math.wikipedia.org: A vision for a collaborative semi-formal, language independent math(s) encyclopedia

J. Corneli and M. Schubotz

[1] Goldsmiths, University of London, London, UK
j.corneli@gold.ac.uk
[2] Universität Konstanz, Konstanz, Germany
moritz.schubotz@uni.kn

1 What has been done?

Wikipedia, the world's largest encyclopedia, treats large portions of mathematics. The Mathematics WikiProject indexes 31444 articles at the time of writing.[1] For some sub-fields of STEM, Wikipedia also contains a significant amount of related domain knowledge. Until recently, the main – and still most familiar – way of representing this knowledge was in encyclopedic human readable articles in different languages, which were to some degree linked to each other. With the implementation of Wikidata this has changed. Now, there is a centralized database that has unique IDs for each semantic concept. From those concepts, there are links to the individual language versions of the Wikipedia articles. In addition, Wikidata contains links between concepts, and properties of the items of various data types, including mathematical formulae [SS16]. In this way, Wikidata forms the backbone of a knowledge graph. The project, then, exists somewhere between two domains: *digital libraries* and *artificial intelligence*.

2 Our plan for the future

We propose to rely on Wikidata, and build "math.wikipedia.org" as a frontend to that. This service will allow users to conveniently store and retrieve formulas, terms, and other mathematical concepts. Building a knowledge base with machine-readable formulas and links between concepts will produce a richer domain for reasoning than we get from existing projects or other mathematical digital library proposals.

In brief, the challenges are: figuring out what will make a large-scale project like this work from a social perspective; getting the basic technologies working; system integration; and knowledge representation. It is a good time to address these issues. Working with an existing popular platform should help bootstrap the social side, and allow us to focus our development effort on domain-specific issues. State of the art technology now exists for mining relevant knowledge from existing knowledge resources. Wikidata and a handful of other projects have proposed knowledge representation formats for mathematics. One current example, *Gamma function* on Wikidata is pictured in Figure 1. The description presented on this page is particularly terse. More details would be needed to bring the Wikidata page closer to feature parity with the Wikipedia pages in English or German. Indeed, better-articulated frameworks would be required to adequately represent more complex mathematical concepts, e.g., *adjunction in a category*.[2] Whatever the concept, the definition and formulas should be accompanied by framing information (e.g., who invented it, where has the concept been discussed?).

[1] https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Mathematics/Count

[2] https://en.wikipedia.org/wiki/Adjoint_functors

3 How will the public benefit from this work?

We foresee the following broad benefits:

Improvements to Wikipedia: If all of this information is available within Wikidata, we and others can use it to make a better Wikipedia using semi-automated methods. For example, article placeholders, autolinking of entries, automatic provision of references, and so on, could all be provided. By using the knowledge graph, such services can be delivered with contextual sensitivity; e.g., to point out that "this article is missing a basic explanation of XYZ.")

Reasoning and computational methods available by default: Having a computational "frame" over the material, we would be able to infer, e.g, "Here is an example of this thing (or, this thing is an example of some other thing) and here's how you compute with it."

Integration with other knowledge bases and AI systems: Thinking of the system as an data consumer, we should be able to use it to index and centrally store concepts from the broader literature. Thinking of the system as a data provider, if the Wikidata representations are done right, we should be able to use this data in AI applications, for example, as a resource to use in automated question-answering.

4 Related work

The NIST Digital Library of Mathematical Functions [Mil13] and the derived Digital Repository of Mathematical Formulae [CMS+14, CSM+15] are inspiring examples of formula-centric storage. Contentmine is an interesting current project that has had success mining knowledge out of research papers (for now, mostly in chemistry and bioscience). It would be interesting to apply the methods to the mathematics and physics content stored in Arxiv. Babar is another project for extracting knowledge out of Wikipedia, specifically [dL12]. In general we have been highly involved with representing mathematics on the web in a format suited for human readers, notably through work with the MathML Association and PlanetMath. Representing mathematical knowledge in a format suitable for both human and machine interaction requires a somewhat different approach. Michael Kohlhase's recent work on a Semantic Multilingual Glossary for Mathematics is relevant prior art.

Acknowledgments Firstly, we must point out that we are not affiliated with the Wikimedia Foundation, though we are long-time contributors to the Wikipedia project and are inspired by its success. At the time of writing this early stage proposal does not (yet) have official status with any agency or funding body. We welcome feedback.

