CICM 2013

J.H. Davenport — J.H.Davenport@bath.ac.uk

8–12 July 2013

Contents

Chapter 1

8 July 2013

1.1 A Machine-Checked Proof of the Odd Order Theorem
Rooster and Butterflies — Assia Mahboubi

"Machine-Checked" = "by a proof assistant". Proof Assistants are heavily used n Computer Science — note that Coq got the ACM Programming Languages Award. Examples of proofs that **need** computers are Four-Colour Theorem [Gon08] and Kepler Conjecture [Hal05].

So lets look at the odd-order theorem [FT63, Note it's a whole volume], and a collective simplification effort resulted in two volumes of LMS Notes [BGC94, Pet00]. Recall 'composition series' and 'simple', and the Jordan–Hölder Theorem which says that the composition sequence is essentially unique. **But**, and unlike prime factors of numbers, the composition series does not uniquely determine the group.

Sketch of proof.

1. Let G be a minimal counter-example, i.e. a simple group of odd order which is not cyclic.
2. Look at maximal strict subgroups of G.
3. "Local Analysis" (volume 1 [BGC94], three pages of notation) shows that there are five possible types: I–V.
4. "Character Theory" rules out all but II.
5. II involves finite fields, so we turn to Galois Theory, which completes the contradiction.

Now done in Coq [GT12, Currently 150kLoC].

What is a proof assistant?

1. A formal logic — think of this as the machine code of the system. Coq has Type Theory.
2. Proof Assistant — think of this as the compiler
3. Proof checker — running the machine code.
4. Libraries — we wrote MathComp.

A textbook features

- set-theoretic flavour
- a common belief in first-order logic is enough
- Shameless use of Axiom of Choice.

  Let A be a square matrix. Then:

$$\det(A) = \sum_{\sigma \in S_n} \prod_i a_{\sigma(i),i} \tag{1.1}$$

  (taken from [Lan65]).

Observe what is missing, e.g. a definition of $n!$

Hence we make strong use of Coq's type theory and its very expressive dependent types, with a special status for computation: `2==3` is a Boolean, and `compute in 2==3` is `false`. Coq's implementation has unification, user-defined notations, implicit arguments. Noted that defining matrices in front of us required quite of a lot of semi-mechanical "mix-ins".

Theorem 1 $\mathrm{tr}(A^T) = \mathrm{tr}(A)$ — Coq lets one define ^{T} appropriately.

Proof is three rewrites and an application.

Q Isabelle?

A I don't know. We didn't particularly want the constructiveness of Coq, and had to work to get around it. Computation *was* vital, and this seems to require constructiveness. Dependent types were very important.

Q What can you do with this proof?

A Good question.

Q How about the full classification?

A That would be much harder, as there is no equivalent of the LMS books.

1.2 Formulations in Cylindrical Algebraic Decomposition — Wilson

Simple example — choice of variable ordering. Note [BD07] for the theory of how much this can matter.

Note equational constraints [McC99]. Which constraint we make "equational" affects the performance. In TTICAD, the same is true, but *for each clause*. There's the `sotd` heuristic [DSS04], but it can be fooled, and doesn't take account of the **real** structure. How complicated is $\mathbf{R}^1$? This leads to our `ndrr` heuristic. Gröbner preconditioning [WBD12] led us to the `TnoI` heuristic.

We now have a combinatorial explosion of the number of choices.

We are aiming towards algorithms which make "near-optimal" choices of detailed formulation.

Q Many people don't know about CAD.

A In the current state, that's probably a good thing. In Maple it can be used in the background without the user being aware.

1.3

1.4 Certification of Bounds of Non-linear Functions: The Template Method — Margon

Note that [Hal05, Hal08] consists of thousands of nonlinear inequalities. Attempt at formal proof [Hal11]. Aim for certificates of positivity and then proof via Coq.

