Neural Discovery in Mathematics: Do Machines Dream of Colored Planes?

Konrad Mundinger, Max Zimmer, Aldo Kiem, Christoph Spiegel, and Sebastian Pokutta

Correspondence: mundinger@zib.de, spiegel@zib.de

Mathematical Discovery through ML

The Goal: Using ML, *prove the existence of objects with certain properties* to advance our understanding of abstract structures.

Case Study: Hadwiger-Nelson Problem / Chromatic Number of the Plane

What is the minimum number of colors c needed to color the Euclidean plane so that no two points at unit distance share the same color?

Lower bounds have been explored using SAT solvers but *new colorings*

$$f:\mathbb{R}^2 \to \{1,\ldots,c\}, \quad f(x)\neq f(y) \quad \text{whenever} \quad \|x-y\|=1 \tag{1}$$

relied on human mathematical intuition.

Our Contribution

A machine learning framework that enables gradient-based exploration of the solution spaces. This led to *several new mathematical results* improving bounds for long-standing open problems.

The Methodology

1. **Probabilistic Reformulation:** Replace discrete colorings $f:\mathbb{R}^2\to\{1,\ldots,c\}$ with probabilistic ones $p:\mathbb{R}^2\to\Delta^c$ and minimize:

$$\mathscr{L}_R(p) = \int_{[-R,R]^2}\int_{\partial B_1(x)} p(x)^T p(y)\,\mathrm{d}\nu(y)\,\mathrm{d}\mu(x). \tag{2}$$

Key Insight

If $\mathscr{L}_R(p)=0$ then $\arg\max(p(x))$ satisfies the original constraint a.e.

2. **Neural Network Approximation:** Use SIREN NNs p_θ as universal function approximators with inherent spectral bias towards structured solutions.

3. **Unsupervised Training:** Sample (x_i,y_i) with $\|x_i-y_i\|=1$ and minimize $\mathscr{L}_R(p)$ using the approximate gradient

$$\nabla_\theta \mathscr{L}_R(p_\theta) \approx \nabla_\theta\left[\frac{1}{n}\sum_{i=1}^{n} p_\theta(x_i)^T p_\theta(y_i)\right]. \tag{3}$$

4. **Formalization:** Trained p_θ only provides numerical evidence. We extract formal colorings through mathematical analysis or automated procedures.

Minor modifications adapt the framework to different problem variants.

Links for more

code · paper · application · IOL Lab

Appl. 1: Almost Coloring the Plane

Question: How much of the plane needs to be removed so that we can color the remainder?

History: Pritikin (1998) and Parts (2020) established 99.985% can be colored with $c=6$ colors and Parts (2020) established best bounds for $c\in\{1,\ldots,5\}$.

Results: Adding an additional "removal" color, we solve

$$\mathscr{L}_R^\lambda(p_\theta) = \mathscr{L}_R(p_\theta^c) + \lambda \int_{[-R,R]^2} p_\theta(x)_{c+1}\,\mathrm{d}\mu(x). \tag{4}$$

Re-discovered known constructions for $c\neq 5$, but:

Theorem

96.26% of the plane can be colored with 5 colors.

