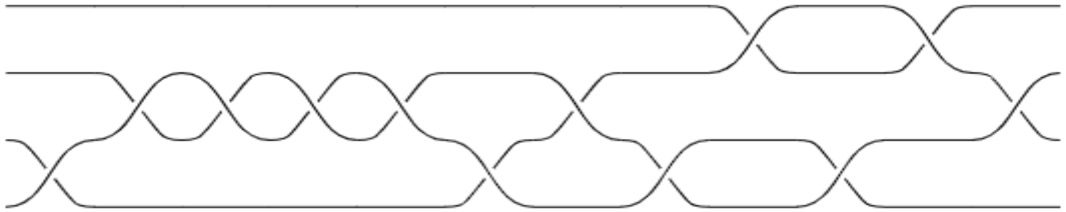
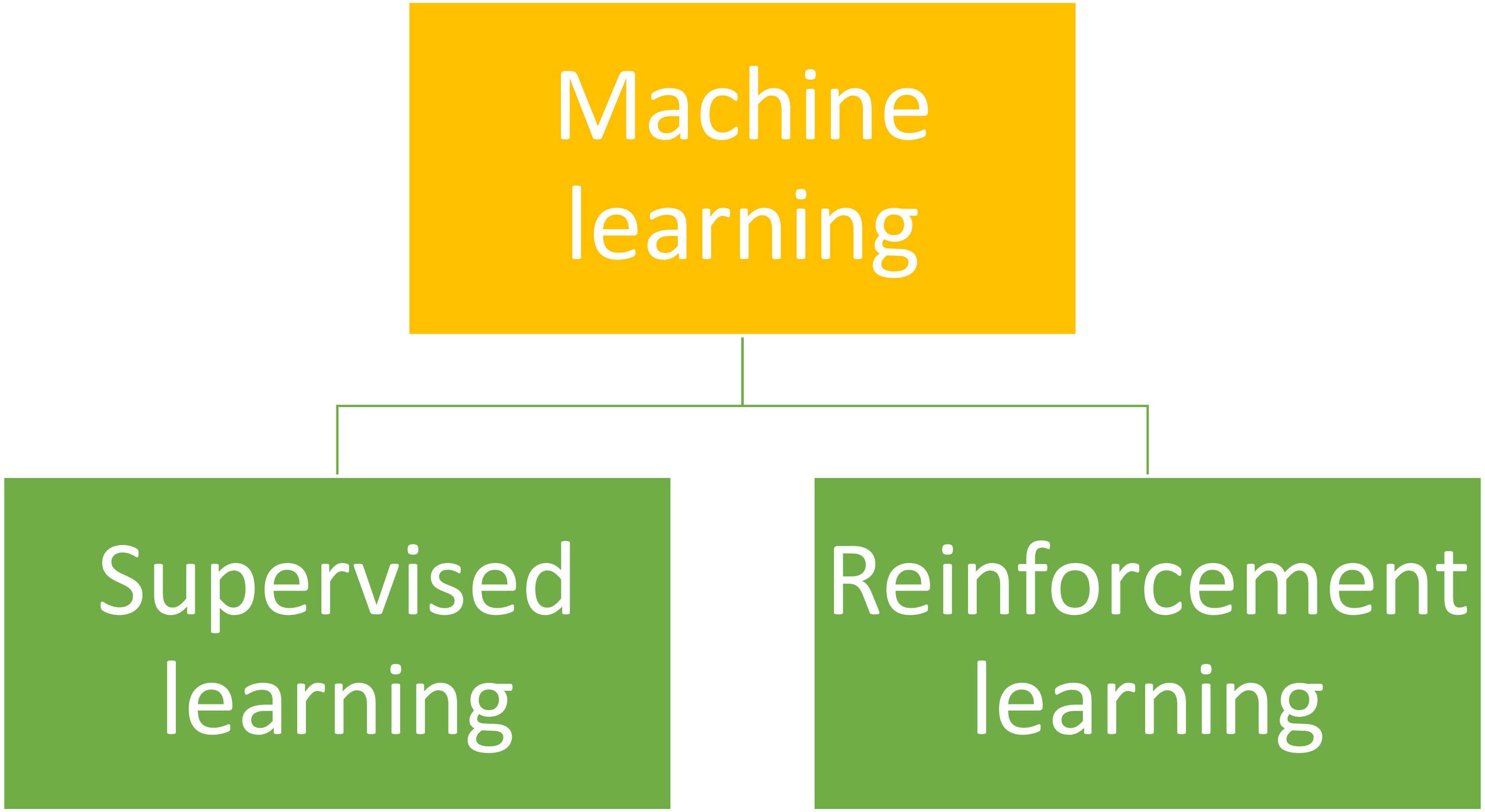
Using machine learning to solve mathematical problems and to search for examples and counterexamples in pure maths research