Lemma 1 (typical example) $\forall x \in [3, 64] 2\pi - 2x \arcsin(\cos(0.797\ldots < \ldots$

1.4.1 Methdology

1. Let $f^* = \min_{k \in K} f(k)$.
2. Let f_{sa} be a semi-algebraic approximation to f with $f(x) \geq f_{sa}(x)$.
3. Let f_{pop} be a "proof of positivity" approximation to f_{sa} with $f_{sa}(x) \geq f_{pop}(x)$.
4. Show $f^*_{pop} > 0$ and then $f^* \geq f^*_{sa} \geq f^*_{pop} > 0$.

Our new strategy is via max-plus estimators. Build an abstract syntax tree with semi-algebraics as the leaves and transcendental operations and dyadics as the internal nodes. End up wth proof-of-positivity in $3n$ variables ($2n$ extra 'lifting' variables). Usng "sum-of-squares" can get $(2n)^{2d}$. Worst case (done?) 6 transcendentals and 5.7 hours. Have an external progam OCaml/C to produce the SoS decomposition, whch Coq then verifies. There's a problem with the floating-point in the SoS, though. Workaround via an ϵ-polynomial.

Q Is log semi-algebraic?

A No — confused debate.

1.5 Verifying and Platform for Digital Imaging: a multi-tool strategy — Heras

Part of the ForMath project. La Rioja is using formal algebraic topology libraries to biomedical maging (looking for reliable software).

1.5.1 Tools and Methods

Fiji/ImageJ are Java programs used life scences and biomedicine. They are extensible via plugins. This is a big multi-author system not desiggne to be formalised. Note also that "native methods" (such as log) are a problem.

Krakatoa/Why An enviroment for proving correctness of Java programs annotated wth JML.

Coq/ACL2 used as backup theorem provers when Why fails. But in practice Why often fails because of under-specification at the Krakatoa stage.

Coq2ACL2 is a Proof General extension. Transforms the Coq from Why to ACL2.

XLL Xmall Logical Language translates Isabelle/HOL into ACL2 as well.

1.5.2 Verification

Demonstration, getting stuck in Why on a post-condition.

We have 2-d images of a 3-d reality. Use Fiji's *maximal projection methods* to do the reconstructon.

1.5.3 Conclusions

We re-use I2EA environment to combine systems. Fiji is unsound, but the errors are mnor and can easily be fixed. We would like to reconstruct the ACL2 proofs into Coq.

Q How do you know Why only generates first-order output?

A "Because it does" — we've never seen higher-order ones.

Q You said "reals" in Coq.

A Yes, **R** not floating-point.

1.6 The Formalization of Syntax-Bsed Mathematical Algorithms Using Quotation and Evaluation — Farmer

Lecture delivered by video recording.

Example — symbolic differentiation (and associated simplification rules).

A "Syntax Framework" is a means of describing the interplay between syntax and semantics. The members of L_{syn} denote objects on D^M_{syn}. Evaluation maps L_{syn} to L_{obj} . Converse is a quotation operation. Need $E(Q(f)) = f$ always. So, in out theory T, we can reason indirectly via L_{syn}, or directly. But Q and E cannot be defined in T (apparently x is free in $xx + 3$, but not in its syntactic representation). Need a "replete" syntactic ramework. Examples: Lisp (smplifed semantics, Robinson's work, and my Chiron.

In ths, we have ditectly readon about L_{Poly} in T. We can write the sentence that says O_{pd} has the same computational behaviour as `PolyDiff`. n T we can directly raon about the syntax of expressions in L.

But there are problems.

Evaluation Cannot be defined via Gödel diagonalisation: $LIAR = Q(\neg LIAR)$

Quotation x is not free in $Q(x + 3)$, butt is free in $E(Q(x + 3)) = x + 3$. This causes problems with substitution, free variables etc.

Extension If we add a constant to L, this is not a conservative extension (details on the slides, but JHD didn't follow).

Interpretation Let Φ be an interpretation of T in T'. Then Φ has to be injective on constants.

Hence the global approach needs modification to its logic — MathScheme projetc.

1.7 A Universal Machine for Biform Theory Graphs — Rabe

Three aspects.

