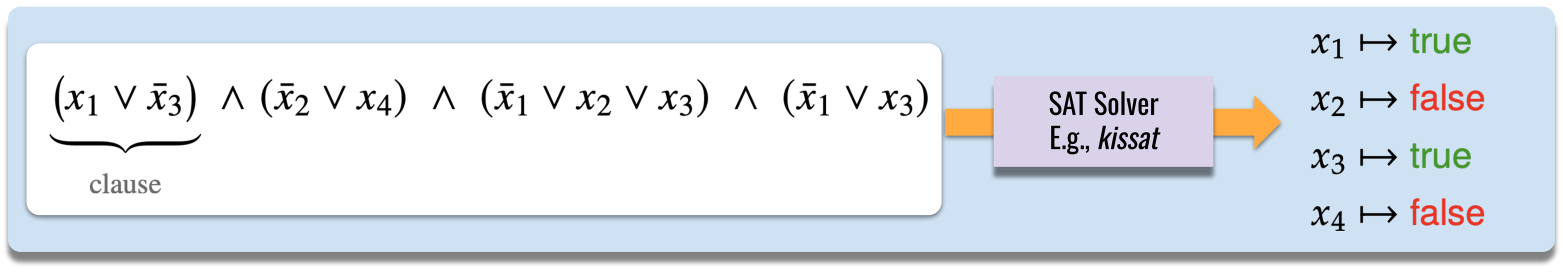
SAT Solving for Discrete Mathematics

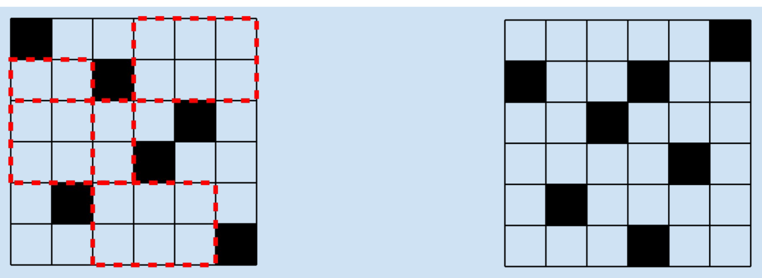
Bernardo Subercaseaux (bersub@cmu.edu), Carnegie Mellon University

SAT solving?



Satisfiability of propositional formulas is the quintessential NP-complete problem. The NP-hard part says that solving it efficiently in practice is a challenging task, but fortunately, decades of progress on both software and hardware allow modern solvers to tackle instances with millions of variables and clauses. The NP-completeness part, and especially the fact that it was the first problem proven to be so, suggest that SAT is a good problem to naturally express a variety of NP-complete problems.

In mathematics, especially discrete mathematics, we study finite objects, and often look for examples, counterexamples, and constructions. Finding these is often a computationally challenging task, which is where SAT solving comes into play. While many have heard of SAT solvers being used to find the best bounds on e.g., Ramsey numbers, a lesser known idea is that they can be helpful in many areas of research, as a way to search for counterexamples to conjectures we make along our research, and even suggesting new conjectures to look at.



(a) Coloring 6 cells misses several rectangles. (b) 7 colored cells suffice.