References

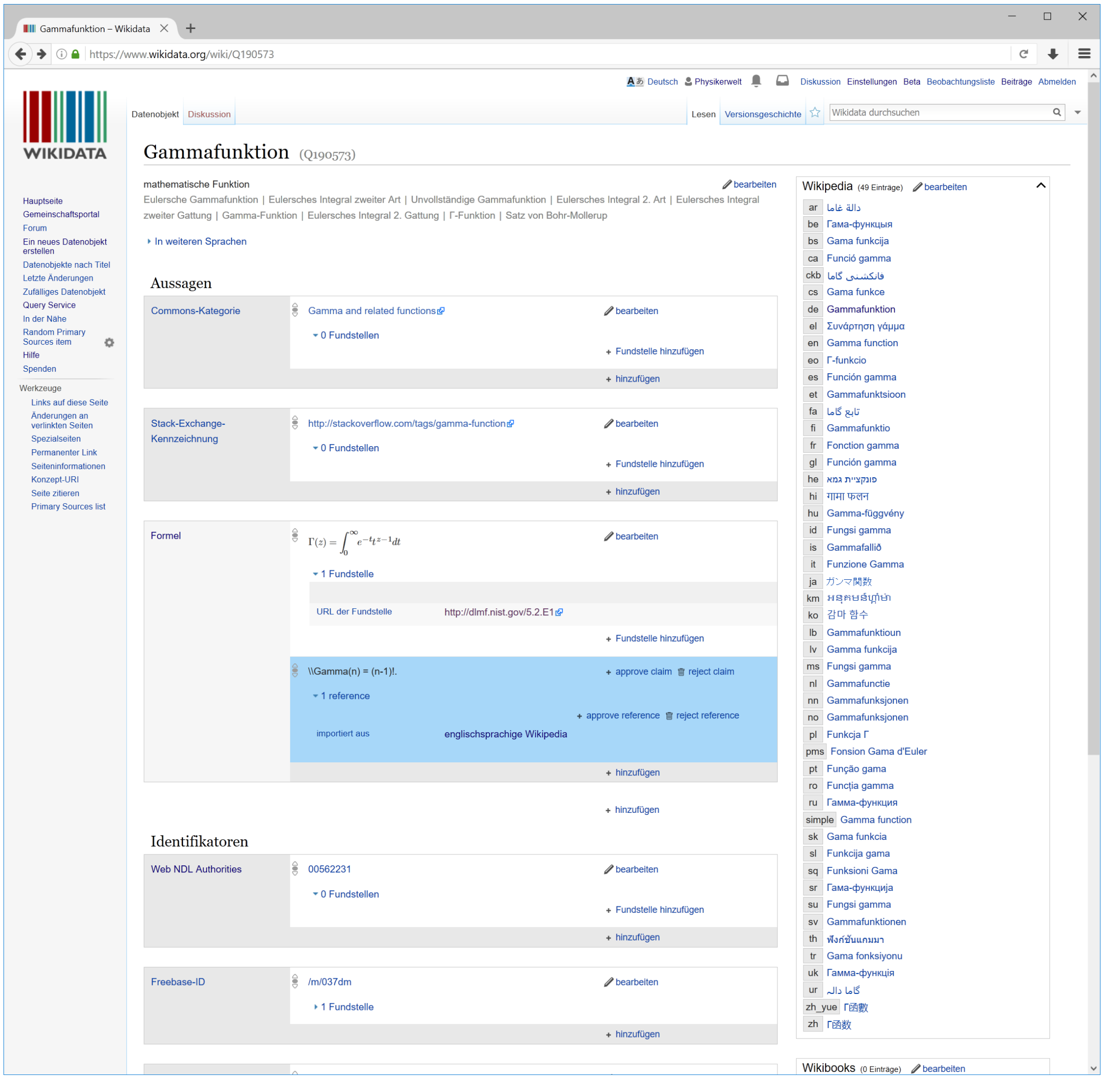
[CMS+14] H. S. Cohl, A. McClain M, B. V. Saunders, M. Schubotz, and J. C. Williams. Digital Repository of Mathematical Formulae. In *CICM*, LNCS, 2014.