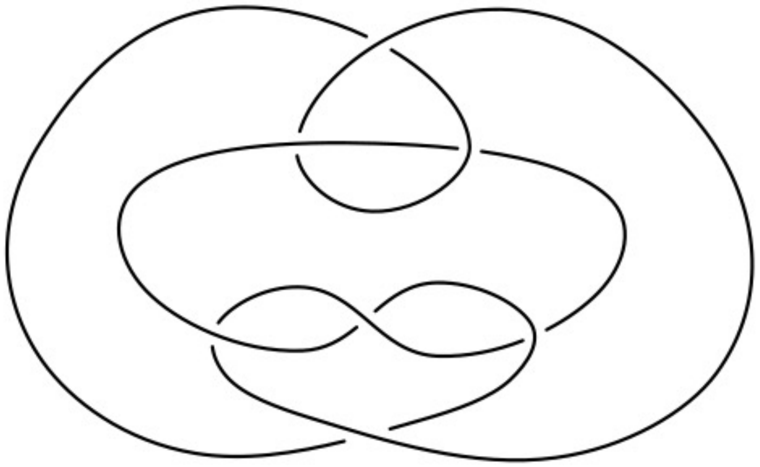
Alexei Vernitski

University of Essex

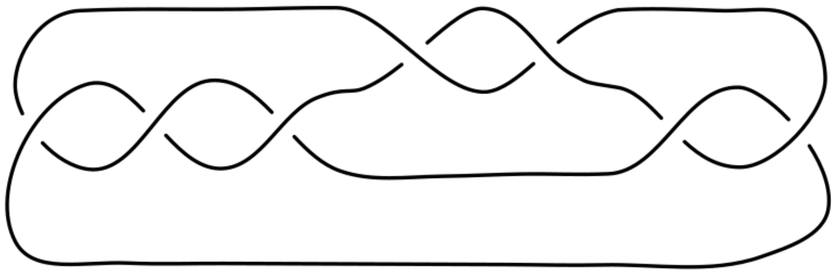
What is machine learning?



