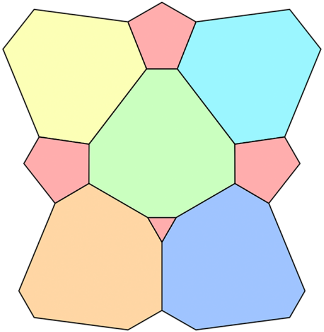
Neural Discovery in Mathematics

Do Machines Dream of Colored Planes?

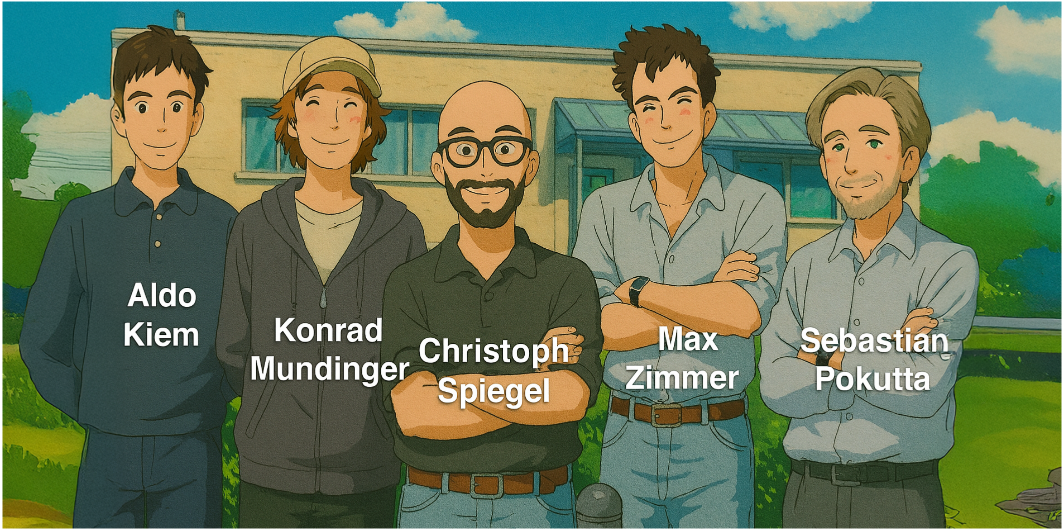
ICML 2025 in Vancouver

K. Mundinger, M. Zimmer, A. Kiem, **C. Spiegel**, S. Pokutta

17th of July 2025



About us...


Aldo Kiem
Konrad Mundinger
Christoph Spiegel
Max Zimmer
Sebastian Pokutta

Neural Discovery in Mathematics

How can a picture be a proof?

> Each such problem has two sides: one is the construction of an extremal structure, the other is the proof of its optimality. In this course we are putting extra emphasis on explicit constructions of extremal graphs, which do not customarily feature in standard treatments of the field. These constructions often require useful tools from algebra, geometry, or discrete Fourier analysis; the other main objective of these notes is to highlight them.

p.7 of LET'S BE EXPLICIT! lecture notes by Tibor Szabó, July 2024

The Hadwiger Nelson problem

Question. What is the smallest number of colors sufficient for coloring the plane in such a way that no two points of the same color are a unit distance apart?

Considering the infinite graph with vertex set $\mathbb{E}^2$ and edges $\{x, y\}$ for any $\|x - y\| = 1$, we are studying the **chromatic number of the plane** $\chi(\mathbb{E}^2)$.

Theorem (Aubrey D.N.J. de Grey, 2018; USD 1000 problem of Erdős)

There is a unit distance graph with chromatic number 5 and therefore

$$\chi(\mathbb{E}^2) \geq 5.$$

Upper bounds are given by explicit colorings $g : \mathbb{E}^2 \to [c] := \{1, \ldots, c\}$, usually derived through tessellations using simple polytopal shapes, which give

$$\chi(\mathbb{E}^2) \leq 7.$$

The Hadwiger Nelson problem

Question. What is the smallest number of colors sufficient for coloring the plane in such a way that no two points of the same color are a unit distance apart?

Considering the infinite graph with vertex set $\mathbb{E}^2$ and edges $\{x, y\}$ for any $\|x - y\| = 1$, we are studying the **chromatic number of the plane** $\chi(\mathbb{E}^2)$.

Theorem (Aubrey D.N.J. de Grey, 2018; USD 1000 problem of Erdős)

There is a unit distance graph with chromatic number 5 *and therefore*

$$\chi(\mathbb{E}^2) \geq 5.$$

Upper bounds are given by explicit colorings $g : \mathbb{E}^2 \to [c] := \{1, \dots, c\}$, usually derived through tessellations using simple polytopal shapes, which give

$$\chi(\mathbb{E}^2) \leq 7.$$

The Hadwiger Nelson problem

Question. What is the smallest number of colors sufficient for coloring the plane in such a way that no two points of the same color are a unit distance apart?

Considering the infinite graph with vertex set $\mathbb{E}^2$ and edges $\{x, y\}$ for any $\|x - y\| = 1$, we are studying the **chromatic number of the plane** $\chi(\mathbb{E}^2)$.

