Application for a Dagstuhl Seminar
Math-in-the-Middle: An API for Mathematics

Jacques Carette and Wolfram Decker and Viviane Pons and Florian Rabe

April 11, 2017

Abstract

[1] [2] In the last decades an ecosystem of open source mathematical and scientific software has emerged. This software is developed by overlapping international communities of researchers in both computer science and mathematics, by teachers, engineers, and amateurs. EdN:1 EdN:2

Current systems range from specialized libraries (e.g. MPIR, LinBox) to thematic systems (e.g. GAP in group theory, PARI in number theory, Singular in algebraic geometry, xcas) to general purpose systems like Mathemagix or Sage.

Moreover, there is a wealth of databases of mathematical objects like the Online Encyclopedia of Integer Sequences, the L-functions and Modular forms Database, the Small Groups Library, the Online Atlas of Finite Simple Groups, and many more.

Interactive computing environments Jupyter and knowledge management systems MathHub can serve as friendly interfaces for these systems.

It remains very difficult to connect domain specific, optimised systems into higher-level systems and avoid the expensive duplicate development of libraries.

We propose that an appropriate basis for a solution is found in abstract logic and knowledge representation; Projects to design knowledge representation languages, and frameworks for system interoperability have been initiated in the past: OpenMath, SCIENCE, and OpenDreamKit, to name a few.

None of them has reached a critical mass of supported systems.

In this seminar, we bring together agents from relevant communities in order to develop tools, best practices, and open standards, for sharing algorithms and data across computational mathematics systems, in order to reach that critical mass.

[3] [4] [5] EdN:3 EdN:4 EdN:5

Contents

[1] EdNote: @all revise

[2] EdNote: about 250 words

[3] EdNote: deadline is April 15th for the Initial Submission, April 22nd for the final (Markus check again!)

[4] EdNote: See http://www.dagstuhl.de/programm/dagstuhl-seminare/antrag/ for details about how to write proposal

[5] EdNote: Should Proof assistants/proofs/verification be a thing here?

[6] EdN:6

[7] EdN:7

1 Basic Information about the Seminar

1.1 Title

The title *Math-in-the-Middle: An API for Mathematics* means to emphasize the vision of a universal environment for scientific and mathematical computation. We hope this will provide a joint goal for the currently disparate communities. [8] EdN:8

1.2 Organizers

[9] EdN:9

- PD Dr. Florian Rabe
  Jacobs University Bremen - Computer Science
  P.O. Box 750 561 28725 Bremen Germany
  Phone: + 49 421 200 3051
  Email: f.rabe@jacobs-university.de
  Homepage: http://kwarc.info/people/frabe/
- PD Dr. Viviane Pons
  Université Paris-Sud - Laboratoire de Recherche en Informatique
  Bat 650 Ada Lovelace 91405 Orsay Cedex France
  Email: viviane.pons@lri.fr
  Homepage: https://www.lri.fr/~pons/

1.3 Type of event, duration, and size

We propose a 5-day Dagstuhl Seminar with 45 participants.

1.4 Topics

[10] EdN:10

- databases / information retrieval
- semantics / formal methods
- sw-engineering

[6] EdNote: Comment by JC: I don't think it is necessary to re-focus, but rather to make sure that there are clear "new" topics to discuss, with a clear hope that the discussion can lead to real progress. So if the workshop is to be about interoperability of (some) systems that focus on 'computation', it seems that a good starting point would be a list, articulated by some of the principal stakeholders of said systems, of what kinds of computations they would like to borrow from another system AND what currently prevents them from just doing it.

I, of course, entirely agree with the basic idea "that systems should share knowledge better". But you and I are 'solution providers' in this area, where the solutions are implemented in systems which otherwise do not "do" enough to be attractive at large. So we need to work with the system builders to understand what their actual problems are, and bridge the gap. Or so I think. It seems we need equal involvement (in the workshop definition) from some of the principal system builders involved in OpenDreamKit to get a balanced proposal.

[7] EdNote:

[8] EdNote: markus: I find this title a bit bland. I am a fan of catchy titles though.