Supervised learning: the training phase

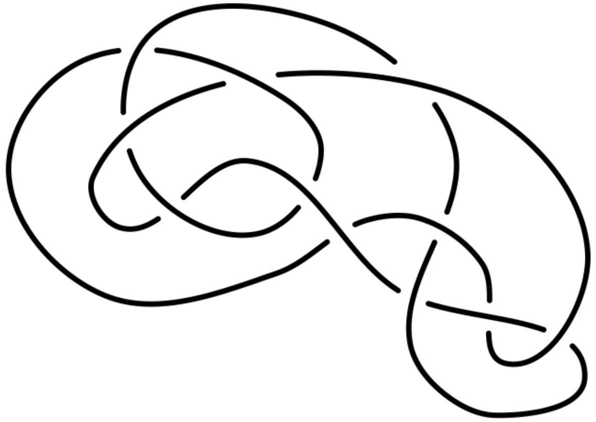


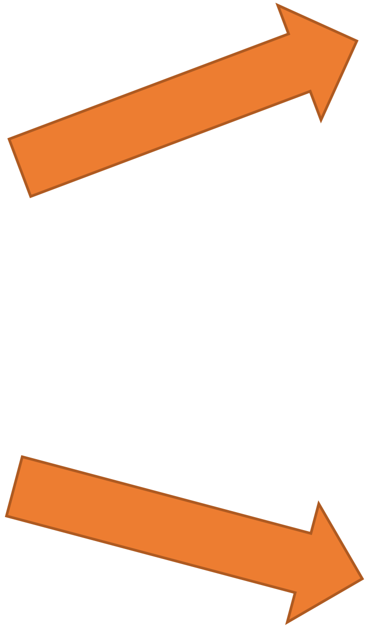
Not knotted



Knotted

Supervised learning: the testing phase

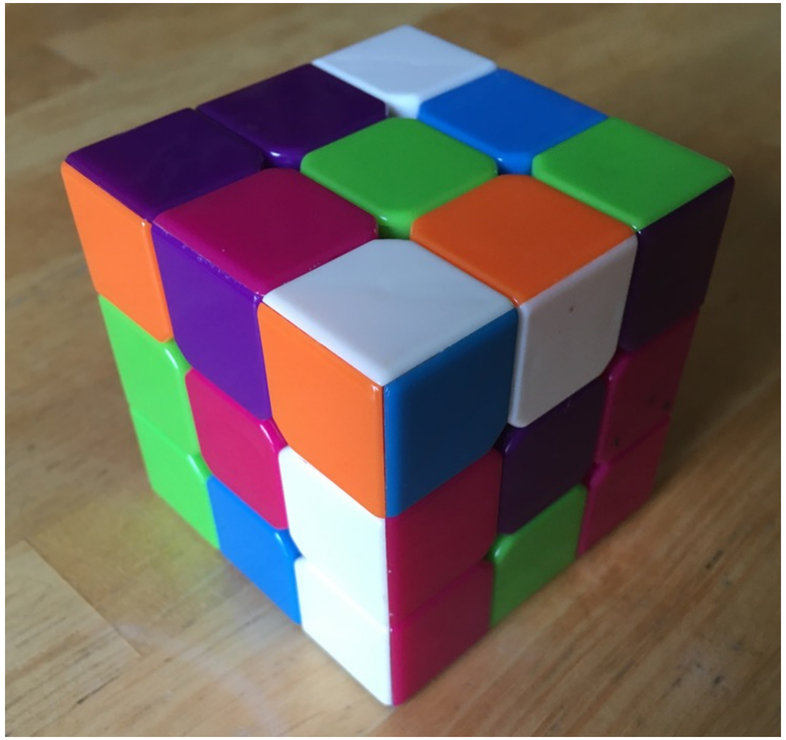




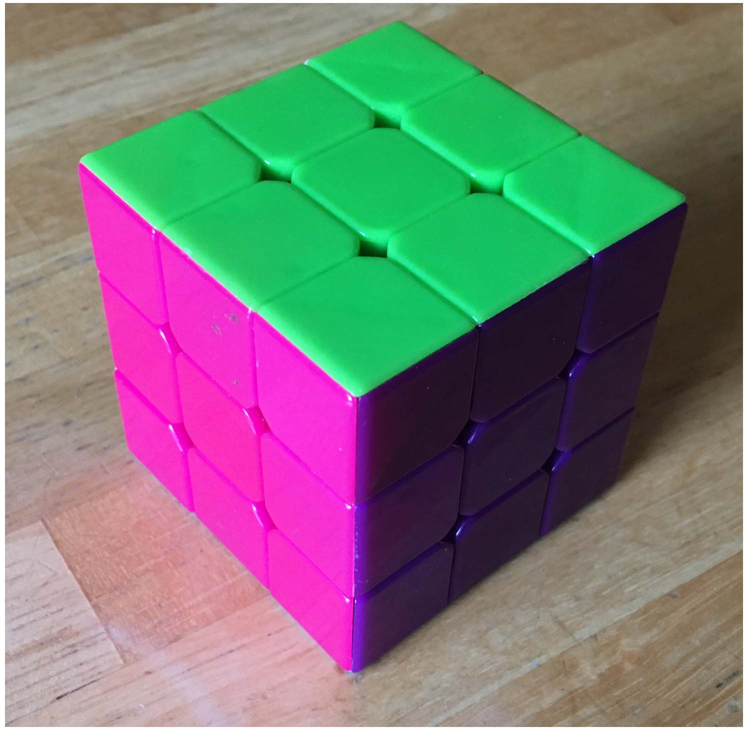
Not knotted?

Knotted?

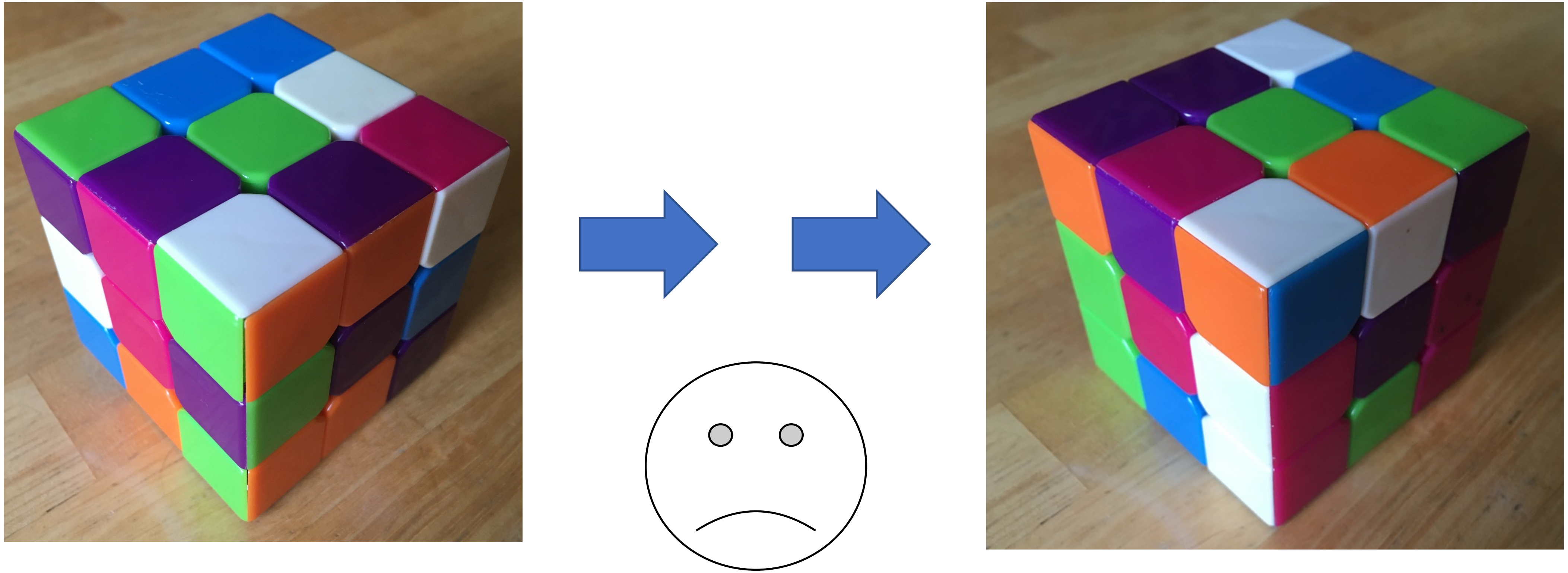
Reinforcement learning: making some random moves



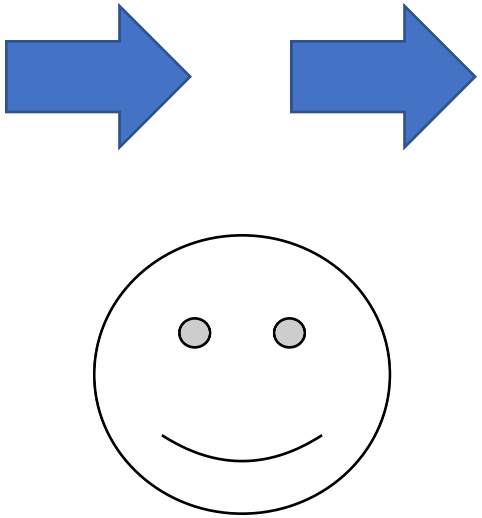
Reinforcement learning: making some other random moves

