Women in Logic Online & Network Mathematics

Valeria de Paiva
Topos Institute, Berkeley

March 20, 2024

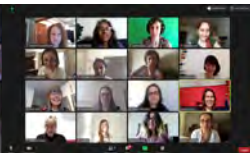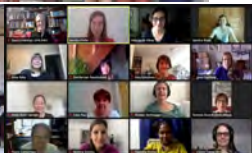# Thank you Agata and Josephine!

Outline

- Women in Logic: why, what, and how?
- Network Mathematics
- Research Prototypes:
    - Parmesan
    - MathAnnotator
    - MathGloss
- What's Next?

Women in Logic Collective

WiL Workshops

Our flagship event: Women in Logic workshop

- This year is our 8th WiL event, in Tallinn, Estonia!
- The first WiL happened in Reykjavik, Iceland @LiCS in 2017
- 2018 Oxford, UK @ FLoC
- 2019 Vancouver, CA @ LiCS
- 2020 Paris, FR @Summer of LOVE Online
- 2021 Rome, IT @LiCS, Online
- 2022 Haifa, IS @FLoC
- 2023 Rome, IT @FSCD

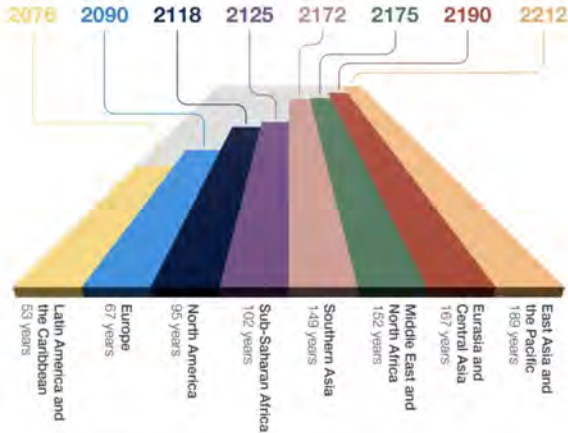website https://www.womeninlogic.org/

Why Women in Logic?

- World reasons: It may take close to 300 years to achieve full gender equality (UNWomen, `https://bit.ly/3IHN1Pi`)

- The gender gap is real. It will take 132 years (WEF `https://bit.ly/4af9vCQ`) to to achieve equality. We are getting worse — in 2020, we needed 100 years to close the gender gap. We've lost a whole generation.

- Education is closer to equality, but the majority of children who are out of school are girls

The world

Global Gender Gap Index 2023

At current pace, when are the regions likely to close the gender gap?

WORLD ECONOMIC FORUM

2076 | 2090 | 2118 | 2125 | 2172 | 2175 | 2190 | 2212

Latin America and the Caribbean — 68 years
Europe — 67 years
North America — 95 years
Sub-Saharan Africa — 102 years
Southern Asia — 149 years
Middle East and North Africa — 152 years
Eurasia and Central Asia — 167 years
East Asia and the Pacific — 189 years
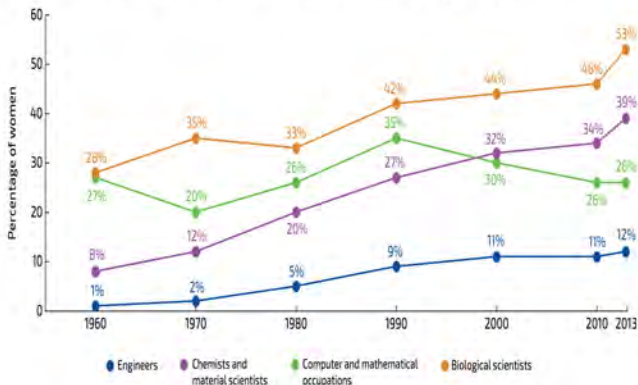
Why Women in Logic?

- Labour markets of the future: Women make up almost half (49.3%) of total employment across non-STEM occupations, but just 29.2% of all STEM workers.
- Female STEM graduates entering STEM employment increasing with every cohort. But no retention!
- Women currently account for 29.4% of entry-level STEM workers; yet for leadership roles such as VP and C-suite, representation drops to 17.8% and 12.4%
- Women working in AI approximately 30%, roughly 4 percentage points higher than it was in 2016, while total workers increased six times between 2016 and 2022 (WEF https://bit.ly/3TmRzzr)

Why Women in Logic?

"Solving the equation: The Variables for Women's Success in Engineering and Computing" C. Corbett, C. Hill, 2015

FIGURE 1. WOMEN IN SELECTED STEM OCCUPATIONS, 1960-2013

Why Women in Logic?

