Graphical structure of unsatisfiable boolean formulas

PhD Dissertation Summary

Vaibhav Karve

PhD advisor: Anil Hirani

Department of Mathematics

Using the tools of topology and graph theory, we introduce a new variant on the classical logic and computer science problem of boolean satisfiability (k-SAT). k-SAT asks if there exists a truth assignment that satisfies a given boolean formula. Our variant deals with multi-hypergraphs instead of boolean formulas and uses truth assignments on vertices instead of variables. We call this novel decision problem GRAPHSAT. Historically, k-SAT (for $k \geq 3$) was the first problem that was proven to be NP-complete, independently by Cook [1] and Levin [2], making it central to the study of algorithms and computational complexity. We shed new light on k-SAT by analyzing GRAPHSAT. In particular, we provide a complete list of multi-hypergraphs that obstruct satisfiability.

We introduce a projection map $\pi : \text{CNF} \to \text{MHGraph}$ that maps boolean formulas in conjunctive normal form (CNF) to multi-hypergraphs. This map projects several CNFs to the same multi-hypergraph. We then extend the notion of satisfiability to the space MHGraph by defining an element G to be satisfiable if and only if $\pi^{-1}(G)$ is nonempty and every element of $\pi^{-1}(G)$ is satisfiable. This lumps together CNFs into classes of sentences that get mapped to the same G under π. Thus, by studying properties of unsatisfiable graphs we are studying entire classes of boolean formulas. This results in a reduction of granularity compared to studying CNFs directly. However, we argue that the map π does capture the essential features of any formula and offers a complete picture of the topological structure of its boolean constraints. For example, we show that this notion of satisfiability is closed under subgraphing as well as under the operation of considering topological-minors. The family of satisfiable multi-hypergraphs is also closed under the action of a variety of higher-dimensional analogues of well-known graph operations. We explore these analogous graph operations and analyze their effects vis-à-vis the topological properties of multi-hypergraphs.

Robertson and Seymour, in their seminal papers [3] proved that a graph family that is closed under the graph minor operation always has a finite obstruction set. This result does not apply to multi-hypergraphs. However, we obtain an obstruction set for the family of satisfiable multi-hypergraphs for the GRAPHSAT decision problem. We enumerate this complete list of obstructions with the aid of software written in Python. Figure 1 below shows the complete list of obstructions to 2-SAT. We obtain a similar but much larger list of obstructions for 3-SAT as shown in Figure 2.

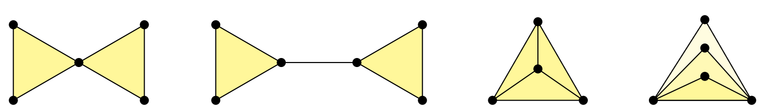

Figure 1: The complete set of minimal unsatisfiable simple graphs.

We first present a brute-force algorithm for naïvely solving GRAPHSAT with a computational complexity of $O(\exp(v+3e))$ for a multi-hypergraph of order v and size e. Next, we improve upon this by introducing a graph-rewriting algorithm that performs local rewrites at a vertex of degree d to yield a final complexity of $O(\exp(v+3e-2d))$. The graph-rewriting algorithm exploits the bivalence of boolean logic in conjunction with well-studied graph operations like edge subdivision, graph homeomorphisms, and edge contractions, to create an abstract rewriting system that can efficiently reduce multi-hypergraphs to their normal forms while preserving several different notions of equi-satisfiability extended to graphs.

Additionally, we present a Python package (also named `graphsat`) as an easily-extensible library capable of running several different algorithms on boolean formulas, simple graphs, multi-graphs, hypergraphs, and multi-hypergraphs. Inspired by the principles of reproducible research, all calculations in this dissertation are replicable using this package. We have released `graphsat` as an open-source package in the hope that it can be used for future work in this field.