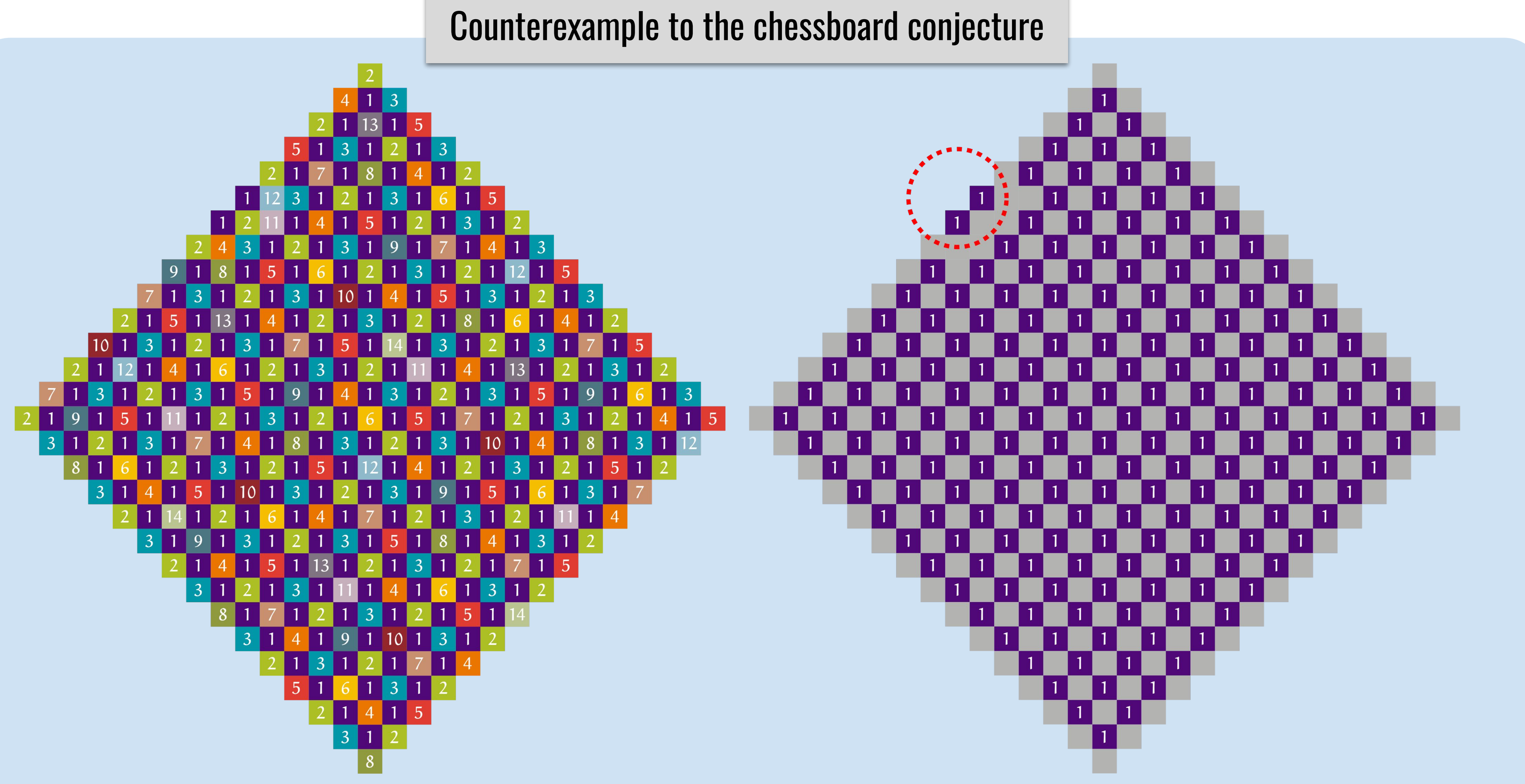
Discrete version of Danzer sets:
How many cells of an *n x n* grid need to be colored so that every subrectangle of area *n* contains a colored cell?
Figure illustrates that *g(6) = 7.*
With SAT I found that g can decrease! g(29) > g(30).
Question: does this happen infinitely often?

Graph Coloring

Since the celebrated *Four Color Theorem* (Appel and Haken, 1976), graph coloring has become very fertile ground for computational mathematics.