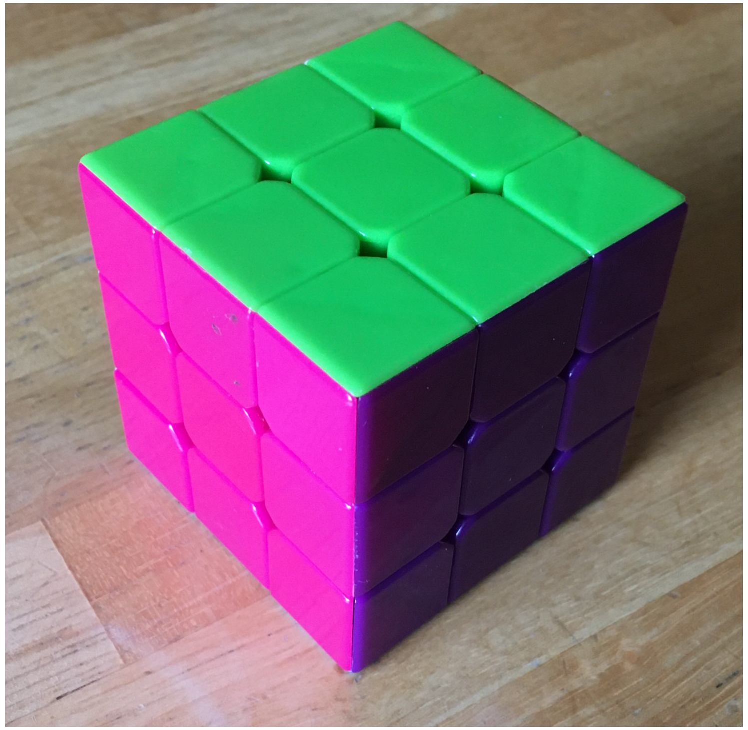


We want less of these moves, whatever they were



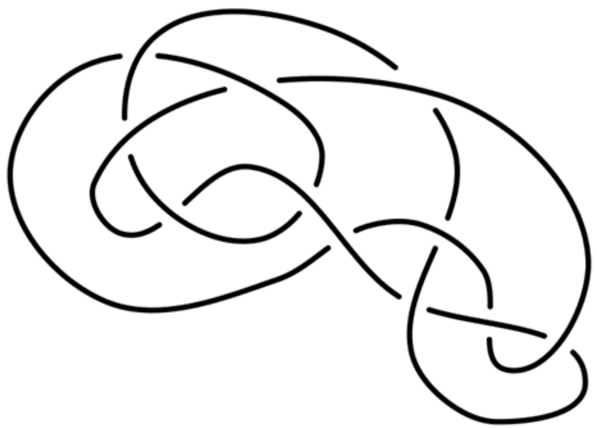
We want more of these moves, whatever they were





Machine learning

Supervised learning
- learning on examples
- used in classification

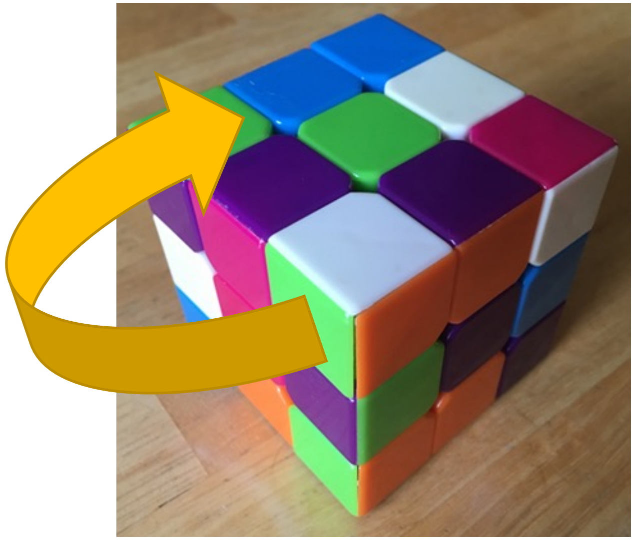


Reinforcement learning
- learning from rewards
- used to develop desirable multi-step behaviour

Machine learning aided by deep learning

For each meaningful move, there are too many positions (about 10^{18}) in which it can be usefully applied.

It is impossible to store them all.