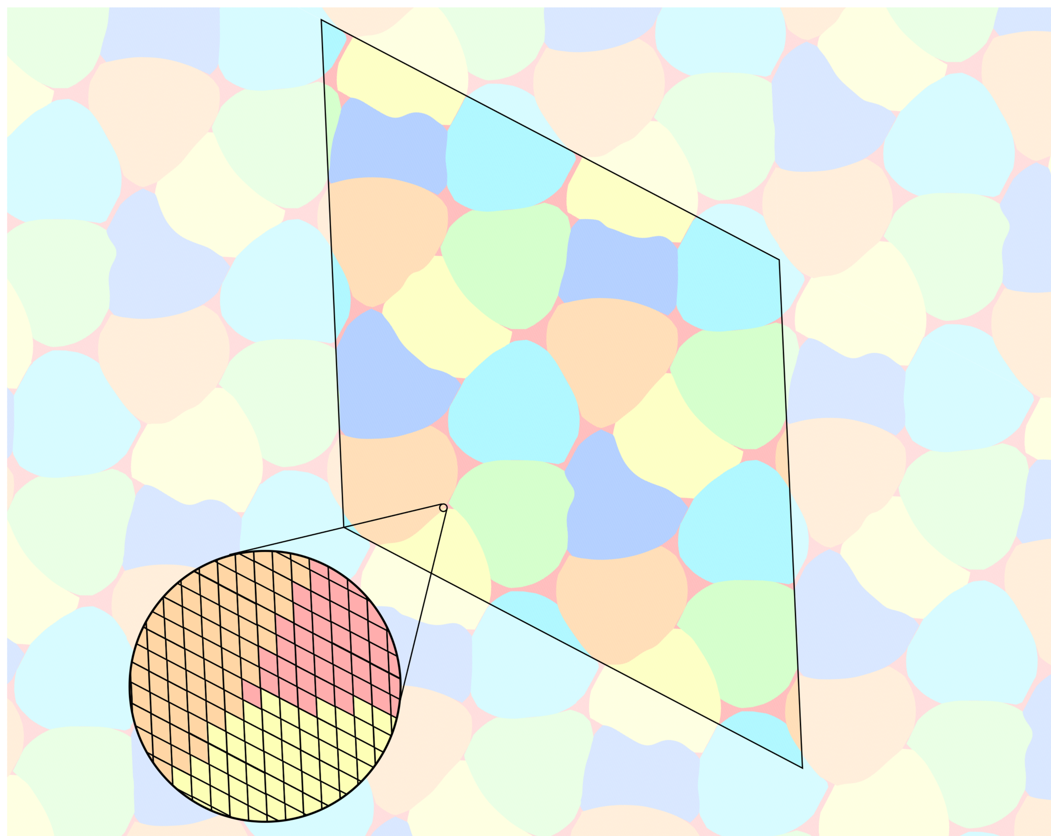

Fig. 2: Almost five-coloring: formal construction covering 96.26% of the plane.

Formalization:

1. Extract periodic structure via tiling vectors v_1, v_2.
2. Enforce periodicity by prepending change-of-basis mapping $x\mapsto M^{-1}x \bmod 1$ and retrain.
3. Discretize finely using parallelogram-structure.
4. Iteratively fix unit-distance conflicts through a minimum vertex cover problem.
5. Resolve all remaining conflicts by removing pixels.

Is this the best possible?

# colors	1	2	3	4	5	6
prior best	77.04%	54.13%	31.20%	8.25%	4.01%	0.02%
our result	77.13%	54.29%	31.51%	8.52%	**3.74%**	0.04%

Table 1: Fraction of plane requiring removal (lower is better)

Appl. 2: Avoiding Different Distances

Question: Can we avoid distance d_i in color i?

History: Soifer calls determining which "type" $(1,1,1,1,1,d)$ is realized by a coloring extremely difficult. Previously known range was

$$0.415 \approx \sqrt{2}-1 \le d \le 1/\sqrt{5} \approx 0.447.$$

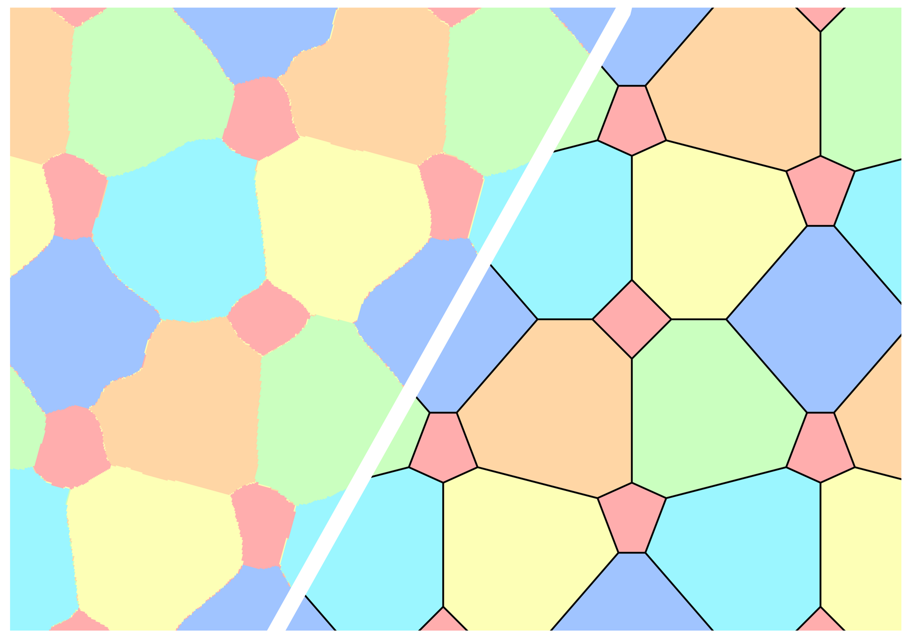

Figure 1: Coloring suggested by NN (L) and formalized version (R).