During my PhD work, using advanced SAT solving techniques, we proved that the *packing chromatic number of the infinite grid is* 15. A packing ***k***-coloring of a graph ***G = (V, E)*** is a function ***f: V → {1, ..., k}*** such that if two different vertices ***u, v*** get a common color ***c***, then their distance in ***G*** is at least ***c + 1.***

Counterexample to the chessboard conjecture

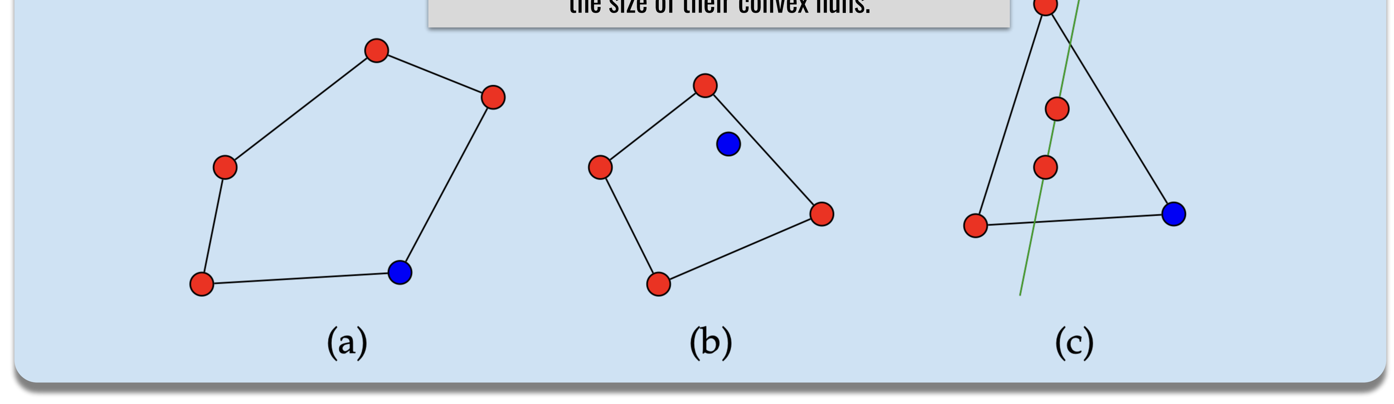


Discrete Geometry

While it might seem at first that propositional logic and Euclidean geometry live in different universes, it turns out that SAT solving has been very effective in discrete geometry. The key intuition for this is that many problems in discrete geometry don't rely on the exact coordinates of the points, but rather in combinatorial properties about how the points are placed, which can be captured by boolean variables.

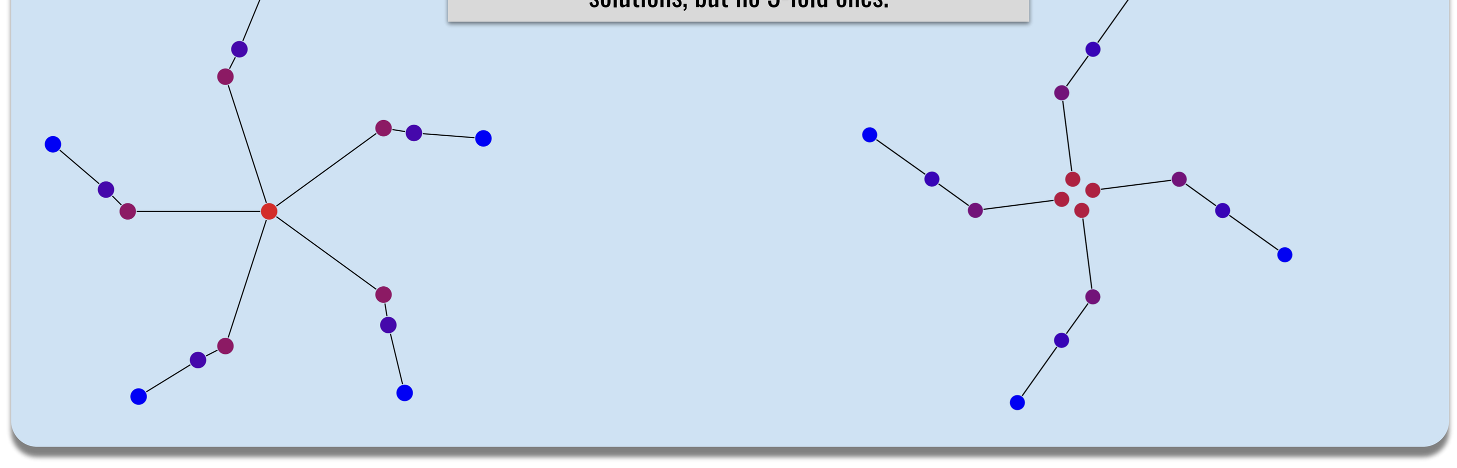
In the 1930s, Klein proved that every set of 5 points in the plane, without 3 in a line, contain a convex quadrilateral. Her proof, which culminated in the *Happy Ending Theorem,* illustrates well how problems in discrete geometry can be solved by considering high-level cases on the structure of pointsets.

