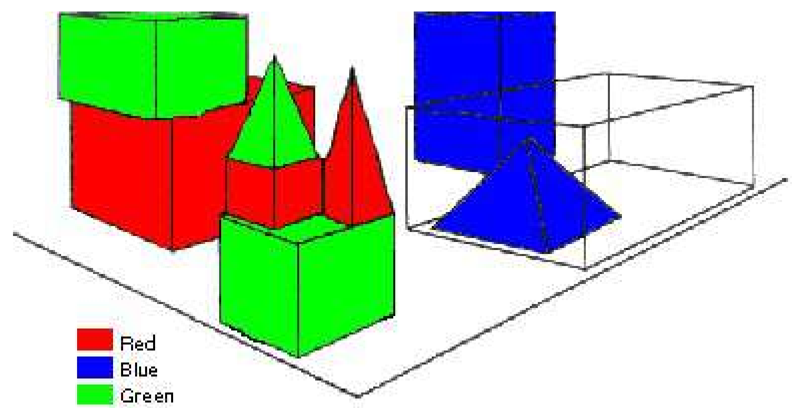
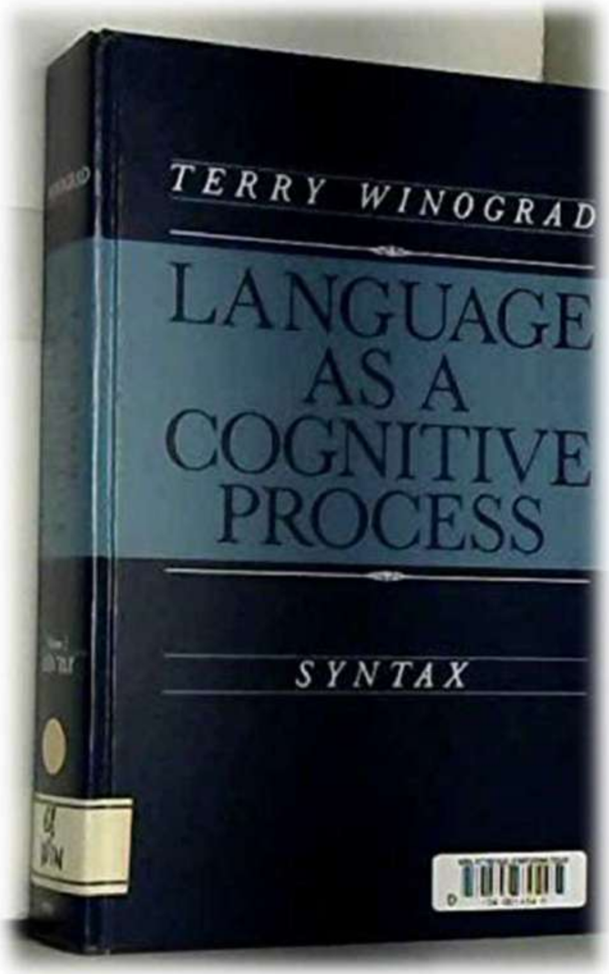
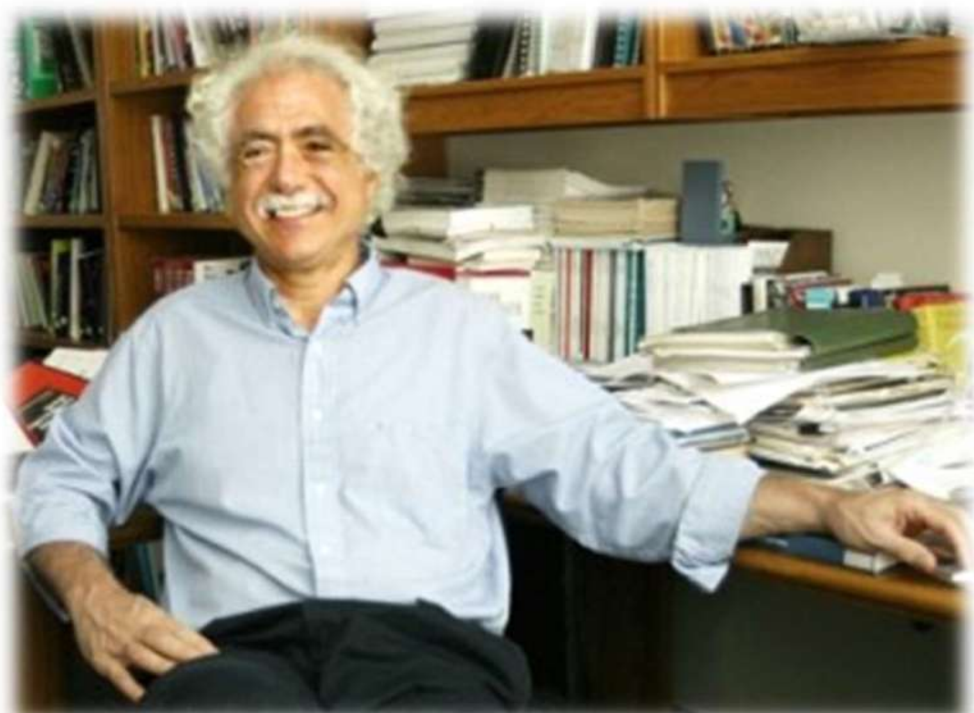
Person: Pick up a big red block.
Computer: OK.
Person: Grasp the pyramid.
Computer: I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.
Person (changing their mind): Find a block which is taller than the one you are holding and put it into the box.
Computer: BY "IT", I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING.
Computer: OK.
Person: What does the box contain?
Computer: THE BLUE PYRAMID AND THE BLUE BLOCK.
Person: What is the pyramid supported by?
Computer: THE BOX.
Person: How many blocks are not in the box?
Computer: FOUR OF THEM.
Person: Is at least one of them narrower than the one which i told you to pick up?
Computer: YES, THE RED CUBE.
Person: Is it supported?
Computer: YES, BY THE TABLE.
Person: Can the table pick up blocks?
Computer: NO.
Person: Can a pyramid be supported by a block?
Computer: YES.
Person: Can a pyramid support a pyramid?
Computer: I DON'T KNOW.
Person: Stack up two pyramids.
Computer: I CAN'T.

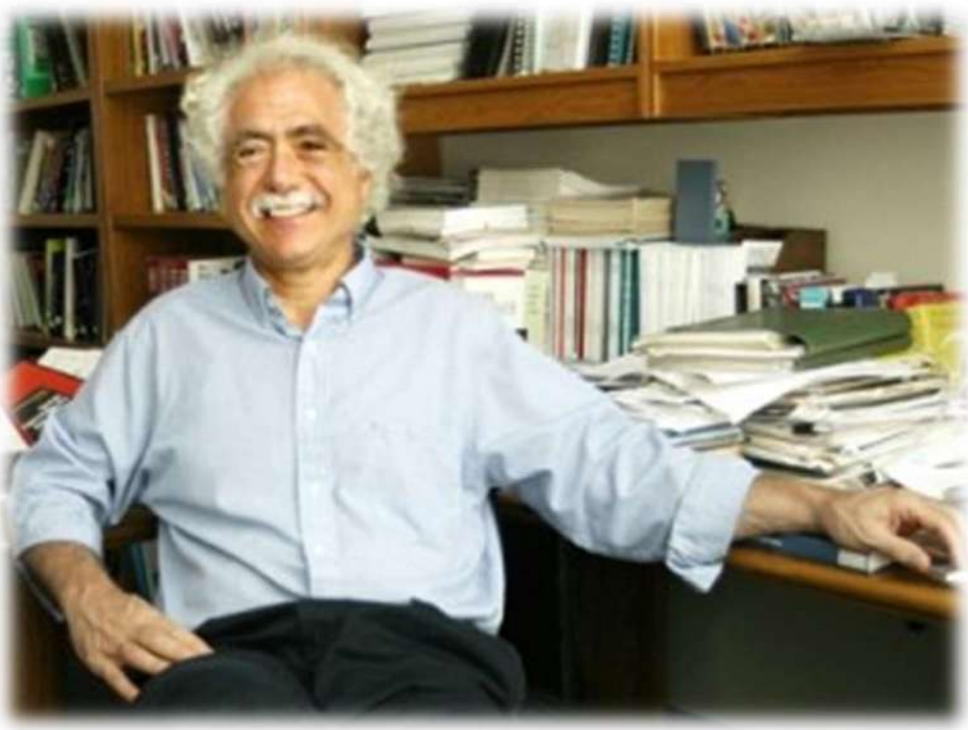


Person: Will you please stack up both of the red blocks and either a green cube or a pyramid?
Computer: OK.
Person: Which cube is sitting on the table?
Computer: THE LARGE GREEN ONE WHICH SUPPORTS THE RED PYRAMID.
:



SHRDLU (1971)

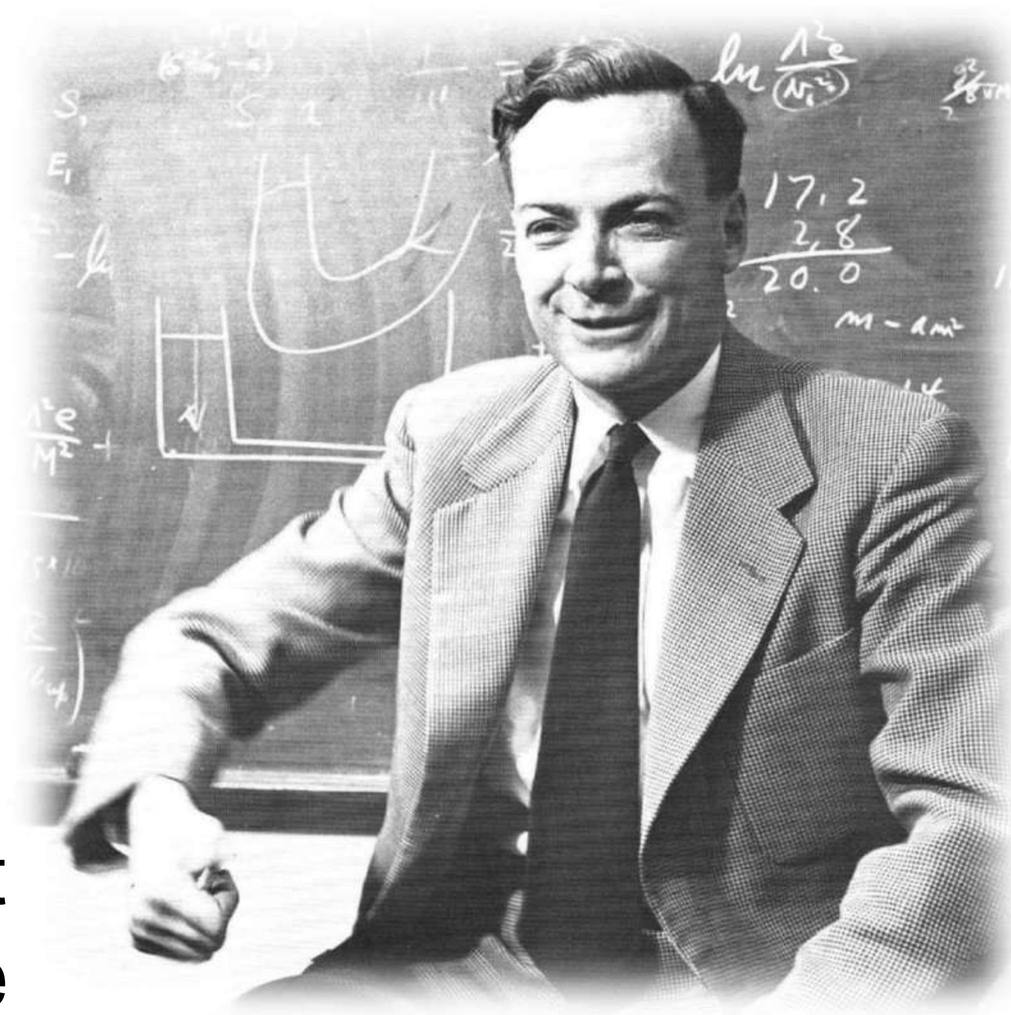




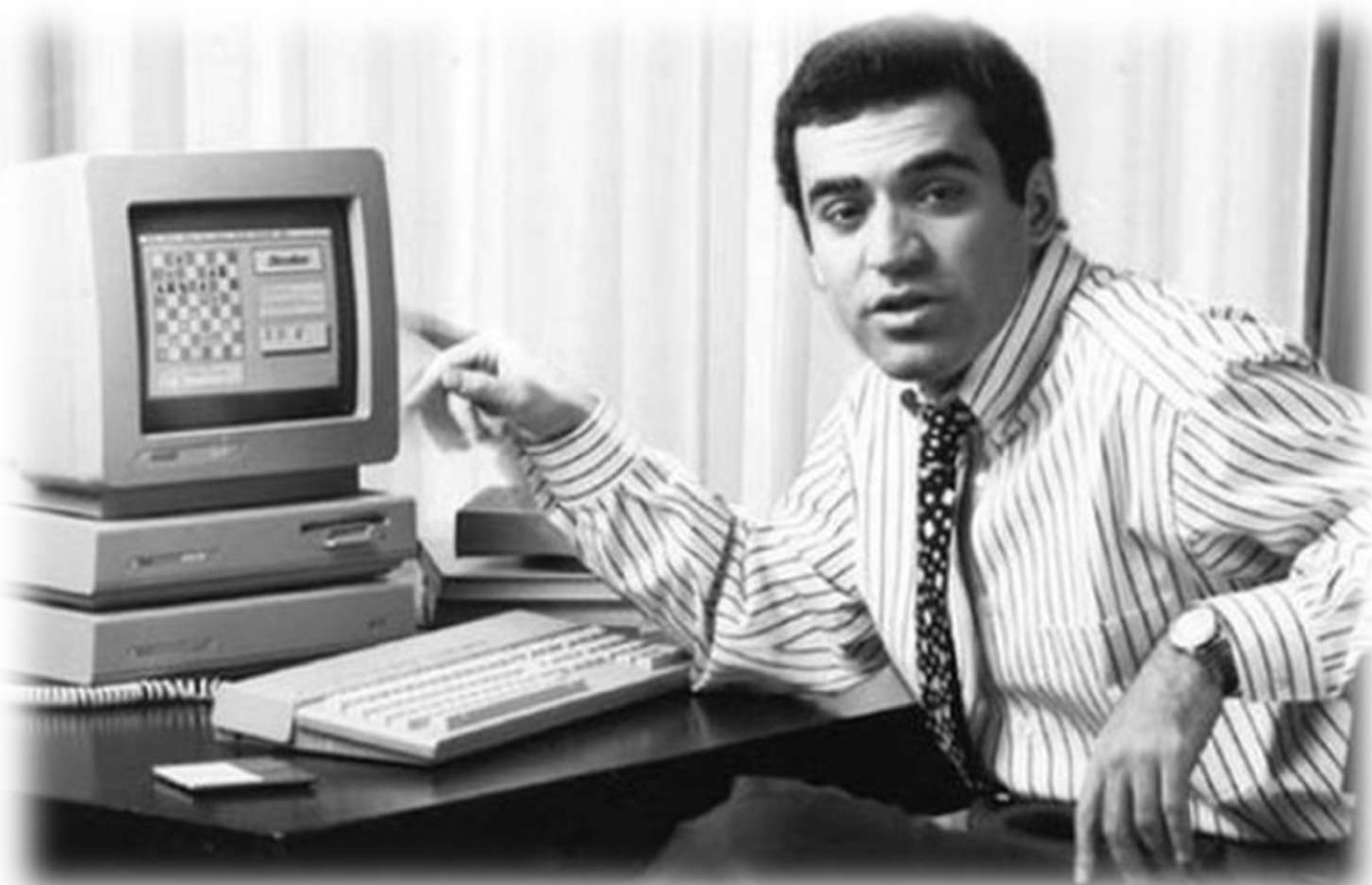
1997


Google

"… a fun analogy in trying to get some idea of what we're doing in trying to understand nature, is to imagine that the gods are playing some great game like chess… and you don't know the rules of the game, but you're allowed to look at the board, at least from time

