Berlin Mathematics Research Center
MATH+

Do Machines Dream of Colored Planes?

Christoph Spiegel

joint work with


Konrad Plato
BMS


Max Zimmer
BMS


Aldo Kiem
BMS / MATH+


Sebastian Pokutta
ZIB / TUB / MATH+

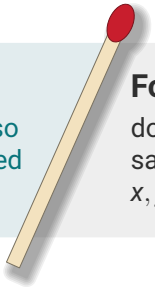
ZIB

A seemingly simple question …

… but a famous open problem in discrete geometry. Posed in the 1950s by 18-year-old Edward Nelson, popularized in the 1960s by Paul Erdős, known as *Hadwiger–Nelson*.

How many colors do you need

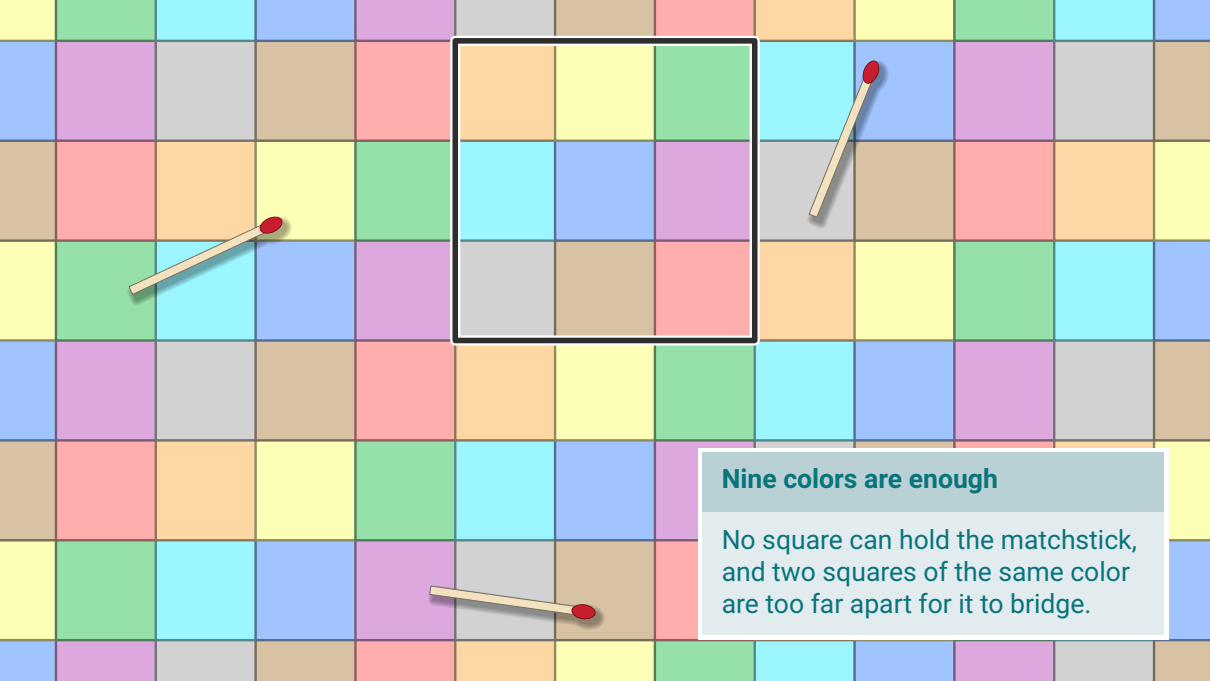
to color an (unbounded) sheet of paper so that the two ends of a matchstick dropped onto it never touch the same color?