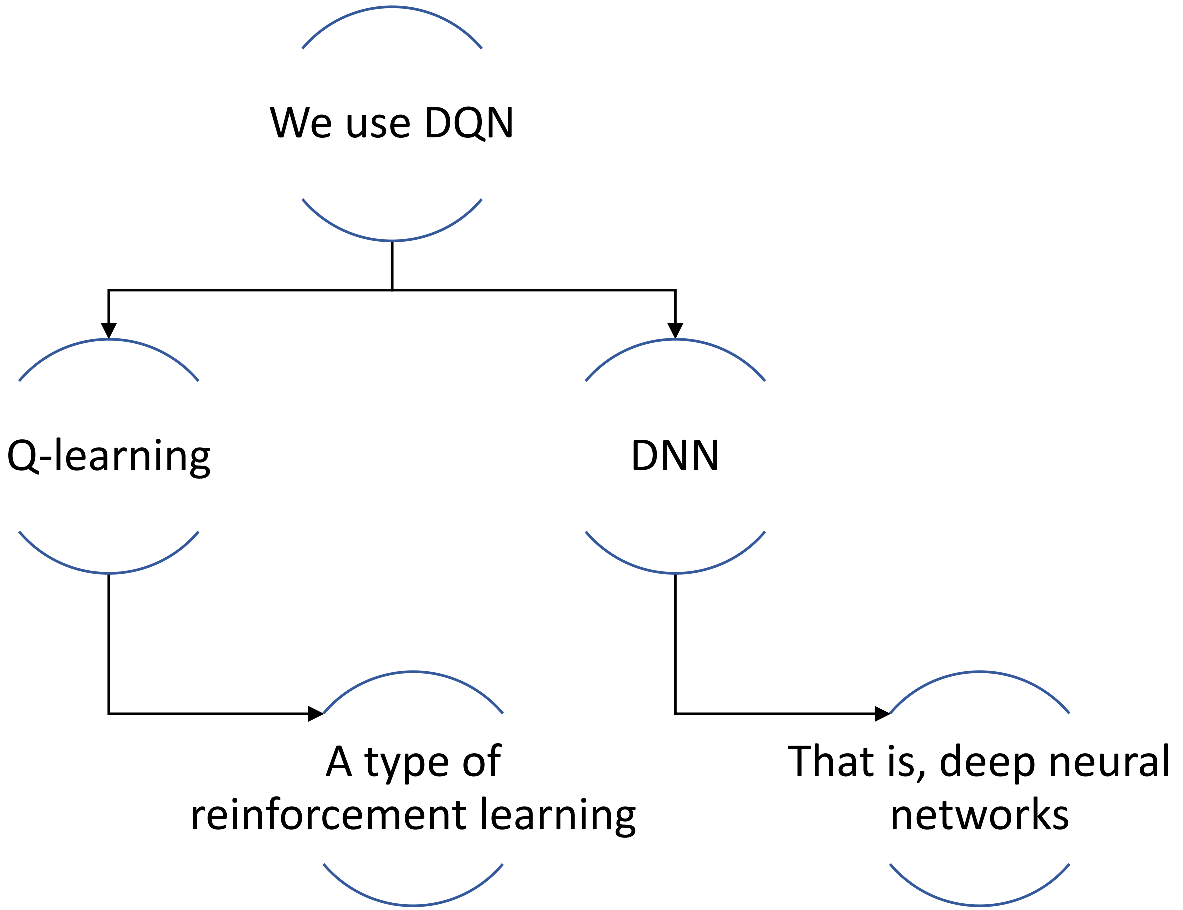


TensorFlow playground:
examples of supervised learning with neural networks

https://playground.tensorflow.org/

An example from our research

Khan, Abdullah, Alexei Lisitsa, and Alexei Vernitski. "Reinforcement learning algorithms for the Untangling of Braids." In The International FLAIRS Conference Proceedings, vol. 35. 2022.



Some neural networks are better than others

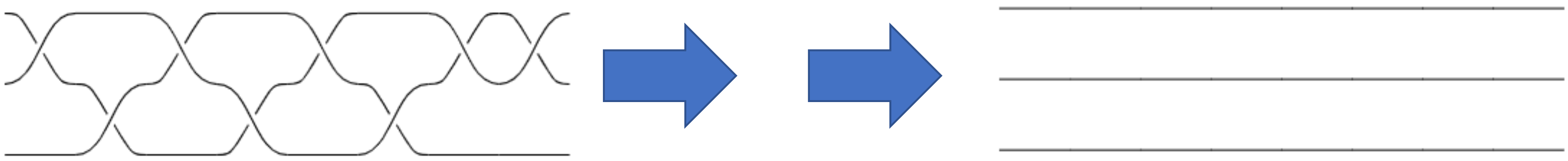
- In our current research we found some natural mathematical problems which ‘traditional’ deep neural networks (known as feed-forward neural networks or multiple-layer perceptrons) cannot solve (in the scenario of supervised learning).
- But certain advanced types of neural networks (called graph convolutional neural networks) can solve these mathematical problems.

Untangling braids

Lisitsa, Alexei, Mateo Salles, and Alexei Vernitski. "An application of neural networks to a problem in knot theory and group theory (untangling braids)." arXiv preprint arXiv:2206.05373 (2022).

Our computer code is here: https://github.com/mateosi98/Braids

It is reinforcement learning



Which braids to untangle?

Use random braids (as long as they are equivalent to the trivial braid)

- Lisitsa, Alexei, Mateo Salles, and Alexei Vernitski. "An application of neural networks to a problem in knot theory and group theory (untangling braids)." arXiv preprint arXiv:2206.05373 (2022).

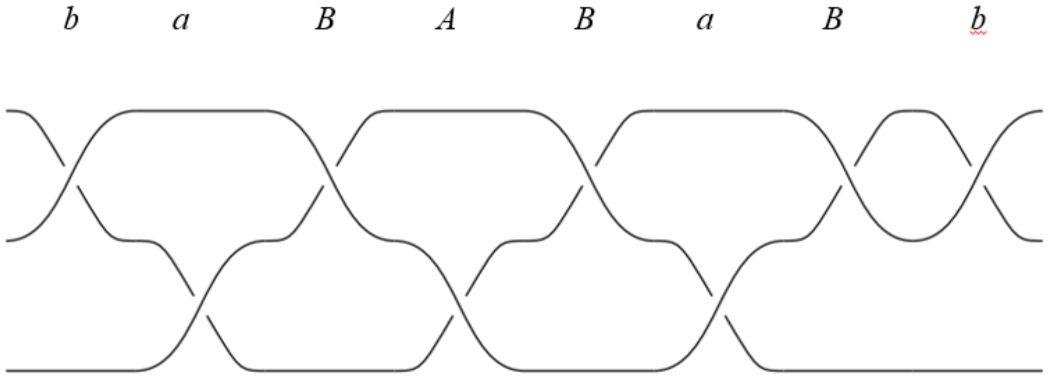
Use deliberately constructed tangled braids

- Khan, Abdullah, Alexei Lisitsa, and Alexei Vernitski. "Reinforcement learning algorithms for the Untangling of Braids." In The International FLAIRS Conference Proceedings, vol. 35. 2022.
- Khan, Abdullah, Alexei Vernitski, and Alexei Lisitsa. "Untangling braids with multi-agent q-learning." In 2021 23rd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), pp. 135-139. IEEE, 2021.