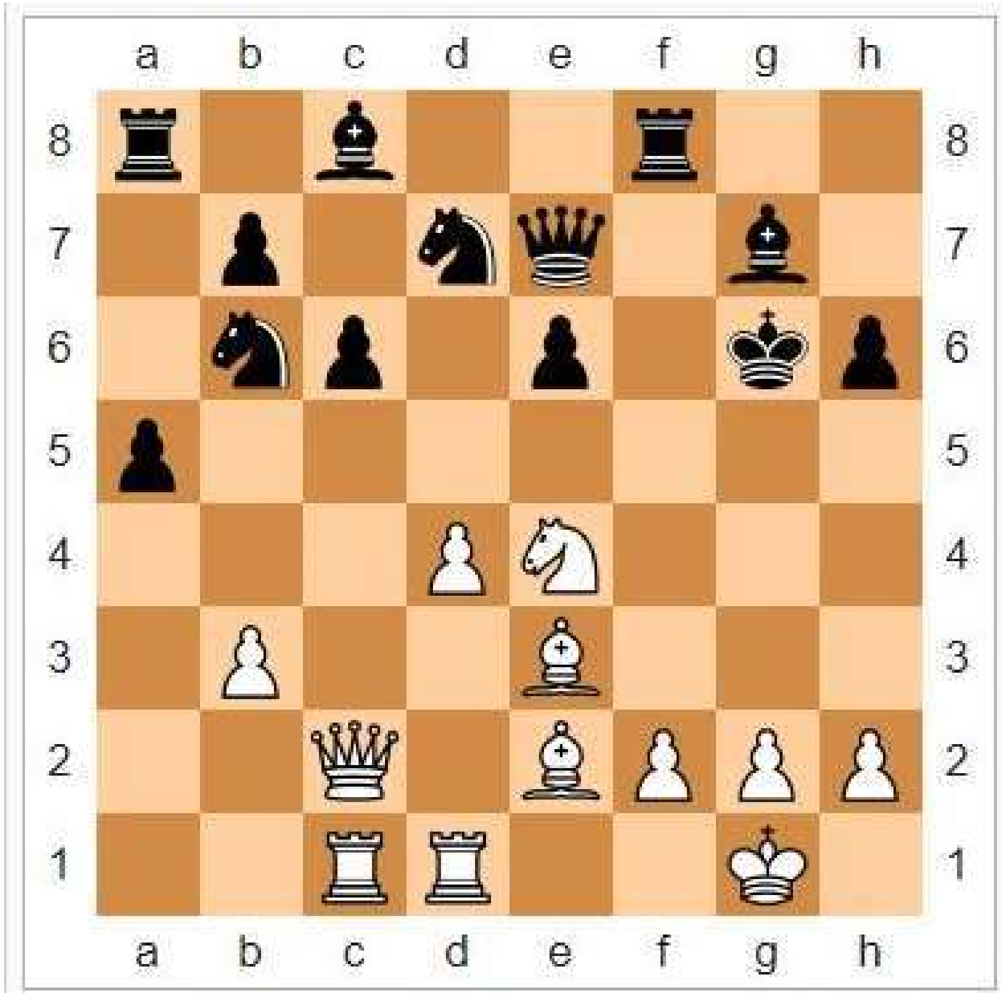


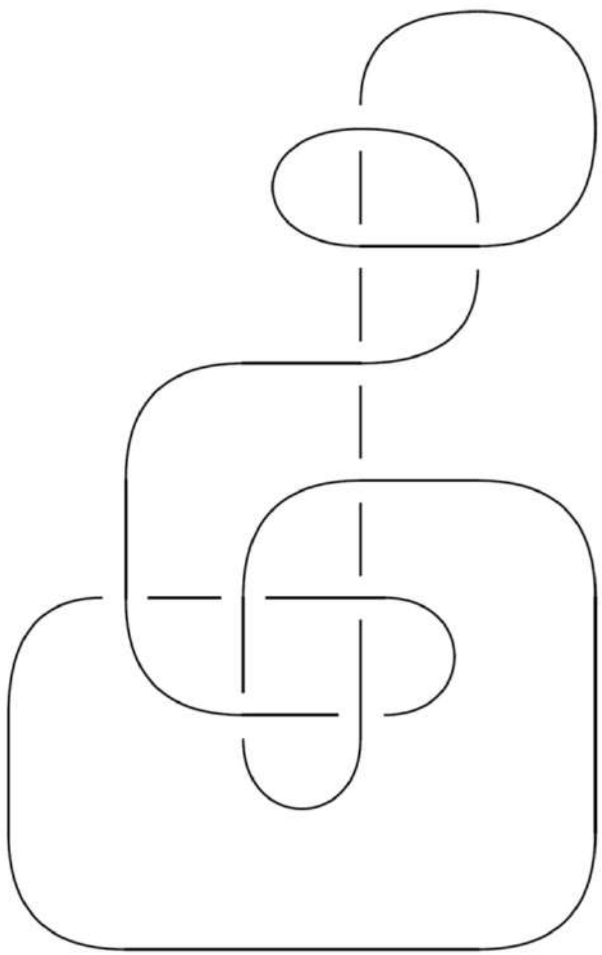
to time… and from these observations you try to figure out what the rules of the game are."



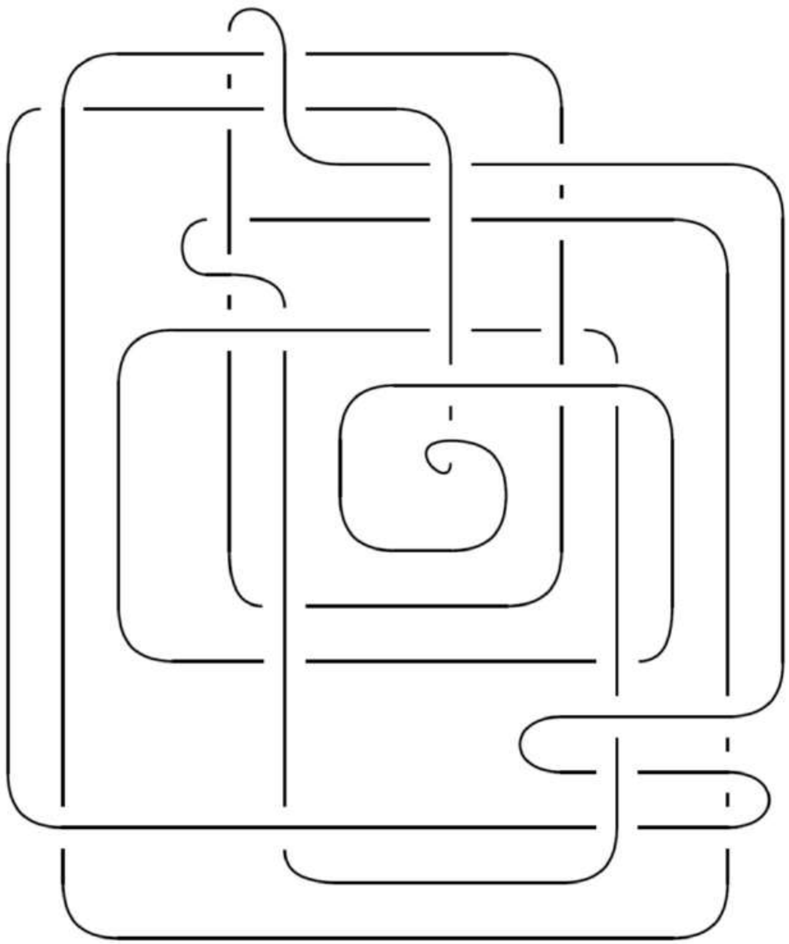
1992

1996


a
b
c
d
e
f
g
h
8
7
6
5
4
3
2
1



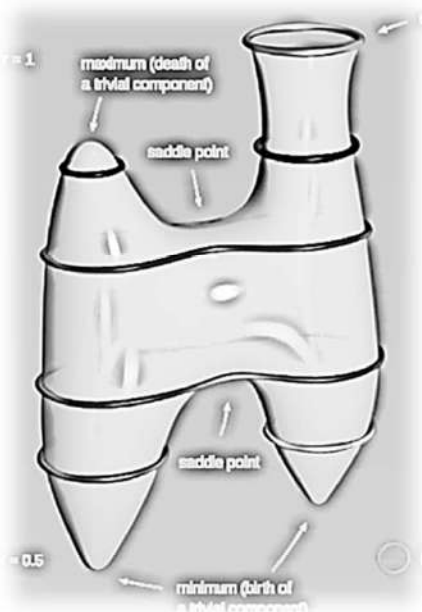
10 crossings



30 crossings

FIBERED KNOTS AND POTENTIAL COUNTEREXAMPLES TO THE PROPERTY 2R AND SLICE-RIBBON CONJECTURES

ROBERT E. GOMPF, MARTIN SCHARLEMANN, AND ABIGAIL THOMPSON



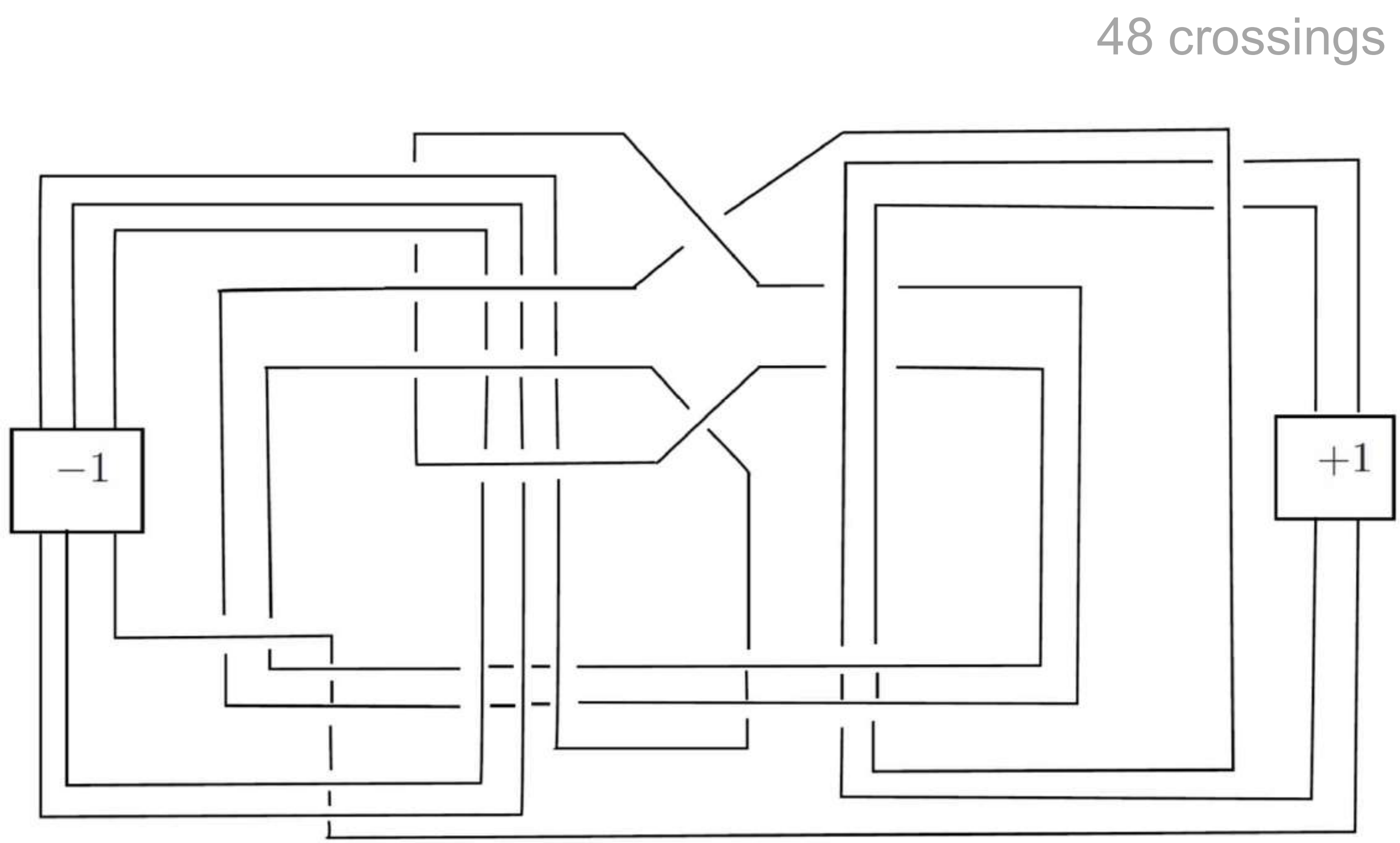
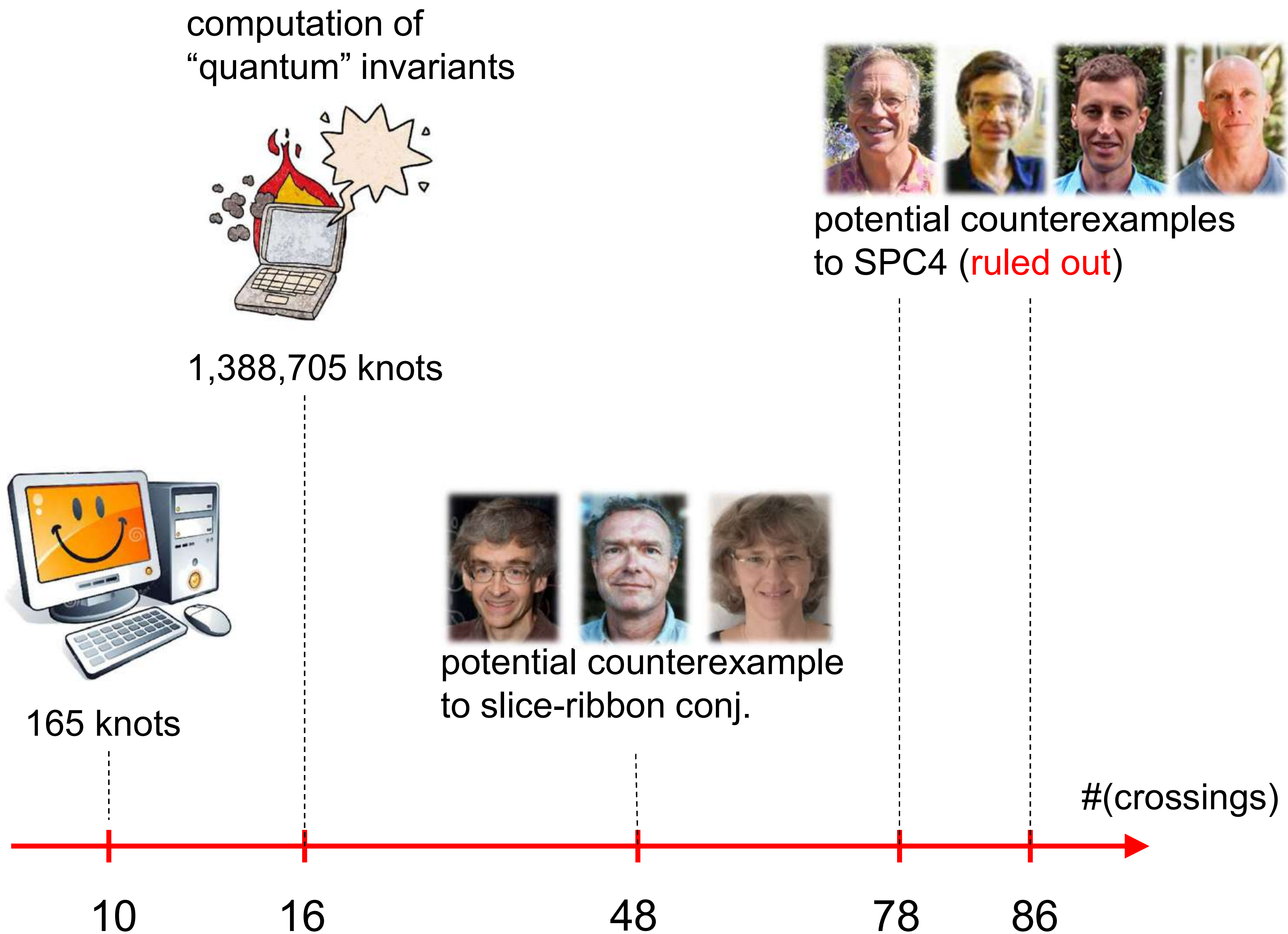


FIGURE 2. A slice knot that might not be ribbon


computation of
“quantum” invariants
1,388,705 knots
165 knots
potential counterexamples
to SPC4 (ruled out)
potential counterexample
to slice-ribbon conj.
#(crossings)
10
16
48
78
86

Steve Smale

$n =$	1	2	3	4	5	6	7	8	9	10	11	12
TOP	1	1	1	1	1	1	1	1	1	1	1	1
PL	1	1	1	?	1	1	1	1	1	1	1	1
DIFF	1	1	1	?	1	1	28	2	8	6	992	1

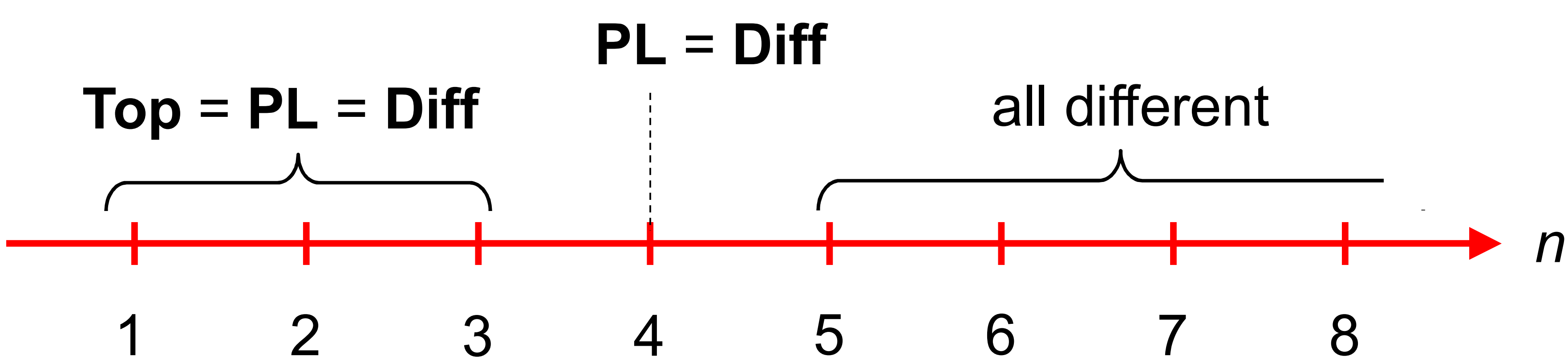
Number of homotopy *n*-spheres in each category.