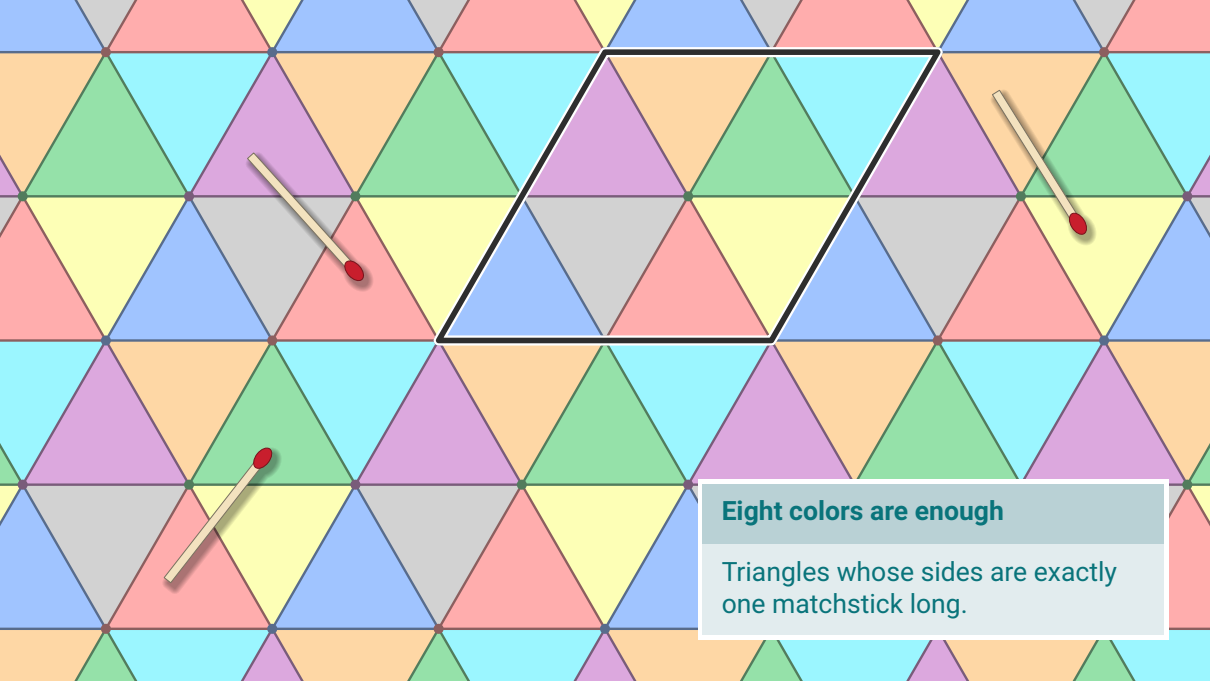


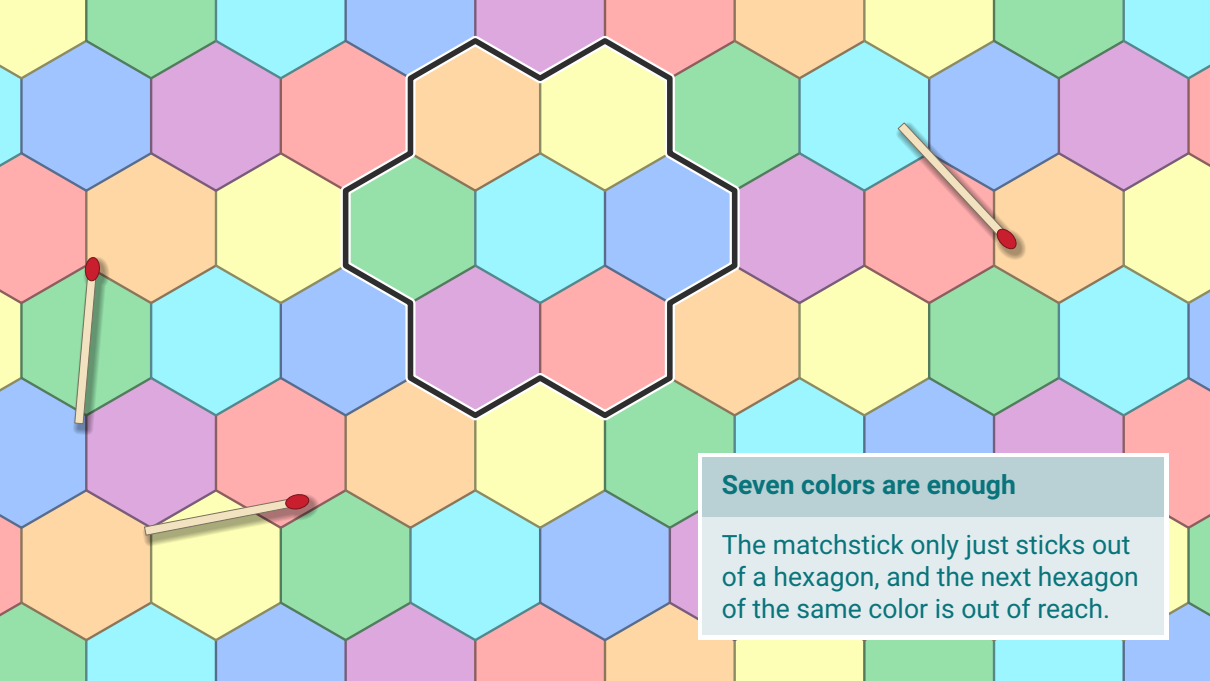
For which $c \in \mathbb{N}$

does there exist $g\colon \mathbb{R}^2 \to [c]$ satisfying $g(x) \neq g(y)$ for any $x, y \in \mathbb{R}^2$ with $\|x - y\| = 1$?

For 75 years, such colorings were largely designed by hand. Here we explore if a neural network can learn to find them.


Nine colors are enough
No square can hold the matchstick, and two squares of the same color are too far apart for it to bridge.


Eight colors are enough
Triangles whose sides are exactly one matchstick long.


Seven colors are enough
The matchstick only just sticks out of a hexagon, and the next hexagon of the same color is out of reach.

So how many colors do you need?

1 2 3 4 5 6 7 8 9

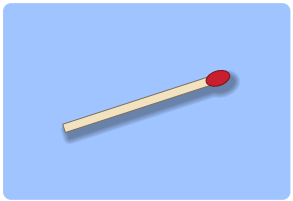
known since the 1950s

Seven are enough. Can it be done with fewer?

So how many colors do you need?

1	2	3	4	5	6	7	8	9

known since the 1950s

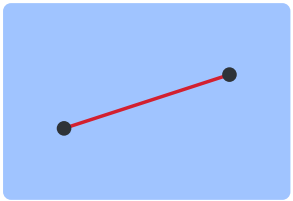


One color is not enough
shown through 1 matchstick

So how many colors do you need?

1	2	3	4	5	6	7	8	9

known since the 1950s



One color is not enough

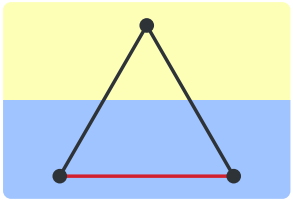
shown through 1 matchstick

The edge is a unit distance graph with chromatic number 2

So how many colors do you need?

1	2	3	4	5	6	7	8	9

known since the 1950s



Two colors are not enough

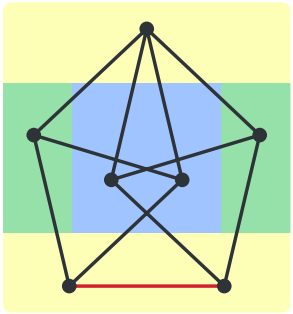
shown through 3 matchsticks

The triangle is a unit distance graph with chromatic number 3

So how many colors do you need?

1	2	3	4	5	6	7	8	9

1961 *known since the 1950s*



Three colors are not enough

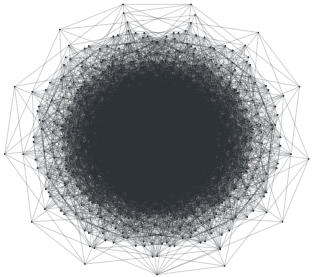
shown through 11 matchsticks

The Moser spindle is a unit distance graph with chromatic number 4

Moser & Moser 1961

So how many colors do you need?

1	2	3	4	5	6	7	8	9
		1961	*2018*			*known since the 1950s*		



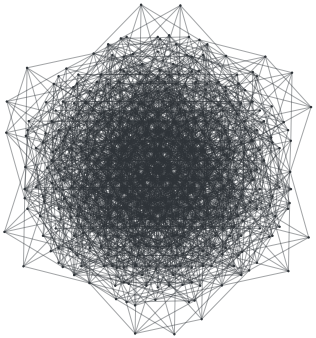
Four colors are not enough
shown through 7,877 matchsticks

Aubrey de Grey found the first unit distance graph with chromatic number 5

de Grey 2018

So how many colors do you need?

1	2	3	4	5	6	7	8	9
		1961	*2018*			*known since the 1950s*		



Four colors are not enough

shown through ~~7,877~~ 2,504 matchsticks

A 2018 Polymath project simplified the construction of de Grey

Heule 2019

So how many colors do you need?

1	2	3	4	5	6	7	8	9
		1961	*2018*	*2026*		*known since the 1950s*		



Five colors are not enough

shown without an explicit construction

Result #158 in OpenAI's October release of 721 manuscripts establishes that five colors do not suffice

OpenAI 2026

So how many colors do you need?

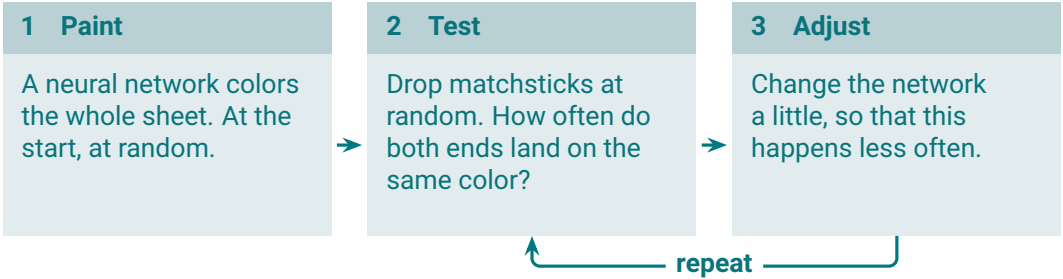
1	2	3	4	5	6	7	8	9
		1961	*2018*	*2026*	*open*	*known since the 1950s*		

Six or seven? Open for more than 75 years.

Colorings were largely designed by hand.

We turn it into a learning problem.

What if a machine looked for the coloring?



What if a machine looked for the coloring?

Implicit representation

Model a probabilistically relaxed coloring through a neural network

$$p_\theta : \mathbb{R}^2 \to \Delta_c.$$

→

2 Test

Drop matchsticks at random. How often do both ends land on the same color?

→

3 Adjust

Change the network a little, so that this happens less often.

repeat

What if a machine looked for the coloring?