US gov Employment report 2022-2032
https://www.bls.gov/news.release/ecopro.nr0.htm

> The computer and mathematical occupational group is
> projected to experience much faster than average employment
> growth. The expanding digital presence of businesses and
> consumers will produce a massive amount of data that these
> workers collect, organize, analyze, and protect. Employment
> in occupations such as data scientists, statisticians, and
> information security analysts is projected to grow more than
> 11 times the average rate.

'Leaking pipeline' needs to be addressed!

Our community

In 2013 Orna Kupfermann first woman to chair LiCS had some dismaying news

Women at LiCS (Logic in Computer Science),
Data from 2013, thanks to Orna Kupferman!

Chairs: 27 men (96%), 1 women (4%)
Organizing committee: 21 men (81%), 5 women (19%)
Advisory board: 15 men (100%), 0 woman (0%)
Typical PC: 30 men (88%), 4 women (12%)
Invited speakers MFPS-LICS-CSF: 6 men (86%), 1 woman (14%)
[with special session speakers: 13 men, 1 woman].
Tot Awards: 8 men (100%), 0 women (0%)
Tot awards ever: 41 men (95%), 2 women (5%)
Kleene awards ever: 18 men (95%), 1 women (5%)

Authors of accepted papers last year: 5% women.
Submissions this year: 445 men (96.3%), 17 women (3.7%)

Beginnings

2016 Celebration of 30 years of LiCS, New York.

Impromptu talk

#MeToo movement

What can we do?

• One day workshops to celebrate women, by talking about our work in Logic!

• Two Invited Speakers and an open call for contributed short talks

• Co-located with LiCS, FLoC, and FSCD

8 years, big growth during pandemic, new challenges

Steering committee

Organizers

WiL 2024 organizers: Daniele Nantes, Tephilla Prince and Elaine Pimentel

Many great Invited Speakers and PC members!

Find our Blog, Facebook Group, Data Collection, Directories, Twitter/X, mailing list, and Slack.

Join us!

Better math for society

Jeff Dean (@🖼) ✓
@JeffDean

I signed this letter along with many others. Math education is clearly deeply important, for AI and more generally.

Jelani Nelson ✓ @minilek · Mar 6

@elonmusk and @sama may not agree on much of late, but do agree AI is built on strong math foundations, including algebra and calculus, applauding @UofCalifornia for recent clarifications on math requirements for admission.

Many industry leaders signed:...
Show more

Strong Math Foundations
are Important for AI

Better math for practitioners

140,000

Growth of Mathematics Literature
1940–2017

120,000

100,000

80,000

60,000

40,000

20,000

0

1930 1940 1950 1960 1970 1980 1990 2000 2010 2020 2030

Journals   All
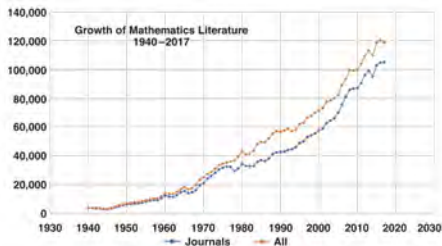
E. Dunne, "Looking at the mathematics literature", Notices of the American Mathematical Society, vol. 68, no. 2 pp. 171–175, 2019. ams.org/journals/notices/202002/rnoti-p171.pdf

- Need infrastructure for organizing and communicating results
- Avoid wasted search time, duplication of research, and missed connections between fields (Quanta Magazine Nov2019)
- AI trained to read science papers can predict future discoveries, Nature (2019).
- Want to do same for Math literature. How?

Network Mathematics

How do we encode, connect, and share knowledge in the 21st Century?

- Make math **more accessible** to everyone: students, mathematicians, scientists, interested laypeople
- Making use of incredibly fast-improving NLP tools
- Extract math concepts from texts & link to ontologies
- Concepts build a Math **Knowledge Graph**
- Also useful for formalization in proof assistants/theorem provers (AI to Assist Mathematical Reasoning, National Academies, June 2023)

Reading Mathematics

When you read some mathematics you're not familiar with...

We define the notion of a (torsor) for an (inverse semigroup), which is based on (semigroup actions), and prove that this is precisely the structure classified by the (topos) associated with an (inverse semigroup). Unlike in the (group) case, not all (set-theoretic torsors) are (isomorphic): we shall give a complete description of the (category of torsors)...

Can we create a personal *index* of things to check automatically on the fly?