John Milnor

The generalized Poincare conjecture:

late 1950s

- **Top**: true for all *n*
- **PL**: true for all $n \neq 4$ ($n = 4$ currently not known)
- **Diff**: true for n = 1, 2, 3, 5, and 6



Generalized Poincare conjecture:

Every homotopy 4-sphere is diffeomorphic to the standard 4-sphere.

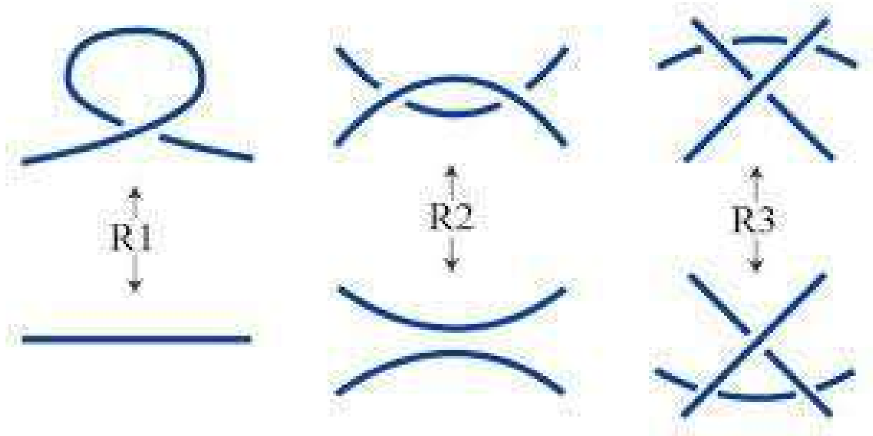
Theorem: If one finds a pair of knots which satisfy the following three properties:

- K and K′ have the same 0-surgery
- K is not slice
- K′ is slice

then the smooth 4-dimensional Poincare conjecture is false.

- Is it knotted?

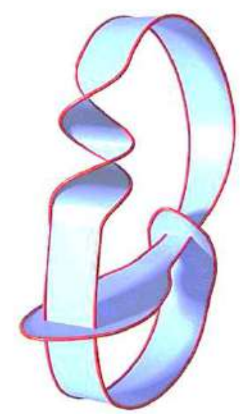
S.G., J.Halverson, F.Ruehle, P.Sulkowski



- Is it ribbon? Is it slice?

S.G., J.Halverson, C.Manolescu, F.Ruehle

(SPC4, slice-ribbon, ...)



- Is it Andrews-Curtis trivial?

work in progress

Combinatorial group theory

Conjecture [J.Andrews and M.Curtis '65]:

Every balanced presentation of the trivial group

$$\langle x_1, \ldots, x_n \mid r_1, \ldots, r_n \rangle$$

can be reduced to the trivial presentation

$$\langle x_1, \ldots, x_n \mid x_1, \ldots, x_n \rangle$$

by a sequence of Andrews-Curtis (Nielsen) moves:

$$r_i, r_j \mapsto r_i r_j, r_j$$

$$r_i \mapsto r_i^{-1}$$

$$r_i \mapsto x_j^{\pm 1} r_i x_j^{\mp 1}$$

"handle slides"

*$$\langle x_1, \ldots, x_n \mid r_1, \ldots, r_n \rangle \leftrightarrow$$

$$\leftrightarrow \langle x_1, \ldots, x_n, x_{n+1} \mid r_1, \ldots, r_n, x_{n+1} \rangle$$

"handle cancellation"

---

* generalized

- No counterexamples with relations of total length <13
- Believed to be false
- Many potential counterexamples, e.g.

$$\langle\, x, y \mid xyx = yxy,\, x^{n+1} = y^{n} \,\rangle \quad n \geq 3$$

S.Akbulut, R.Kirby (1985)

- Validating any of these, disproves the following

<u>Conjecture ("Generalized Property R"):</u>

If surgery on an n-component link L yields the connected sum $(S^1 \times S^2)^{\#n}$, then L is obtained from the 0-framed unlink by a sequence of handle slides.

R.Gompf, M.Scharlemann, A.Thompson (2010)

- A handle decomposition of a homotopy sphere without 3-handles gives a balanced presentation of the trivial group

- AC moves = Kirby moves (without introducing 3-handles)

↓

- A potential counterexample to AC gives a potential counterexample to SPC4

Theorem:

R.Gompf (1991)

$$\langle\, x, y \mid xyx = yxy,\; x^5 = y^4 \,\rangle$$

gives a standard 4-sphere.

THE COMPLEXITY OF BALANCED PRESENTATIONS AND THE ANDREWS–CURTIS CONJECTURE

MARTIN R. BRIDSON

Hard AC presentations

Theorem A. *For $k \geq 4$ one can construct explicit sequences of k-generator balanced presentations $\mathcal{P}_n$ of the trivial group so that*

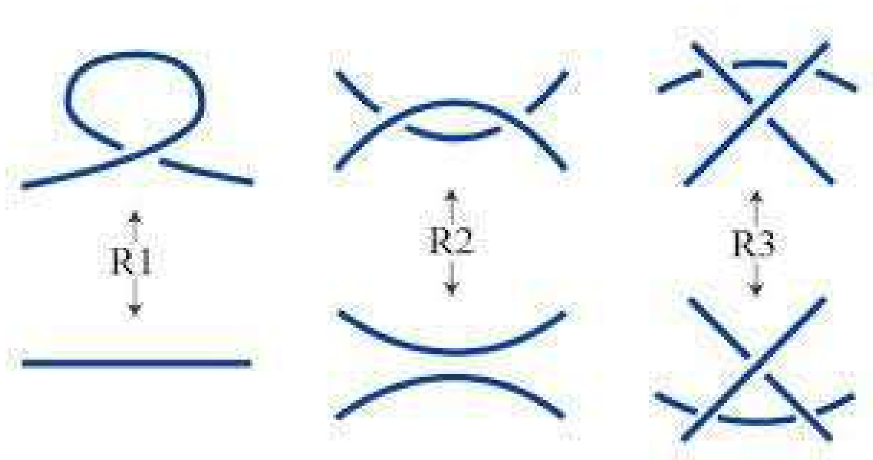
(1) the presentations $\mathcal{P}_n$ are AC-trivialisable;
(2) the sum of the lengths of the relators in $\mathcal{P}_n$ is at most $24(n+1)$;
(3) the number of (dihedral) AC moves required to trivialise $\mathcal{P}_n$ is bounded below by the function $\Delta(\lfloor \log_2 n \rfloor)$ where $\Delta : \mathbb{N} \to \mathbb{N}$ is defined recursively by $\Delta(0) = 2$ and $\Delta(m+1) = 2^{\Delta(m)}$.

7.4. **An Example.** Let me close by writing down an explicit presentation to emphasize that the explosive growth in the length of AC-trivialisations begins with relatively small presentations. Here is a balanced presentation of the trivial group that requires more than 10^{10000} AC-moves to trivialise it. We use the commutator convention $[x, y] = xyx^{-1}y^{-1}$.

$$\langle a, t, \alpha, \tau \mid [tat^{-1}, a]a^{-1}, \quad [\tau\alpha\tau^{-1}, \alpha]\alpha^{-1}, \quad \alpha t^{-1}\alpha^{-1}[a,\ [t[t[ta^{20}t^{-1},\ a]t^{-1},\ a]t^{-1},\ a]], \quad a\tau^{-1}a^{-1}[\alpha,\ [\tau[\tau[\tau\alpha^{20}\tau^{-1},\ \alpha]\tau^{-1},\ \alpha]\tau^{-1},\ \alpha]] \rangle.$$

- Is it knotted?

S.G., J.Halverson, F.Ruehle, P.Sulkowski



Hard unknots

- Is it ribbon? Is it slice?

S.G., J.Halverson, C.Manolescu, F.Ruehle

(SPC4, slice-ribbon, ...)

Hard ribbon knots

- Is it Andrews-Curtis trivial?

work in progress

Hard AC presentations

Winograd schemas:

The trophy would not fit in the brown suitcase because it was too big (*small*).

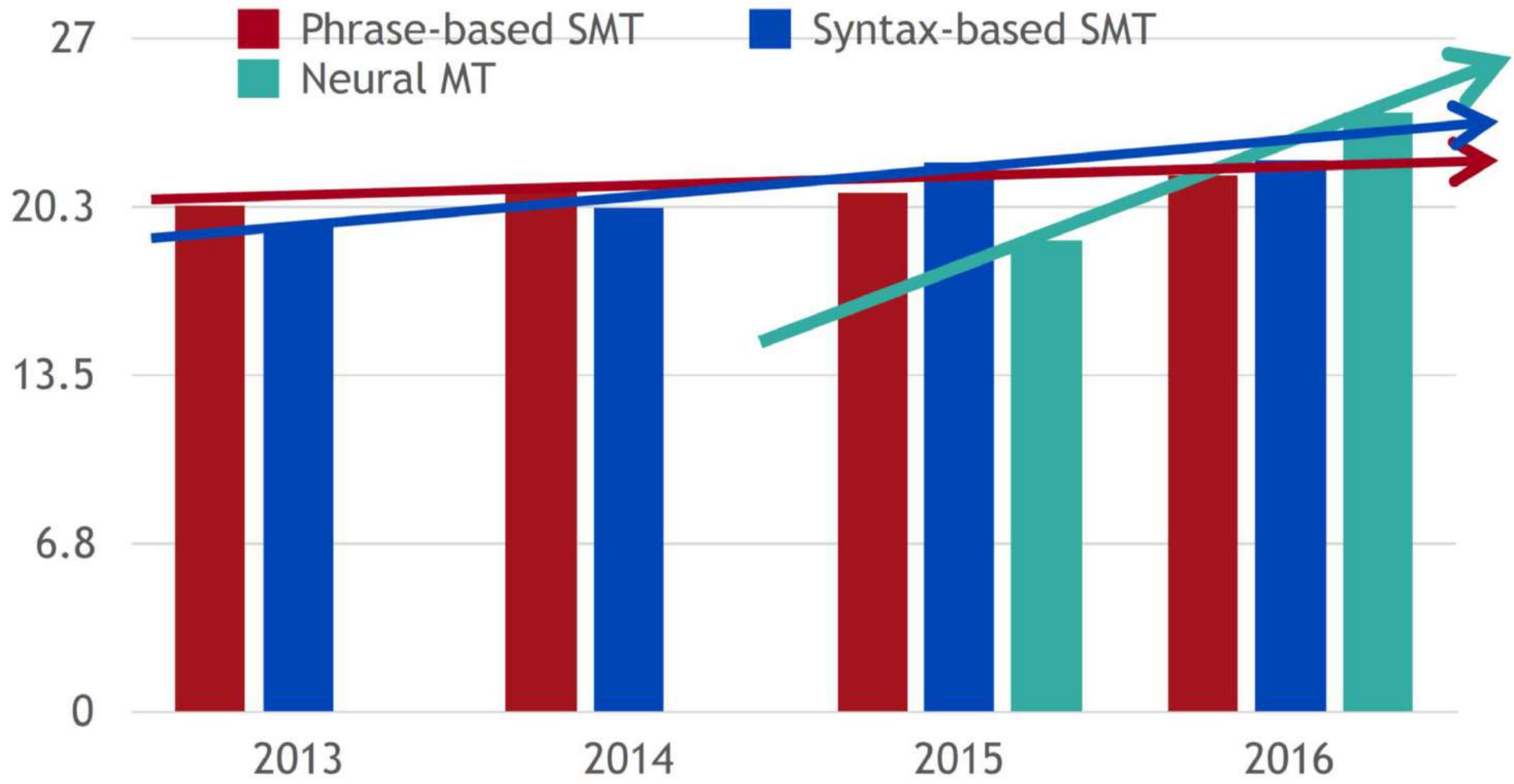
What was too big (*small*)?

The town councilors refused to give the demonstrators a permit because they feared (*advocated*) violence.

Who feared (*advocated*) violence?

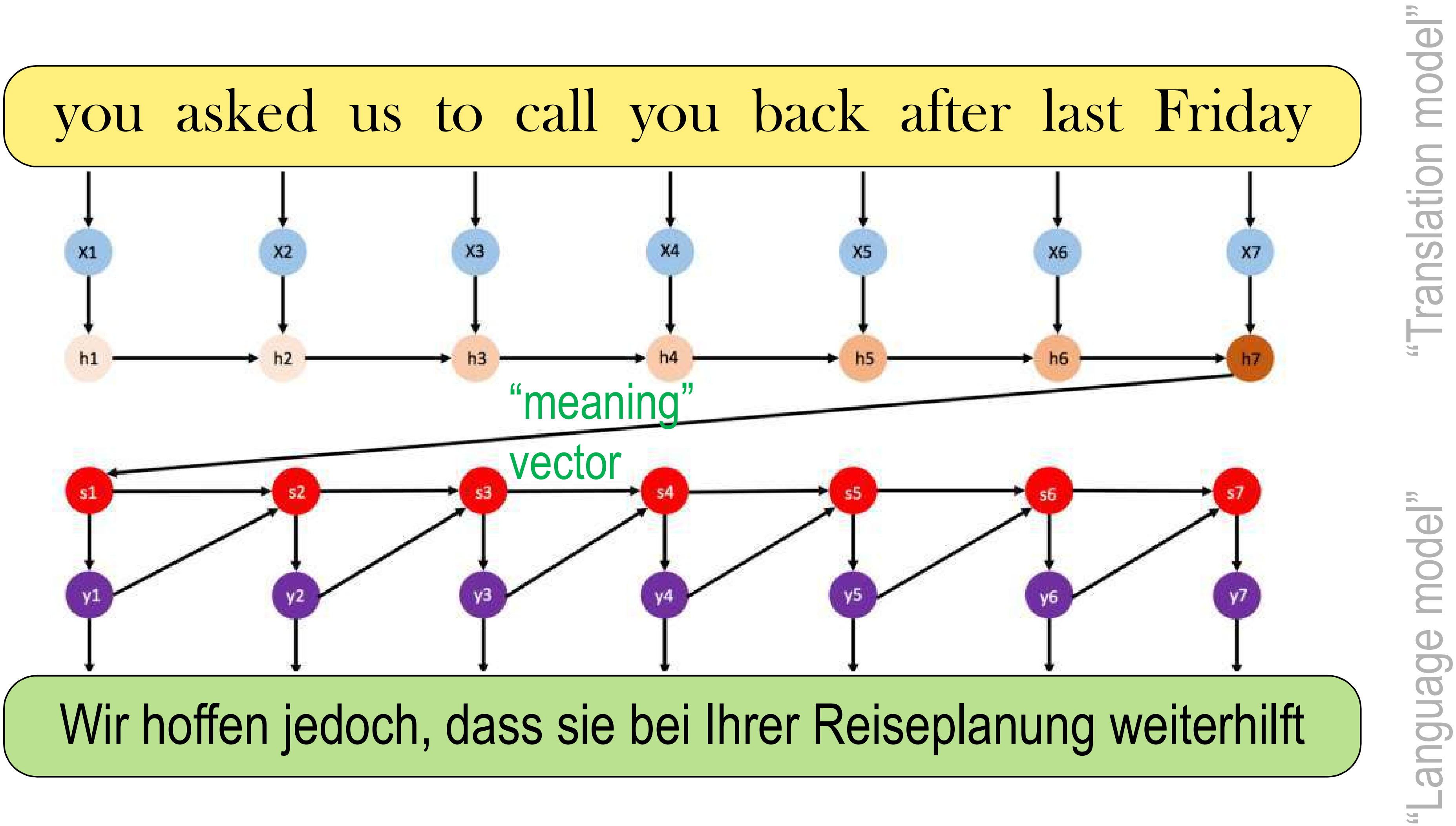
MT progress over time

[Edinburgh En-De WMT newstest2013 Cased BLEU; NMT 2015 from U. Montréal]