Implicit representation

Model a probabilistically relaxed coloring through a neural network

$$p_\theta : \mathbb{R}^2 \to \Delta_c.$$

→

Probabilistic sampling

Sample pairs of points satisfying $\|x - y\| = 1$ to estimate the loss

$$\mathcal{L}(\theta) = \mathbb{E}\big[\, p_\theta(x)^\top p_\theta(y) \,\big].$$

→

3 Adjust

Change the network a little, so that this happens less often.

repeat

What if a machine looked for the coloring?

Implicit representation

Model a probabilistically relaxed coloring through a neural network

$$p_\theta \colon \mathbb{R}^2 \to \Delta_c.$$

→

Probabilistic sampling

Sample pairs of points satisfying $\|x - y\| = 1$ to estimate the loss

$$\mathcal{L}(\theta) = \mathbb{E}\big[\, p_\theta(x)^\top p_\theta(y) \,\big].$$

→

Gradient step

Use backpropagation to take gradient descent steps

$$\theta \leftarrow \theta - \eta \, \nabla_\theta \, \widehat{\mathcal{L}}(\theta).$$

repeat

What if a machine looked for the coloring?

Implicit representation

Model a probabilistically relaxed coloring through a neural network

$$p_\theta : \mathbb{R}^2 \to \Delta_c.$$

→

Probabilistic sampling

Sample pairs of points satisfying $\|x - y\| = 1$ to estimate the loss

$$\mathcal{L}(\theta) = \mathbb{E}\big[\, p_\theta(x)^\top p_\theta(y) \,\big].$$

→

Gradient step

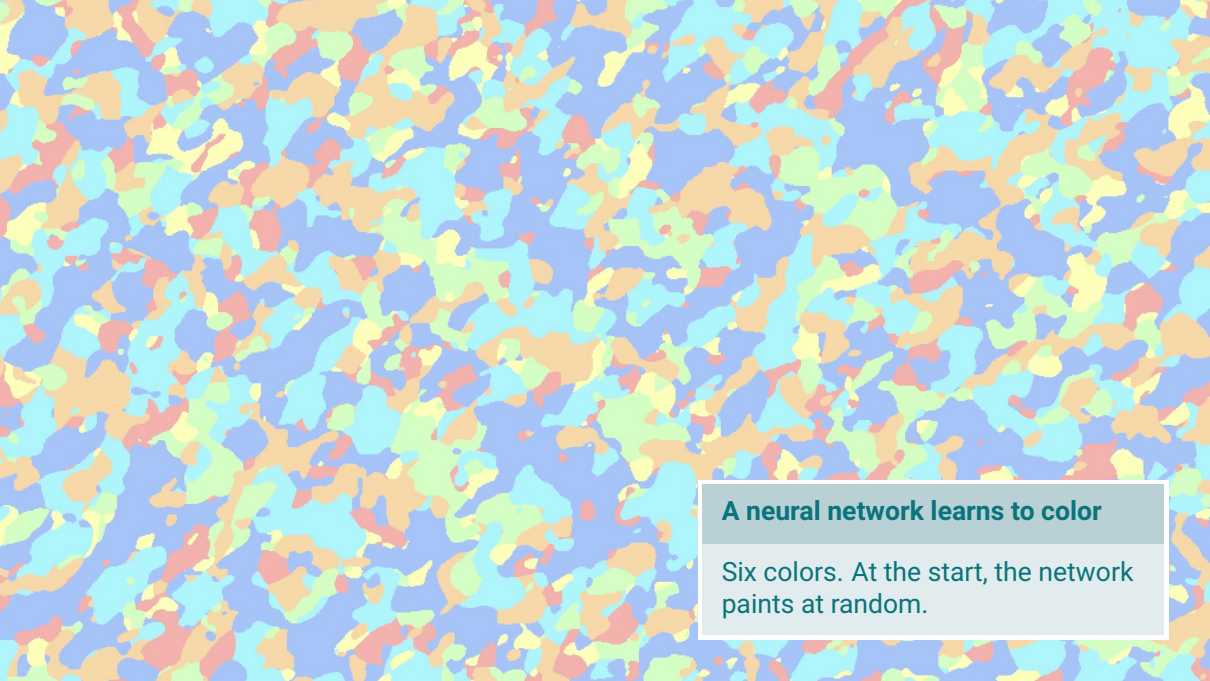
Use backpropagation to take gradient descent steps

$$\theta \leftarrow \theta - \eta \, \nabla_\theta \, \widehat{\mathcal{L}}(\theta).$$

repeat

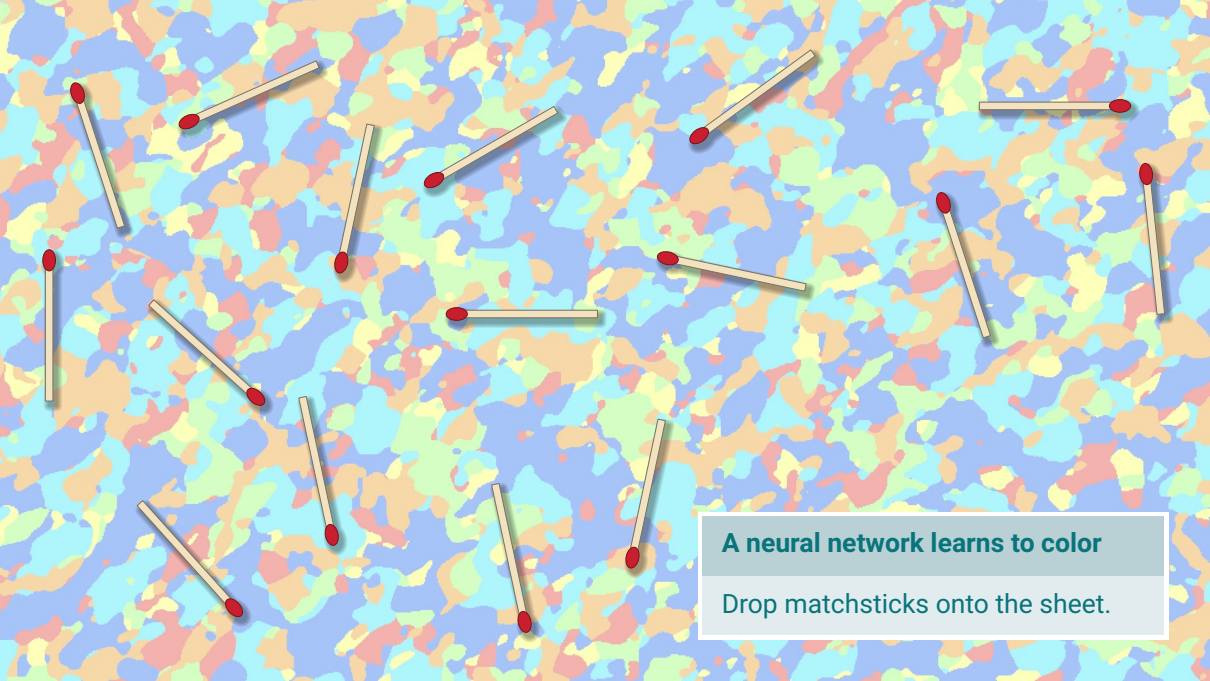
Universal approximation
says a neural network can represent the coloring.

