MICHAEL AT THE GATES OF ...

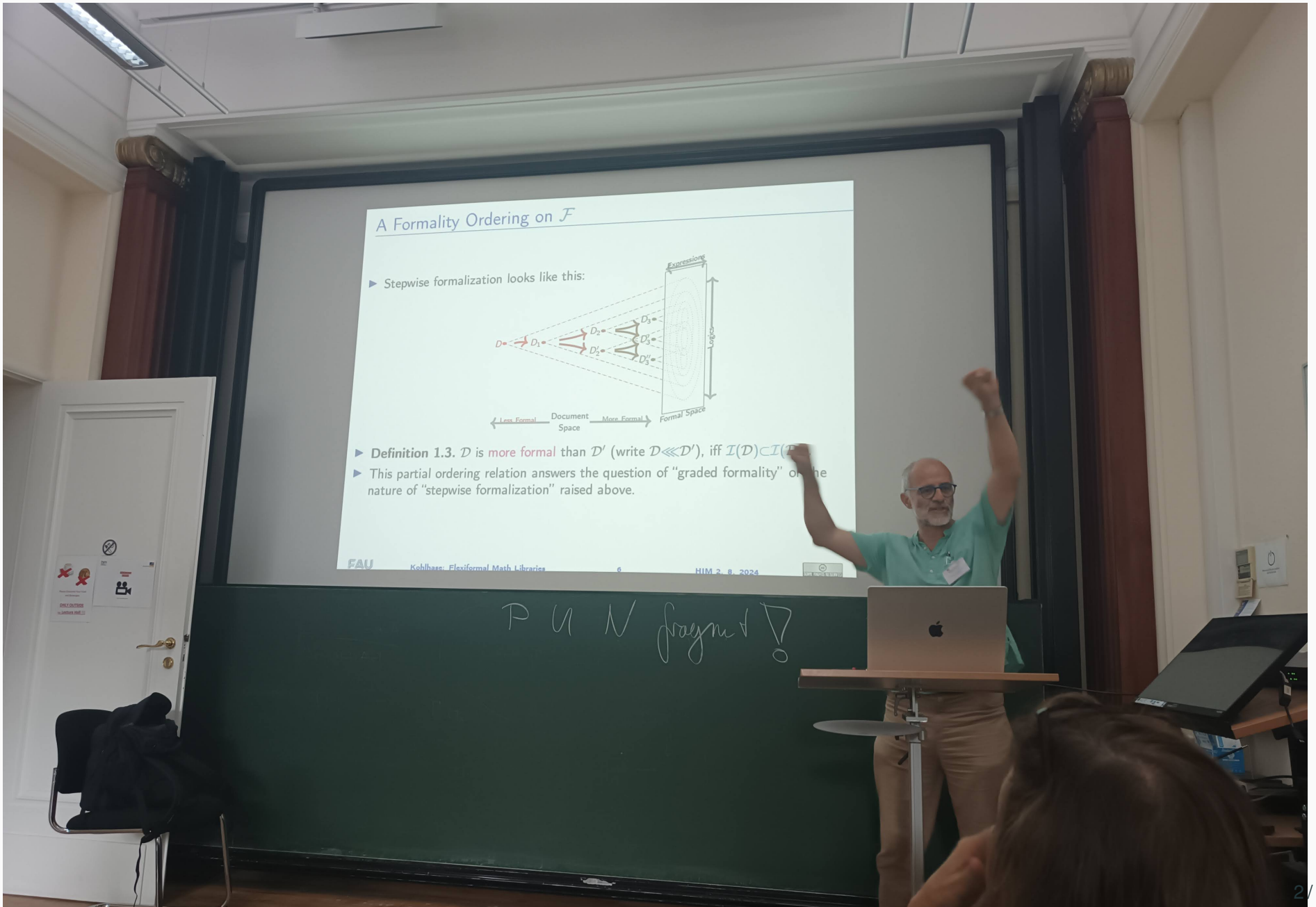
(some recollections and where it all went)