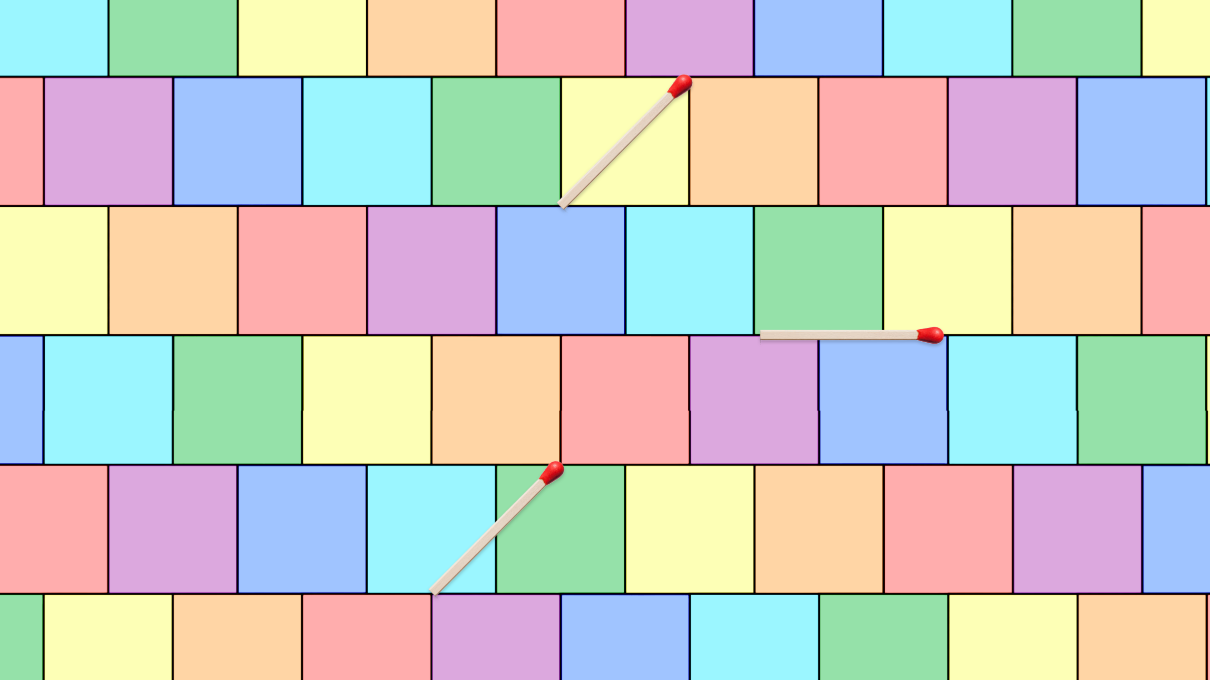
Theorem (Aubrey D.N.J. de Grey, 2018; USD 1000 problem of Erdős)

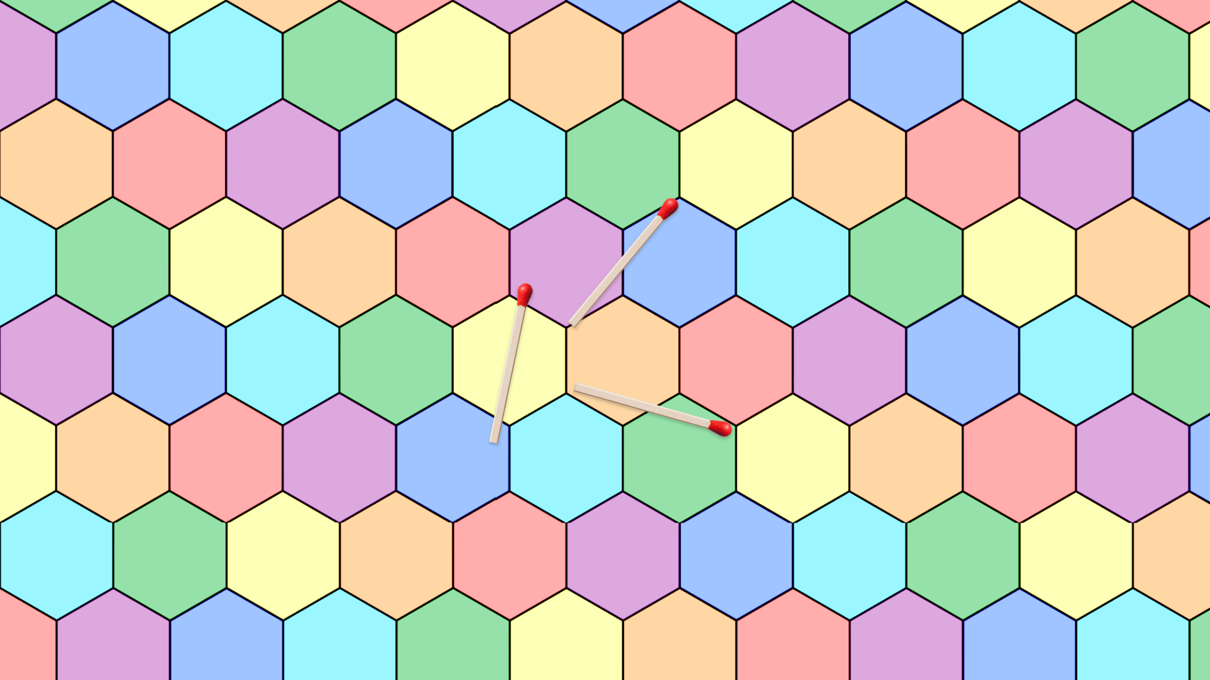
There is a unit distance graph with chromatic number 5 *and therefore*

$$\chi(\mathbb{E}^2) \geq 5.$$

Upper bounds are given by explicit colorings $g : \mathbb{E}^2 \to [c] := \{1, \ldots, c\}$, usually derived through tessellations using simple polytopal shapes, which give

$$\chi(\mathbb{E}^2) \leq 7.$$





Constructions through Implicit Representation

Question. Can we use computers to find colorings $g : \mathbb{E}^2 \to [c]$ so that

$$\left\{x \in \mathbb{E}^2 \mid g(x) = g(y) \text{ for any } y \in B_1(x)\right\} = \emptyset?$$