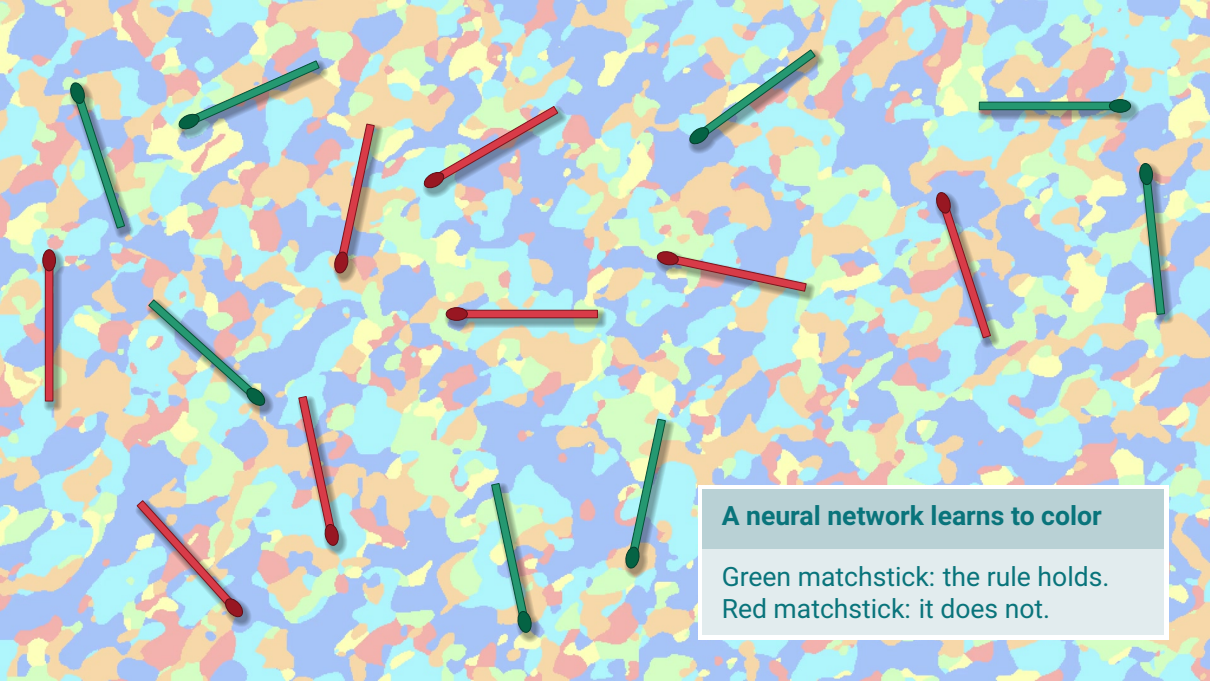
Experience in ML
says gradient descent will probably find it.

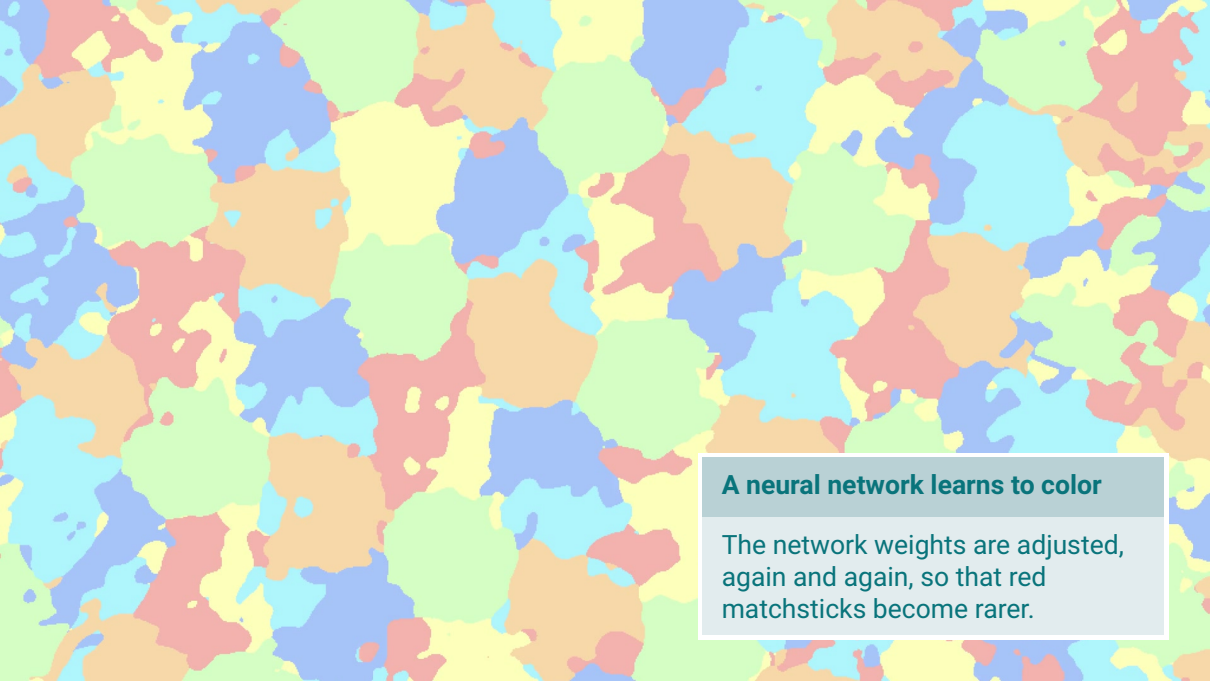


A neural network learns to color

Six colors. At the start, the network paints at random.

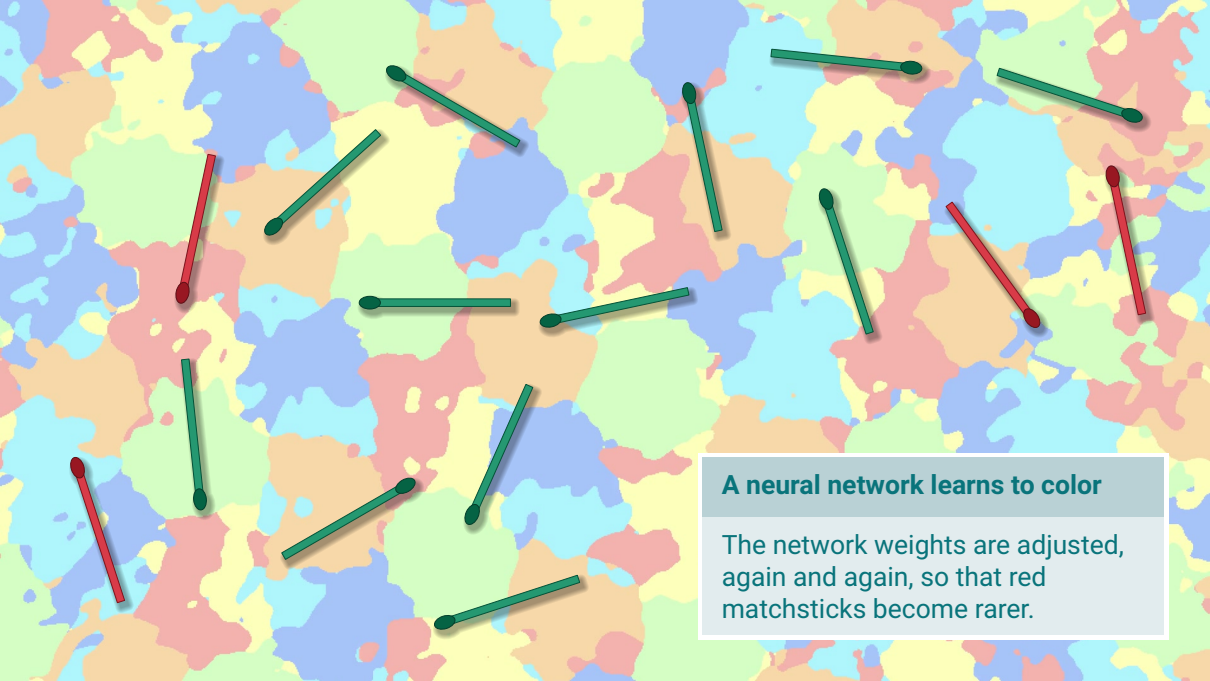

A neural network learns to color
Drop matchsticks onto the sheet.