[CSM+15] H. S. Cohl, M. Schubotz, M. A. McClain, B. V. Saunders, C. Y. Zou, A. S. Mohammed, and A. A. Danoff. Growing the DRMF with Generic LaTeX Sources. In *CICM*, 2015.

[dL12] Pierre Raymond de Lacaze. BABAR: Wikipedia Knowledge Extraction, 2012.

[Mil13] Bruce R. Miller. Three years of DLMF: web, math and search. In J. Carette and D. Aspinall, editors, *CICM*, volume 7961 of *LNCS*, 2013.

[SS16] M. Schubotz and A. P. Sexton. A smooth transition to modern mathoid-based math rendering in wikipedia with automatic visual regression testing. In M. Kohlhase, editor, *WiP at CICM*, Aachen, 2016.



Gammafunktion (Q190573)

mathematische Funktion — bearbeiten

Eulersche Gammafunktion | Eulersches Integral zweiter Art | Unvollständige Gammafunktion | Eulersches Integral 2. Art | Eulersches Integral zweiter Gattung | Gamma-Funktion | Eulersches Integral 2. Gattung | Γ-Funktion | Satz von Bohr-Mollerup

In weiteren Sprachen

Aussagen

Commons-Kategorie: Gamma and related functions — bearbeiten
0 Fundstellen
+ Fundstelle hinzufügen
+ hinzufügen

Stack-Exchange-Kennzeichnung: http://stackoverflow.com/tags/gamma-function — bearbeiten
0 Fundstellen
+ Fundstelle hinzufügen
+ hinzufügen

Formel: $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$ — bearbeiten
1 Fundstelle
URL der Fundstelle: http://dlmf.nist.gov/5.2.E1
+ Fundstelle hinzufügen

\\Gamma(n) = (n-1)!. — + approve claim, reject claim
1 reference
+ approve reference, reject reference
importiert aus: englischsprachige Wikipedia
+ hinzufügen

+ hinzufügen

Identifikatoren

Web NDL Authorities: 00562231 — bearbeiten
0 Fundstellen
+ Fundstelle hinzufügen
+ hinzufügen

Freebase-ID: /m/037dm — bearbeiten
1 Fundstelle
+ hinzufügen

Wikipedia (49 Einträge) bearbeiten

ar	دالة غاما
be	Гама-функцыя
bs	Gama funkcija
ca	Funció gamma
ckb	فانکشنی گاما
cs	Gama funkce
de	Gammafunktion
el	Συνάρτηση γάμμα
en	Gamma function
eo	Γ-funkcio
es	Función gamma
et	Gammafunktsioon
fa	تابع گاما
fi	Gammafunktio
fr	Fonction gamma
gl	Función gamma
he	פונקציית גמא
hi	गामा फलन
hu	Gamma-függvény
id	Fungsi gamma
is	Gammafallið
it	Funzione Gamma
ja	ガンマ関数
km	អនុគមន៍ហ្គាម៉ា
ko	감마 함수
lb	Gammafunktioun
lv	Gamma funkcija
ms	Fungsi gamma
nl	Gammafunctie
nn	Gammafunksjonen
no	Gammafunksjonen
pl	Funkcja Γ
pms	Fonsion Gama d'Euler
pt	Função gama
ro	Funcția gamma
ru	Гамма-функция
simple	Gamma function
sk	Gama funkcia
sl	Funkcija gama
sq	Funksioni Gama
sr	Гама-функција
su	Fungsi gamma
sv	Gammafunktionen
th	ฟังก์ชันแกมมา
tr	Gama fonksiyonu
uk	Гамма-функція
ur	گاما دالہ
zh_yue	Γ函數
zh	Γ函数

Wikibooks (0 Einträge) bearbeiten

Figure 1: Screenshot of the WikiData page of the gamma function https://www.wikidata.org/wiki/Q190573, with German language localization. On the top of the page is the title and the id. Below there is a natural language description and a list of aliases. In the block on the right there are links to the various Wikipedia pages about the gamma function. One of the blocks on the left contains a definition that was manually entered, and originates from the NIST Digital Library of Mathematical Functions. The formula in the blue box, $\Gamma(n) = (n-1)!$, was automatically extracted by Yash Nager (master student at Universität Konstanz) from the English Wikipedia pages and currently waits for human verification. Links to external identifiers and other related resources are provided.