Source: http://www.meta-net.eu/events/meta-forum-2016/slides/09_sennrich.pdf

seq2seq: Encoder + Decoder



Need “context” vectors

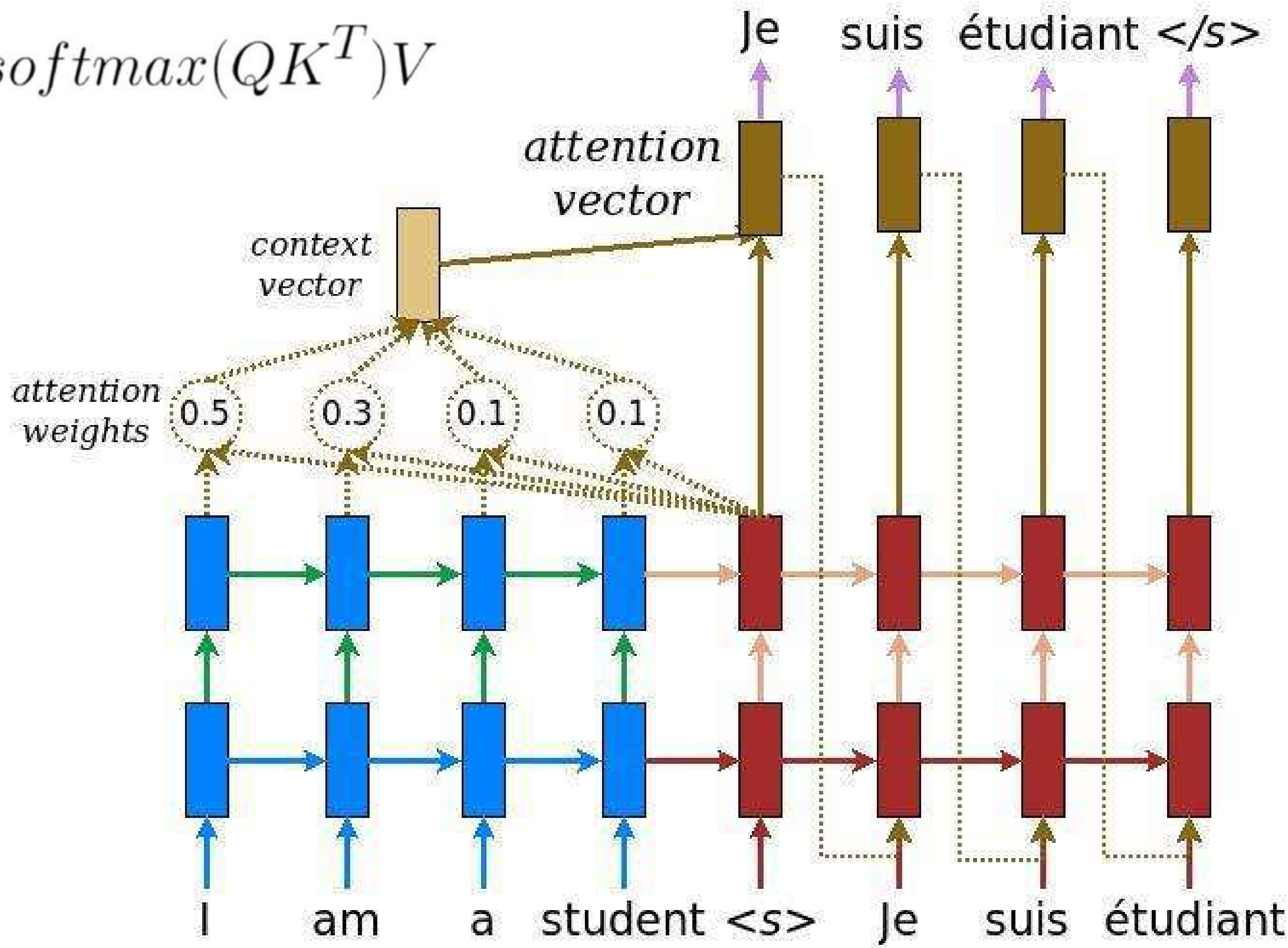

汀步路，
小心滑倒！
Slip and fall down carefully!


AIRPORTS AUTHORITY OF INDIA
फर्श पर खाना सख़्त मना है ।
—आदेश से
EATING CARPET
STRICTLY PROHIBITED
—By Order

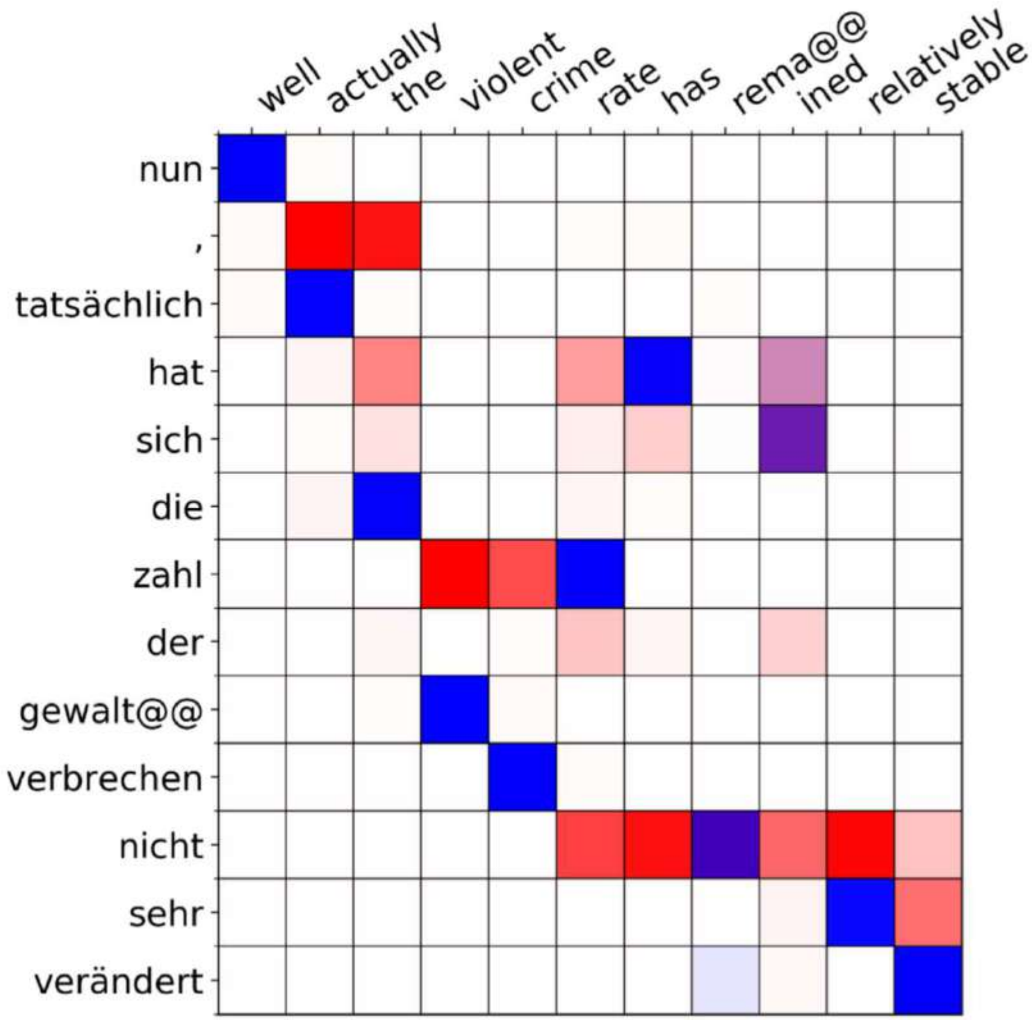

请勿食用
Please don't be edible
外带食物
Exterior Girdle food


Bengali
Russian
Portuguese
Arabic
Hindi
Spanish
Mandarin
German
Italian
French
Translate

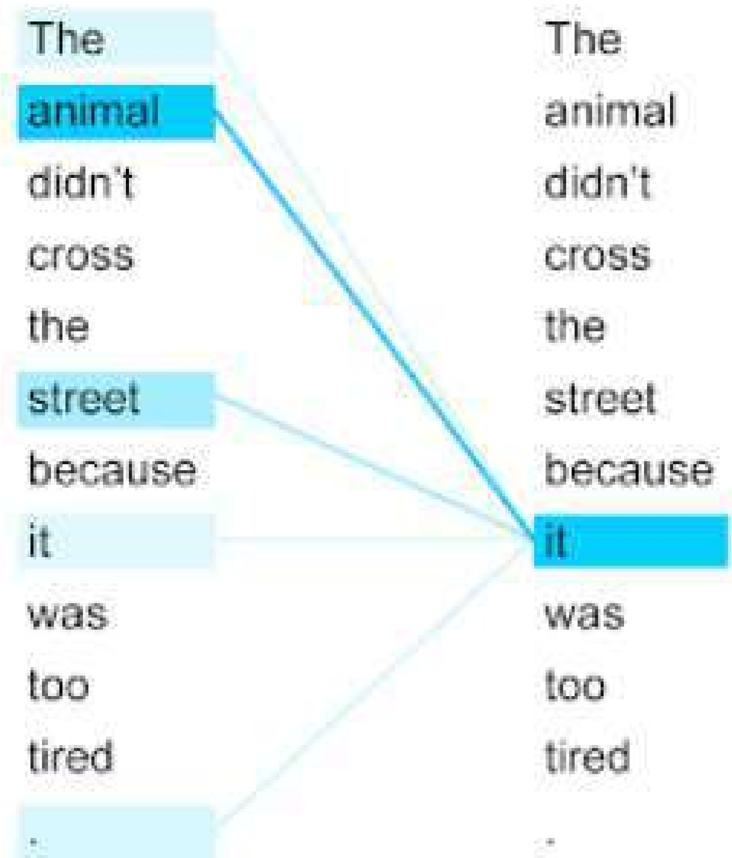
$$softmax(QK^T)V$$



Source: https://github.com/tensorflow/nmt



variational attention (blue)
vs prior alignment (red)



self-attention

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

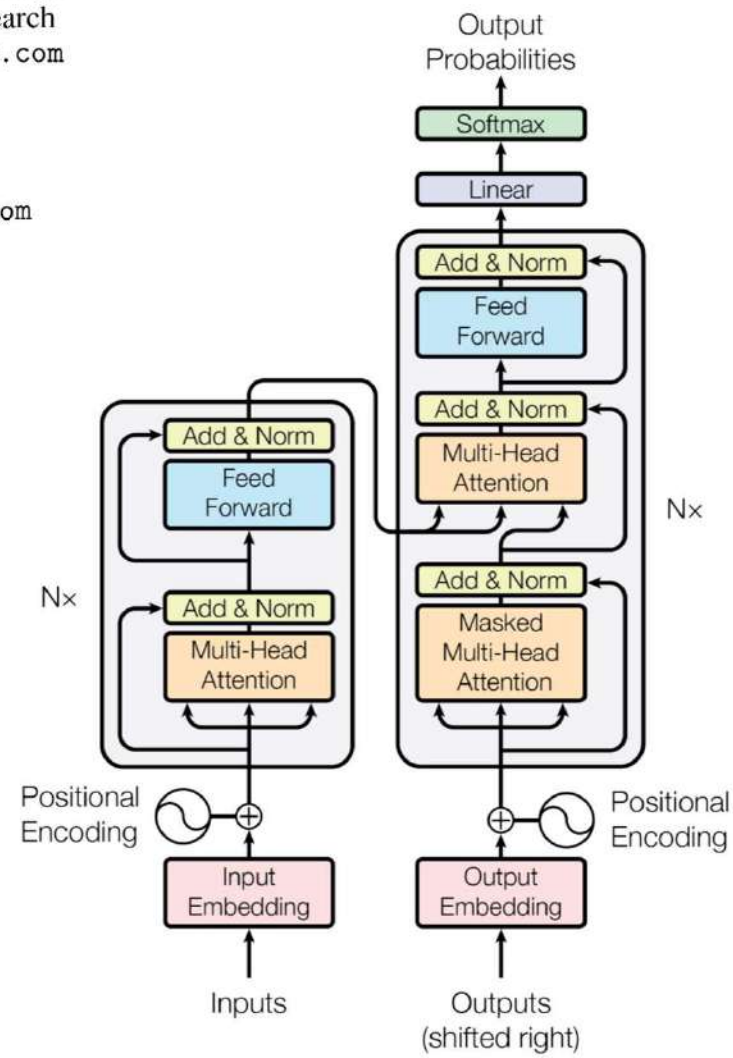
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.



Sentence	Google Translate	Transformer
The cow ate the hay because it was **delicious**.	La vache mangeait le foin parce **qu'elle** était délicieuse.	La vache a mangé le foin parce **qu'il** était délicieux.
The cow ate the hay because it was **hungry**.	La vache mangeait le foin parce **qu'elle** avait faim.	La vache mangeait le foin parce **qu'elle** avait faim.
The women stopped drinking the wines because they were **carcinogenic**.	Les femmes ont cessé de boire les vins parce **qu'ils** étaient cancérogènes.	Les femmes ont cessé de boire les vins parce **qu'ils** étaient cancérigènes.
The women stopped drinking the wines because they were **pregnant**.	Les femmes ont cessé de boire les vins parce **qu'ils** étaient enceintes.	Les femmes ont cessé de boire les vins parce **qu'elles** étaient enceintes.
The city councilmen refused the female demonstrators a permit because they **advocated** violence.	Les conseillers municipaux ont refusé aux femmes manifestantes un permis parce **qu'ils** préconisaient la violence.	Le conseil municipal a refusé aux manifestantes un permis parce **qu'elles** prônaient la violence.
The city councilmen refused the female demonstrators a permit because they **feared** violence.	Les conseillers municipaux ont refusé aux femmes manifestantes un permis parce **qu'ils** craignaient la violence	Le conseil municipal a refusé aux manifestantes un permis parce **qu'elles** craignaient la violence.*

Lukasz Kaiser, 2017

“The Transformer” are a Japanese [[hardcore punk]] band.
==Early years==
The band was formed in 1968, during the height of Japanese music history. Among the legendary [[Japanese people|Japanese]] composers of [Japanese lyrics], they prominently exemplified Motohiro Oda's especially tasty lyrics and psychedelic intention. Michio was a longtime member of the every Sunday night band PSM. His alluring was of such importance as being the man who ignored the already successful image and that he municipal makeup whose parents were – the band was called Jenei.<ref>http://www.separatist.org/se_frontend/post-punk-musician-the-kidney.html</ref> From a young age the band was very close, thus opting to pioneer what ...

:

=== 1981-2010: The band to break away ===
On 1 January 1981 bassist Michio Kono, and the members of the original line-up emerged. Niji Fukune and his [[Head poet|Head]] band (now guitarist) Kazuya Kouda left the band in the hands of the band at the May 28, 1981, benefit season of [[Led Zeppelin]]'s Marmarin building. In June 1987, Kono joined the band as a full-time drummer, playing a ...

Lukasz Kaiser, 2017

REFORMER: THE EFFICIENT TRANSFORMER

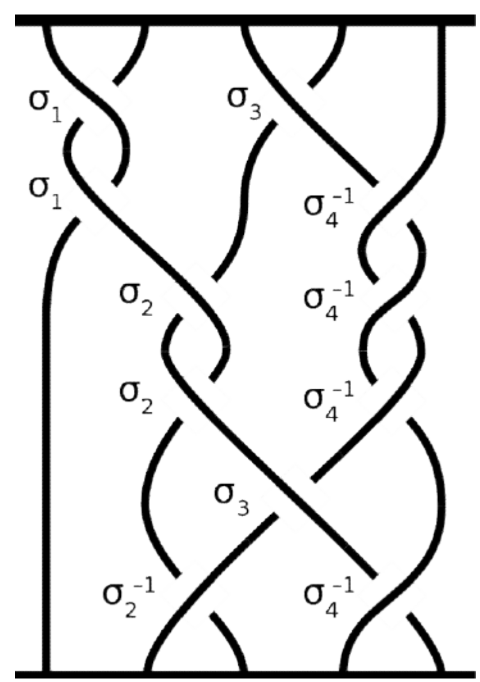
Nikita Kitaev* **Łukasz Kaiser*** **Anselm Levskaya**
U.C. Berkeley & Google Research Google Research Google Research
kitaev@cs.berkeley.edu {lukaszkaiser,levskaya}@google.com