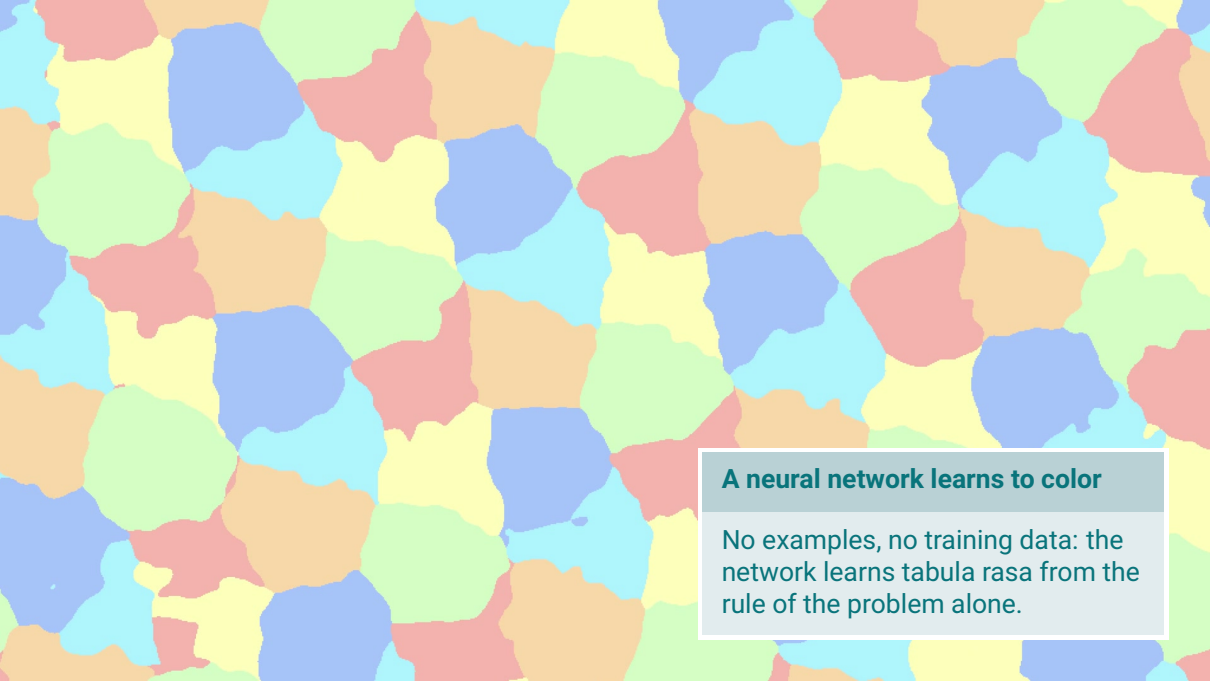

A neural network learns to color
Green matchstick: the rule holds.
Red matchstick: it does not.



A neural network learns to color

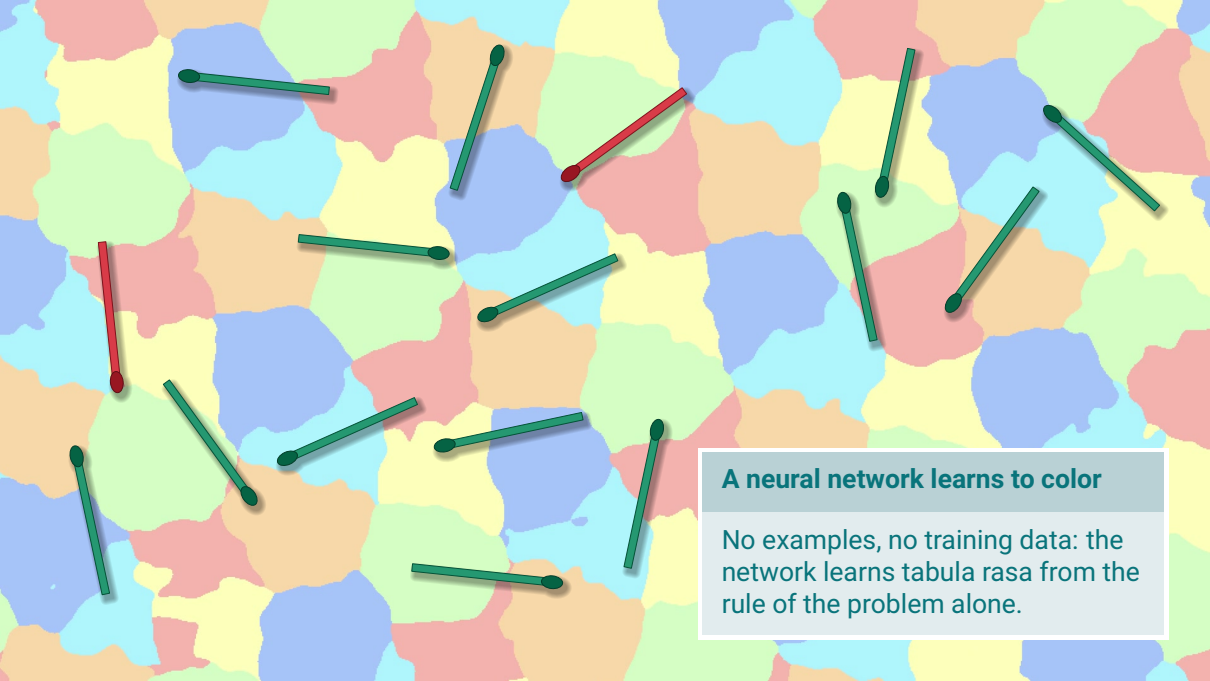
The network weights are adjusted, again and again, so that red matchsticks become rarer.