Declarative such as OpenMath. Relatively easy: $+ : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$

Deductive Need to make a lot of decisions.

Computational Even more decisions.

There are actually declaratve components in each of the other two systems.We want these to be refinements of the abstract declarative system. In other words, every ssystem has its (implicit) meta-logic.

We start with MMT and make it computation-aware. See [RK13].

Case study, OpenMath as meta-system, with CDs as specifications. Translate this to a computational system Scala. The translations are done mostly auotmatically, leaving the programmer to focus on the declarative theory on MMT *and/* implementing the abstract classes in Scala.

1.8 Calculemus Business Meeting

1.8.1 Business

Windsteiger chaired the meeting, and JHD ended up taking the notes.

1.8.2 Membership

Track Chairs: 2011 (Farmer), who therefore comes off, 2012 (Dos Reis), 2013 (Windsteiger). It was proposed that Davenport chaired 2014. In fact this is the Trustees decision, so it is now a *fait accompli*

The elected trustees had a vacancy. There had been two candidates: Asperti and van der Hoeven. The election had not been run, however. The secretary pleads *incuria*. It was noted that Sorge had been elected in 2011.

It was noted that the Charter still referred to six programme chairs and six elected — JHD moved, and MK seconded, that the Charter be amended to conform with reality (three of each). **GdR**

We still need more candidates: MK nominated FR, seconded JC. SMW nominated, and JHD seconded, Elena Smirnova. WW nominated JC, seconded JHD.

Nominations would remain open until the end of CICM, then we would need an election. **GdR**

1.8.3 Calculemus/CICM

RR was the Calculemus member of CICM, and the trustees wished him to continue. Again, the Charter needed fixing. WW was agreed as his alternate. It should also allow for an alternate to the delegate, especially at the meeting itself. **GdR**

1.8.4 Publicity

It was noted (again) that publicity was important, and that this needed to be done at the Calculemus/MKM/... level *as well as* just by CICM.

1.8.5 General

MK queried the extent to which Calculemus still needed its own organisation. JHD commented, based on his experience as General Chair of 2011, that the existence of separate tracks, with chairs and advertised PCs, and track-specific publicity, was good. No-one dissented.

It had been noted that the Trustees had no Chair, only a Secretary. This had been intended to resemble Stalin's position as Secretary of the Communist Party, but probably wasn't clear.

SMW noted that SYNASC was an example of a similar conference with multiple areas, but had one organising committee. JHD supported the SYNASC argument. MK noted that individual tracks had fluctuated widely in popularity.

JHD proposed, and no-one dissented from

> The meeting mandated the trustees to investigate abolishing the separate administrative structure of Calculemus, while retaining the academic identity. **Trustees**

It was noted that the (membership) mailing lists were in various states of disarray. It probably only made sense these days to have a CICM mailing list, but this should be left to the Calculemus and CICM Trustees. **Trustees**

1.8.6 Calculemus 2013

WW reported. There were 15 submissions initially, but 2 were withdrawn. There were 3 or 4 referees per paper, and 2 papers/referee. Of the 13, we accepted 5 (38% acceptance rate). The EasyChair gap between accept and reject was 1.7 (!). He felt, and all agreed, that this had led to an excellent programme.

WW was thanked by acclamation.

SMW noted that he had seen a system whereby "conflict of interest" was merely managed by being silent, and co-authorship was handled out-of-band by a co-Chair. JC noted that over-zealous conflict declaration in EasyChair also made the General Chair's task more difficult.

1.8.7 A.O.B.

SMW asked that the candidate's statement explicitly addressed the "to do list" formulated above. This met with general support.

Chapter 2

9 July 2013; Systems and Projects

5-minute talks; no questions.

2.1 A Dynamic Geometry System based on Gröbner Cover

Uses Geogebra with Sage[1]. Computes the algebraic description of the geometric locus. Obtain a Gröbner Cover of the equations for the loci of the curves. hence information about special points, which get lost by elimination theory, or drowned in extra points.