Josef Urban

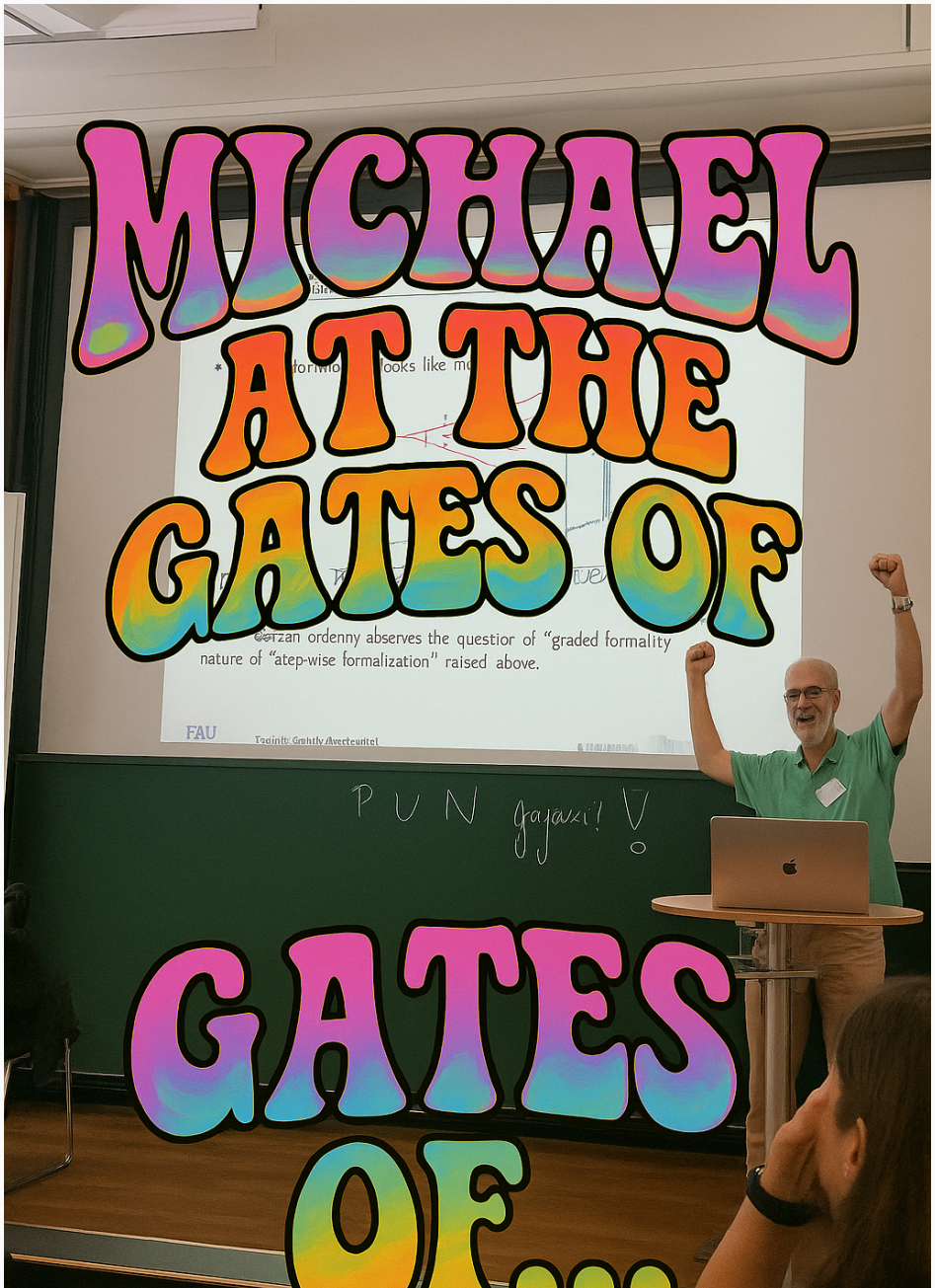
Czech Technical University in Prague

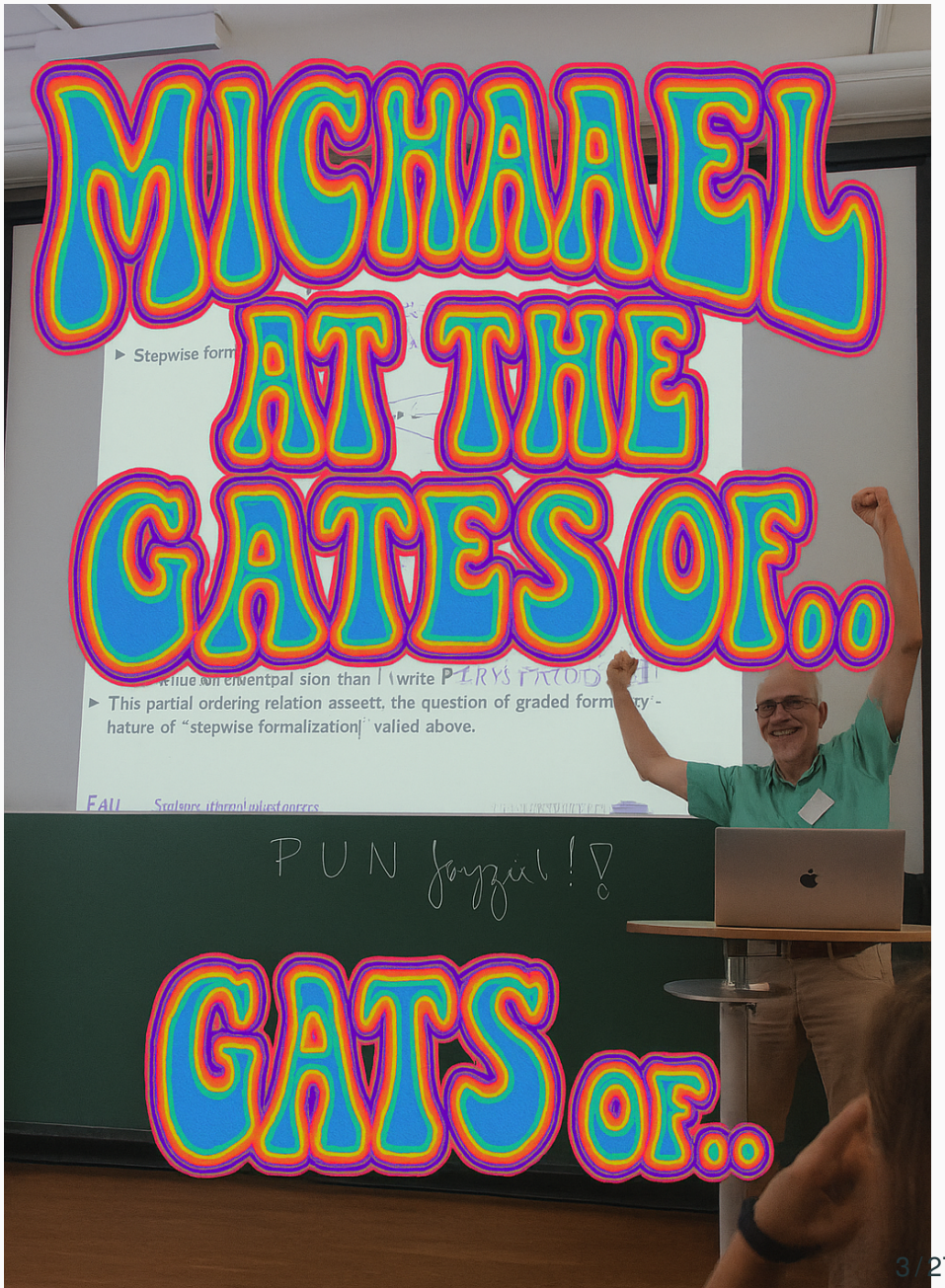
MiKo 60
May 17, 2025, Erlangen

Michael ...


A Formality Ordering on $\mathcal{F}$
Stepwise formalization looks like this:
Expressions
Logics
Less Formal
Document Space
More Formal
Formal Space
Definition 1.3. $\mathcal{D}$ is more formal than $\mathcal{D}'$ (write $\mathcal{D} \ll \mathcal{D}'$), iff $\mathcal{I}(\mathcal{D}) \subset \mathcal{I}(\mathcal{D}'$
This partial ordering relation answers the question of "graded formality" of the nature of "stepwise formalization" raised above.
FAU
Kohlhase: Flexiformal Math Libraries
HIM 2. 8. 2024
P U N fragen!

... at the Gates ...


MICHAEL
AT THE
GATES OF
GATES
OF...


MICHAEL
AT THE
GATES OF..
GATS OF..

... of ...

- Math Knowledge Management ?
- Calculemus ?
- Higher-Order ATP ?
- Omega ?
- Proof Planning ?
- OmDoc/XML/sTeX/OpenMath ?
- Flexiformalization ?
- LaTeXML / ArxMLiv ?
- MathWiki / PlanetMath / eMath3.0 ?
- MathWebSearch ?
- QED / Big Math ?
- Math Software ?
- Library Pluralism ?
- Alignments / Autoformalization ?
- Conjecturing ?
- Licensing ?
- ... ?

