

# MathGloss: Linked Undergraduate Math Concepts

Lucy Horowitz    Valeria de Paiva

EuroProofNet Workshop 2023

September 7, 2023



**TOPOS**  
INSTITUTE

# What is a group?

# What is a group?

According to Wikipedia,

**Definition** [ [edit](#) ]

The axioms for a group are short and natural... Yet somehow hidden behind these axioms is the [monster simple group](#), a huge and extraordinary mathematical object, which appears to rely on numerous bizarre coincidences to exist. The axioms for groups give no obvious hint that anything like this exists.

[Richard Borcherds](#), *Mathematicians: An Outer View of the Inner World*<sup>[4]</sup>

A group is a non-empty [set](#)  $G$  together with a [binary operation](#) on  $G$ , here denoted " $\cdot$ ", that combines any two [elements](#)  $a$  and  $b$  of  $G$  to form an element of  $G$ , denoted  $a \cdot b$ , such that the following three requirements, known as **group axioms**, are satisfied:<sup>[5][6][7][a]</sup>

But before we even get there,  $\mathbb{Z}$ .

# What is a group?

According to Wikipedia,

## Definition [\[ edit \]](#)

The axioms for a group are short and natural... Yet somehow hidden behind these axioms is the [monster simple group](#), a huge and extraordinary mathematical object, which appears to rely on numerous bizarre coincidences to exist. The axioms for groups give no obvious hint that anything like this exists.

[Richard Borcherds](#), *Mathematicians: An Outer View of the Inner World*<sup>[4]</sup>

A group is a non-empty [set](#)  $G$  together with a [binary operation](#) on  $G$ , here denoted " $\cdot$ ", that combines any two [elements](#)  $a$  and  $b$  of  $G$  to form an element of  $G$ , denoted  $a \cdot b$ , such that the following three requirements, known as [group axioms](#), are satisfied:<sup>[5][6][7][a]</sup>

But before we even get there,  $\mathbb{Z}$ .

According to the nLab,

## 1. Definition

Classically, a [group](#) is a [monoid](#) in which every element has an [inverse](#) (necessarily unique). When written with a view toward [group objects](#) (see [Internalization](#) below), one should rather say that a group is a monoid together with an inversion operation.

An [abelian group](#) is a group in which moreover the order in which two elements are multiplied is irrelevant.

## 2. Delooping

To some extent, a group "is" a [groupoid](#) with a single object, or more precisely a [pointed](#) groupoid with a single object.

# What is a group?

According to Lean,

# The Situation

# The Situation

Many resources, many audiences

# The Situation

Many resources, many audiences

How can people from different backgrounds make sense of them?



# The Situation

Many resources, many audiences

How can people from different backgrounds make sense of them?

NYT calls assistants "proof whiners" for not understanding definitions ("AI is Coming for Mathematics, Too")

# The Situation

Many resources, many audiences

How can people from different backgrounds make sense of them?

NYT calls assistants "proof whiners" for not understanding definitions ("AI is Coming for Mathematics, Too")

How can we improve communication between mathematician and computer?

Organize the various resources for undergraduate-level math

Organize the various resources for undergraduate-level math  
Use this organization to:

Organize the various resources for undergraduate-level math

Use this organization to:

make math more accessible to everyone: students,  
non-mathematician scientists, "laypeople," etc.

Organize the various resources for undergraduate-level math

Use this organization to:

- make math more accessible to everyone: students, non-mathematician scientists, "laypeople," etc.

- bridge the language gap between natural math and formal

## Solution: MathGloss

Our table, combining terms from four (more to come!) different corpora and linking them to Wikidata. So far, 906 terms in total.

# Example



# Example

(almost) all notes  
from my math major  
@UChicago, atomized  
and linked

Theorems are not  
yet included in  
MathGloss

758 definitions  
total, 500 mapped  
manually to Wikidata

Concepts the French government expects its undergrads to know, organized by subject

Terms are translated into English

Where possible, a pointer is given to the corresponding entry in Lean

543 terms total, 369 mapped to Wikidata with wikimapper

Pagetitles from the nLab,  
a higher-math wiki, minus  
people

18k+ pages, 5377 mapped  
to Wikidata with  
Wikimapper

We restrict to those terms  
which are in other corpora

Similar to MathGloss but  
with cross-language rather  
than cross-resource linking

305 terms in total, some  
languages don't have their  
word included for every  
concept

Terms are already mapped  
to Wikidata



A simple Python package that takes a page title and spits out a Wikidata ID

A simple Python package that takes a page title and spits out a Wikidata ID

But there is the common problem of math words having the same spelling as regular English words



A simple Python package that takes a page title and spits out a Wikidata ID

But there is the common problem of math words having the same spelling as regular English words

e.g. group, eld, ring...

A simple Python package that takes a page title and spits out a Wikidata ID

But there is the common problem of math words having the same spelling as regular English words

e.g. group, eld, ring...

wikimapper often maps to "disambiguation" pages, when Wikipedia itself knows that there are multiple senses of the same word

A simple Python package that takes a page title and spits out a Wikidata ID

But there is the common problem of math words having the same spelling as regular English words

e.g. group, eld, ring...

wikimapper often maps to "disambiguation" pages, when Wikipedia itself knows that there are multiple senses of the same word

We circumvent this by using Wikipedia's system: add (mathematics) or (topology) etc. to terms, and then map

# Where else?

# Where else?

Other online databases

# Where else?

Other online databases

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

## Where else?

Other online databases

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

TAC stands for "Theory and Applications of Categories," a journal in category theory

# Where else?

Other online databases

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

TAC stands for "Theory and Applications of Categories," a journal in category theory

We took the 755 abstracts from papers and processed the text using spaCy (an NLP library for Python) and extracted "concepts:"

nouns, compounds, adjective-noun phrases  
selected for frequency of appearance



## Where else?

Other online databases

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

TAC stands for "Theory and Applications of Categories," a journal in category theory

We took the 755 abstracts from papers and processed the text using spaCy (an NLP library for Python) and extracted "concepts:"

nouns, compounds, adjective-noun phrases  
selected for frequency of appearance

Not included in this iteration of MathGloss

# Future work

Pick up spaCy again, along with the problem of "definition extraction" from open-source textbooks

Pick up spaCy again, along with the problem of "definition extraction" from open-source textbooks

Expand to more online resources with even more viewpoints

Pick up spaCy again, along with the problem of "definition extraction" from open-source textbooks

Expand to more online resources with even more viewpoints

Make the site a bit easier to navigate!

- Pick up spaCy again, along with the problem of "definition extraction" from open-source textbooks
- Expand to more online resources with even more viewpoints
- Make the site a bit easier to navigate!
- Make the mappings easy to update

## Some References

(see <https://mathgloss.github.io/MathGloss/>)



The nLab: [ncatlab.org](https://ncatlab.org)



Undergraduate mathematics in mathlib:  
<https://leanprover-community.github.io/undergrad.html>



Tim Hosgood's Dictionary: <https://thosgood.com/maths-dictionary/>



wikimapper: <https://github.com/jcklie/wikimapper>



Roberts, Siobhan. "AI Is Coming for Mathematics, Too":  
<https://www.nytimes.com/2023/07/02/science/ai-mathematics-machine-learning.html>