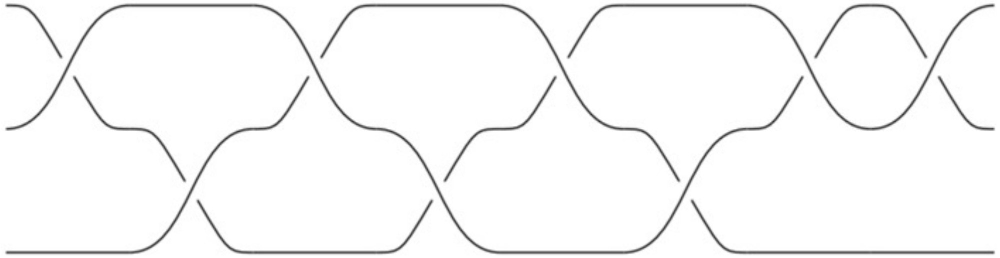
What other decisions we needed to make

How braids are presented

A braid as a word



A braid as a matrix



$$\begin{pmatrix} -1 & 0 & 1 & 0 & 1 & 0 & 1 & -1 \\ 0 & -1 & 0 & 1 & 0 & -1 & 0 & 0 \end{pmatrix}$$

What moves we have

- For example, do we use all six versions of the third Reidemeister move, or only some of them?
- Do we fix the length of a braid, or do we allow it to become shorter or longer?

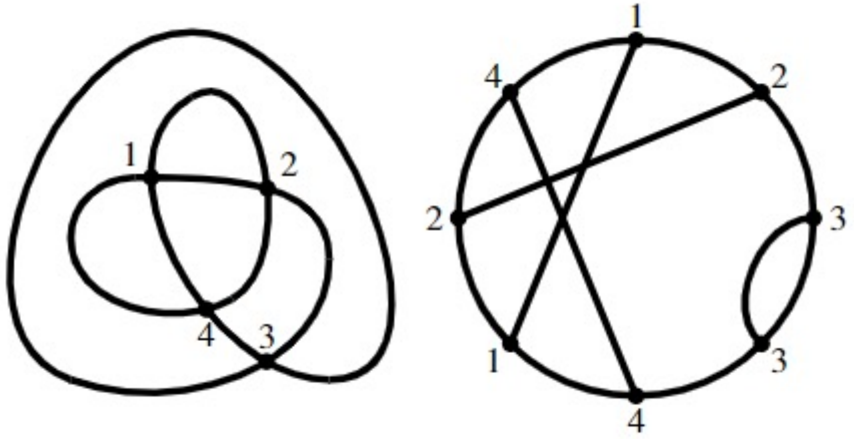
Do we tell the agent at what positions in the current braid what Reidemeister moves can be applied, or do we allow the agent to learn this the hard way?

What size of the neural network do we need?

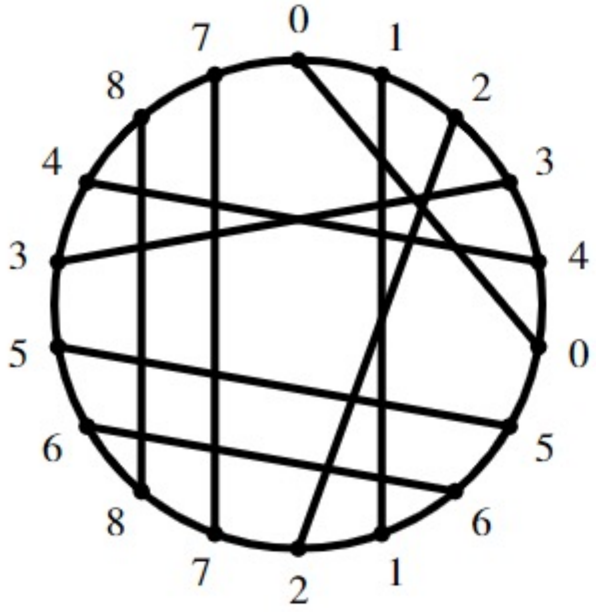
Gauss diagrams and errors in recent publications

Khan, Abdullah, Alexei Lisitsa, and Alexei Vernitski. "Training AI to Recognize Realizable Gauss Diagrams: The Same Instances Confound AI and Human Mathematicians." In Proceedings of the 14th International Conference on Agents and Artificial Intelligence, pp. 990-995. SCITEPRESS-Science and Technology Publications, 2022.

Gauss diagram of a closed planar curve



Our example of an unrealisable Gauss diagram



Khan, Abdullah, Alexei Lisitsa, Viktor Lopatkin, and Alexei Vernitski. "Circle graphs (chord interlacement graphs) of Gauss diagrams: Descriptions of realizable Gauss diagrams, algorithms, enumeration." arXiv preprint arXiv:2108.02873 (2021).

- a documented error in several (at least two) published papers,
- which can be conveniently disproved on relatively small and numerous counter-examples.
- An opportunity to investigate how machine learning compares with human reasoning.