Illustration of Klein's proof
The configurations are split into cases according to the size of their convex hulls.



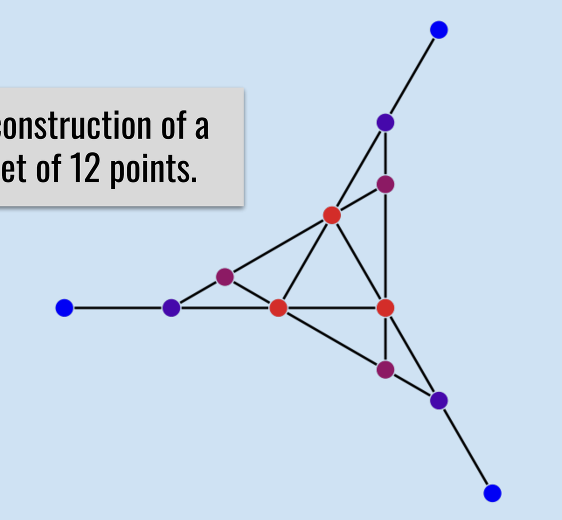
Our most recent work in this area, to be presented here at CICM, has made significant progress in dealing with **symmetry** when looking for geometric constructions with SAT solvers. For example, we were able to complete catalogue the rotational symmetries of sets of 16 points, no 3 of them on a line, that do not contain a convex 6-gon.

Erdős-Szekeres problem.
For *k = 6*, there are 4-fold and 5-fold symmetric solutions, but no 3-fold ones.

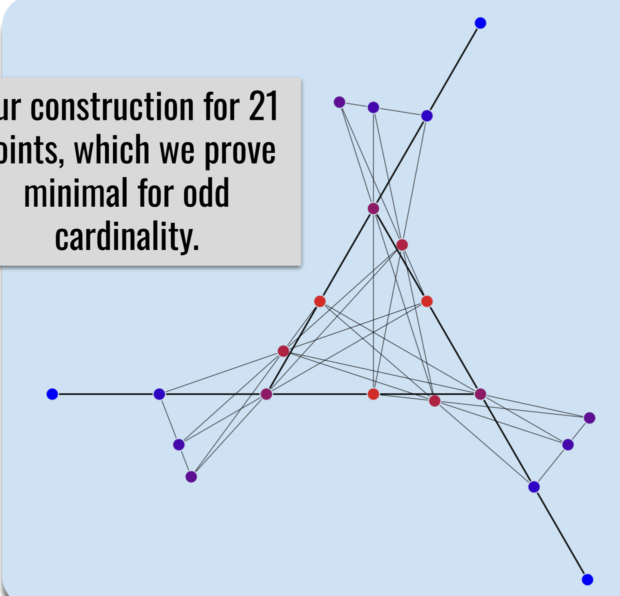


Another fascinating problem is discrete geometry is whether there exists sets of points such that any line passing through them is very *unbalanced*. More formally, a pointset *S* is said to be *k*-everywhere-unbalanced if for any line *L* touching two points of *S*, the number of points of *S* above *L* and the number of points below *L* differ by at least *k*. It is not yet known whether a 3-everywhere-unbalanced pointset exists.