Idea. Consider a probabilistic relaxation to functions $p : \mathbb{E}^2 \to \Delta_c$ minimizing the loss

$$L_R(p) := \int_{[-R,R]^2} \int_{\partial B_1(x)} p(x)^T p(y) \; \mathrm{d}y \; \mathrm{d}x. \tag{1}$$

Challenge. Can we find a *parameterized and (easily) differentiable* family p_θ, i.e., an implicit representation, and optimize Equation (1) over θ through gradient descent?

Constructions through Implicit Representation

Question. Can we use computers to find colorings $g : \mathbb{E}^2 \to [c]$ so that

$$\left\{x \in \mathbb{E}^2 \mid g(x) = g(y) \text{ for any } y \in B_1(x)\right\} = \emptyset?$$

Idea. Consider a probabilistic relaxation to functions $p : \mathbb{E}^2 \to \Delta_c$ minimizing the loss

$$L_R(p) := \int_{[-R,R]^2} \int_{\partial B_1(x)} p(x)^T p(y) \; \mathrm{d}y \; \mathrm{d}x. \tag{1}$$

Challenge. Can we find a *parameterized and (easily) differentiable* family p_θ, i.e., an implicit representation, and optimize Equation (1) over θ through gradient descent?

Constructions through Implicit Representation

Question. Can we use computers to find colorings $g : \mathbb{E}^2 \to [c]$ so that

$$\left\{x \in \mathbb{E}^2 \mid g(x) = g(y) \text{ for any } y \in B_1(x)\right\} = \emptyset?$$

Idea. Consider a probabilistic relaxation to functions $p : \mathbb{E}^2 \to \Delta_c$ minimizing the loss

$$L_R(p) := \int_{[-R,R]^2} \int_{\partial B_1(x)} p(x)^T p(y) \, \mathrm{d}y \, \mathrm{d}x. \tag{1}$$

Challenge. Can we find a *parameterized and (easily) differentiable* family p_θ, i.e., an implicit representation, and optimize Equation (1) over θ through gradient descent?

How to parameterize p_θ?

A few candidates: Taylor or Fourier Series, Splines, Wavelets, SVM, RBF, ...
but we landed on simple feed-forward Neural Networks with sine activations.

Algorithm. We used *batched gradient descent* and *Monte Carlo sampling* to minimize $L_R(\theta)$: sample $x^{(i)} \in [-R, R]^2$ and $y^{(i)} \in \partial B_1(x^{(i)})$ and update parameters through

$$\theta_{k+1} = \theta_k - \alpha_k \sum_{i=1}^{m} \nabla_\theta\, p_\theta(x^{(i)}) \cdot p_\theta(y^{(i)})/m.$$

Let's see what happens if we try to color the plane with six colors...

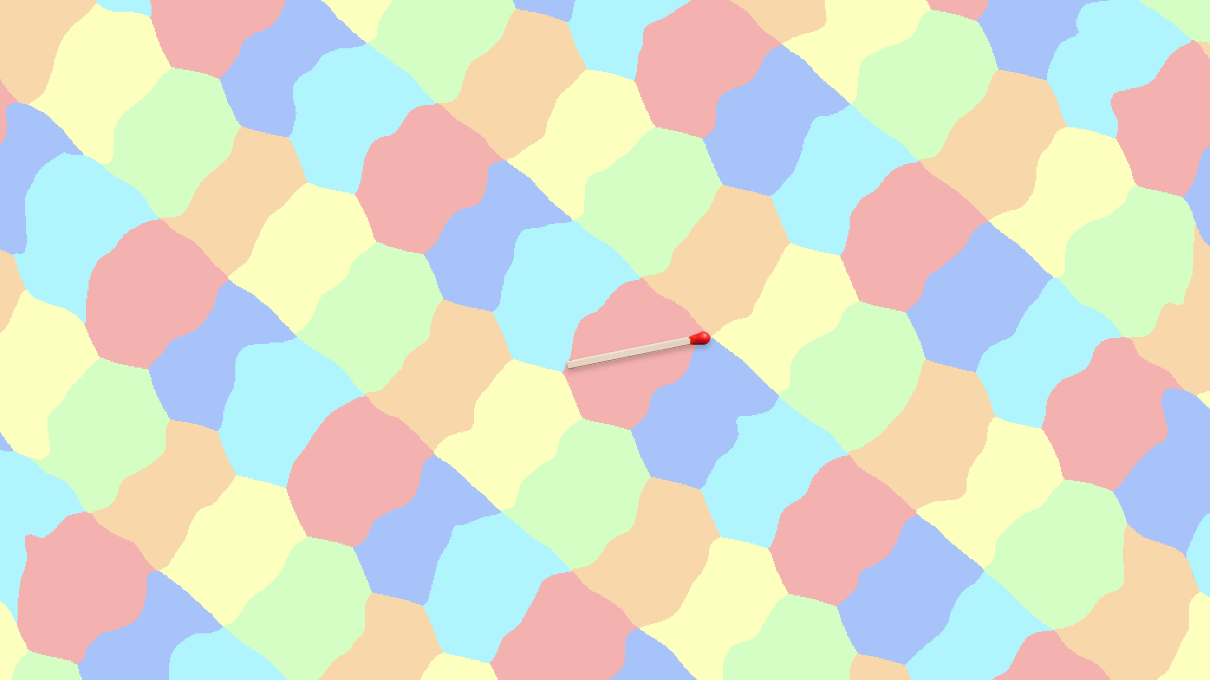
How to parameterize p_θ?