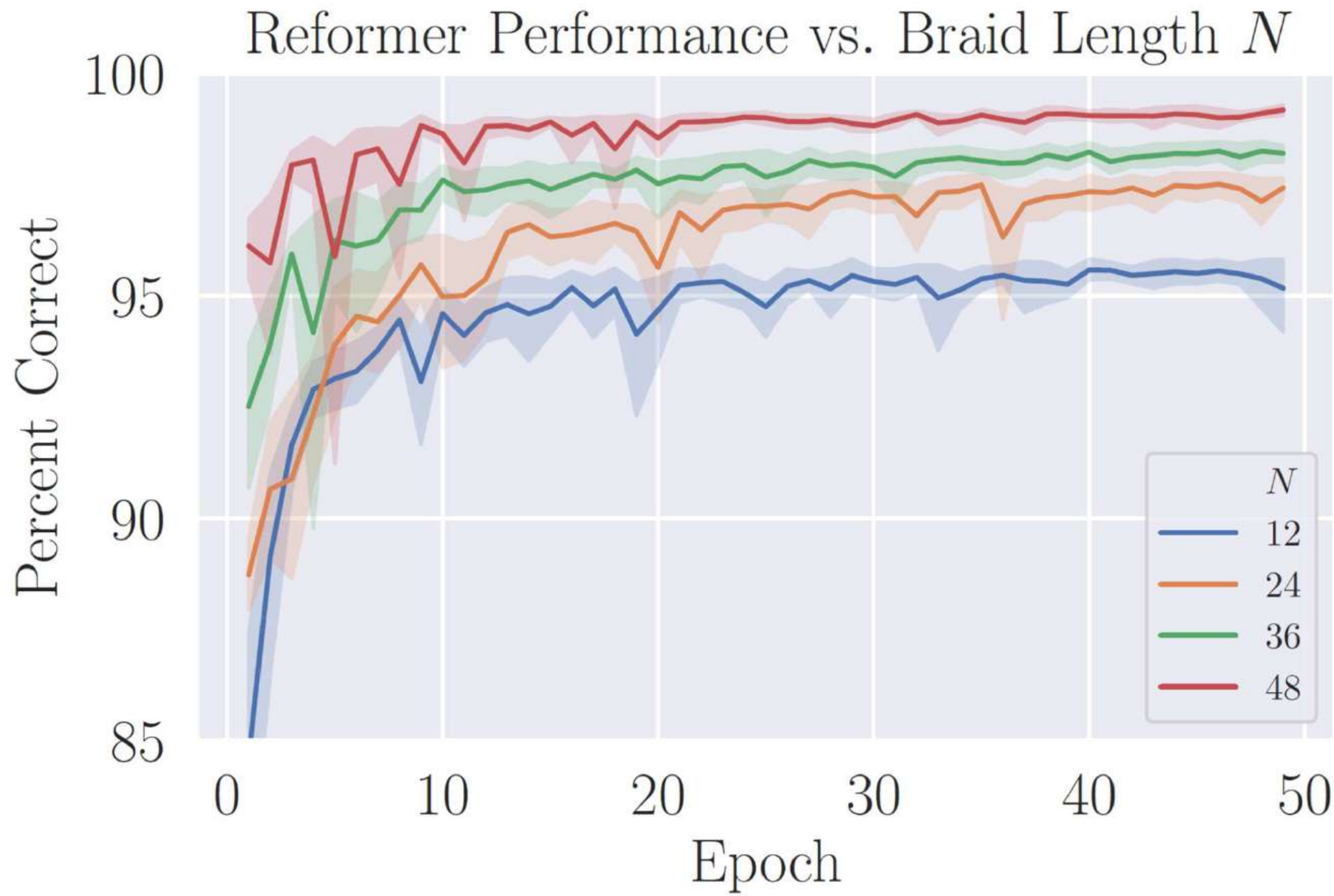
ABSTRACT

Large Transformer models routinely achieve state-of-the-art results on a number of tasks but training these models can be prohibitively costly, especially on long sequences. We introduce two techniques to improve the efficiency of Transformers. For one, we replace dot-product attention by one that uses locality-sensitive hashing, changing its complexity from $O(L^2)$ to $O(L \log L)$, where L is the length of the sequence. Furthermore, we use reversible residual layers instead of the standard residuals, which allows storing activations only once in the training process instead of N times, where N is the number of layers. The resulting model, the Reformer, performs on par with Transformer models while being much more memory-efficient and much faster on long sequences.

Learning to Unknot

Sergei Gukov[1], James Halverson[2,3], Fabian Ruehle[4,5], Piotr Sułkowski[1,6]





Reformer performance on UNKNOT as function of braid length. Performance increases with N.

The Computational Complexity of Knot and Link Problems

Joel Hass [*], Jeffrey C. Lagarias [†]and Nicholas Pippenger [‡]

February 1, 2008

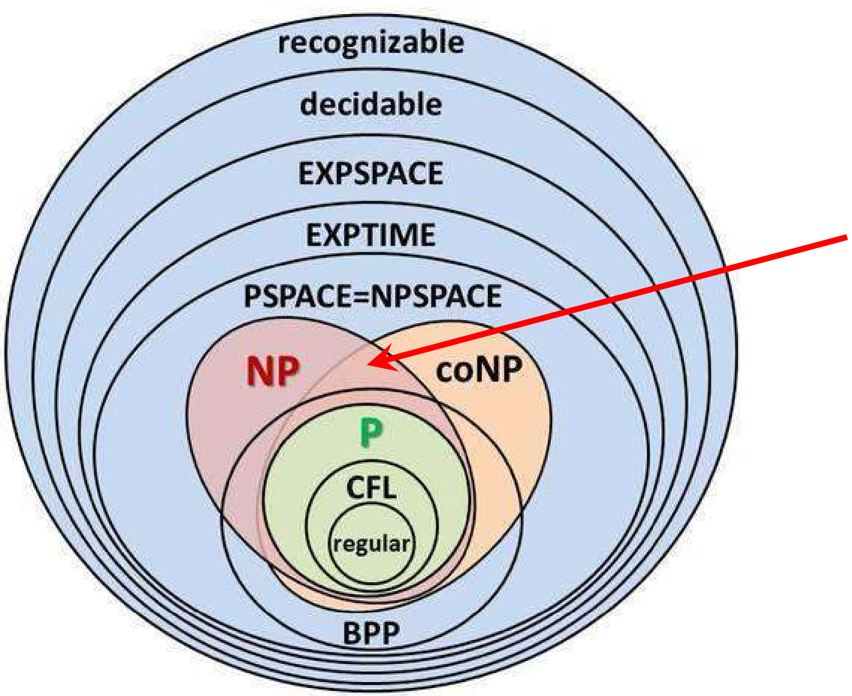
Abstract

We consider the problem of deciding whether a polygonal knot in 3-dimensional Euclidean space is unknotted, capable of being continuously deformed without self-intersection so that it lies in a plane. We show that this problem, UNKNOTTING PROBLEM is in **NP**. We also consider the problem, UNKNOTTING PROBLEM of determining whether two or more such polygons can be split, or continuously deformed without self-intersection so that they occupy both sides of a plane without intersecting it. We show that it also is in **NP**. Finally, we show that the problem of determining the genus of a polygonal knot (a generalization of the problem of determining whether it is unknotted) is in **PSPACE**. We also give exponential worst-case running time bounds for deterministic algorithms to solve each of these problems. These algorithms are based on the use of normal surfaces and decision procedures due to W. Haken, with recent extensions by W. Jaco and J. L. Tollefson.

Knottedness is in NP, modulo GRH

Greg Kuperberg*
Department of Mathematics, University of California, Davis, CA 95616

Given a tame knot K presented in the form of a knot diagram, we show that the problem of determining whether K is knotted is in the complexity class NP, assuming the generalized Riemann hypothesis (GRH). In other words, there exists a polynomial-length certificate that can be verified in polynomial time to prove that K is non-trivial. GRH is not needed to believe the certificate, but only to find a short certificate. This result complements the result of Hass, Lagarias, and Pippenger that unknottedness is in NP. Our proof is a corollary of major results of others in algebraic geometry and geometric topology.



$$\text{Unknottedness} \in \mathsf{NP} \cap \mathsf{coNP}$$

$$\text{integer} \overset{?}{=} \text{product of two primes}$$

:

THE EFFICIENT CERTIFICATION OF KNOTTEDNESS AND THURSTON NORM

MARC LACKENBY

1. Introduction

How difficult is it to determine whether a given knot is the unknot? The answer is not known. There might be a polynomial-time algorithm, but so far, this has remained elusive. The complexity of the unknot recognition problem was shown to be in NP by Hass, Lagarias and Pippenger [10]. The main aim of this article is to establish that it is in co-NP. This can be stated equivalently in terms of the Knottedness decision problem, which asks whether a given knot diagram represents a non-trivial knot.

Theorem 1.1. *Knottedness is in NP.*

In some sense, this result is not new. It was first announced by Agol [1] in 2002, but he has not provided a full published proof. In 2011, Kuperberg gave an alternative proof of Theorem 1.1, but under the extra assumption that the Generalised Riemann Hypothesis is true [19]. In this paper, we provide the first full proof of the unconditional result.

Combined with the theorem of Hass, Lagarias and Pippenger [10], Theorem 1.1 gives the following corollary.

Corollary 1.2. *If either of the decision problems* Unknot recognition *or* Knottedness *is NP-complete, then NP = co-NP.*

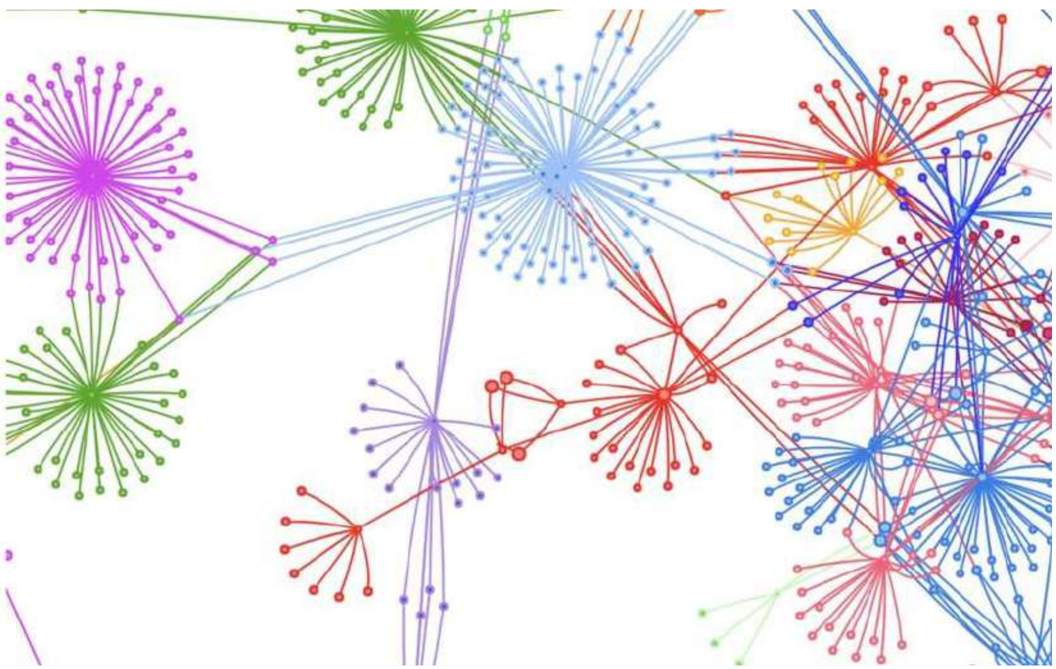
This is because if any decision problem in co-NP is NP-complete, then the complexity classes NP and co-NP must be equal. Since this is widely viewed not to be the case (see Section 2.4.3 in [7] for example), then it seems very unlikely that either of these decision problems is NP-complete.

The Unbearable Hardness of Unknotting*

Arnaud de Mesmay[1], Yo'av Rieck[2], Eric Sedgwick[3], and Martin Tancer[4]

Abstract

We prove that deciding if a diagram of the unknot can be untangled using at most k Riedemeister moves (where k is part of the input) is **NP**-hard. We also prove that several natural questions regarding links in the 3-sphere are **NP**-hard, including detecting whether a link contains a trivial sublink with n components, computing the unlinking number of a link, and computing a variety of link invariants related to four-dimensional topology (such as the 4-ball Euler characteristic, the slicing number, and the 4-dimensional clasp number).



cf. connected components of a graph:

- Not finite
- Not explicitly presented

arXiv:1907.05981v1 [math.GT] 13 Jul 2019

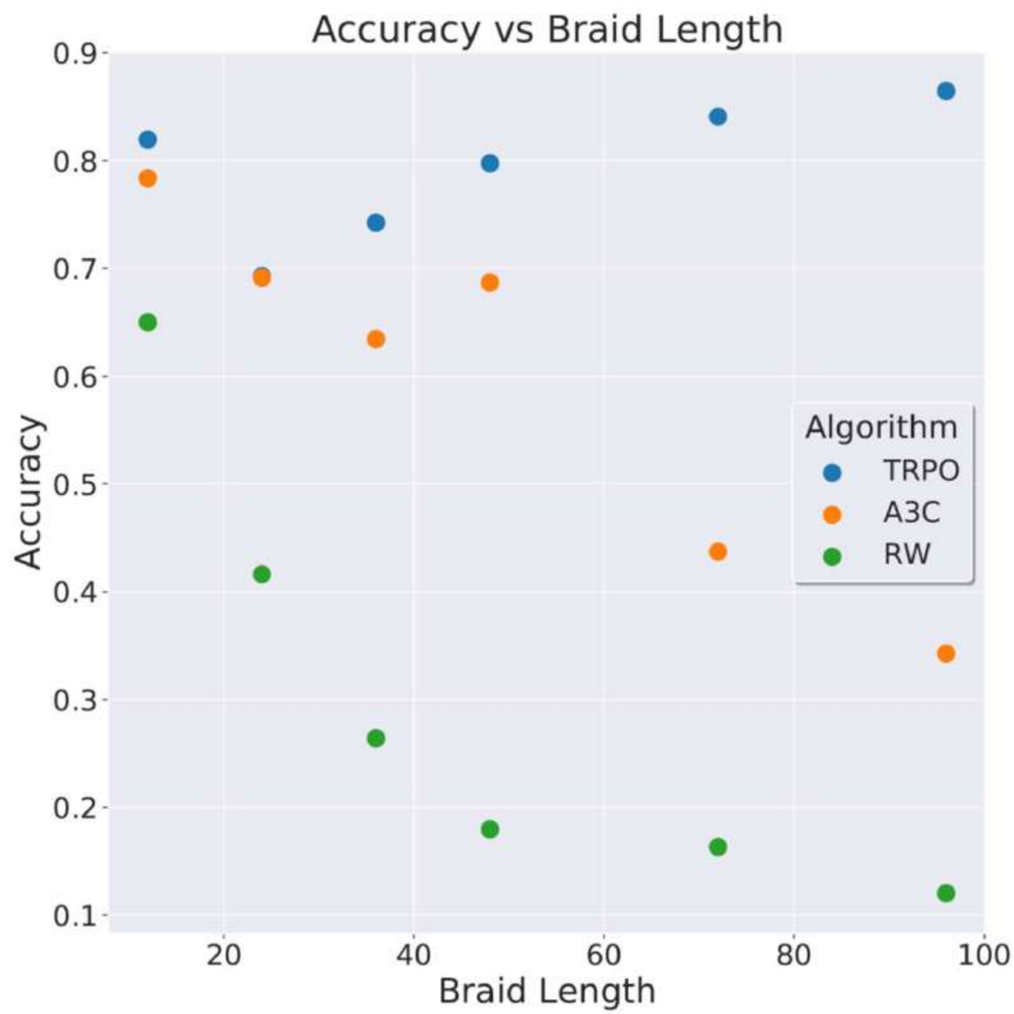
Coloring invariants of knots and links are often intractable

Greg Kuperberg*
University of California, Davis

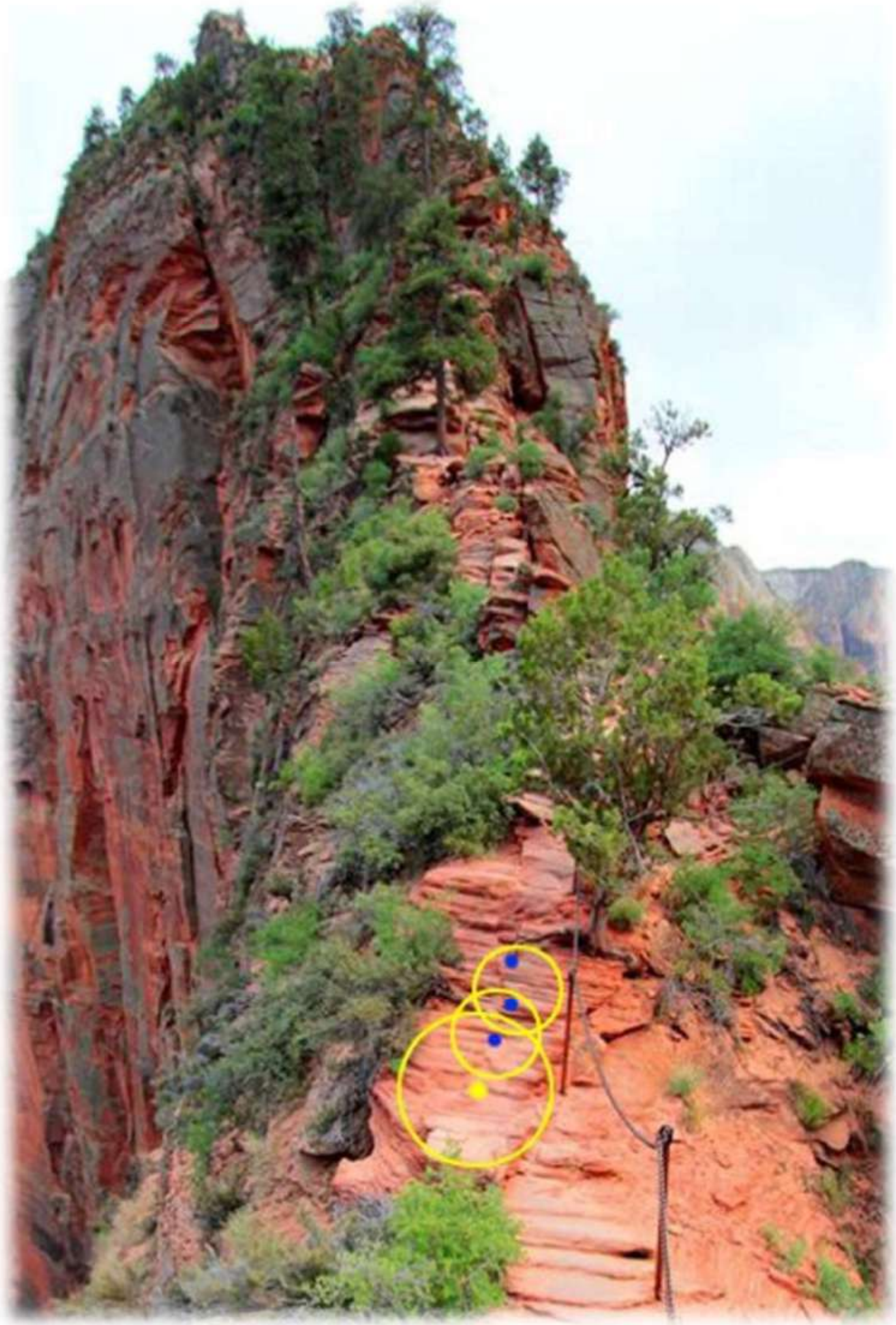
Eric Samperton†
University of California, Santa Barbara