Results: Through the modified loss

$$\sum_{k=1}^{c}\int_{[-R,R]^2}\int_{\partial B_{d_k}(x)} p_\theta(x)_k\, p_\theta(y)_k\,\mathrm{d}\nu_k(y)\,\mathrm{d}\mu(x) \tag{5}$$

we discovered two novel colorings that extended range significantly. *First improvement in 30 years!*

Theorem

$(1,1,1,1,1,d)$ can be realized for $0.35\le d\le 0.65$.

Is this the best possible?

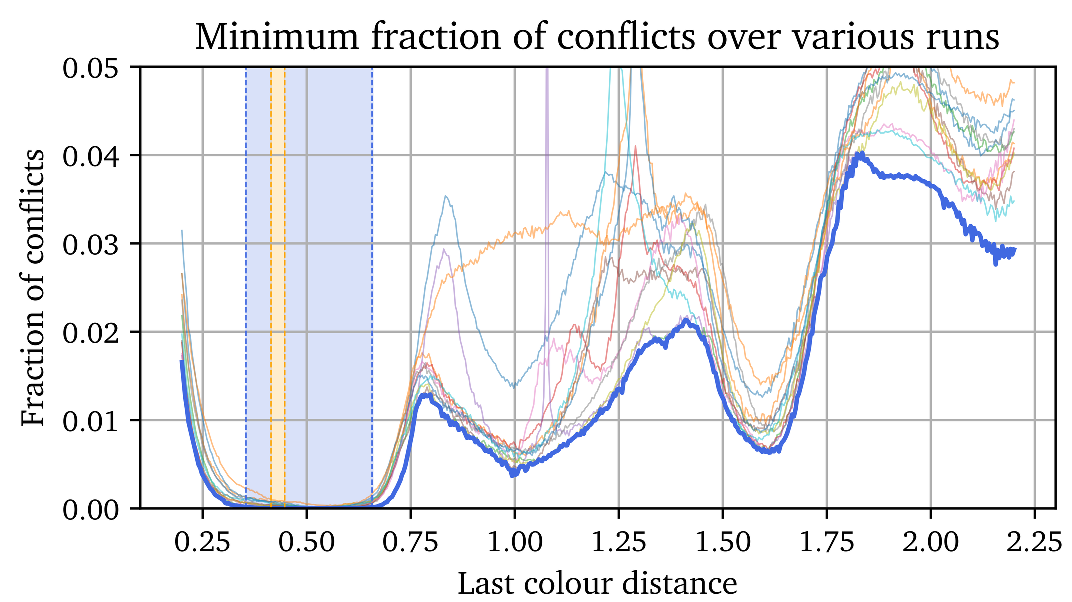

Fig. 4: Conflict rate vs. distance d revealing extended valid ranges

Erdős also asked for the *polychromatic number of the plane*, i.e., the smallest c for which some type $(d_1,\ldots,d_c)$ can be realized. We found *no evidence that the current bound of 6 can be improved.*

Appl. 2 (cntd.)