[9] EdNote: @all: add yourself

[10] EdNote: @all: we have to select 1-3 from a list of topics; strangely none fits very well
markus: I commented out the ones that fit least, cutting it down to 3 that seem to fit reasonably well

1.5 Keywords

[11] EdN:11

- computer algebra
- computation
- reusability
- interoperability

1.6 Proposed Seminar Dates

[12] EdN:12

Block-out Dates:

•

Preferred Dates:

•

2 Description of the Seminar

3 Introduction

From their earliest days, computers have successfully been used in mathematics to perform complicated or tedious calculations more reliably, or at all, to make tables, to prove theorems (famously the four colour theorem), to explore new theories.

With computers and open source software becoming widely, and cheaply, available to everyone, the last decades have seen the emergence of open-source software in research communities; the spectrum ranges from special purpose Excel worksheets, to more sophisticated libraries such as MPIR (for high performance integer arithmetic) or LinBox (linear algebra), to topical systems such as PARI in number theory, Singular in algebraic geometry, or GAP in group theory. Some packages are written and embedded in popular general purpose programming languages, wheras some systems feature their own programming language.

Addionally, there is a wealth of databases of mathematical objects: the online Atlas of Finite Simple Groups, the Small Groups Library, the L-functions and Modular Forms Database, FindStat, the Regular Graphs library, the Online Encyclopedia of Integer Sequences, and many more. [13] EdN:13

One major problem posed by the multitude of systems with different communities, developers, and focuses, is that they do not compose well. The current state of the art in interfacing between systems is to understand the conventions and internals of the systems involved sufficiently to write a bespoke interface, and to translate data and representations of objects semantically correctly.

As an example, the group-centric computer algebra system GAP includes custom interfaces to Singular, the Graph Isomorphism Solvers Nauty and Bliss, separately with different representations of graphs, the advanced coset enumerator ACE, Holt's automatic groups programs KBMAG, and more.

Sage, a free general purpose open-source mathematics software with a mission to provide an alternative to MAGMA, Maple, Mathematica, and Matlab, includes GAP as a subsystem. The interface between Sage and GAP consists of hand-written translation routines, which for

[11] EdNote: @all: revise/extend

[12] EdNote: @all: Dagstuhl will pick the date, typically early 2018, but it could be be several months earlier or later; please add your preferences

[13] EdNote: markus: I might be focusing this too much on pure maths/combinatorics?

example have to interface SAGE's integers with GAP's integers, and SAGE's permutations with GAP's.

This method of interfacing systems is tedious, time consuming, and error-prone.

Ideally there would be a single, well-defined, standard that defiens how to efficiently and correctly communicate data between systems. Such a standard would have to be universal enough to cover most of computational mathematics, easy to use, and efficient.

(maybe itemize? what about composability/transitivity?

- **Universality**
- **Usability**
- **Efficiency**

)

One cannot expect every developer of mathematics software to change it to the needs of this standard, we need to work towards composability.[14] EdN:14

4 The Math-In-The-Middle Approach

To address the issues introduced in the preceding section, we propose to establish the Math-in-the-Middle based on OMDoc/MMT approach as described in **??**.

The idea is to flexibly formalise mathematical concepts in a very general logical framework, and to encourage authors to describe their systems in this framework, ultimately establishing a common *meaning space* in which data can be automatically, efficiently, and semantically correctly translated between systems. Parts of the process of description and discovery can in many cases be automated, as is witnessed by our work on GAP, SAGE, and .

The vision is to establish this Math-in-the-middle approach as the universal API for as many computer mathematics software packages as possible, without forcing these packages to change their software significantly.

Working on WP6 of OpenDreamKit we realised more and more that it is necessary to communicate and work together in one place

The seminar will consist of few talks that introduce and showcase our initial successes working on WP6 of ODK involving Sage, GAP, LMFDB, the core concepts of MMT, and many workgroup sessions to learn from each other about developers needs and about barriers to entry and use.

5 Questions, Issues, Objectives, Results

The central goal of this seminar is to systematically identify the current obstacles to universality, usability, and efficiency, and to sketch out future solutions.

To achieve this goal, we will

- Bring together developers and users of (open source) mathematics software, logicians, and knowledge representation experts.
- Get an overview of the current trends and developments in open source mathematics software.
- Promote MMT/Math-in-the-Middle
- Train mathematicians in the art of MMT, and get an idea of how MMT will be used by developers of domain specific software.

[14] EDNOTE: Markus: This needs more substance

- Share perspectives and best practices, build a joint vision, and seek venues for tighter collaboration.
- Encourage participants to get involved in the standardisation process and to provide MMT interfaces to their software.