Alon's construction of a 2-EU set of 12 points.

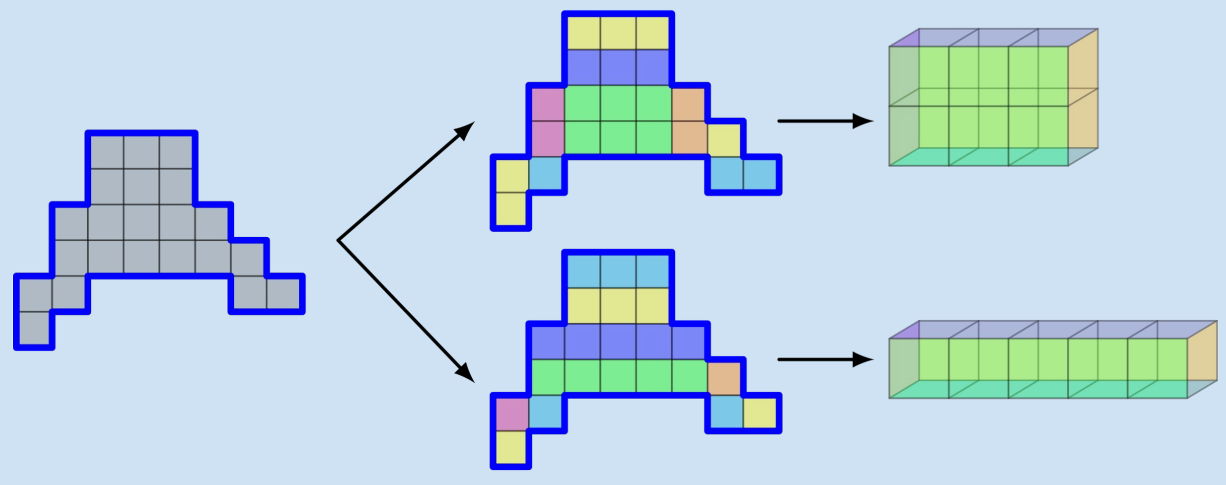


Our construction for 21 points, which we prove minimal for odd cardinality.



(Un)folding Boxes

Illustration of a 2-foldable net



Open problem: what is the smallest net (in area) that can fold into 3 non-isomorphic boxes?

The answer was conjectured to be 4, but proved it is at least 60. We are able to find 2-foldable nets of areas up to 150, while previous approaches could only reach area 90.

From a technical perspective, this work is particularly interesting for replacing **global constraints** (connectivity, acyclicity) for **local constraints.** This allows for a much more compact encoding, and moreover, to find contradictions faster when the solver considers a partial solution that cannot lead to a satisfying solution.

Encoding Theory

More recently, my focus has been on making SAT encodings less of an art and more of a science. This involves studying both the possibilities and limits of SAT encodings, and specially those that serve as building blocks of more complicated encodings.

Theorem

Consider the following problems:
minimum independent set, minimum vertex cover, maximum clique, and c-coloring.
Each of these can be encoded with only $O(n^2 / \lg n)$ clauses.

Naively, these problems would be encoded with ***O(n*** ^{2}***)*** clauses: consider the independent set problem, and variables x_v representing that vertex ***v*** is selected. Then for each edge ***{u, v}*** a clause $(\neg x_u \vee \neg x_v)$ would be added.

The main idea behind this theorem is a connection with graph theory, leveraging "*biclique coverings*". I leveraged the following result of Chung, Erdős, and Spencer (1983): every graph can be written as a union of bicliques whose total weight is $O(n^2 / \lg n)$.

Example of a biclique encoding