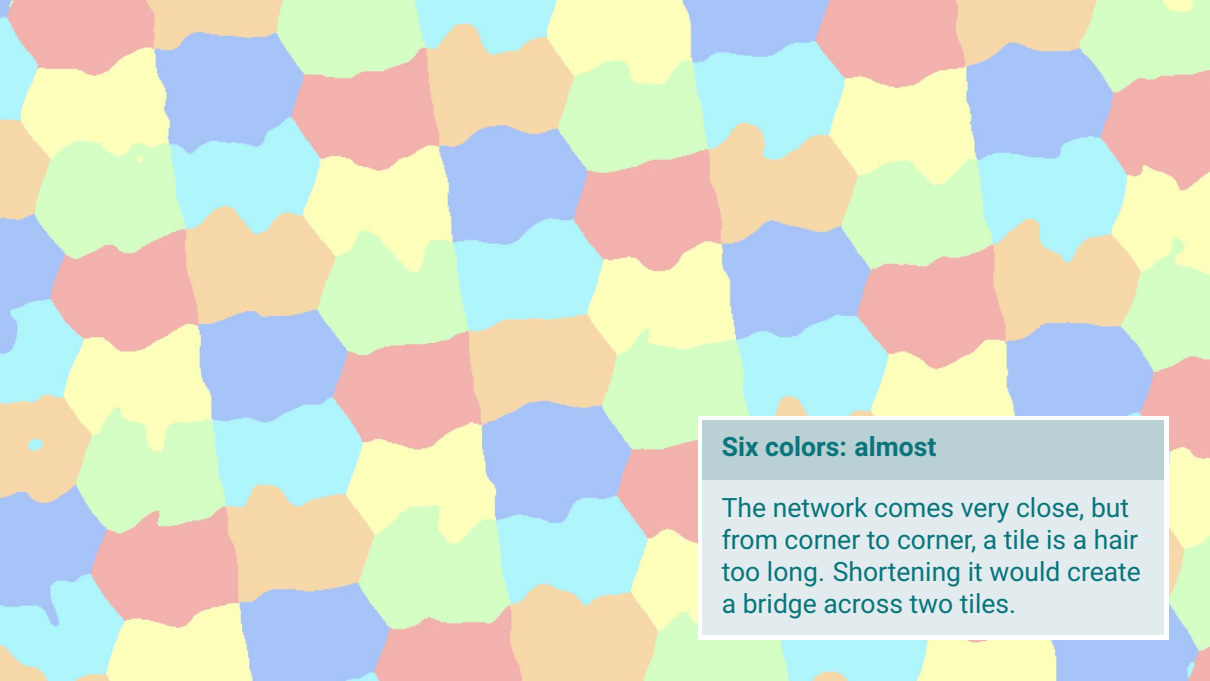

A neural network learns to color
The network weights are adjusted, again and again, so that red matchsticks become rarer.



A neural network learns to color

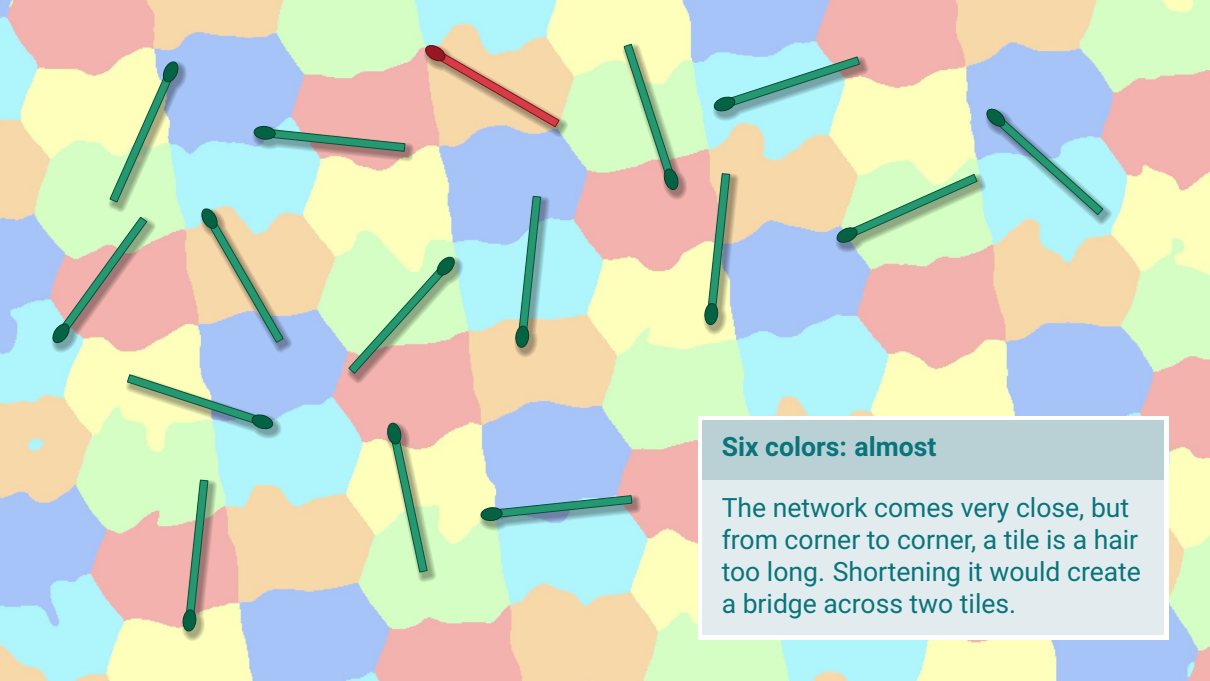
No examples, no training data: the network learns tabula rasa from the rule of the problem alone.


A neural network learns to color
No examples, no training data: the network learns tabula rasa from the rule of the problem alone.



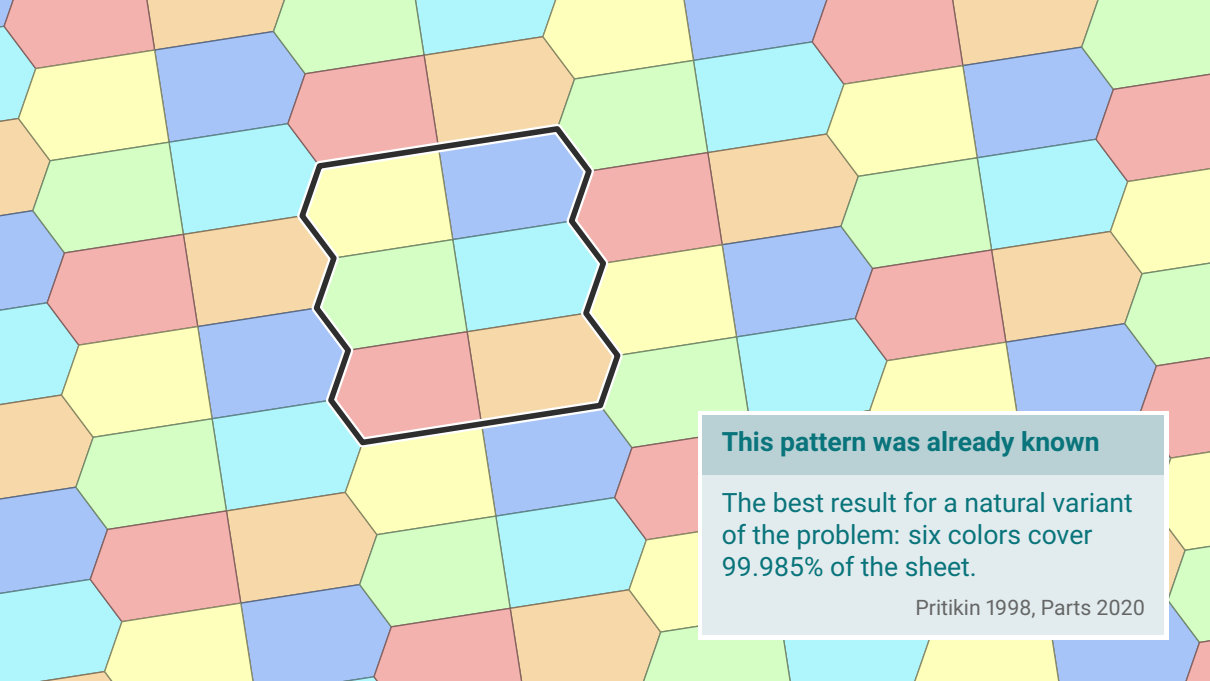
Six colors: almost

The network comes very close, but from corner to corner, a tile is a hair too long. Shortening it would create a bridge across two tiles.