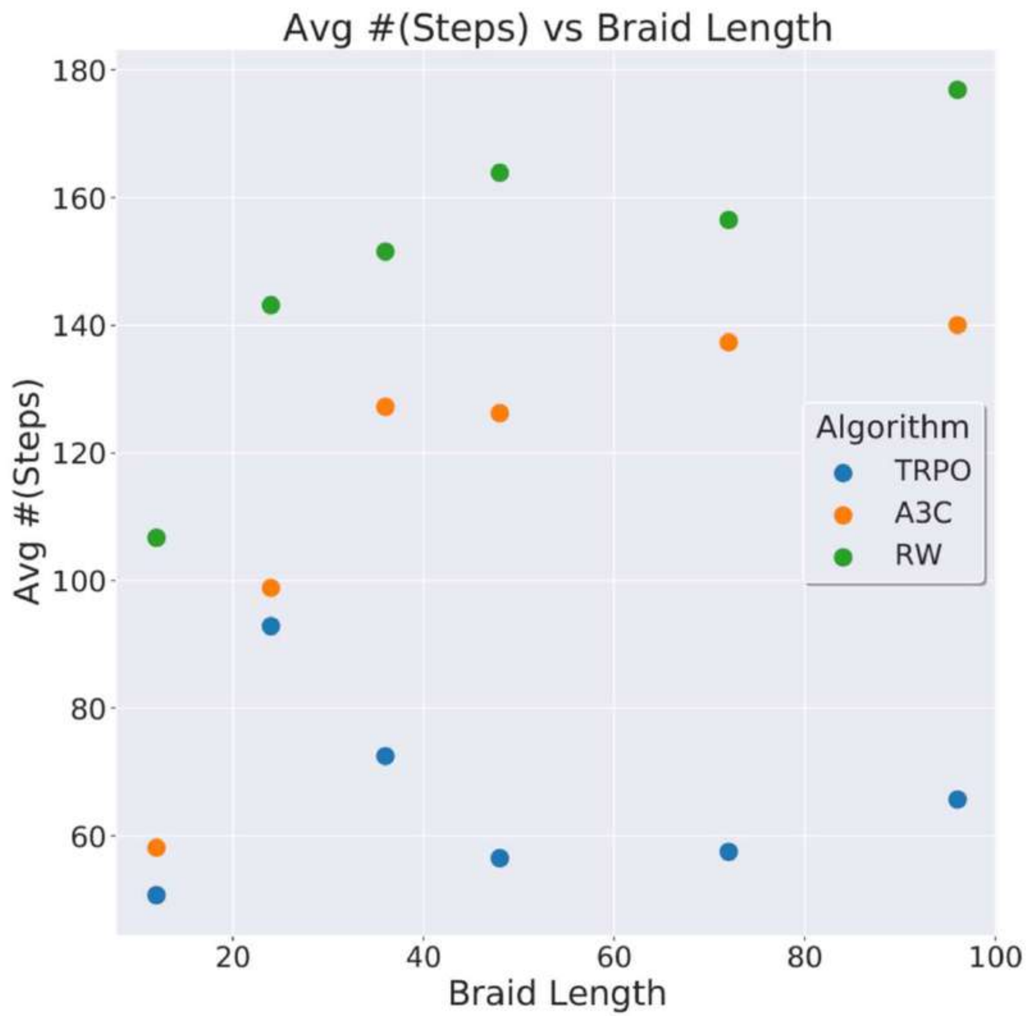
(Dated: July 16, 2019)

Let G be a nonabelian, simple group with a nontrivial conjugacy class $C \subseteq G$. Let K be a diagram of an oriented knot in S^3, thought of as computational input. We show that for each such G and C, the problem of counting homomorphisms $\pi_1(S^3 \setminus K) \to G$ that send meridians of K to C is almost parsimoniously #P-complete. This work is a sequel to a previous result by the authors that counting homomorphisms from fundamental groups of integer homology 3-spheres to G is almost parsimoniously #P-complete. Where we previously used mapping class groups actions on closed, unmarked surfaces, we now use braid group actions.

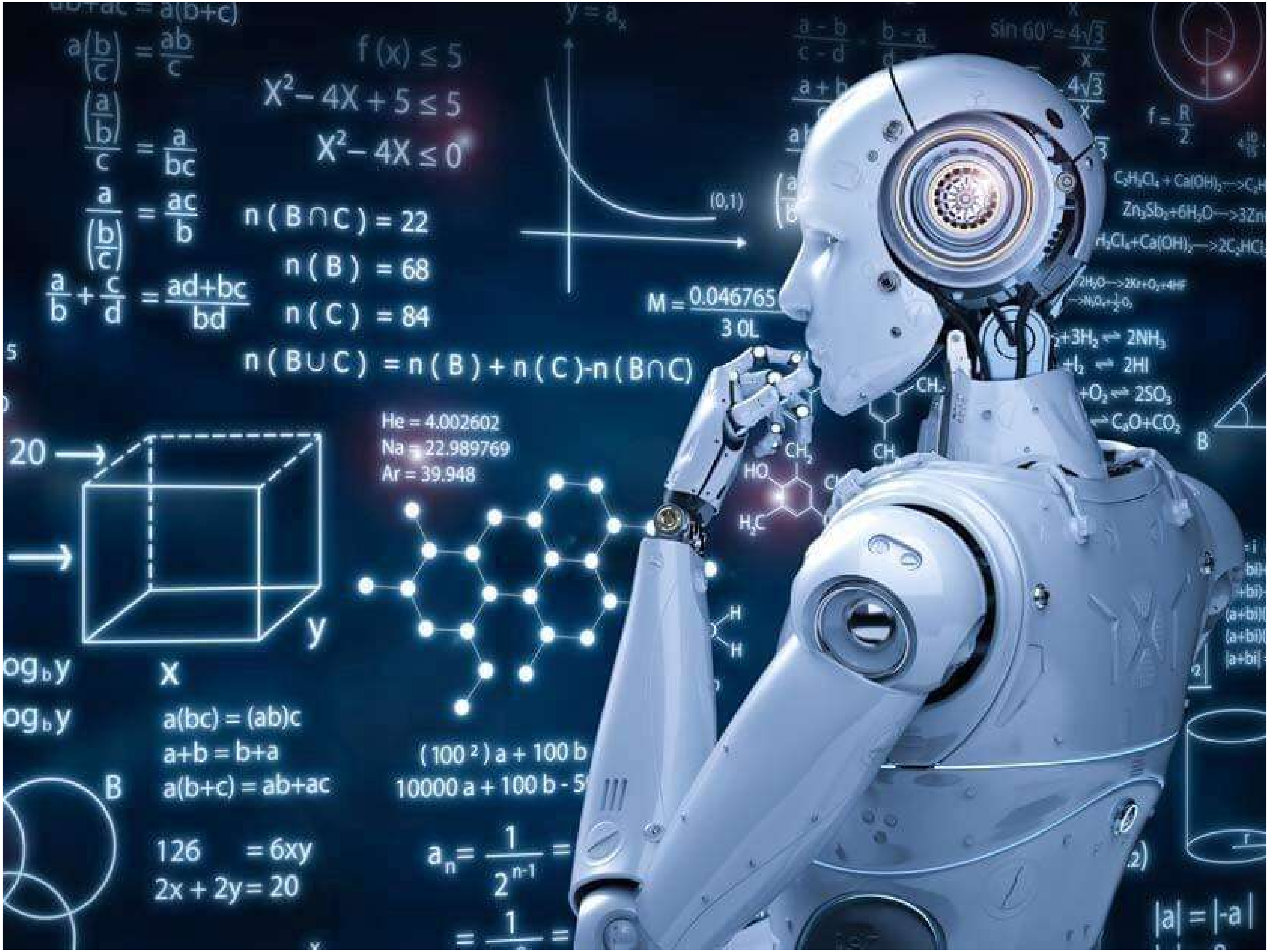


Fraction of unknots whose braid words could be reduced to the empty braid word as a function of initial braid word length.

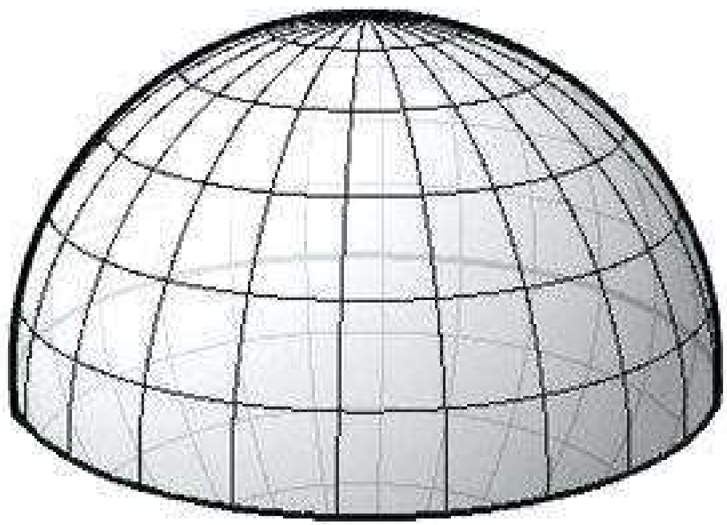





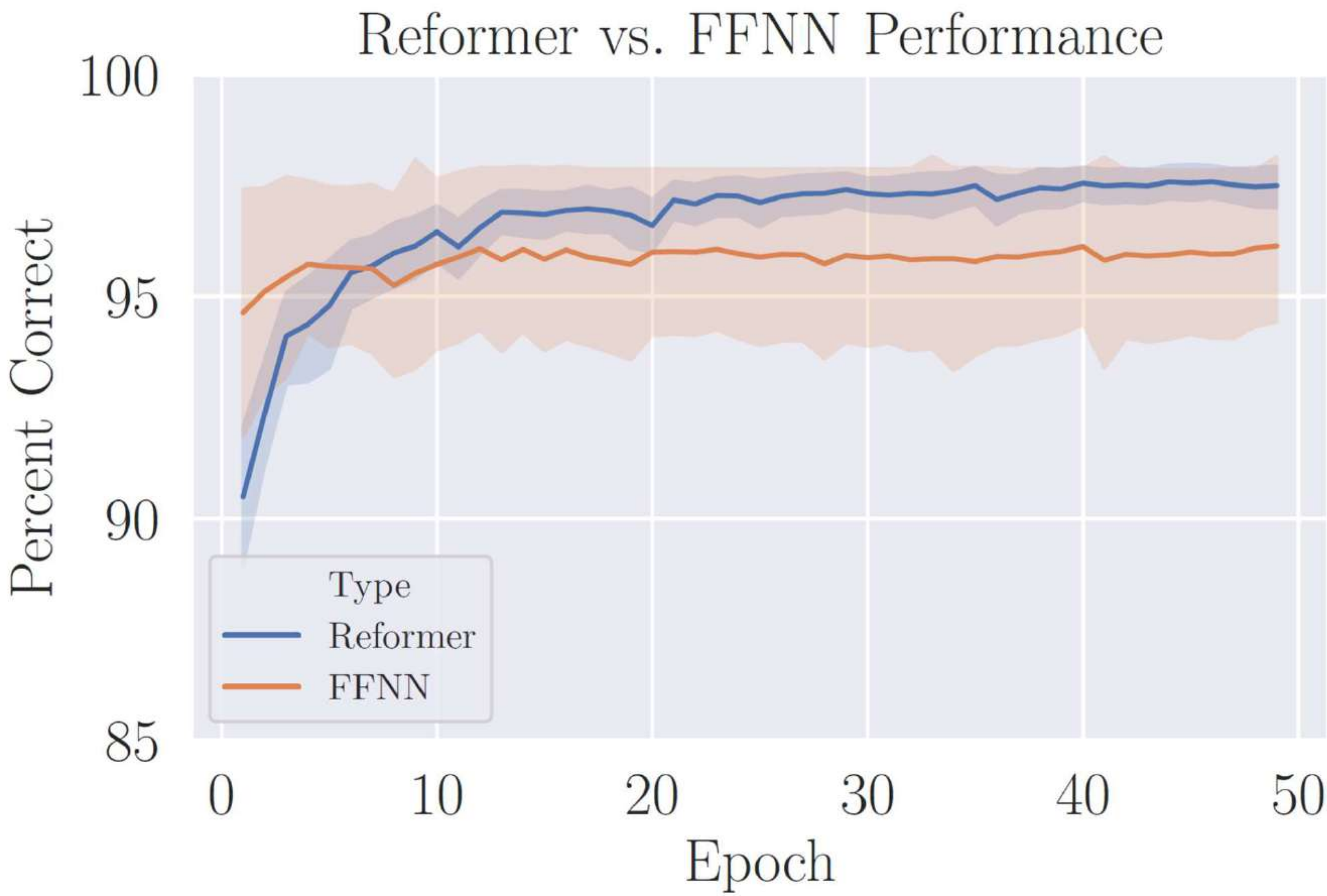
Average number of actions necessary to reduce the input braid word to the empty braid word as a function of initial braid word length.



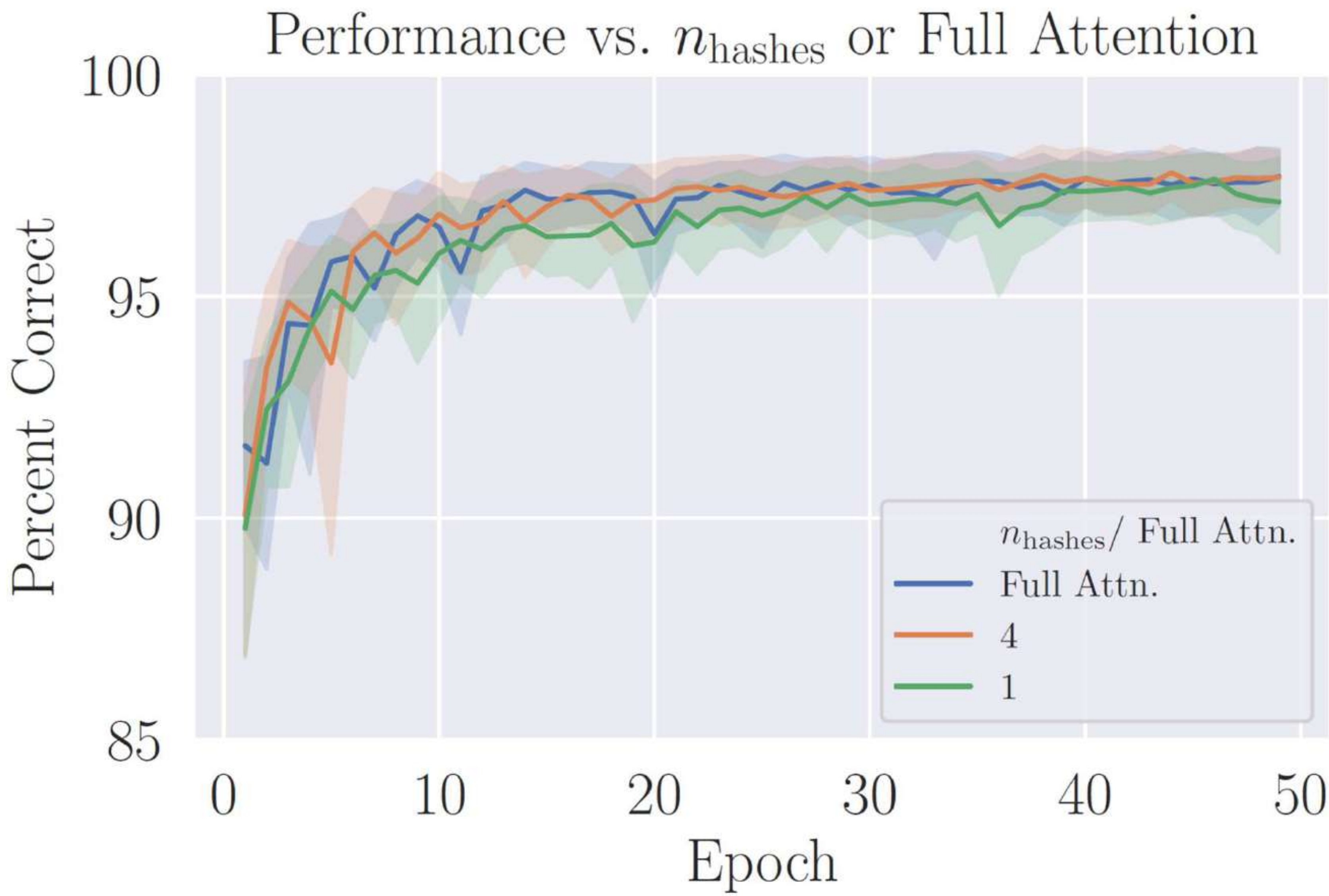
f(x) ≤ 5
X²– 4X + 5 ≤ 5
X²– 4X ≤ 0
n(B∩C) = 22
n(B) = 68
n(C) = 84
n(B∪C) = n(B) + n(C)-n(B∩C)
M = 0.046765 / 3 0L
He = 4.002602
Na = 22.989769
Ar = 39.948
a(bc) = (ab)c
a+b = b+a
a(b+c) = ab+ac
126 = 6xy
2x + 2y = 20
(100²) a + 100 b
|a| = |-a|