A few candidates: Taylor or Fourier Series, Splines, Wavelets, SVM, RBF, ...
but we landed on simple feed-forward Neural Networks with sine activations.

Algorithm. We used *batched gradient descent* and *Monte Carlo sampling* to minimize $L_R(\theta)$: sample $x^{(i)} \in [-R, R]^2$ and $y^{(i)} \in \partial B_1(x^{(i)})$ and update parameters through

$$\theta_{k+1} = \theta_k - \alpha_k \sum_{i=1}^{m} \nabla_\theta \, p_\theta(x^{(i)}) \cdot p_\theta(y^{(i)})/m.$$

Let's see what happens if we try to color the plane with six colors...



Unfortunately this coloring was already known...

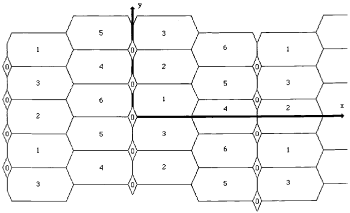
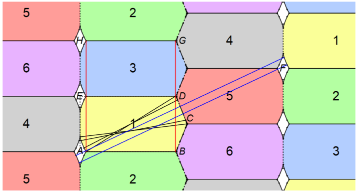
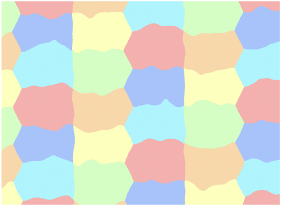


FIG. 3. A good 7-coloring of $(\mathbb{R}^2, 1)$.





Theorem (Pritikin 1995; refined by Parts 2020)

99.985% of the plane can be colored with 6 colors while avoiding unit distances.

This implies that any unit distance graph with chromatic number 7 must have order $\geq 6\,993$.

But the principle works! Can we study some variants of the original problem?

Neural Discovery in Mathematics

Application 1: Almost succeeding...

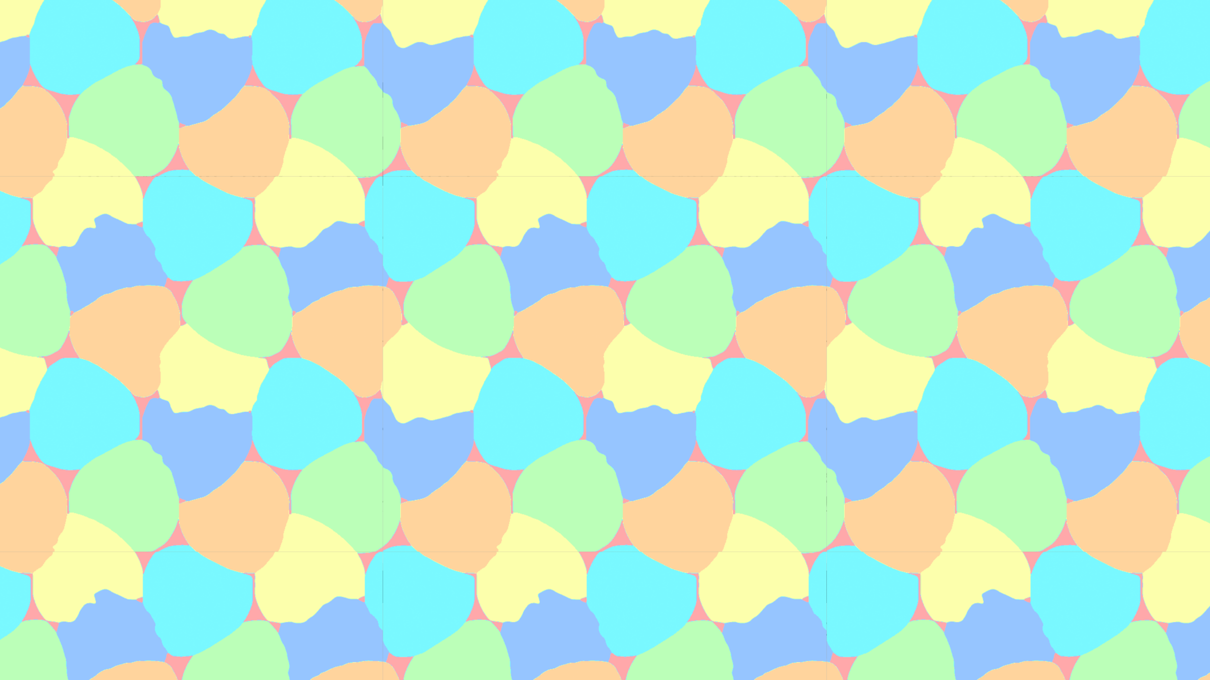
Question. What is the smallest percentage of the plane that needs to be removed so that we can color the rest with $1, 2, \ldots, 6$ colors without monochromatic conflicts?