A part of this research (the results concerning 2-SAT [4]) has also been verified in the Lean Theorem Prover as an experiment in formal mathematics. Lean is an open-source theorem prover and programming language that aims to bridge the gap between interactive and automated theorem proving by using automated tools and methods in a structure that supports construction of fully-specified axiomatic proofs. This experiment pushes the burden of peer-review and mathematical trust away from the result and onto Lean's well-tested mathematical kernel. In the process of formalizing our research, we introduced definitions of graphs, subgraphs, multi-hypergraphs, boolean formulas in conjunctive normal form, 2-SAT, and GRAPHSAT to Lean's math library . All code for this formalization has been released as an open-source Lean package named `leansat`.

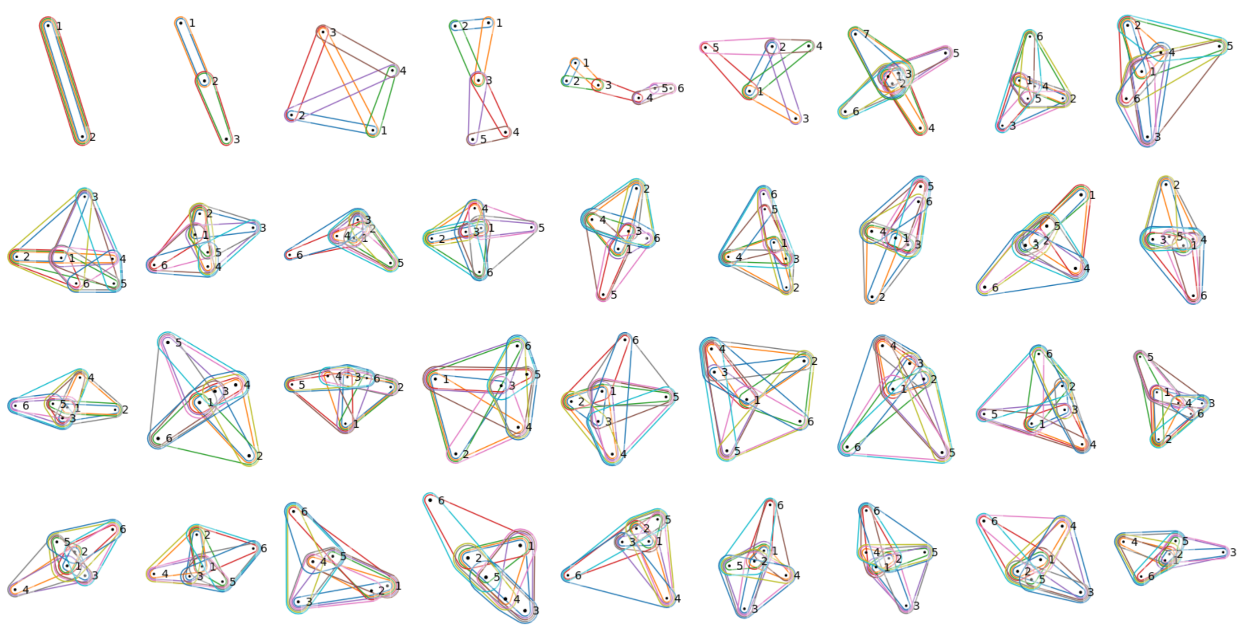

Figure 2: The first few elements of the complete set of minimal unsatisfiable multi-hypergraphs.

References

[1] Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, STOC '71, pages 151–158, New York, NY, USA, 1971. ACM.

[2] L. A. Levin. Universal enumeration problems. *Problemy Peredači Informacii*, 9(3):115–116, 1973.

[3] Neil Robertson and P.D. Seymour. Graph minors XX. Wagner's conjecture. *Journal of Combinatorial Theory, Series B*, 92(2):325 – 357, 2004. Special Issue Dedicated to Professor W.T. Tutte.

[4] Vaibhav Karve and Anil N. Hirani. The complete set of minimal simple graphs that support unsatisfiable 2-cnfs. *Discrete Applied Mathematics*, 283:123 – 132, 2020.