosenblatt and R. T. Witter. Counterfactual fairness is basically demographic parity. In *AAAI Conference on Artificial Intelligence*, 2023.
- [8] R. T. Witter. Backgammon is hard. In *International Conference on Combinatorial Optimization and Applications*, 2021.

- [9] R. T. Witter\* and A. Lyford. Applications of graph theory and probability in the board game ticket to ride. In *International Conference on the Foundations of Digital Games*, 2020.

*As is the custom in theoretical computer science, authors are listed in alphabetical order unless otherwise specified with an asterisk.*