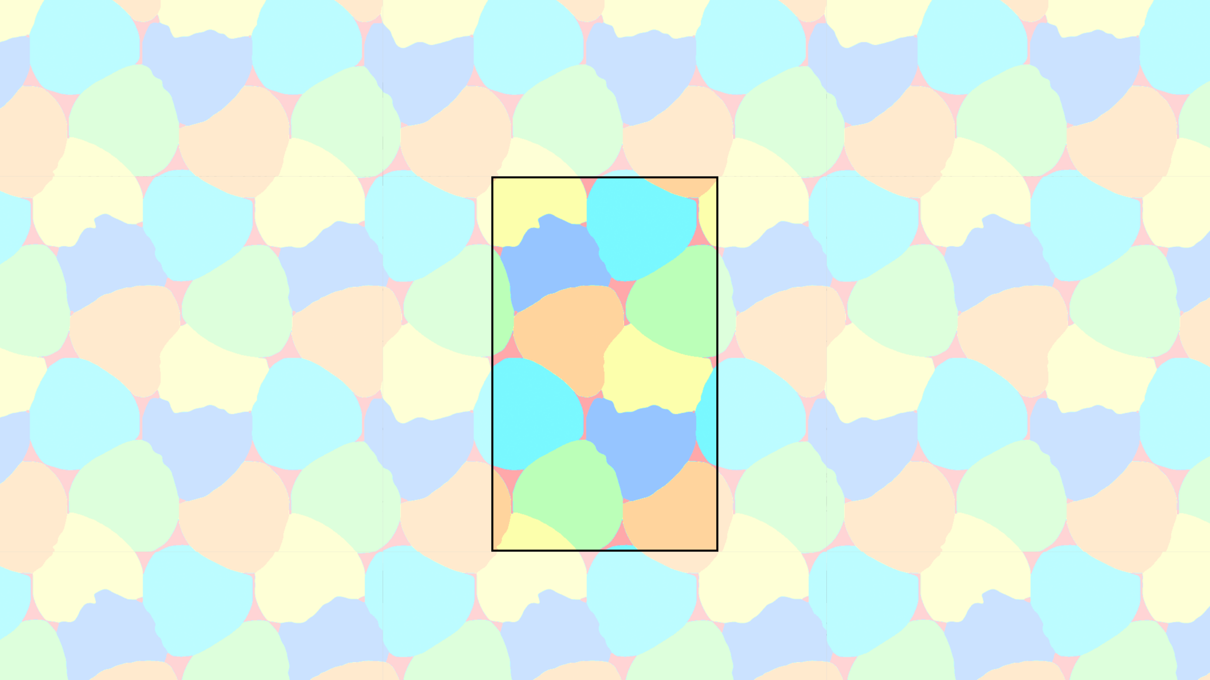
# colors	**1**	**2**	**3**	**4**	**5**	**6**
prior	77.04%	54.13%	31.20%	8.25%	4.01%	0.02%
numerics	77.07%	54.21%	31.34%	8.29%	**3.60%**	0.03%
formalized	77.13%	54.29%	31.51%	8.52%	**3.74%**	0.04%

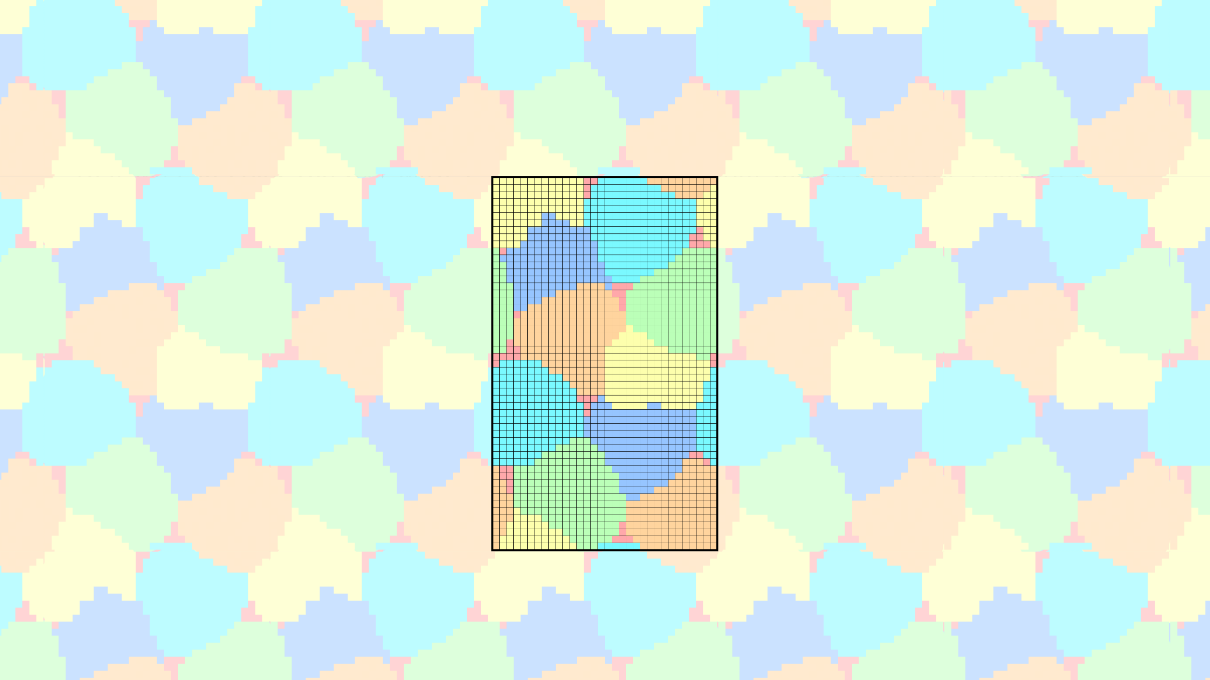
All previously best known values are due to Paarts (2020) building on work of Pritikin (1998) for 6 colors and Croft (1967) for 1, 2, 3, and 4 colors.