My first (?) contact with Michael - Mizar XML

```
Date: Thu, 14 Mar 2002 15:02:20 -0500
From: Michael Kohlhase <Michael_Kohlhase@asuka.mt.cs.cmu.edu>
Reply-To: Michael Kohlhase <kohlhase+@cs.cmu.edu>
To: mizar-forum@mizar.uwb.edu.pl

Hi,

[JU:]
> The first stage that is currently discussed, is using XML as an internal
> format for various Mizar files. Once this is done, we can think of expor
> to other formats, be it HELM or OMDoc. I did not find any DTDs on the
> MOWGLI page ... maybe I missed something?

[MK:]
I am not sure that the HELM format is a good format to transform into,
since it is very oriented towards the internal data format in Coq. I
understood Claudio's remark as "first use some XML-format for Mizar" just
as Coq uses the HELM format, which I think is very reasonable.
```

Mizar XML : the MKM Dream/Utopia (?)

- Export all (formal) math libraries into compatible formats
- Load them into a single big system
- Do joint search, retrieval, indexing, reasoning , ..., over all of them
- Is this even possible ?
- (People like Claudio heavily involved ...)
- Will we ever have a single system for all of code ?
- An eternal fight between Chaos (disruption) and Order (integration) ?
- My “small” project: make the Mizar Math Library more accessible
- Just a tiny piece in these big visions ...
- My first visit to MKM in Bremen in 2005 (met Florian, spoke about XML-ization)
- Where it all went: Lean (a single system) vs MMT/OMDoc/Deducti/ ?

Earlier almost contacts - Omega etc

- Late 1990s: I was playing with IMPS (little theories)
- Tried symbolic machine learning (ILP) over it
- The guiding idea: learn heuristics for particular situations
- 2000: interested also in the Omega system
- Mostly to look at its proof planning methods
- But also to see how all the TP modules there interact
- Interacted with Volker Sorge
- Perhaps being the first person who managed to install Omega outside SB?
- Today: lots of ML for high-level proof sketches (planning?)
- ... also lots of research on little theories (involving also Michael)
- quite a few larger ATP/AI metasystems (e.g., my MaLARea since 2006)
- however no really big equivalent of the big Omega ambition?

CALCULEMUS for me, MPTP and AITP

- Lots of my stays at the Mizar lab funded by CALCULEMUS (thanks Saarbruecken people!)
- My first paper on translating Mizar for ATPs targeted CALCULEMUS
- It didn't go as planned ...
- But I got quite a bit of support from people like Volker Sorge and Michael
- Ultimately, the first MPTP paper published at MKM 2003 in Bertinoro
- TODAY: MPTP has been the most researched AITP corpus
- used for devel/eval of AI/TP systems for over 20 years now
- For more, see my MPTP20 talk from 2023:
  `http://grid01.ciirc.cvut.cz/~mptp/MPTP20.pdf`

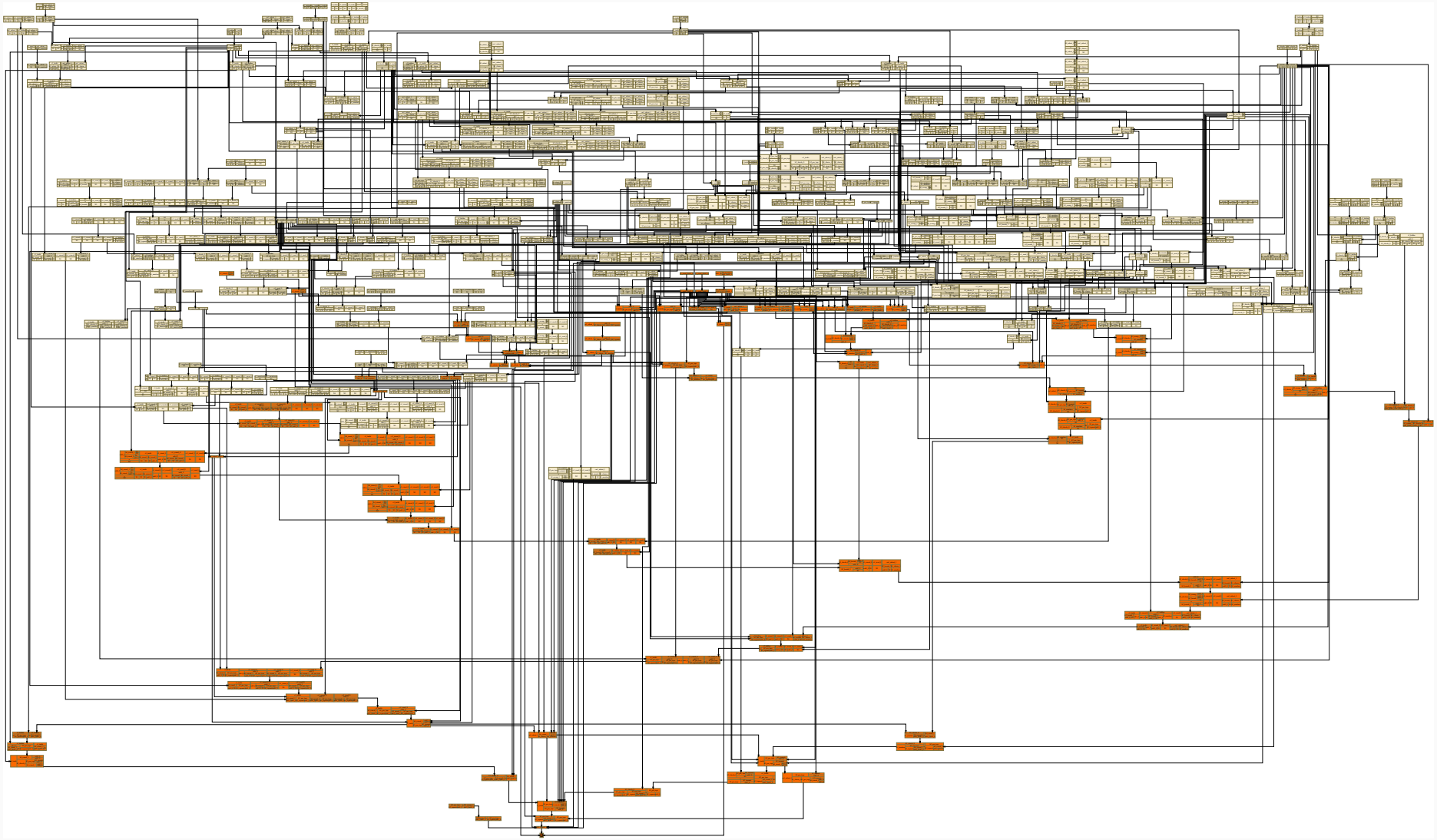
Feedback prove/learn loop for ENIGMA on Mizar data

- Done on 57880 Mizar problems recently
- Serious ML-guidance breakthrough applied to the best ATPs
- Ultimately a 70% improvement over the original strategy in 2019
- From 14933 proofs to 25397 proofs (all 10s CPU - no cheating)
- Went up to 40k in more iterations and 60s time in 2020
- 75% of the Mizar corpus reached in July 2021 - higher times and many runs: `https://github.com/ai4reason/ATP_Proofs`