Network Mathematics

- To improve the organization and dissemination of math via tools as search engines, knowledge graphs, glossaries, etc
- **Parmesan** (with Collard and Subrahmanian, NIST) two preprints on mathematical concept extraction for Education, arxiv 2307.06699, 2208.13830, two workshop pres, recently accepted for COLING/LREC 2024
  https://bit.ly/ParmesanMath
- **MathChat** (Gao, Kovalev, Moss), using chatGPT/GPT4 to extract terms, arXiv:2309.00642
  https://bit.ly/math_annotator
- **MathGloss** (Lucy Horowitz, Chicago), connecting open source electronic glossaries on the web
  https://mathgloss.github.io/MathGloss/database
- other projects starting MathNLI, MathKG, Math-causality

Extracting mathematical concepts

- Want to use AI/NLP tools off-the-shelf. Need textual corpora.
- Construct corpora and find tools for extracting math concepts.
- What linguists have for "extracting math concepts"?
- Terminology extraction for identifying fundamental mathematical concepts can be:
  Named Entity Recognition (NER),
  Keyword Extraction (KE)
  or Definition Extraction (DE)
- No ATE system dedicated to mathematics
- but a collection of rule-based and neural network systems to begin with

Mathematical Corpora

- Corpora prepared, 755 abstracts from the open source journal *Theory and Application of Categories*, around 2020, at https://github.com/ToposInstitute/tac-corpus.

- The abstracts were processed with the BERT version of spaCy, from 2022.

- We also use the corpus of nLab entries, already processed with spaCy and available at https://github.com/ToposInstitute/nlab-corpus.

- Use 'sanitized' sentences, i.e. sentences of medium length and with no LATEX, after LateXML (Miller, NIST).

Baselines and Gold Standards

- Checking baselines is HARD, as they do NOT exist.
- No human annotated/golden standard datasets of math concepts
- there was an ACL 2017 Competition on **scientific** concepts. one small golden dataset, later expanded to a SciERC corpus. little transfer?
- Results with *silversets* (Collard et al 2022) very low
- Video https://youtu.be/-ZhZjMn1Zpk

Parmesan: mathematical concept extraction for education

- A different tackle: let users ask for concepts! in
  https://arxiv.org/abs/2307.06699.
- Parmesan does entity linking and collocation search for CT,
  from a human learner perspective
- uses LaTeXML converter (NIST) and spaCy
- Term Extractors used:
  DyGIE++ (SOTA 2019),
  TextRank, mwetoolkit,
  SpERT.PL(2021) and PL-Marker (2022)
- Last two recent neural term extractors, trained for science.
  little transfer, see tables in preprint.
- Definition extraction using linguistic tools

Parmesan demo: (to appear LREC 2024)
http://www.jacobcollard.com/parmesan2/

PARsing Math Entities Search And Navigation (Parmesan)

Generative AI for Mathematical Knowledge (Moss et al)

- Still want a **WordNet for mathematics**, new team with three mathematicians

- Instead of using an ATE system we want to see how chatGPT/GPT4 can do in the task of extracting concepts

- Concepts can produce indices for monographs, books, and self-study more easily

- Make sure the concepts extracted are correct and as complete as possible

- Understand opportunities and limitations of large language models – Old English has similar problems with lack of golden datasets

- Need to annotate a set toevaluate chatGPT

Math Vernacular

- Many mathematical terms are NOT math concepts:
  characterization, conjecture, consequence, counter-example, paper, etc

- Traditional issues of terminologies:
  - plurals or not? (*presheaves or presheaf*?),
  - adjective without a noun? (*2-categorical version*?),
  - adjective is related to noun/verb, which one?
    (*interpolated function, interpolate, interpolation*?)

- constructions with two or more adjectives, add sub-expressions
  differential graded category, graded category?

- expressions with prepositions *sheaf of germs of analytic funcs*

- proper nouns as possessives *Shanin's method, Lagrange interpolation*

Annotation guidelines

- Treat math concepts as black-boxes, as much as possible
- Use singular for concepts
- convention: Terms like *theorem, corollary, conjecture, paper* are not math concepts
- convention: Mathematicians' names are not mathematical concepts
- If a long span is a concept, add important subspans e.g. *enriched accessible category, accessible category*
- avoid prepositions inside of concepts *sheaf of germs of analytic functions* we add instead *sheaf, germ, analytic function*

Experimenting with chatGPT

- Three experiments: 100 sentences Pilot, 436TAC, 55KnLab
- for Pilot100 we have three mathematicians and chatGPT is the fourth annotator
- For 436TAC only one annotator and chatGPT, plus adjucation
- For 55KnLab no human evaluation, so far
- Many 'minor mistakes' were made by the **annotators**!
- NLI annotation tool repurposed

chatGPT prompts

Given the following Context, extract the words that denote Math concepts.
Be sure to make the concept words singular. For example when we see 'functors' in a sentence, we would extract 'functor' rather than 'functors' or when we see 'categories' we would like to extract 'category' instead of 'categories'!