Theorem

96.29% of the plane can be 5-colored with no monochromatic unit distance pairs.







Application 2: Going off-diagonal...

Definition. A c-coloring *realizes* $(d_1, \ldots, d_c)$ if color i does not contain distance d_i.

Problem (The continuum of six-colorings; Soifer in Nash and Rassias' *Open Problems in Mathematics*)

Determine the set of d for which $(1, 1, 1, 1, 1, d)$ can be realized.

Soifer (1991) found a coloring for $d = 1/\sqrt{5}$. Hoffman and Soifer (1993) also found one for $d = \sqrt{2} - 1$. Both of these are part of a family that covers any

$$0.414 \approx \sqrt{2} - 1 \leq d \leq 1/\sqrt{5} \approx 0.447.$$

Theorem (Mundinger, Pokutta, S., Zimmer 2024)

The set of realizable types contains the interval $[0.354, 0.553]$.

Published in Geombinatorics Quarterly, XXXIV

Application 2: Going off-diagonal...

Definition. A c-coloring *realizes* $(d_1, \ldots, d_c)$ if color i does not contain distance d_i.

Problem (The continuum of six-colorings; Soifer in Nash and Rassias' *Open Problems in Mathematics*)

Determine the set of d for which $(1,1,1,1,1,d)$ can be realized.

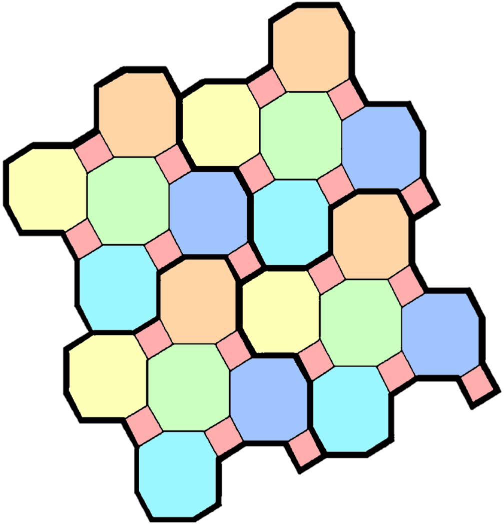
Soifer (1991) found a coloring for $d = 1/\sqrt{5}$. Hoffman and Soifer (1993) also found one for $d = \sqrt{2} - 1$. Both of these are part of a family that covers any

$$0.414 \approx \sqrt{2} - 1 \leq d \leq 1/\sqrt{5} \approx 0.447.$$

Theorem (Mundinger, Pokutta, S., Zimmer 2024)

The set of realizable types contains the interval $[0.354, 0.553]$.

Published in Geombinatorics Quarterly, XXXIV



Application 2: Going off-diagonal...

Definition. A c-coloring *realizes* $(d_1, \ldots, d_c)$ if color i does not contain distance d_i.

Problem (The continuum of six-colorings; Soifer in Nash and Rassias' *Open Problems in Mathematics*)

Determine the set of d for which $(1, 1, 1, 1, 1, d)$ can be realized.

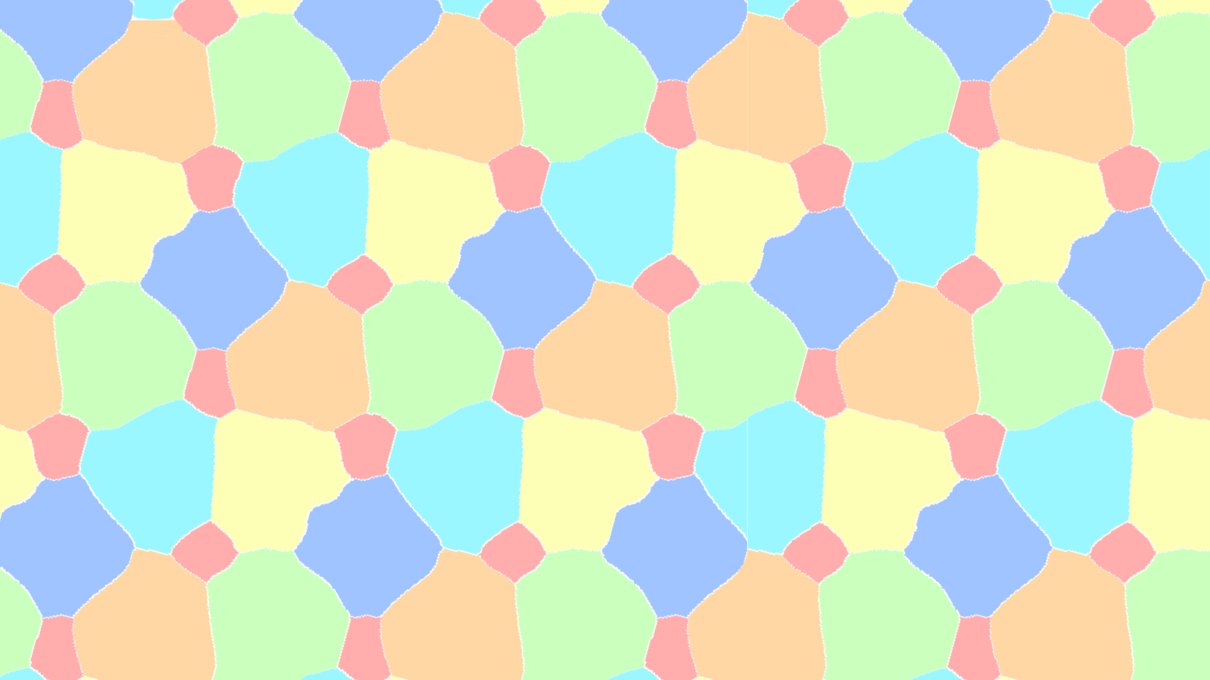
Soifer (1991) found a coloring for $d = 1/\sqrt{5}$. Hoffman and Soifer (1993) also found one for $d = \sqrt{2} - 1$. Both of these are part of a family that covers any