Some of the upcoming major challenges are:

- Enabling **computations involving multiple systems**, as transparently as possible;
- Keeping the development efforts manageable as the size and complexity of software systems increase;
- Training a new generation of users and developers.
- Combining the above to remove usability barriers, in particular via **unified user interfaces**, and **Virtual Research Environments** that groups of users can setup to collaborate on data, software, computations, or knowledge;

A key step is to strengthen collaborations *between* communities, in order to:

- Seek for opportunities for collaboration or outsourcing of components to save on development efforts;
- Share expertise and best practices;
- Improve cross-systems development workflows.

Research Questions Participants will be asked to give short talks that specifically address the following research questions from the perspective of their field:

- Why are current systems not more interoperable? What design changes are necessary to increase interoperability in the future?
- What are the current approaches towards interoperability? How successful or promising are they?
- How can correctness be guaranteed in a distributed setting? Should there be a single universal checker (which would be hard to agree on) or many decentral ones (which may preclude interoperability)?
- How can we design interchange languages that naturally subsume existing (and future!) formal systems?
- Should a logical framework permit the definition of any logical system? Or do the logics currently implemented have points in common that could be hard wired into the framework itself?
- How reasonnable is it to propose a single universal proof format? Or do we need different formats for different families of proof systems and a partial interoperability between the formats? How should a proof format be evaluated (generality, conciseness, efficiency of proof-checking, ...)?
- How should universal proof library be exchanged? Is Web technology sufficient or do we need specific tools to organize data bases of proofs?
- How can we practically and reliably relate individual systems with their representation in an interchange format or a logical framework? How can two systems agree on the meaning of an exchanged theorem and thus trust each other?

Impact on the Research Community By challenging participants to address research questions concerning universality, we do not only raise awareness of the importance of these issues. We also help identify the key steps towards *proving in the large* and *universality* of proofs. This will allow the development of a common objective and framework for interoperable and reusable proof development that is crucial for realizing the full potential of formal mechanizations.

This seminar with the associated Dagstuhl proceedings will provide an overview of the problem, the state of the art of current solutions and the active researchers pursuing them, and the most promising ideas for future solutions. It will collect and strengthen the small, often-disparate communities that currently work towards universality, e.g., in the very different PxTP (Proof eXchange for Theorem Proving) and LFMTP (dedicated to logical frameworks and meta-languages) workshops.

The seminar will not only allow for cross-fertilization between

- research on logical frameworks, proof formats, logics, proof engineering, mathematics formalization, and program verification,
- foundational research on these topics and application or system-oriented approaches.

It will also structure and streamline future collaboration, e.g., by kicking off new workshops or large international grant proposals.

5.1 Relation to Previous Dagstuhl Seminars

6 List of Potential Participants

We have identified [15] relevant researchers with a particular focus on EdN:15

- including leading experts from all involved communities,
- bringing together researchers with particular interest in universality and interoperability.

[16] EdN:16

We carefully selected participants to cover the following areas:

- computation systems
  - symbolic computation
  - exact computation
  - numerical computation
- related systems
  - user interfaces
  - databases
  - knowledge bases
- applications
  - mathematical databases
  - scientific computation
  - industrial applications
- system integration
  - integration frameworks
  - individual system connections

Moreover, the participants include[17] EdN:17

- XXX from Germany, XXX from France, XXX from UK, XXX from the rest of Europe, XXX from North America
- XXX junior, XXX female (one of them co-organizer), and XXX industrial researchers.

[15] EdNote: add number

[16] EdNote: @all: revise areas and subareas depending on what people we want; the topics listed here will be referenced in the spreadsheet containing the suggested participants

[17] EdNote: add numbers at the end

7 Information on the Organizers

[18] EdN:18

7.1 Brief presentation of the organizers

Florian Rabe has developed the MMT framework and the LF module system and is the main contributor of the LATIN atlas. He has extensive expertise in individual deduction systems including major case studies regarding Mizar, HOL Light, PVS, and TPTP.

Viviane Pons is a researcher in Combinatorics, leader of the dissemination work package of the OpenDreamKit H2020 project. She has been involved in the development of SageMath. She has been organizing Sage Days and given many Sage interventions and tutorial.