	$\mathcal{S}$	$\mathcal{S} \odot \mathcal{M}_9^0$	$\mathcal{S} \oplus \mathcal{M}_9^0$	$\mathcal{S} \odot \mathcal{M}_9^1$	$\mathcal{S} \oplus \mathcal{M}_9^1$	$\mathcal{S} \odot \mathcal{M}_9^2$	$\mathcal{S} \oplus \mathcal{M}_9^2$	$\mathcal{S} \odot \mathcal{M}_9^3$	$\mathcal{S} \oplus \mathcal{M}_9^3$
solved	**14933**	16574	20366	21564	22839	22413	23467	22910	23753
$\mathcal{S}$%	+0%	+10.5%	+35.8%	+43.8%	+52.3%	+49.4%	+56.5%	+52.8%	+58.4
$\mathcal{S}+$	+0	+4364	+6215	+7774	+8414	+8407	+8964	+8822	+9274
$\mathcal{S}-$	-0	-2723	-782	-1143	-508	-927	-430	-845	-454

	$\mathcal{S} \odot \mathcal{M}_{12}^3$	$\mathcal{S} \oplus \mathcal{M}_{12}^3$	$\mathcal{S} \odot \mathcal{M}_{16}^3$	$\mathcal{S} \oplus \mathcal{M}_{16}^3$
solved	24159	24701	25100	**25397**
$\mathcal{S}$%	+61.1%	+64.8%	+68.0%	**+70.0%**
$\mathcal{S}+$	+9761	+10063	+10476	+10647
$\mathcal{S}-$	-535	-295	-309	-183

Can you do this in 4 minutes? (359-step ATP proof)



Can you do this in 4 minutes? (human-written code)

```
theorem Th31: :: BORSUK_5:31
 for A being Subset of R^1
 for a, b being real number st a < b & A = RAT (a,b) holds
 Cl A = [.a,b.]
proof
 let A be Subset of R^1; :: thesis:
 let a, b be real number ; :: thesis:
 assume that
 A1: a < b and
 A2: A = RAT (a,b) ; :: thesis:
 reconsider ab = ].a,b.[, RT = RAT as Subset of R^1 by NUMBERS:12, TOPMETR:17;
 reconsider RR = RAT /\ ].a,b.[ as Subset of R^1 by TOPMETR:17;
 A3: the carrier of R^1 /\ (Cl ab) = Cl ab by XBOOLE_1:28;
 A4: Cl RR c= (Cl RT) /\ (Cl ab) by PRE_TOPC:21;
 thus Cl A c= [.a,b.] :: according to XBOOLE_0:def 10 :: thesis:
 proof
  let x be set ; :: according to TARSKI:def 3 :: thesis:
  assume x in Cl A ; :: thesis:
  then x in (Cl RT) /\ (Cl ab) by A2, A4;
  then x in the carrier of R^1 /\ (Cl ab) by Th15;
  hence x in [.a,b.] by A1, A3, Th16; :: thesis:
 end;
 thus [.a,b.] c= Cl A :: thesis:
 proof
  let x be set ; :: according to TARSKI:def 3 :: thesis:
  assume A5: x in [.a,b.] ; :: thesis:
  then reconsider p = x as Element of RealSpace by METRIC_1:def 13;
  A6: a <= p by A5, XXREAL_1:1;
  A7: p <= b by A5, XXREAL_1:1;
  per cases by A7, XXREAL_0:1;
   suppose A8: p < b ; :: thesis:
    now :: thesis:
     let r be real number ; :: thesis:
     reconsider pp = p + r as Element of RealSpace by METRIC_1:def 13, XREAL_0:def 1;
     set pr = min (pp,((p + b) / 2));
     A9: min (pp,((p + b) / 2)) <= (p + b) / 2 by XXREAL_0:17;
     assume A10: r > 0 ; :: thesis:
     p < min (pp,((p + b) / 2))
     proof
      per cases by XXREAL_0:15;
       suppose min (pp,((p + b) / 2)) = pp ; :: thesis:
        hence p < min (pp,((p + b) / 2)) by A10, XREAL_1:29; :: thesis:
       end;
       suppose min (pp,((p + b) / 2)) = (p + b) / 2 ; :: thesis:
        hence p < min (pp,((p + b) / 2)) by A8, XREAL_1:226; :: thesis:
       end;
      end;
     end;
     then consider Q being rational number such that
     A11: p < Q and
     A12: Q < min (pp,((p + b) / 2)) by RAT_1:7;
     (p + b) / 2 < b by A8, XREAL_1:226;
     then min (pp,((p + b) / 2)) < b by A9, XXREAL_0:2;
     then A13: Q < b by A12, XXREAL_0:2;
     min (pp,((p + b) / 2)) <= pp by XXREAL_0:17;
     then A14: (min (pp,((p + b) / 2))) - p <= pp - p by XREAL_1:9;
     reconsider P = Q as Element of RealSpace by METRIC_1:def 13, XREAL_0:def 1;
     P - p < (min (pp,((p + b) / 2))) - p by A12, XREAL_1:9;
     then P - p < r by A14, XXREAL_0:2;
     then dist (p,P) < r by A11, Th14;
     then A15: P in Ball (p,r) by METRIC_1:11;
     a < Q by A6, A11, XXREAL_0:2;
     then A16: Q in ].a,b.[ by A13, XXREAL_1:4;
     Q in RAT by RAT_1:def 2;
     then Q in A by A2, A16, XBOOLE_0:def 4;
     hence Ball (p,r) meets A by A15, XBOOLE_0:3; :: thesis:
    end;
    hence x in Cl A by GOBOARD6:92, TOPMETR:def 6; :: thesis:
```

ENIGMA - The Rise of Computronium (2021)