Here are some examples:
{*in-context example*}

Also note that we are looking for concepts like modulation, enriched orthogonality, holonomy, localization, variety, but not words shown in daily English sentences like 'future work', 'conclusion', 'this property'!

We don't want a person's name to be extracted as a math concept although we understand that a person's name could be part of the phrase that denotes a math concept.

Now please solve the following problem.

Context: {*math_sentence*}
Concepts:

Table: Prompt template with specialized instructions

Math Concept Annotator

https://gaoq111.github.io/math_concept_annotation/

Contributions

- problems of mathematical term extraction: abstracts in journals possibly not best
- inter-annotator (dis)agreements
- guidelines to standardize the process
- an annotation tool to help humans with ATE
- best practices for prompts to ChatGPT
- discussing whether ChatGPT could be used as an annotator similar to a human expert
- work presented at UI workshop @ CICM2023, arXiv:2309.00642, updates coming!

MathGloss: connecting glossaries

- Many resources, many audiences
    - How can people from different backgrounds make sense of them?
- NYT calls assistants "proof whiners" ("AI is Coming for Mathematics, Too", July 2023)
- How can we help improve communication between mathematicians and computers?
- Organize the various resources for undergraduate-level math
- bridge the language gap between natural math and formal math

https://mathgloss.github.io/MathGloss/database

MathGloss

MathGloss

This table contains Wikidata ID numbers and the corresponding term from each corpus. To see terms from the corpora that were not mapped to Wikidata, click on the links in the headers.

Wikidata ID	Chicago	France UG	nLab	Hosgood
Q14481419		0-1 law		
Q181296	abelian		abelian group	abelian group
Q318598	abelianization		abelianization	
Q20827138	absolute continuity of measure			
Q332504	absolute continuity		absolutely continuous measure	
Q332465		absolute convergence	absolute convergence	absolute convergence
Q120812	absolute value		absolute value (mathematics)	absolute value (mathematics)

Terms from four corpora linking them to Wikidata. So far, 906 terms in total.

Connecting to Lean4

Continuing work

- Expand our processed corpora of mathematical texts, fine tuning the corpora

- Construct open source glossaries of mathematical concepts for all branches of maths (from good sources available)

- Build our knowledge graph for maths and 'complete' it, as much as possible

- Use mathematical English as a lingua franca between different formal systems

Continuing work

- Results over surveyable collections of data on chatGPT/GPT4 are not too good
- With judicious prompting, ChatGPT can find math terms but misses some important ones and 'hallucinates' some (e.g Dress-Mueller Theorem)
- Also it extracts terms that are not "mathematical terms" e.g. *conjecture*
- ChatGPT may be used very easily and gives very acceptable preliminary results and hence it WILL
- Need to measure the quality of these LLMs results

What else?

- Other online databases

- Use NLP to get terms from other types of resources like textbooks:

- Other interactive theorem provers (Coq, Isabelle, Agda) could produce resources like the UG $\Rightarrow$ Lean4.

- Dagstuhl Seminar October 2023, 'Automated Mathematics: Integrating Proofs, Algorithms and Data'
`https://www.dagstuhl.de/en/seminars/`
`seminar-calendar/seminar-details/23401`

- AIM seminar on 'Open-source cyberinfrastructure supporting mathematics research' (December 2023)
`https://aimath.org/pastworkshops/`
`cyberinfrastructure.html`

What's Next?

- Three parallel projects are progressing.
- Each has new milestones
- More corpora
- Golden standards are still part of the plan
- New work with Dumbrova on MathKG – as off-the-shelf as possible
- Several new experiments on MathNLI and (linguistic) causality (not written yet)

Need Funding!

Thanks!

Some References

Collard et al, *Extracting Mathematical Concepts from Text.* In W-NUT 2022, arxiv 2208.13830.

Collard et al, *Parmesan: mathematical concept extraction for education* arxiv 2307.06699.

de Paiva et al, *Extracting Mathematical Concepts with Large Language Models* arXiv:2309.00642

Horowitz, de Paiva, *MathGloss: Linked Undergraduate Math Concepts*, EuroProofNet Workshop 2023.

spaCy Industrial-Strength Natural Language Processing, https://spacy.io/

Universal Dependencies, https://universaldependencies.org/

LateXML, https://math.nist.gov/~BMiller/LaTeXML/