- We use supervised learning to classify Gauss diagrams as realisable or unrealisable
- We conducted experiments with 4 different ways of encoding Gauss diagrams and with varying sizes

“The Same Instances Confound AI and Human Mathematicians”

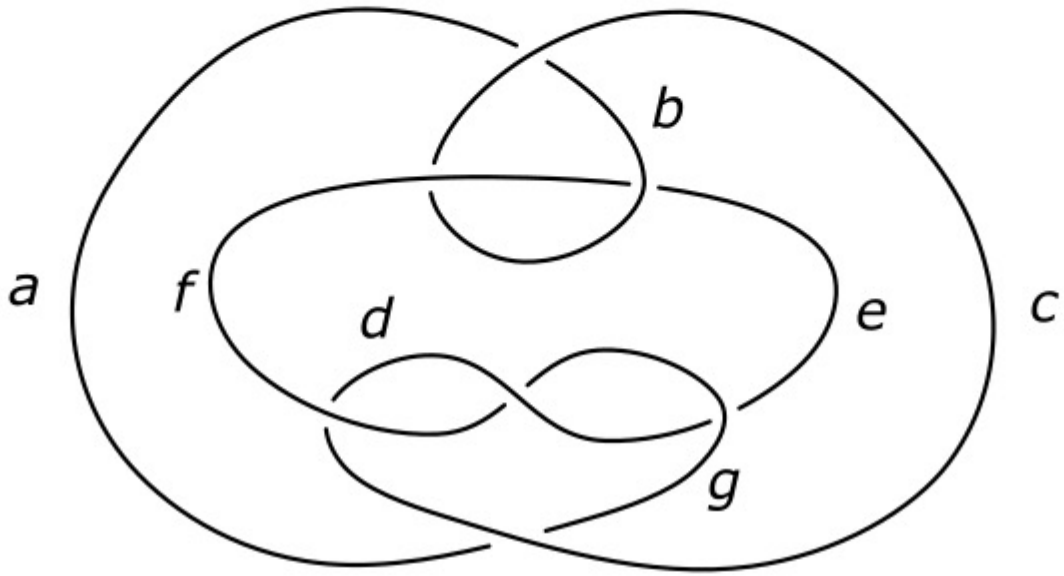
- Accuracy on random Gauss diagrams is about 75%
- Accuracy on our counterexamples is about 20%

- Explicitly including some of our counterexamples in the training set improves accuracy (when tested on our other counterexamples)

Recognising the trivial knot using quandles

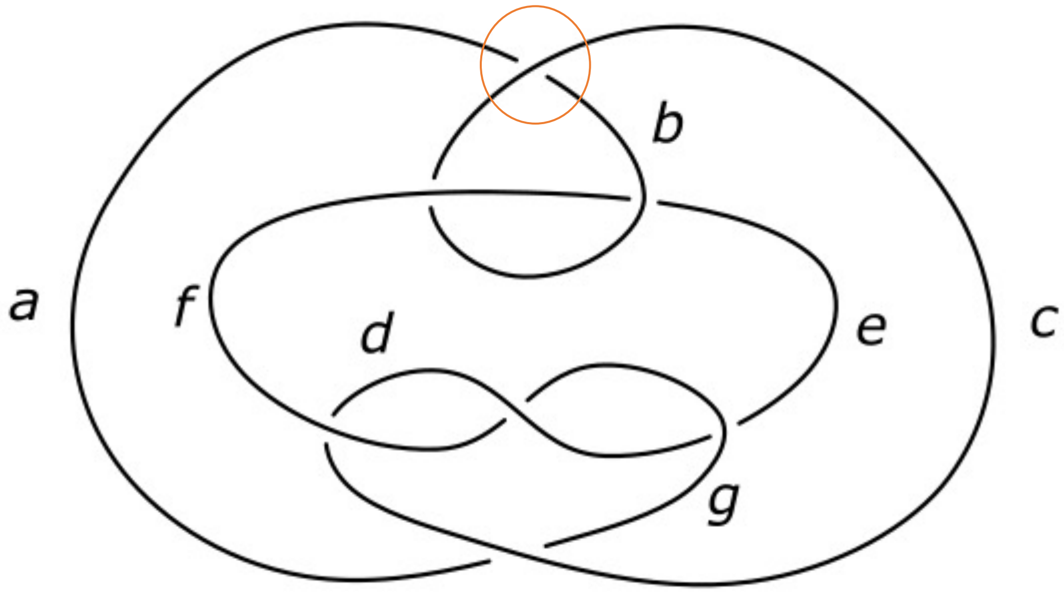
Alexei Lisitsa and Alexei Vernitski. "Using machine learning to detect non-triviality of knots via colorability of knot diagrams." The 7th Conference on Artificial Intelligence and Theorem Proving (AITP 2022).

Denote each arc on a knot diagram by a letter



Fox colouring

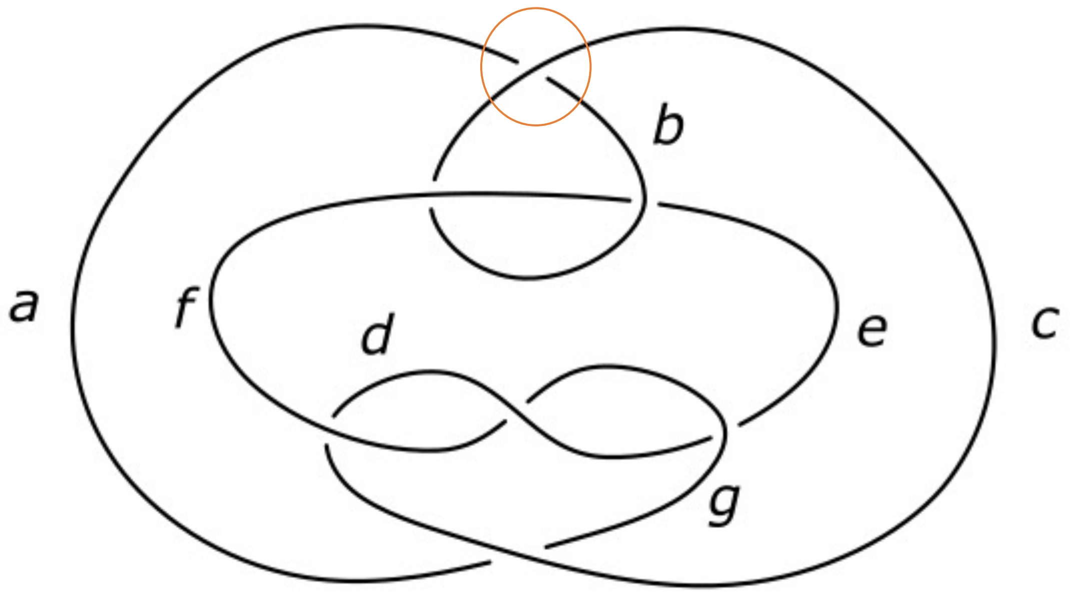
For example, the crossing at the top reads as $a + b = 2c$