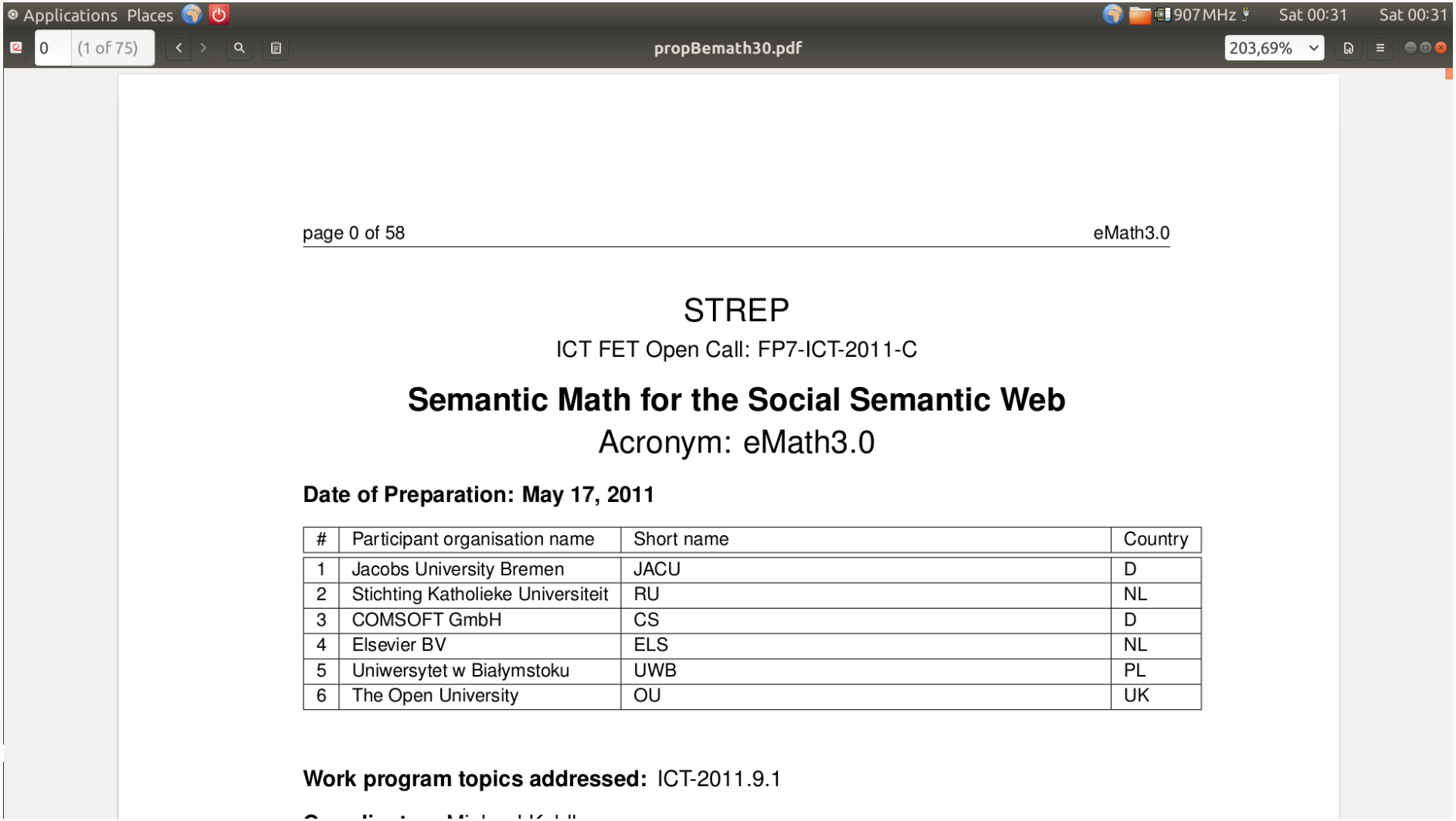
From: Josef Urban <josef.urban@gmail.com>, Date: Jul 26, 2021 at 9:47
Subject: ENIGMA - The Rise of Computronium

I am happy to announce that the ENIGMA system of the E lineage, helped by its Deepire Vampiric cousin, has reached today (July 26th, 2021) the landmark of 75% automatically proved Mizar top-level problems.
[..]
Many of the proofs show that ENIGMA has autonomously (without any human-programmed decision procedures, tactics, and/or dataset preparation/tuning) learned how to routinely perform common mathematical algorithmic tasks such as numeric calculation, matrix manipulation, boolean algebra, integration and differentiation, sequences of standard rewriting and normalization operations in various algebraic theories, etc., and combine them with other reasoning tasks needed for completing the fully formal proofs.
[..]
This suggests that learned guidance combined with efficient search may in near future lead to a new fully declarative problem-solving computing/reasoning architectures applicable to arbitrary computing/reasoning problems without any human engineering.
[..]
Combined with autoformalization this could lead us to large deployment of reasoning machines in science, following Leibniz's Calculemus dream.

MathWikis, eMath, etc

- Wikis were in the air in early 2000s (Wikipedia in 2001)
- The wiki way - incomplete is better than nothing!
- A huge success - Wikipedia perhaps the biggest internet achievement
- Many formal math started to think how to leverage that
- Big MathWiki EU proposal around 2005? (didn't work)
- (I ended up in NL in 2009 thanks to Cezary's MathWiki project)
- The ambitious eMath30 EU proposal led by Michael around 2010 ?
- Flexiformalization and all that: Christoph Lange, Joe Cornelli, ...

eMath30



page 0 of 58 eMath3.0

STREP

ICT FET Open Call: FP7-ICT-2011-C

Semantic Math for the Social Semantic Web

Acronym: eMath3.0

Date of Preparation: May 17, 2011

#	Participant organisation name	Short name	Country
1	Jacobs University Bremen	JACU	D
2	Stichting Katholieke Universiteit	RU	NL
3	COMSOFT GmbH	CS	D
4	Elsevier BV	ELS	NL
5	Uniwersytet w Białymstoku	UWB	PL
6	The Open University	OU	UK

Work program topics addressed: ICT-2011.9.1

eMath30 - early autoformalization futurism

- Tool-chains for computer-assisted/automatic semantic annotation
- (semi-)automated semantization of latex documents.
- (aligned) corpora for evaluation and training of the semantization methods.
- mathematical semantic context for the existing linguistic methods
- integrating them with mathematical semantic pruning and disambiguation methods.
- combining the analysis of natural language with the analysis of mathematics
- large-scale deployment PlanetMath, Wikipedia, arXiv
- TODAY: autoformalization is all the rage, lots of money, DL teams, AGI utopias ...
- people are attacking this in all sorts of ways ...

Side Note: MML Licensing with Michael et al

Licensing the Mizar Mathematical Library

Jesse Alama[1], Michael Kohlhase[2], Lionel Mamane, Adam Naumowicz[3], Piotr Rudnicki[4], and Josef Urban[5,⋆]

[1] Center for Artificial Intelligence, New University of Lisbon
j.alama@fct.unl.pt
[2] Computer Science, Jacobs University
m.kohlhase@jacobs-university.de, lionel@mamane.lu
[3] Institute of Computer Science, University of Bialystok
adamn@math.uwb.edu.pl
[4] Department of Computing Science, University of Alberta
piotr@cs.ualberta.ca
[5] Institute for Computing and Information Sciences; Radboud University Nijmegen
josef.urban@gmail.com

Abstract. The Mizar Mathematical Library (MML) is a large corpus of formalised mathematical knowledge. It has been constructed over the course of many years by a large number of authors and maintainers. Yet the legal status of these efforts of the Mizar community has never been clarified. In 2010, after many years of loose deliberations, the community decided to investigate the issue of licensing the content of the MML, thereby clarifying and crystallizing the status of the texts, the text's authors, and the library's long-term maintainers. The community has settled on a copyright and license policy that suits the peculiar features of Mizar and its community. In this paper we discuss the copyright and license solutions. We offer our experience in the hopes that the communities of other libraries of formalised mathematical knowledge might take up the legal and scientific problems that we addressed for Mizar.