[18] EdNote: add a 0.5-1 page research CV of yourself: overview of an organizer's academic career, especially points out community services and recognitions, list the five most relevant papers

7.2 Florian Rabe - CV

Name	Dr. habil. Florian Rabe, born 1979-09-28
Diploma	Computer science 2004, Universitity of Karlsruhe (distinction)
PhD	Computer science 2008, Jacobs University Bremen (distinction)
Habilitation	Computer science 2014, Jacobs U.
Employment	2008–2014, post-doctoral fellow, Jacobs U.
	2014– DFG-Eigene Stelle, Jacobs U.
Awards and Scholarships	
2005	Best diploma thesis, Computer science faculty
2006	PhD scholarship (1 year), DAAD
2006	Winner Modal Logic $100 Challenge
2007–2008	PhD scholarship, German Merit Foundation
2010	Best Paper Award, MKM conference
2015	Contest Winner "The Future of Logic", UniLog Congress
Membership in Academic Self-governance Committees	
2008 – 2010	Staff Council (Jacobs U.)
2010 - 2012	Provost search committee (Jacobs U.)
2011 - 2012	Constitution committee (Jacobs U.)
2010 – 2013	Board of trustees of MKM interest group
2012 –	Steering committee of CICM conference
Student Advising	14 BSc., 6 MSc., 3 PhD. (some in progress)
Organization	2 conferences, 4 workshops
PC Membership	8 conferences (2 as track chair), 11 workshops (4 as chair)
Third Party Funding	
2009–2012	LATIN (DFG), de-facto PI
2014–2017	OAF (DFG), lead PI
2015–2019	OpenDreamKit (EU Horizon 2020), PI

5 **Important Publications**

- M. Kohlhase and F. Rabe. QED Reloaded: Towards a Pluralistic Formal Library of Mathematical Knowledge. *Journal of Formalized Reasoning*, 2015. accepted pending minor revisions; see http://kwarc.info/frabe/Research/KR_qed_14.pdf.
- F. Rabe. The Future of Logic: Foundation-Independence. *Logica Universalis*, 2015. Winner of the Contest "The Future of Logic" at the World Congress on Universal Logic; to appear; see http://kwarc.info/frabe/Research/rabe_future_15.pdf.
- F. Rabe. How to Identify, Translate, and Combine Logics? *Journal of Logic and Computation*, 2014. doi:10.1093/logcom/exu079.
- F. Rabe and M. Kohlhase. A Scalable Module System. *Information and Computation*, 230(1):1–54, 2013.
- F. Horozal and F. Rabe. Representing Model Theory in a Type-Theoretical Logical Framework. *Theoretical Computer Science*, 412(37):4919–4945, 2011.

7.3 Viviane Pons - CV

Name	Dr. Viviane Pons, born 1985-02-05
Bachelor Degree	Computer science and Mathematics 2006, Univ. Paris-Est (distinction)
Master Degree	Computer science 2010, Univ. Paris-Est (distinction)
PhD	Computer science 2013, Univ. Paris-Est
Employment	2013–2014, post-doctoral researcher, Univ. of Vienna (Austria)
	2014–, Maïtre de conférences en Informatique, Univ. Paris-Sud
Projects	
2015 – 2019	OpenDreamKit (EU Horizon 2020)
	Site leader for Paris-Sud
	Community building / Dissemination Work package leader (WP2)
Organization	4 workshops
Invited speaker / lecturer	
2015	Combinatorics and Sage at EAUMP Summer school on experimental mathematics (Uganda)
2016	Combinatorics Research School at ENS Lyon (France)
2016	ECCO 2016 – Sage Tutorial (Colombia)
2016	Codima School 2016 – Sage and SageMathCloud Tutorial (Scotland)
Other outreaching activities	
2014 – 2015	Experimental mathematics using Sage: talks at SciPy2015 PyConFr2014 and PyCon2015
2016 –	Co-organizer of the PyLadies Paris Chapter
PC Membership	FPSAC 2017 (conference)

Important Publications

- G. Châtel and V. Pons. Counting smaller elements in the tamari and m-tamari lattices. *Journal of Combinatorial Theory, Series A*, 2015. doi:10.1016/j.jcta.2015.03.004.
- V. Pons. Interval structure of the Pieri formula for Grothendieck polynomials. *International Journal of Algebra and Computation*, 2013. doi:10.1142/S0218196713500045.
- V. Pons. Multivariate polynomials in Sage *Sminaire Lotharingien de Combinatoire*, 2011.