Six colors: almost

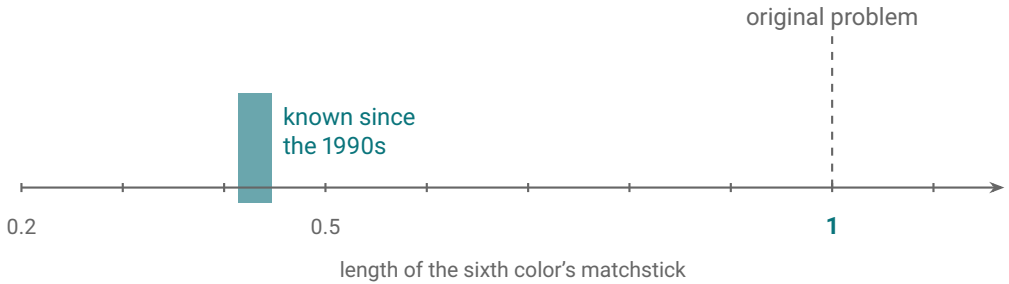
The network comes very close, but from corner to corner, a tile is a hair too long. Shortening it would create a bridge across two tiles.


This pattern was already known
The best result for a natural variant of the problem: six colors cover 99.985% of the sheet.
Pritikin 1998, Parts 2020

Another variant: one color with a rule of its own

For which lengths Soifer 1991

can the sixth color get a matchstick of its own while the other five follow the rule?

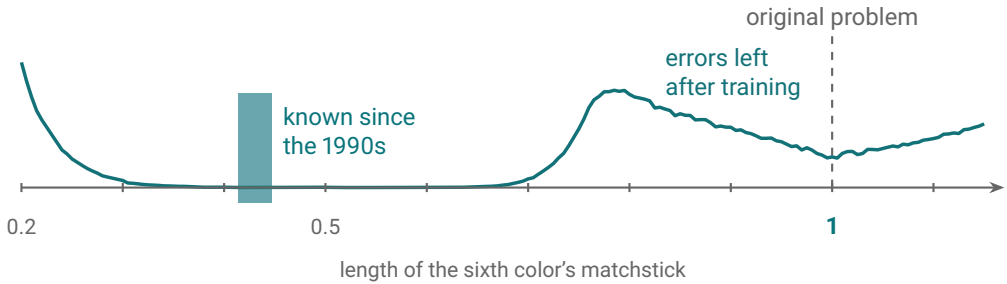


Another variant: one color with a rule of its own

For which lengths

Soifer 1991

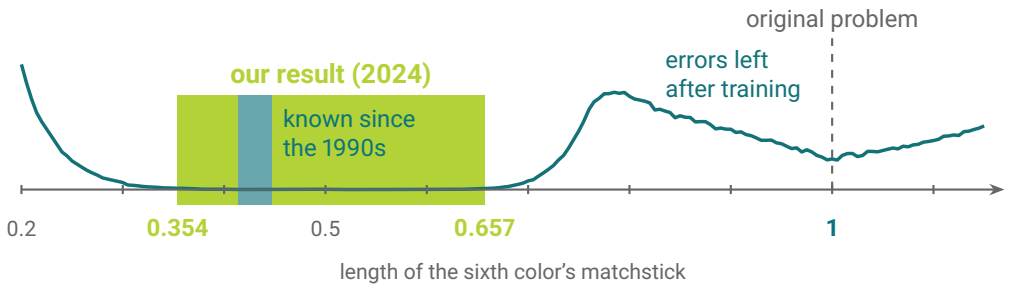
can the sixth color get a matchstick of its own while the other five follow the rule?