Keywords: free culture, open data, free licensing, formal mathematics, mizar

Side Note: MML Licensing with Michael et al

2.2 Two Aspects of Formal Mathematics: Code and Text

Formal mathematical libraries present a number of problems for the license designer. One crucial question for deciding upon a license for the Mizar Mathematical Library: Is a Mizar article more like a journal article, or is it more like a piece of computer code, or is it both? And could existing law be suboptimal in treating code differently from mathematical texts?

Although we are focused on Mizar, we note in passing that other interactive proof assistants, whose communities might want to take up the problem of licensing their content, face this issue in different ways. Definitions and proofs in Coq, for instance, have originally been rather more like computer programs (at least, prima facie) than Mizar texts[4].

We also need to address the issue of what can be done with formal mathematical texts. There is considerable interest in extracting algorithms from proofs of universal-existential theorems. What is the status of such extracted or derived products? Mizar does not, on its face, permit such a straightforward, immediate extraction of algorithms from proofs. There are however several mechanisms which bring Mizar very close to executable code:

1. Formal mathematical formulas (more precisely: clauses) immediately give rise to very real computation in the Prolog language. For example, the Prolog algorithm for reversing lists:

[4] This is however also changing: Coq has become capable of handling mainstream (not just constructive) mathematics recently, has gained a declarative proof mode and one of the aspirations of the Math Components project is to make the Coq presentations accessible to mathematicians.

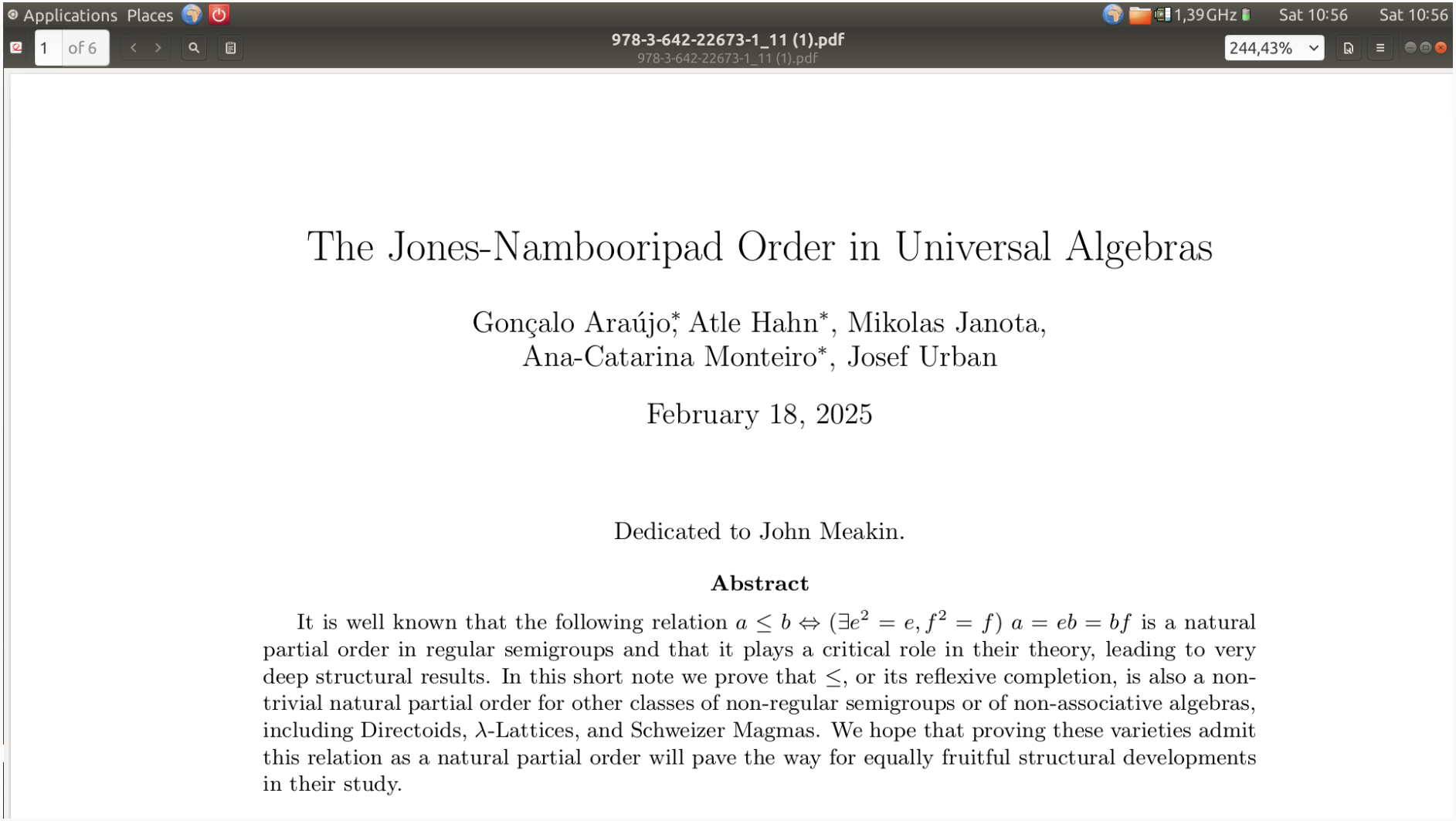
Side Note: MML Licensing with Michael et al

- Problem: Wiki for the Mizar Math Library vs its unclear licence
- Finally the critical mass to seriously think about a good license
- But what is formal math???
- Text? Code? Both?
- GPL? (Linux) Apache, MIT (code) GFDL? CC-BY-SA? (Wikipedia)
- Lots of discussions, also with the FSF people
- Ultimately both GPL-ed as Linux and CC-BY-SA as Wikipedia
- We argued that formal math is both code and text
- TODAY: very relevant in the LLM world - all text is easily turned into code
- Our quote from the paper: *"With sufficiently advanced extraction algorithms (which we are clearly approaching), many documents can become programs."*

LaTeXML etc

- I was impressed by the ArxMLiv project run by Michael et al during my 2010 visit to Bremen
- We used LaTeXML several times in the MathWiki project at RU
- Also for our first autoformalization attempts over Flyspeck with Cezary
- Very recently (2023), we used the ArxMLive dump again:
- And we even got an “automated” math paper out of it
- Use LLMs together with LateXML parsing to mine Arxiv for axioms of algebras
- Automatically (using Prover9) try to prove various properties for various algebras

LaTeXML+LLMs+Prover9 = Math Paper Machine



The Jones-Nambooripad Order in Universal Algebras

Gonçalo Araújo*, Atle Hahn*, Mikolas Janota,
Ana-Catarina Monteiro*, Josef Urban

February 18, 2025

Dedicated to John Meakin.