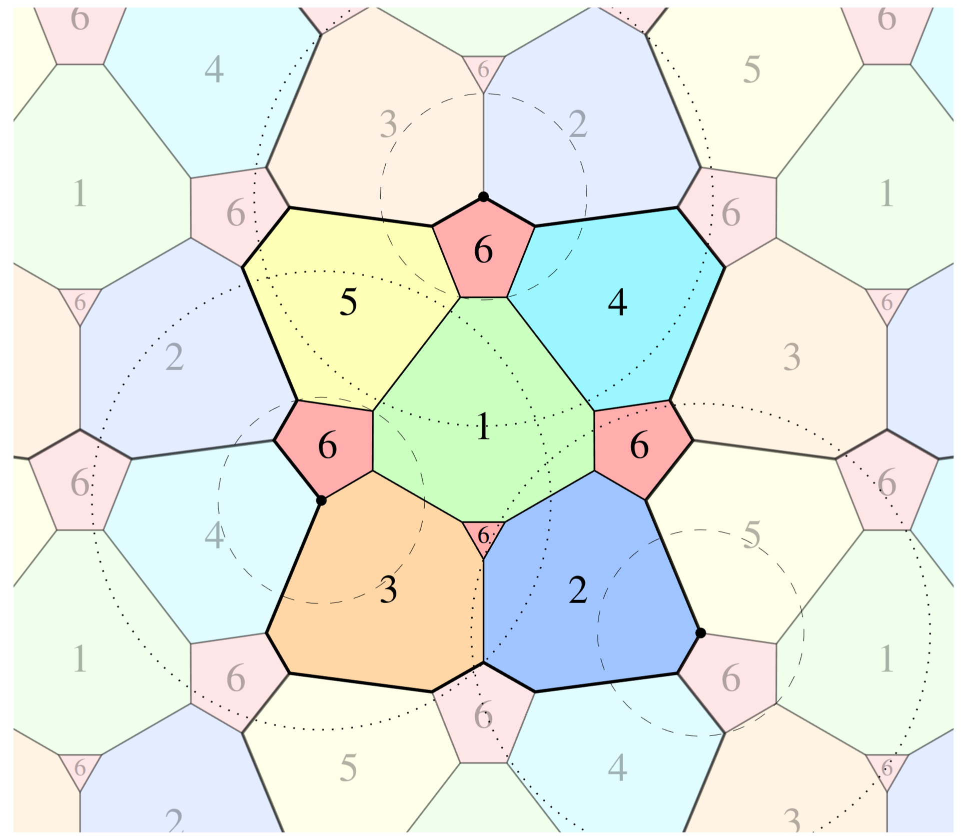

Fig. 5: A 6-coloring of the plane where no red points appear at distance 0.45 and no other color has monochromatic unit-distance pairs.

Appl. 3: Avoiding Triangles

Question: What if we avoid triangles?

History: Conjecture of Erdős et al. states that 3 colors always suffice. Bounds due to Aichholzer & Perz (2019).

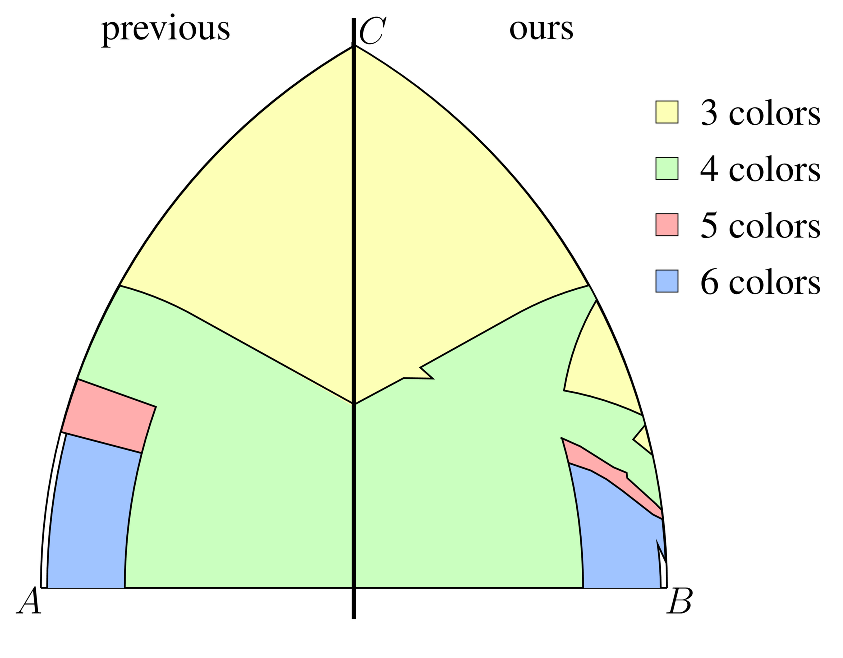

Fig. 6: Classification of triangles by required colors

Outlook

Broader Applications:

- Other Hadwiger-Nelson variants?
- Graph-theoretic problems via graph limits?
- Non-differentiable constraints through adversary?

Our approach demonstrates how ML can drive mathematical discovery and lead to novel insights.