Fox colouring

For example, the crossing at the top reads as $a + b = 2c$

All arcs are treated as numbers modulo a certain number n which is called the determinant of the knot



More complicated relations on crossings

- $g^{-1}fg$ is denoted by $f \rhd g$
- $gfg^{-1}$is denoted by $f \rhd^{-1} g$
- An algebra with two operations $\rhd$, $\rhd^{-1}$ which behave like conjugation in a group is called a *quandle*.

The following are equivalent (for small knot diagrams)

- Checking if the knot diagram represents the trivial knot
- Checking if the determinant is equal to 1
- Checking if the knot diagram cannot be coloured by ‘small quandles’ (from a short list of pre-selected quandles)

We used supervised learning to check

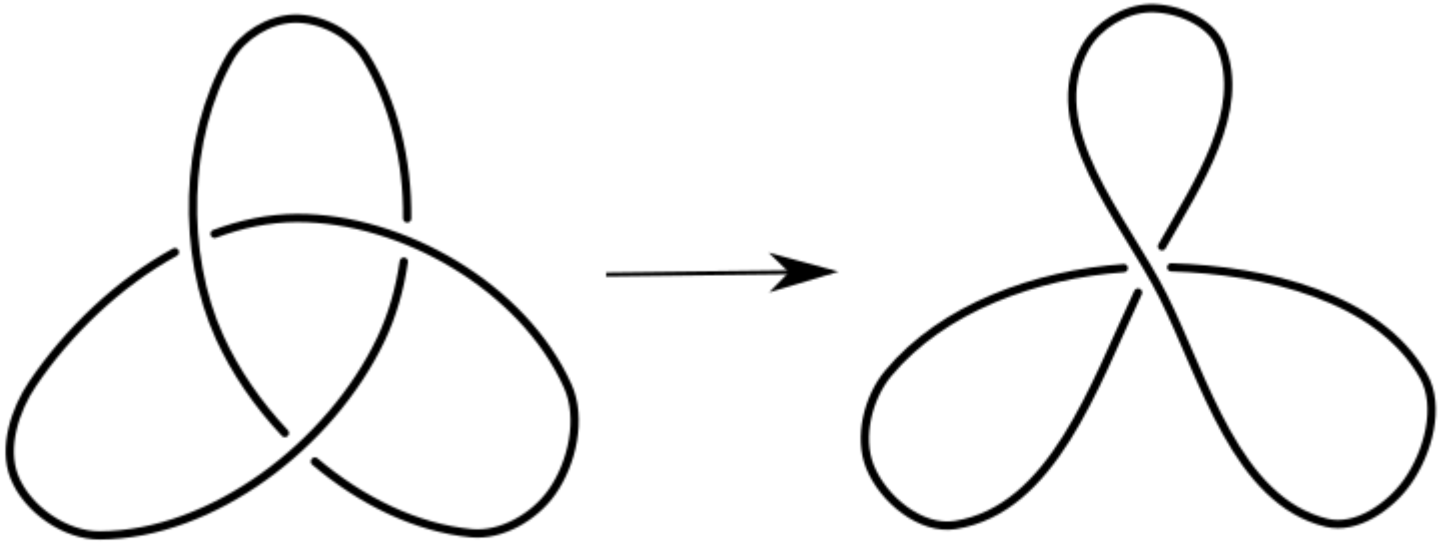
- If the determinant is 1
- If the knot diagram cannot be coloured by small quandles

What types of knot diagrams we used

- We use different types of knot diagrams
  - (these experiments are important to us because we consider the ‘simple’ experiments with supervised learning as a stepping stone towards ‘more complicated’ experiments with reinforcement learning)
- So far, we are especially satisfied with our results on so-called *petal diagrams* of knots

Petal diagrams of knots

All crossings are placed one behind another on the diagram (the diagram below is for illustration only; you need five petals to represent the trefoil knot as a petal diagram)



- With a carefully chosen encoding of a petal diagram, we reach the accuracy of around 96%.
- (We invented a new kind of matrix to represent permutations, instead of permutation matrices.)

‘But your dataset is not balanced!’

- To be exact, in this experiment we deal with petal diagrams of size 7, so there are $7! = 5040$ diagrams.
- Among them,
  - 3934 represent the trivial knot and
  - 1106 represent untrival knots.
- We choose 500 of the former and 500 of the latter and use these 1000 diagrams as the training set.
- We test on a test set which is also a mix of 50% the former and 50% the latter.
- The accuracy is around 96%.

Summary

- We apply the best available machine learning technologies to solving mathematical problems
- This is not research in machine learning
- This is not research in mathematics
- This is research in both