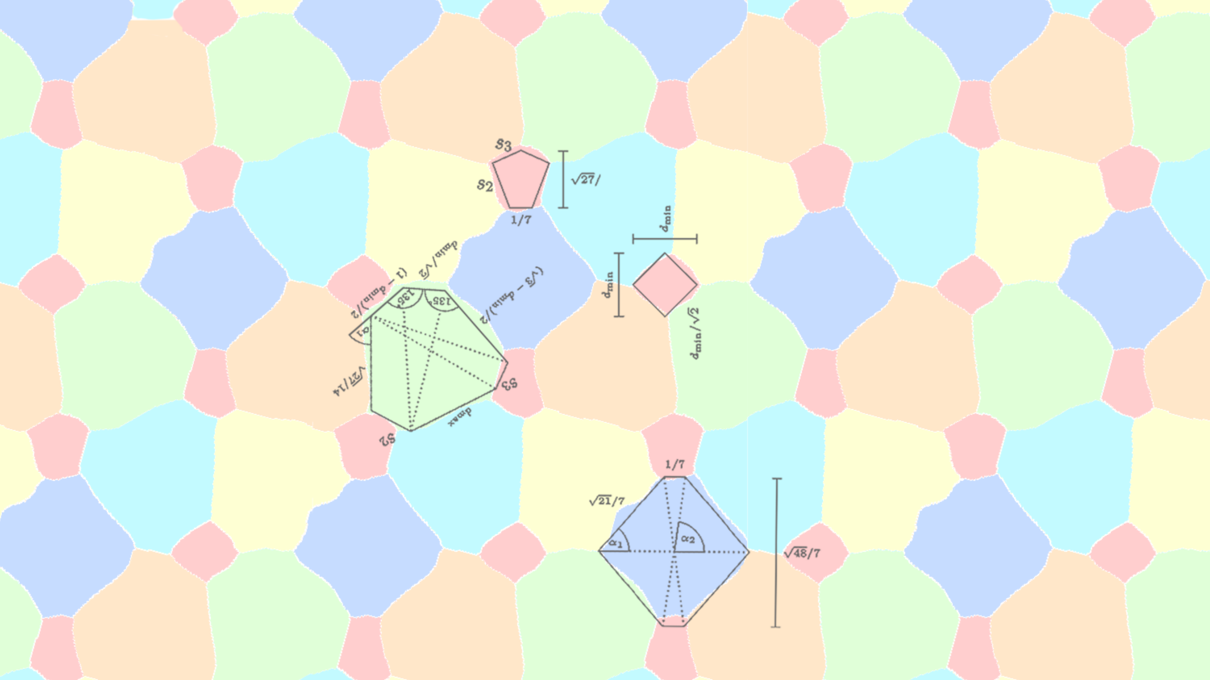
$$0.414 \approx \sqrt{2} - 1 \leq d \leq 1/\sqrt{5} \approx 0.447.$$

Theorem (Mundinger, Pokutta, S., Zimmer 2024)

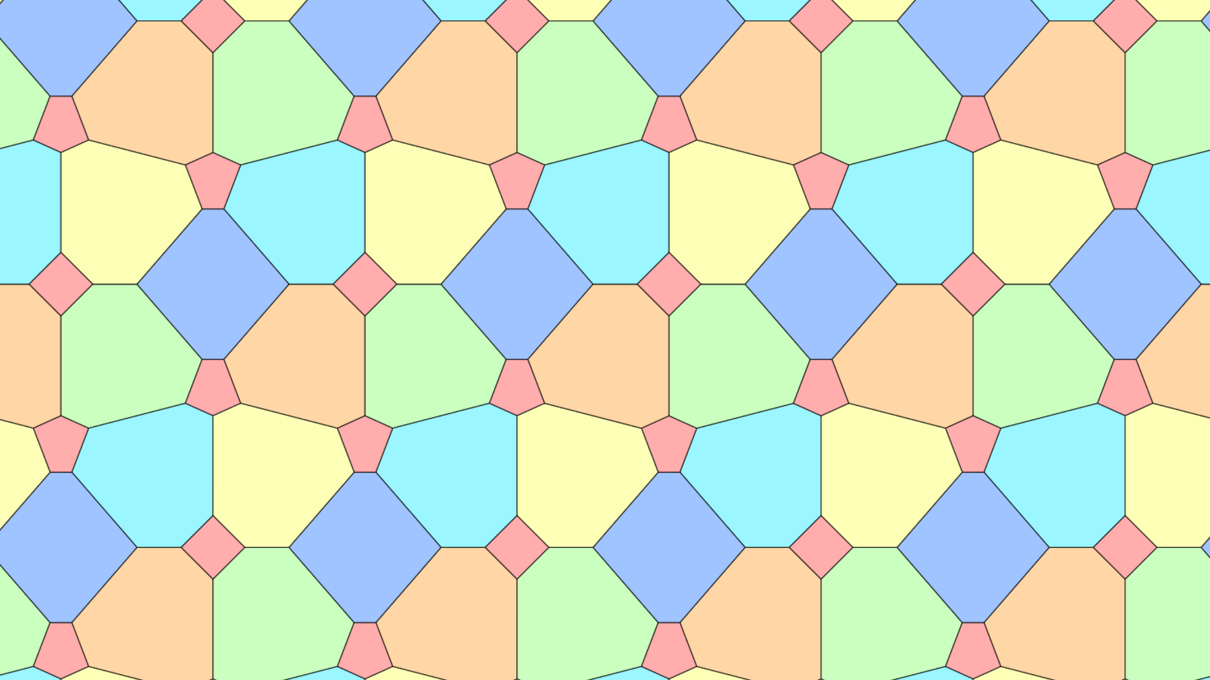
The set of realizable types contains the interval $[0.354, 0.553]$.