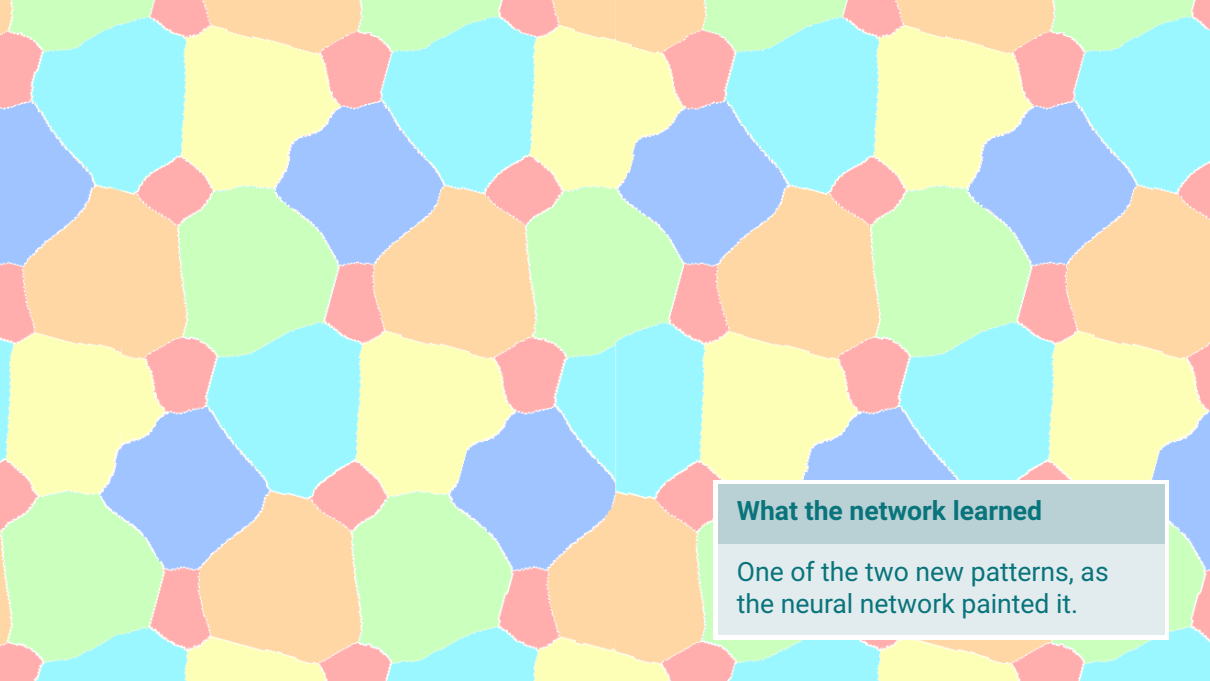


Another variant: one color with a rule of its own

For which lengths Soifer 1991

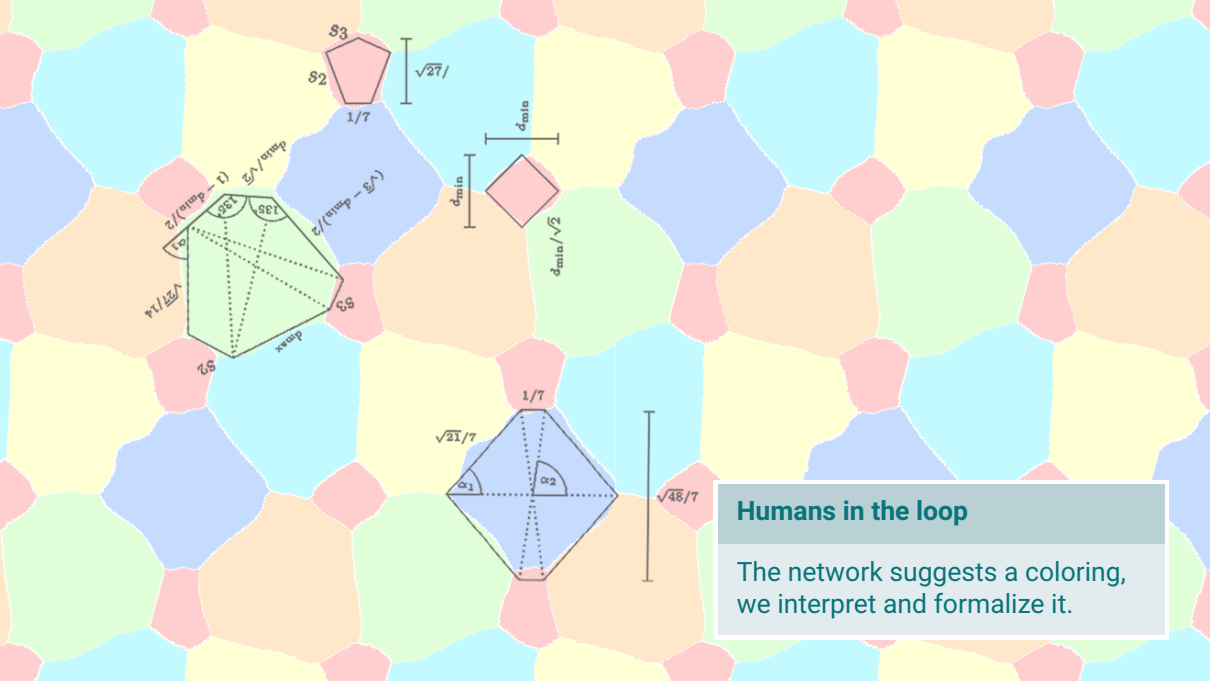
can the sixth color get a matchstick of its own while the other five follow the rule?





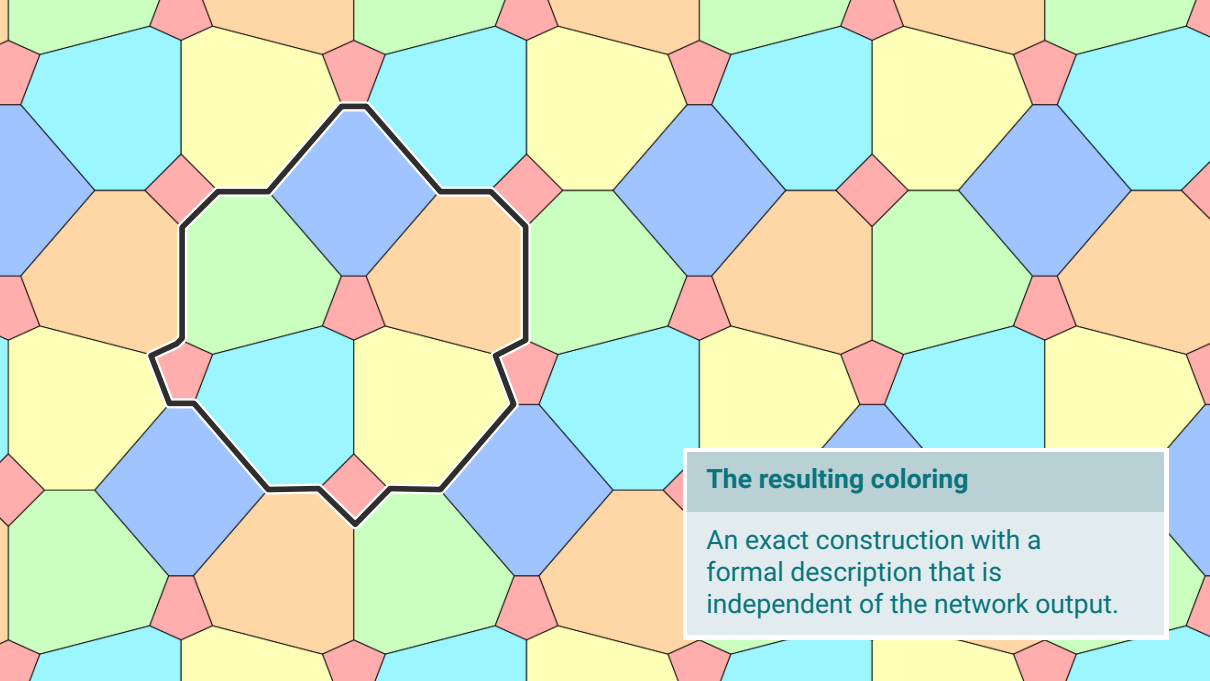
What the network learned

One of the two new patterns, as the neural network painted it.



Humans in the loop

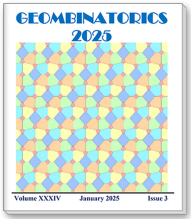
The network suggests a coloring, we interpret and formalize it.



The resulting coloring

An exact construction with a formal description that is independent of the network output.

One result, three audiences

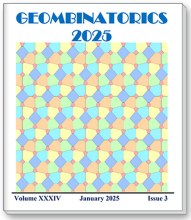


A new statement in discrete geometry

The first progress on this variant of the Hadwiger–Nelson problem **in 30 years**.
And a significant one!

Geombinatorics XXXIV (2024)

One result, three audiences



A new statement in discrete geometry

The first progress on this variant of the Hadwiger–Nelson problem **in 30 years**.
And a significant one!

Geombinatorics XXXIV (2024)

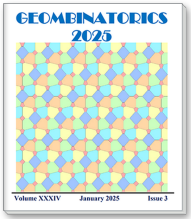


A novel use of machine learning

An oral presentation at ICML 2025, one of the leading ML conferences. Reserved for **top 1 %** of 12,107 submissions.

ICML 2025, Vancouver

One result, three audiences



A new statement in discrete geometry

The first progress on this variant of the Hadwiger–Nelson problem **in 30 years**.
And a significant one!

Geombinatorics XXXIV (2024)



A novel use of machine learning

An oral presentation at ICML 2025, one of the leading ML conferences. Reserved for **top 1 %** of 12,107 submissions.

ICML 2025, Vancouver



An interactive way to communicate science

On touch screens at the **Long Night of the Sciences**, visitors learn about the problem by coloring shapes.

LNdW at ZIB (2025 and 26)

A MATH+ collaboration across fields

Grew out of

EF1-12 2021 – 24
Learning Extremal Structures in Combinatorics



EF1-21 2022 – 25
Scaling up Flag Algebras in Combinatorics



Interplay with

TRR 154 2022 – 26
Mathematical Modelling of Gas Networks



AI4Forest 2023 – 27
AI-Based High-Resolution Forest Monitoring



Continues in

EF-LI-Opt-1 2025 – 28
Neural Generative Models for Algebraic Curves



EF-LI-Opt-3 2025 – 28
Agentic AI in Maths



Here, pure mathematics was the application of machine learning.

BMS students were motivated by a hard problem in discrete geometry, but took what they learned and applied it to gas networks and forest monitoring. **Made possible by MATH+.**