Abstract

It is well known that the following relation $a \leq b \Leftrightarrow (\exists e^2 = e, f^2 = f)\ a = eb = bf$ is a natural partial order in regular semigroups and that it plays a critical role in their theory, leading to very deep structural results. In this short note we prove that $\leq$, or its reflexive completion, is also a non-trivial natural partial order for other classes of non-regular semigroups or of non-associative algebras, including Directoids, λ-Lattices, and Schweizer Magmas. We hope that proving these varieties admit this relation as a natural partial order will pave the way for equally fruitful structural developments in their study.

Further AI/TP topics

- AITP started with Cezary in 2016 (10th anniversary this year!)
- Michael in our first AITP PC!
- Michael (+ Dennis, Florian, Deyan, Cezary): flexi/auto-formalization, alignments, parsing OEIS
- Interesting interactions with the deep-learning maximalists at AITP :-)
- I hope we'll get of more of Michael & Co at AITP
- Lots of other topics in Bonn last year at the Formal trimester (with Michael as a great boss!)

Btw, have you seen our OEIS experiment?

- Apologies to Michael for sometime forgetting their prior work on OEIS conjecturing
- We have been now running our feedback loop for OEIS program invention for 2+ years
- Search-Verify-Learn and repeat
- Starting from scratch (no LLM cheats), now over 800 iterations

Search/Check/Learn feedback loop on OEIS

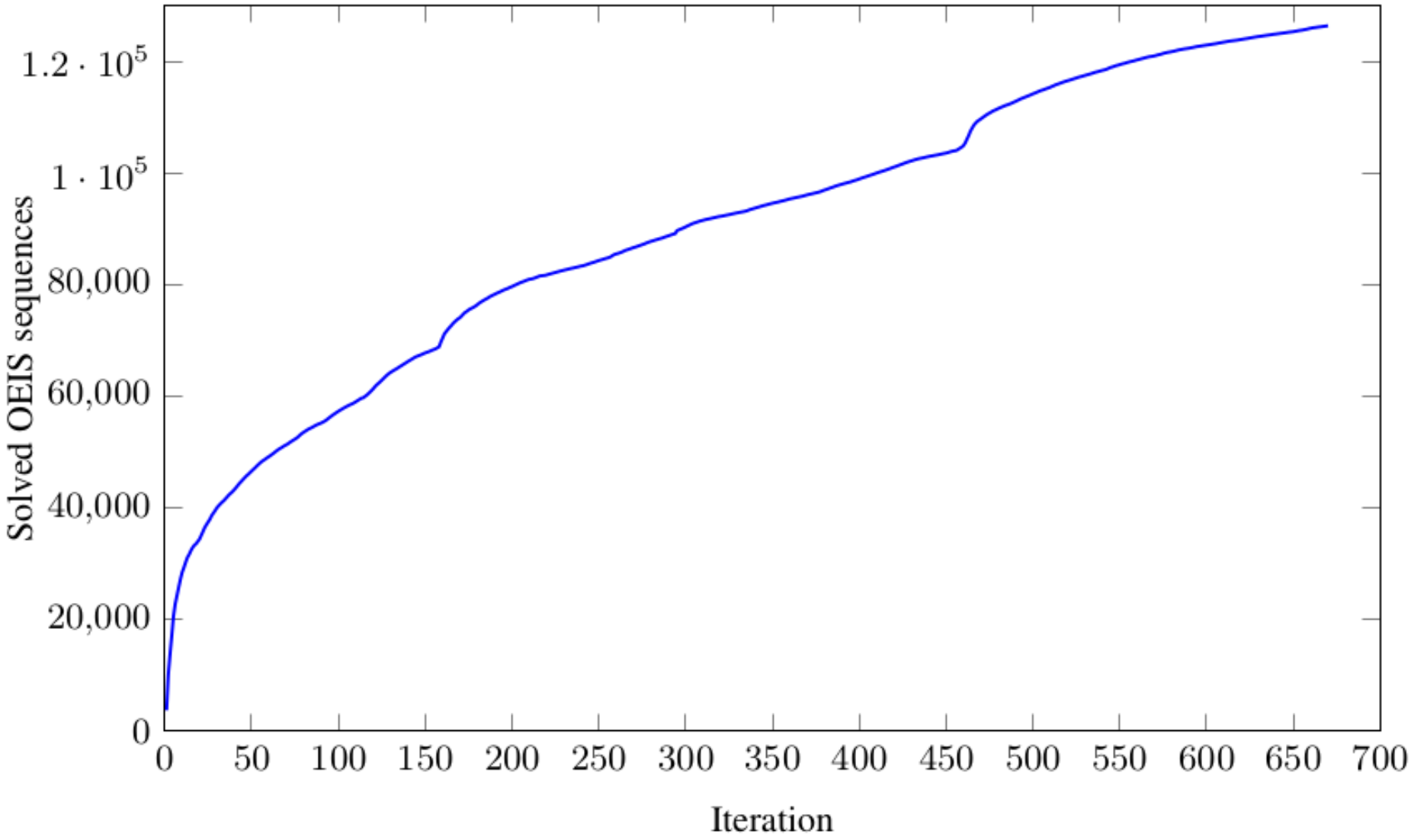
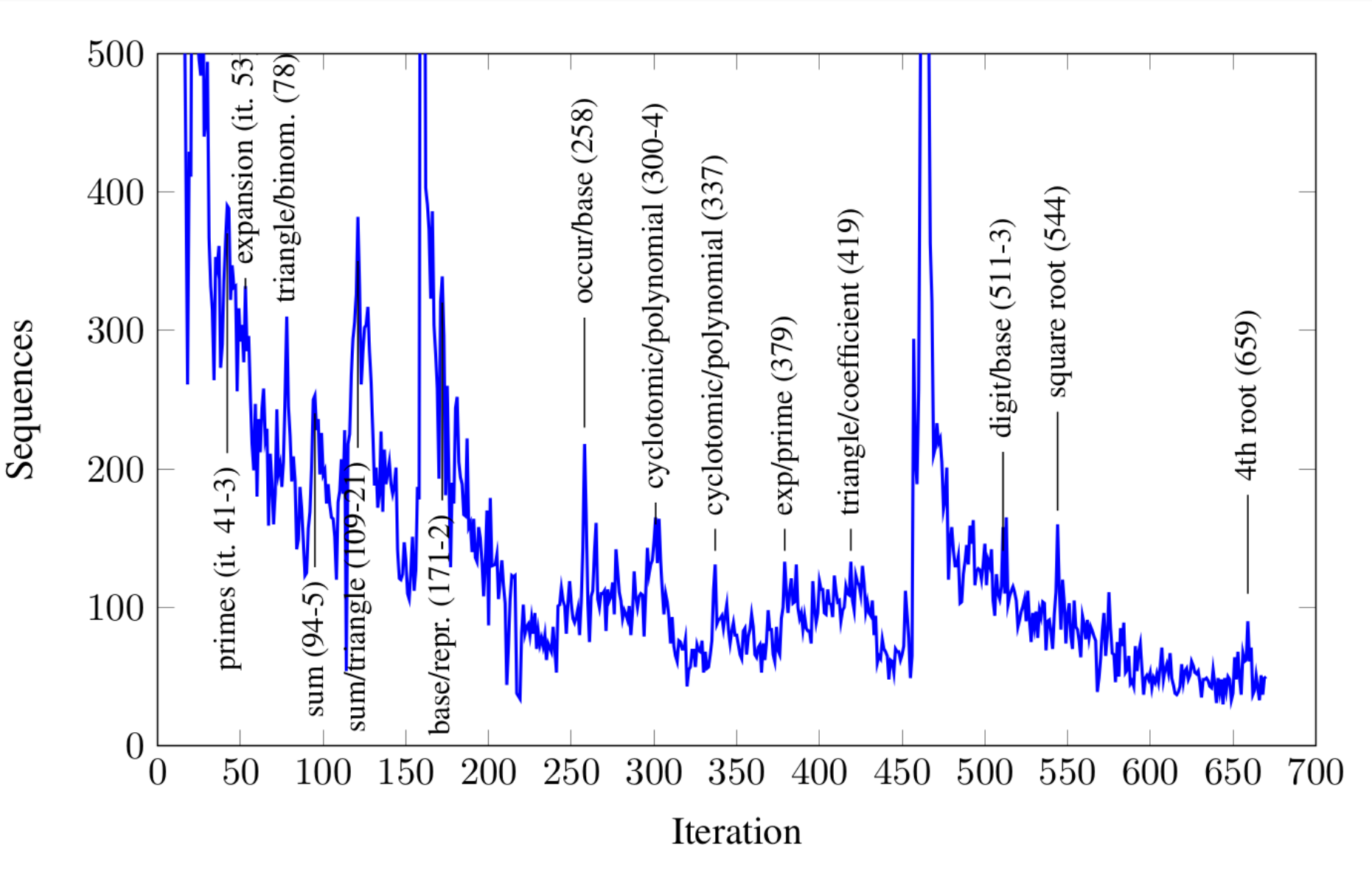
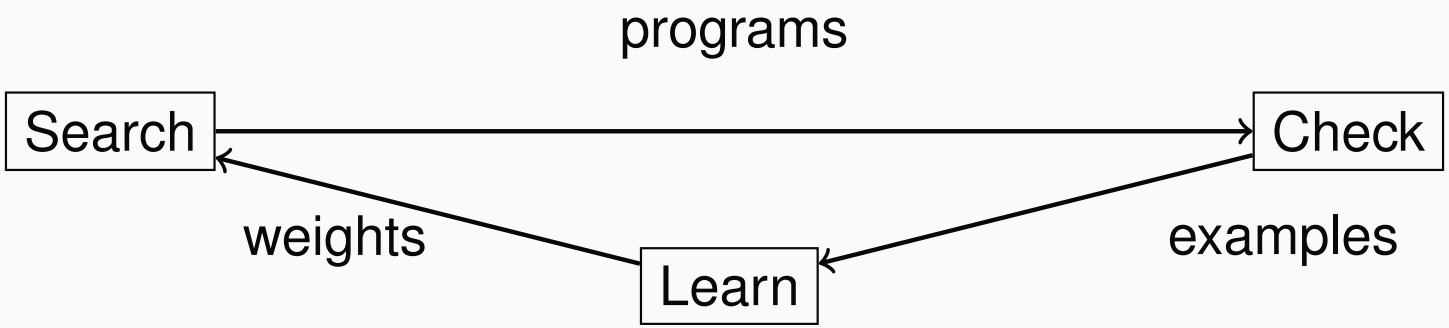