Published in Geombinatorics Quarterly, XXXIV

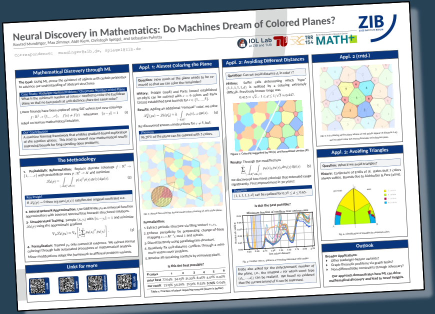




s_3
s_2
$\sqrt{27}/$
1/7
d_{min}
d_{min}
$d_{min}/\sqrt{2}$
1/7
$\sqrt{21}/7$
α_1
α_2
$\sqrt{48}/7$



Thank you for your attention!



Poster #W-401

Today 11:00–13:30

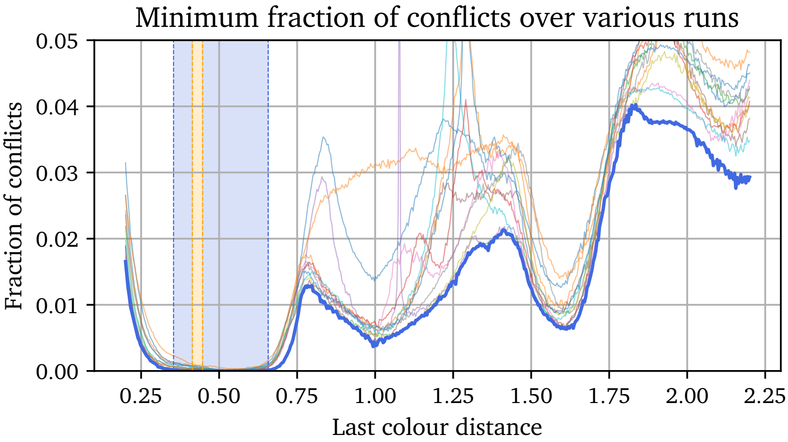
West Exhibition Hall

Join us!

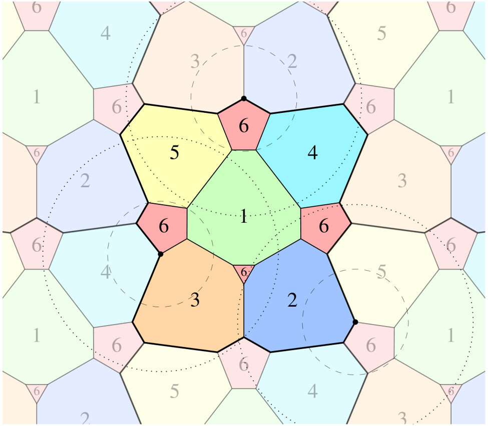
Master students · PhD students · Research scientists · Visiting doctoral fellows

iol.zib.de/join

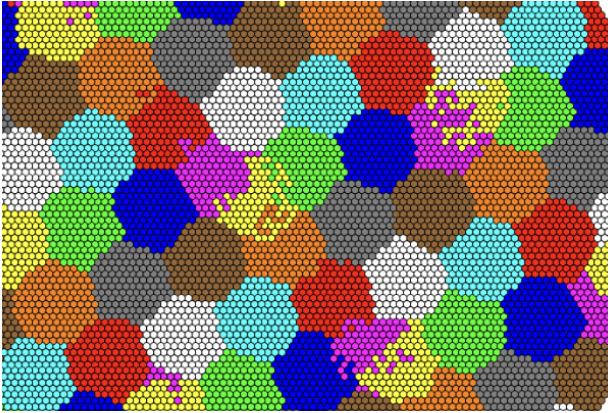
Is the off-diagonal result optimal?



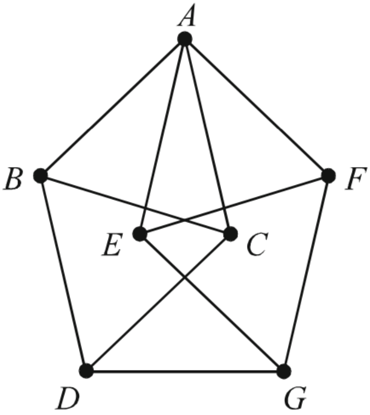
The other off-diagonal coloring



Constructing colorings through SAT solvers



The Moser spindle



Avoiding monochromatic triangles