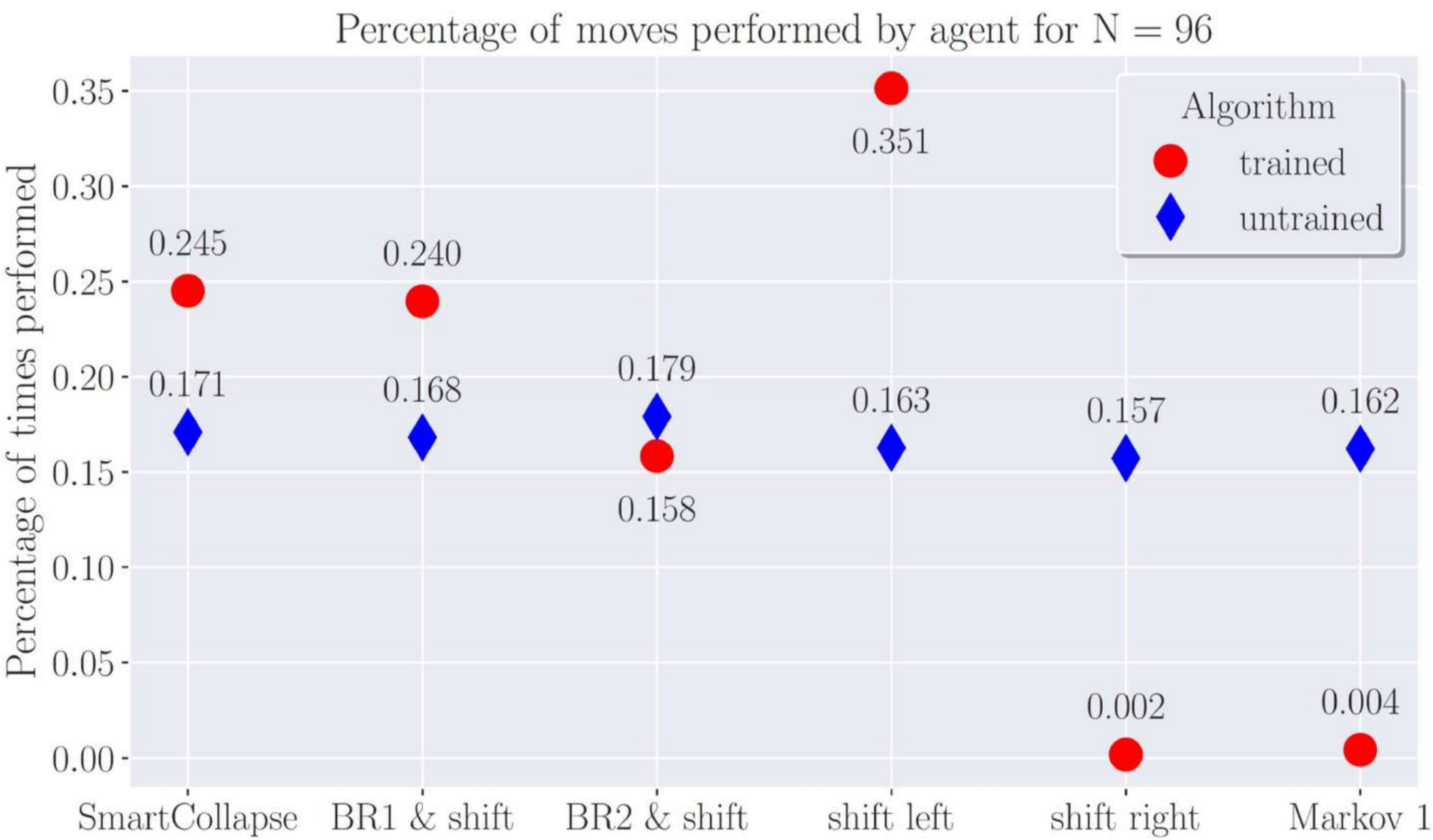
an exotic 4-ball has no smooth radius function
with 3-sphere levels



Performance comparison between reformer and feedforward network (FFNN).

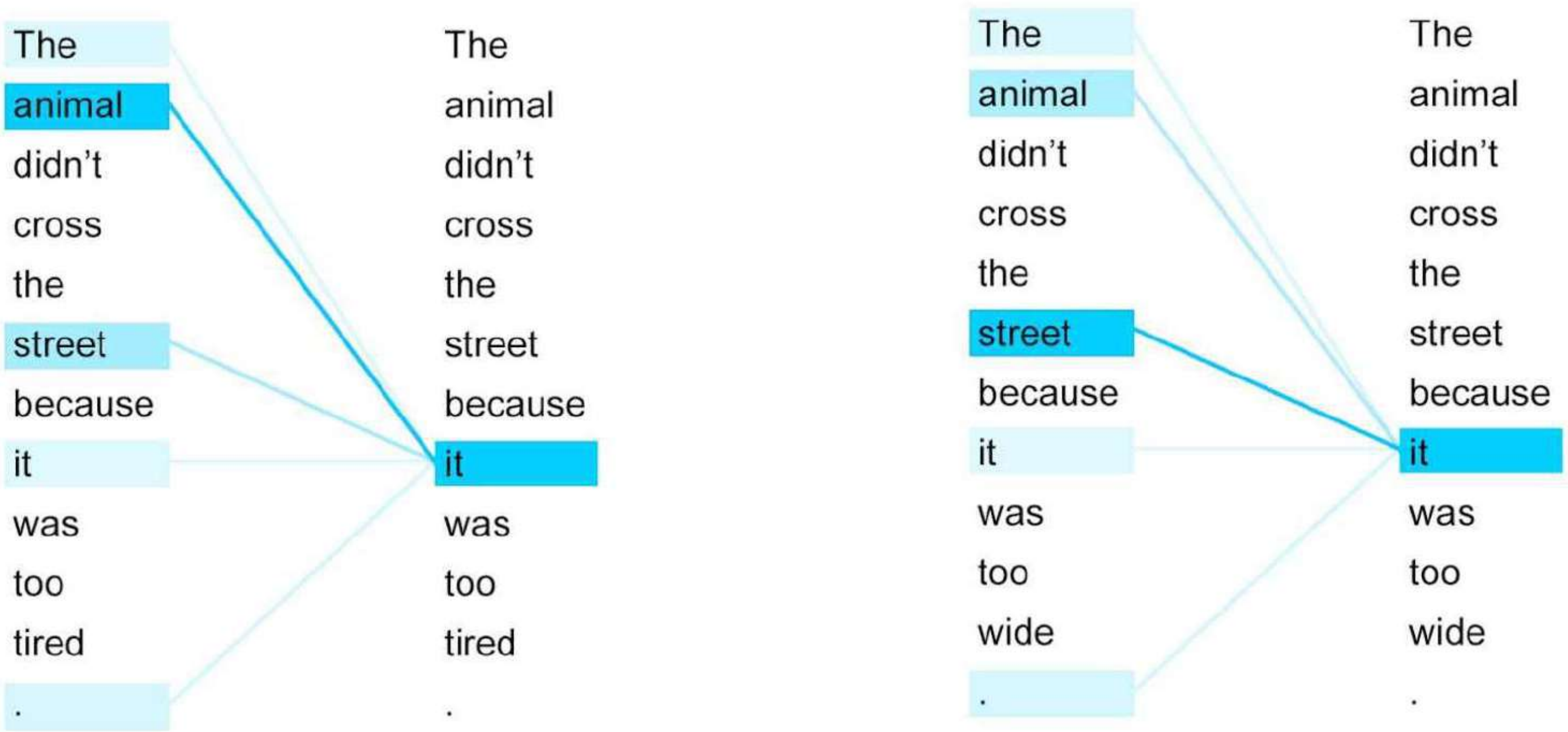


Performance dependence on the number of locality sensitive hashes (LSH).



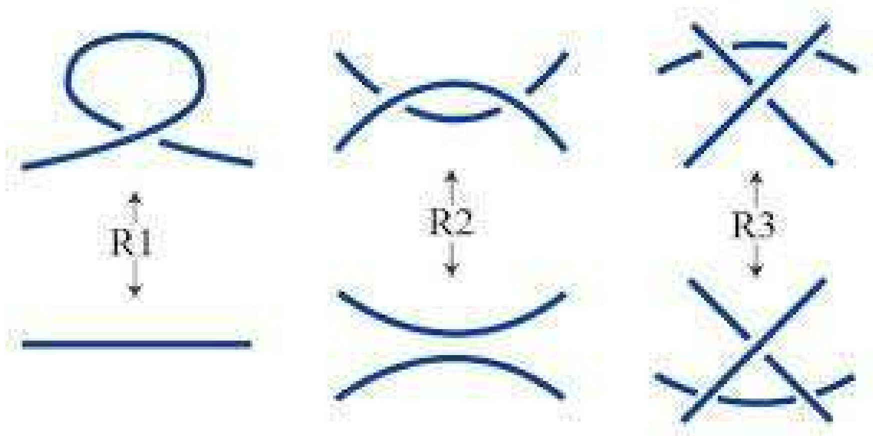
Architectures: Reformer NN, FFNN, … implementation: ChainerRL

Self-Attention



- Is it knotted? **Easy**

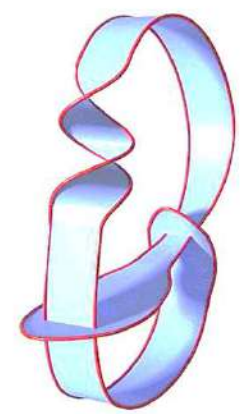
S.G., J.Halverson, F.Ruehle, P.Sulkowski



- Is it ribbon? Is it slice? **Hard**

S.G., J.Halverson, C.Manolescu, F.Ruehle
(see Fabian's talk)

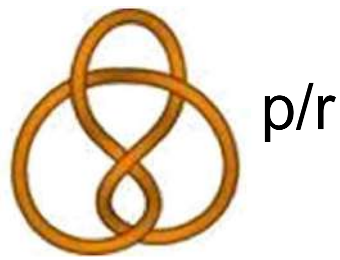
(SPC4, slice-ribbon, ...)



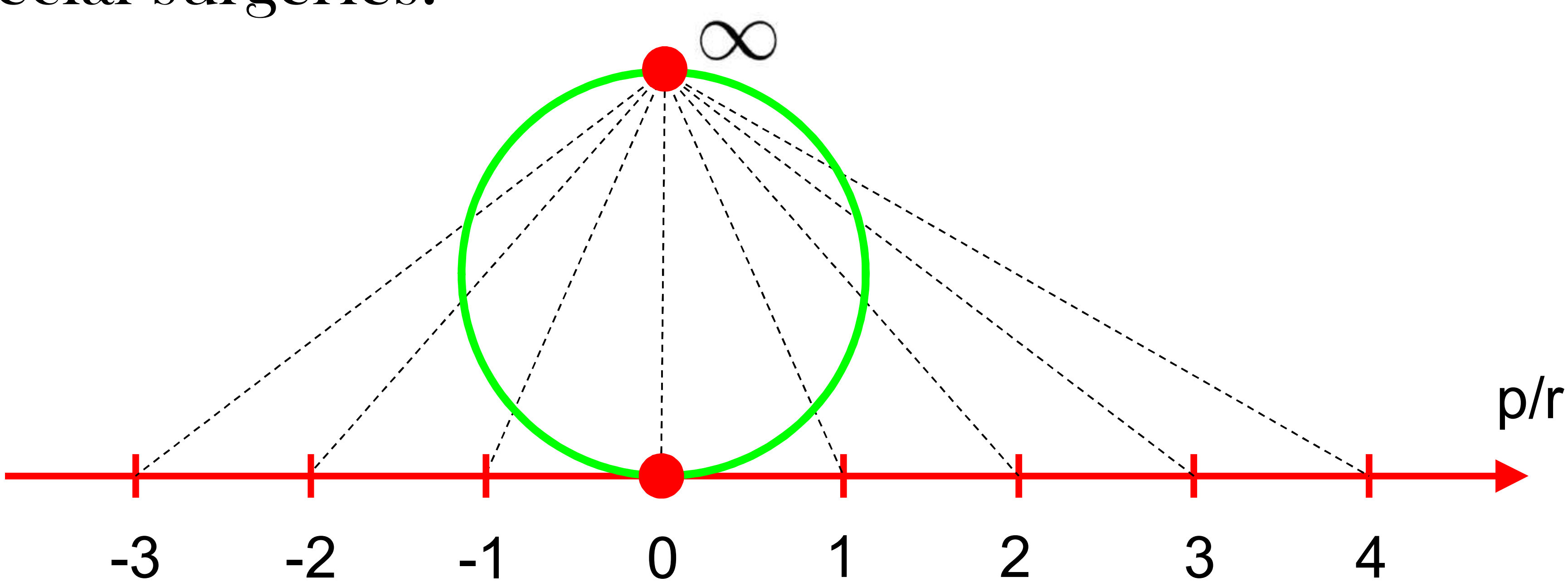
- Is it Andrews-Curtis trivial? **Don't know yet**

Theorem [Lickorish, Wallace]:

Every connected oriented closed 3-manifold arises by performing an integral Dehn surgery along a link in S^3.



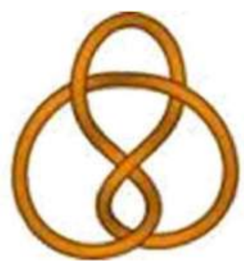
Special surgeries:



Property R

Theorem ("property R" conjecture): D.Gabai (1983)

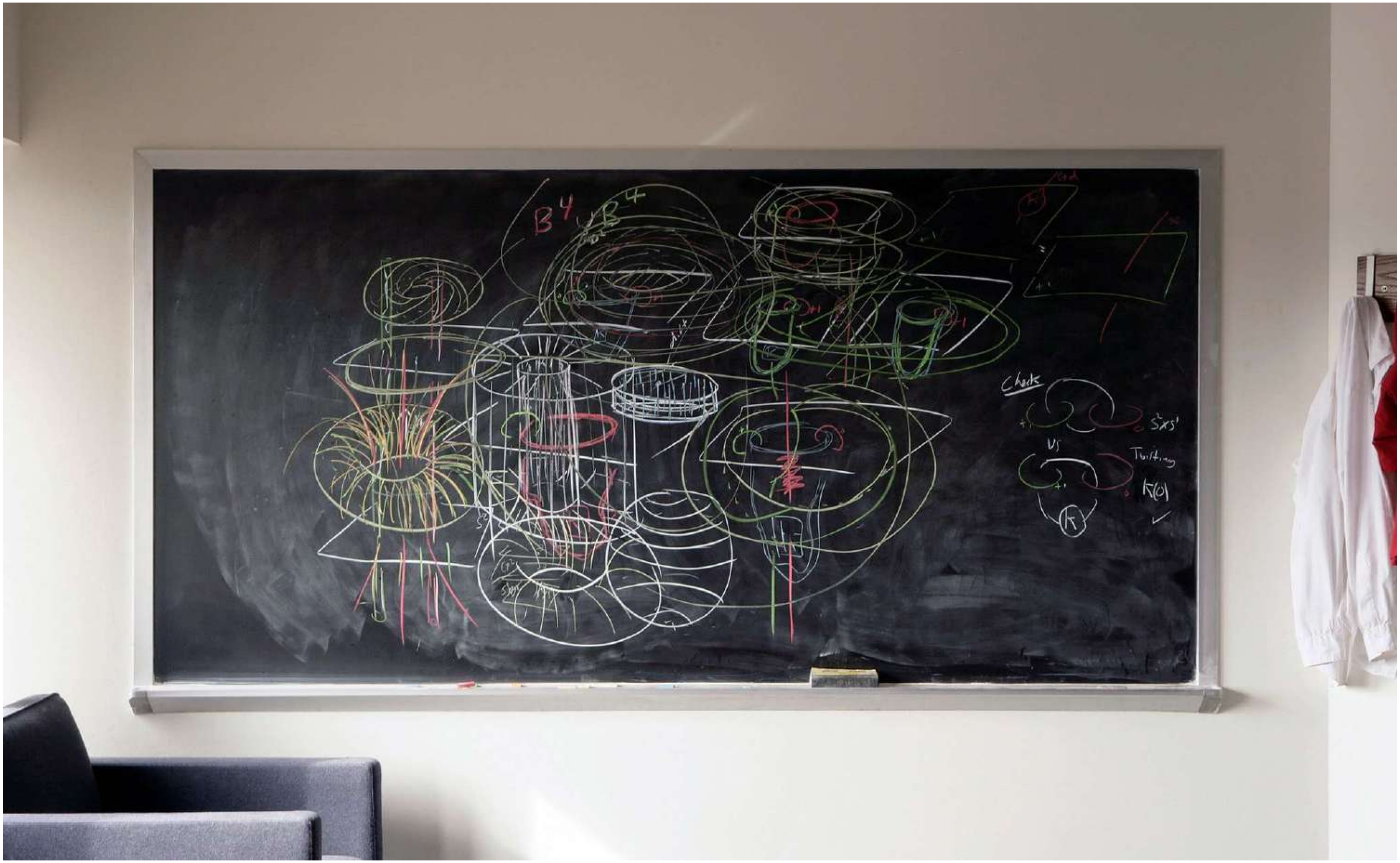
If the 0-surgery on $K \subset S^3$ is homeomorphic to $S^1 \times S^2$, then K is the unknot.