Figure 12: Number y of solved OEIS sequences after x iterations

Some Automatic Technology Jumps



Search-Verify-Train Positive Feedback Loop (OEIS)



- Small Turing-complete DSL for our programs, e.g.:
  $2^{\mathbf{x}} = \prod_{y=1}^{x} 2 = loop(2 \times x, \mathbf{x}, 1)$
  $\mathbf{x}! = \prod_{y=1}^{x} y = loop(y \times x, \mathbf{x}, 1)$
- Analogous to our Prove/Learn feedback loops in learning-guided proving (since 2006 – Machine Learner for Automated Reasoning – MaLARea))
- However, OEIS allows much faster feedback on *symbolic conjecturing*
- 670 iterations and still refuses to plateau - counters RL wisdom?
- Since it interleaves symbolic breakthroughs and statistical learning?
- Cheap: The electricity bill is only $1k-$3k, you can do this at home
- ~4.5M explanations invented: 50+ different characterizations of primes

Some Invented Explanations

- `https://oeis.org/A4578`: Expansion of sqrt(8) in base 3:
  loop2(((y * y) div (x + y)) + y, y, x + x, 2, loop((1 + 2) * x, x, 2)) mod (1 + 2)
- `https://oeis.org/A4001`: Hofstadter-Conway $10k seq: a(n) = a(a(n-1)) + a(n-a(n-1)) with a(1) = a(2) = 1:
  loop(push(loop(pop(x), y-x,pop(x)),x) + loop(pop(x), x-1, x), x - 1, 1)
- `https://oeis.org/A40`: prime numbers:
  2 + compr((loop(x * y, x, 2) + x) mod (2 + x), x)
- `https://oeis.org/A30184`: Expand $\eta(q) * \eta(q^3) * \eta(q^5) * \eta(q^{15})$ in powers of q (elliptic curves):
  loop(push(loop((pop(x) * loop(if (pop(x) mod y) <= 0 then (x - loop(if (x mod (1 + (y + y))) <= 0 then (x + x) else x, 2, y)) else x, y, push(0, y))) + x, y, push(0, x)), x) div y, x, 1)
- `https://oeis.org/A51023`: Wolfram's $30k Rule 30 automaton:
  loop2(y, y div 2, x, 1, loop2(loop2((((y div (0 - (2 + 2))) mod 2) + x) + x, y div 2, y, 1, loop2(((y mod 2) + x) + x, y div 2, y, 1, x)), 2 + y, x, 0, 1)) mod 2
- `https://oeis.org/A2580`: $\sqrt[3]{2}$ Hales's blog: `https://t.ly/tHs1d`

Thanks, Michael!

Thanks Michael for the 20+ years of interactions!
and Happy Birthday! and Have Fun at all the Gates of ...!