The trefoil knot and the figure-8 knot are uniquely characterized by 0-surgery.

D.Gabai (1987)

$$M_3 = S^3_0(K)$$



Chalkboard of David Gabai at Princeton
Jessica Wynne

Conjecture [Akbulut and Kirby '97]:

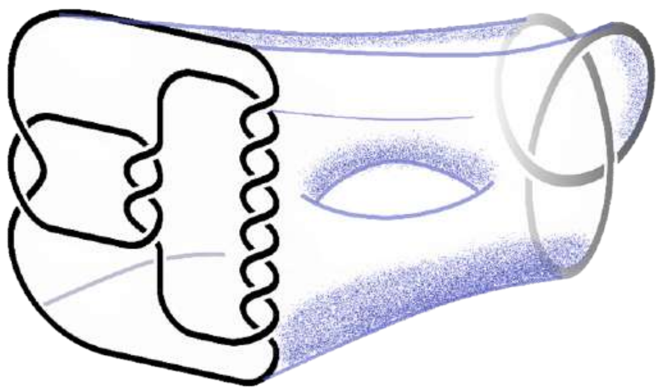
If 0-surgeries on two knots give the same 3-manifold,

$$S^3_0(K) \cong S^3_0(K')$$

then the knots are concordant.

FALSE

P.Kirk, C.Livingston (1999)



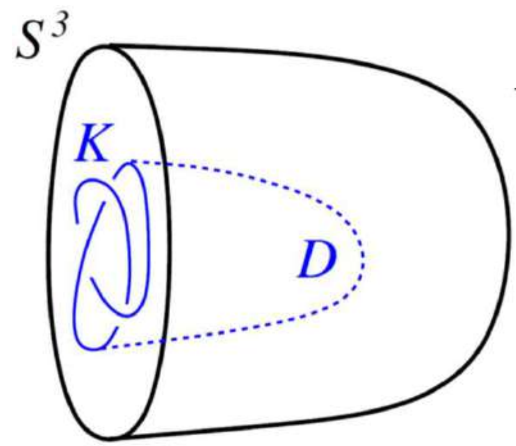
Conjecture:

If 0-surgeries on two knots give the same 3-manifold, then the knots with relevant orientations are concordant.

False if slice-ribbon conjecture is true. FALSE K.Yasui (2015)

Thm: For $M_3 = S^3_0(K)$ at least one of Rokhlin invariants vanishes. M.Hedden, M.H.Kim, T.Mark, K.Park (2018)

Cor: If M_3 is integral homology $S^1 \times S^2$ with two non-trivial Rokhlin invariants, then $M_3 \neq S^3_0(K)$.



Thm: If K is slice, then L.Truong (2021)

$$b_2(M_4) \geq \frac{10}{8}|\sigma(M_4)| + 5$$

where $\partial M_4 = S^3_0(K)$, $b_2(M_4) \neq$ 1, 3, or 23, and M_4 is a two-handlebody (two-handles attached to a 4-ball).

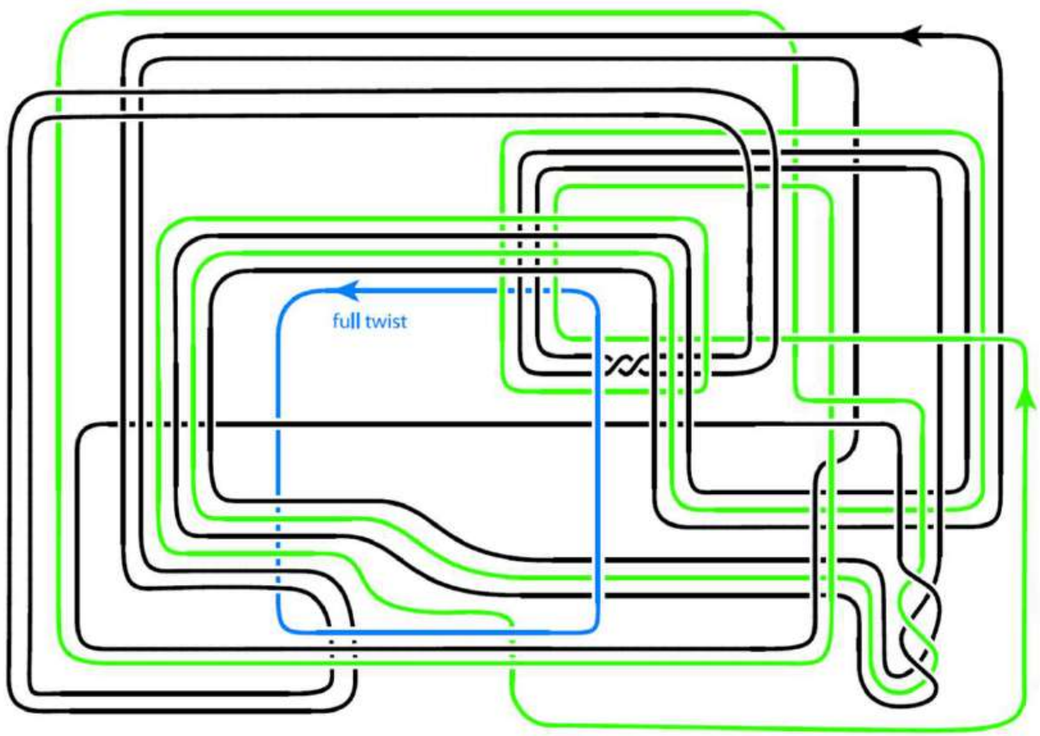
Obstructions to smooth sliceness:

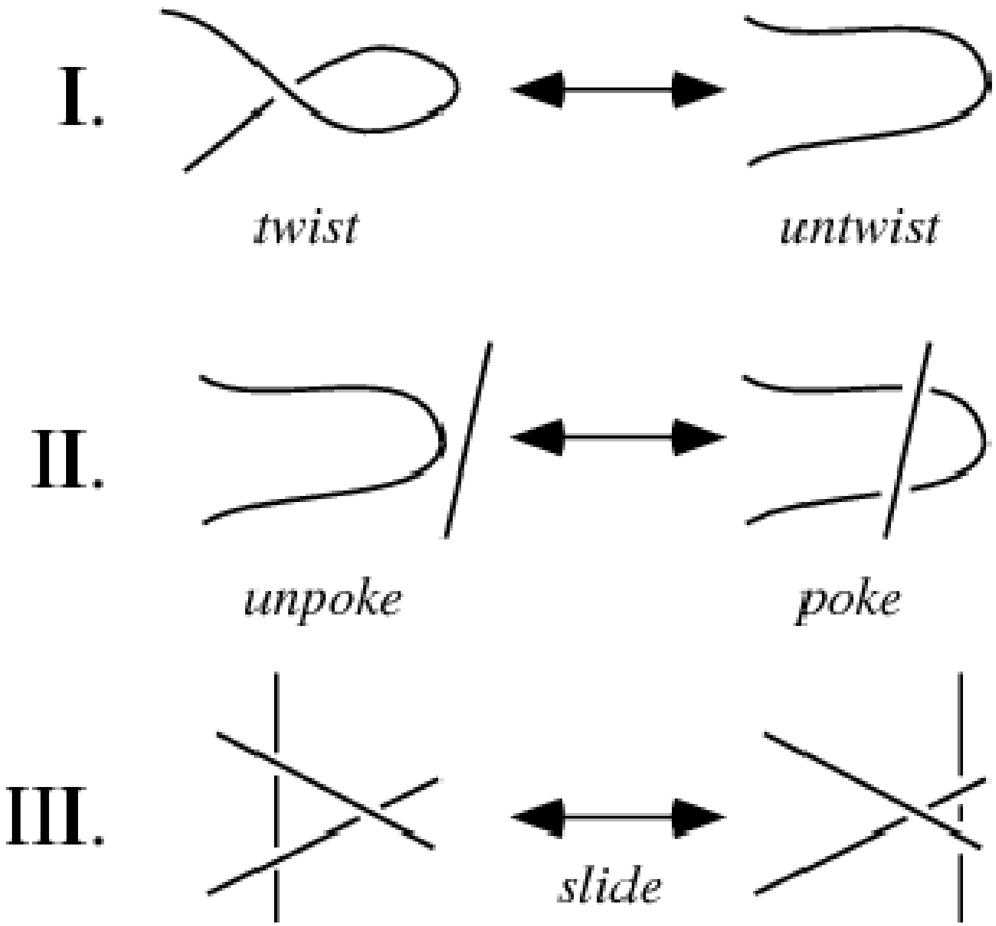
- $\mathrm{Arf}(K)$ Robertello (1965)
- Fox-Milnor condition $\Delta_K(x) = f(x)f(x^{-1})$ (1966)
- Levine-Tristram signature (1969)
- τ Ozsvath-Szabo (2003)
- ε Hom (2014)
- s Rasmussen (2010)
- s_n Lewark-Lobb (2015)
- $\underline{V}_0$ and $\bar{V}_0$ from involutive HF Hendricks-Manolescu (2017)
- $\Upsilon(t)$ Ozsvath-Stipsicz-Szabo (2017)
- φ_j Dai-Hom-Stoffregen-Truong (2019)

$\mathrm{Arf}\ (K)$	OEIS	counts of prime knots with $n = 1$, 2, ... crossings
0	A131433	0, 0, 0, 0, 1, 1, 3, 10, 25, 82, ...
1	A131434	0, 0, 1, 1, 1, 2, 4, 11, 24, 83, ...

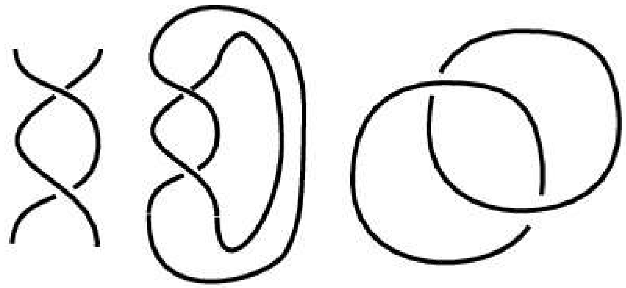
Man and machine thinking about the smooth 4-dimensional Poincaré conjecture.

Michael Freedman
Robert Gompf
Scott Morrison
Kevin Walker

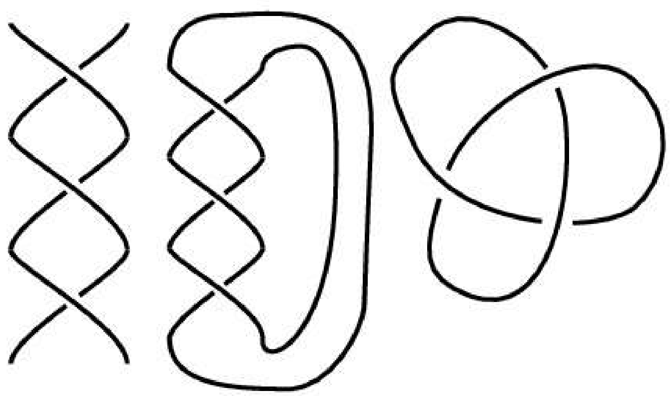




Kurt Reidemeister



Hopf Link



Trefoil Knot

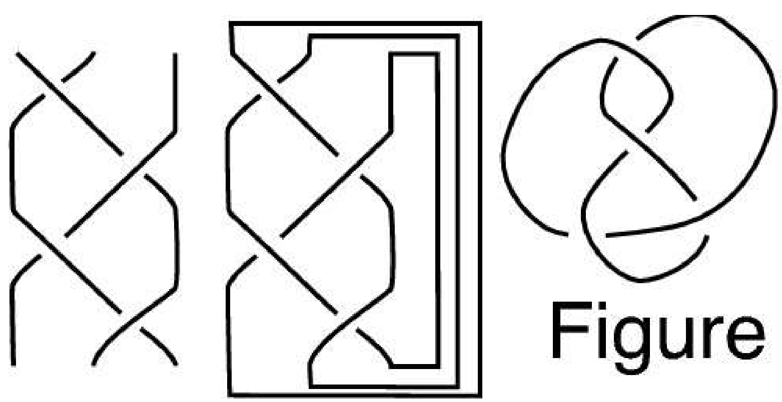
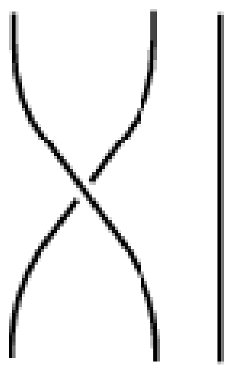
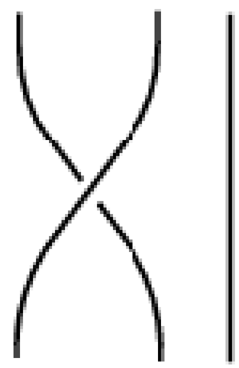


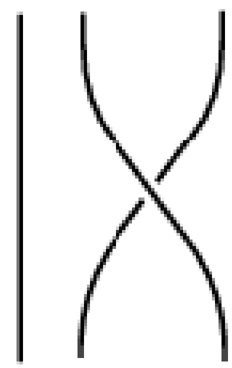
Figure Eight Knot



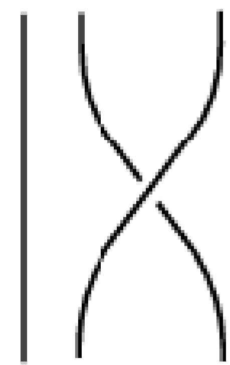
σ_1



σ_1^{-1}

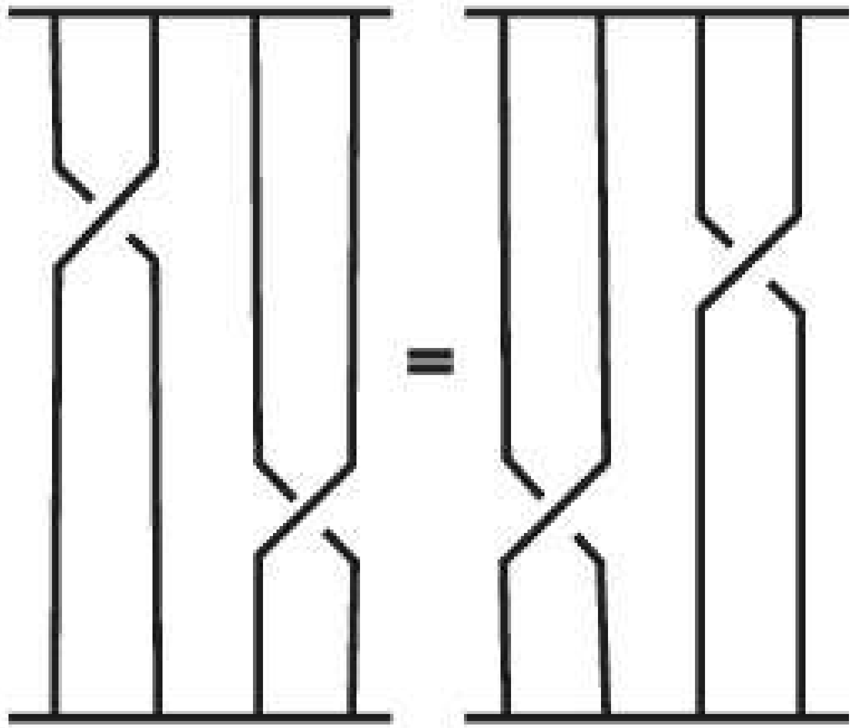


σ_2



σ_2^{-1}

$$\sigma_i \sigma_j = \sigma_j \sigma_i \qquad \text{for } |i - j| > 1$$



$$\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}$$