2.2 Math-Net.ru —

Dgital Archive of Russian Mathematica nowledge: journals (1866–2013), pesns and institutions, video library, manuscript tracking and submission systems. Underpinning CMS. Overall database 80GB, excluding 2TB of video.

2.3 The DeLiVeRMath project — Schöneberg

From FIZ. A very large databaseo f content analysis. Hence we want to do text analysis. Java library and web server. Uses Stanford's key phrase extraction, hacked to understand LaTeX.

[1] I think actually Singular.

2.4 LaTeXML2012 — Ginev

Much better security, new coverage, new libraries (Wikipedia, internationalisation/Unicode). Better support for CSS. `http://latexml.mathweb.org/editor`. Note that DLMF [Nat10] was the inspiration.

2.5 swMath

An information service for Mathematical Software. Suppoted by FIZ, Oberwolfach etc. Thisis a "publication-based" approach. Semi-automatic detection of software-relevant articles. Systematic collection of article references. `http://www.swmath.org`.

2.6 The Web Geometry Laboratory Project — Quaresma

`http://192.168.10.100/WebGeometryLab` on `WGLatCICM`. already in evaluation at schoools. This is collaborative, so one example is "collaborative session to construct an equilateral triangle". Aim to extract personalised information from the students' interactions.

2.7 Software for Evaluation Relevance of Steps in Algebraic Transformations — Prank

In Mathematical Logic, e.g. "find the full DNF of ...". He showed a 10-step derivation, but one student did 269 to get the same result.

2.8 MMT — Rabe

This should be an generic MKM system — foundation-independent. Obviously some customisation is required, e.g. of type-checker. The LaTeX tool generates suitable cross-referenes iin the PDF.

2.9 Web interface for Isabelle — Ring

Note that these packages are typically hard to install, especially on mobiles! Build on Wenzel's Isabelle–Scala — thanks. Proof processng takes place on the server.

2.10 ML4PG in Computer Algebra verification — Heras

Focused on patterns — "I think there is something similar ...". Uses Machine Learning techniques.

2.11 Paral-ITP — Wolff

Pervasive Parallelism for formal content checking. Covers multiple layers: kernel and provers, document model, and front-end technologies (PIDE). Note that Isar is not the only possible interface to Isabelle.

2.12 ForMaRE Project —Lange

Formal Mathematical Reasoning in Economics. Original plan was collaborative game theory [KRW11], but have moved onto "hot topics", such as risk assessment, and 'auctions' — this has proved most successful. Look at ICANN's auction for new TLDs. Most economists use pen-and-paper only. Note also the need for verified software. We have tried to prove existing economics theorems in theorem provers — see MKM talk.

Chapter 3

9 July 2013: Work in Progress

3.1 Tagged PDFs — Ross Moore

Showed an example full of state tables, diagrams etc. Looks OK in a normal browser, then showed Acrobat Pro 11 (far better than previous versions). His example claims (and verifies) PDF/A. Note that Springer's PDFs for this conference claim PDF/A but "There's PDF/A and PDF/A".

Has problems with TEX tables currently. The diagrams are interesting — his software produces arrowheads from several fonts, which complicates things. The mathematics is very much presentation MathML `<mrow>` etc., and `<mo>` tags. Claims accessibilty, and the systtem queries colour contrast (watermarks as it's an examiination paper) and "logical reading order". Note that PDF doens't have "word spaces", just shifts. This needs 'reverse engineering', and explains why we often don't get word spaces when we paste from PDF generally. PDF/A actually allows, and Ross does, but a minimal (0.001pt) genuine space here, which doesn't disturb the typesettiing but *is* present.

There is a tree structure in the PDF that corresponds to the MathML. Ross's progam translates the LATEX, not just into MathML for the PDF, put also a speech rendering which is embedded in the PDF. This is expensve, so his system caches the translations of the individual expressions.

Note that Governments etc. often mandate WCAG, but this applies only to HTML. How does the user know whether he has PDF or HTML at the end of a link? Hence PDF/UA.

3.2 Scaling Feature Based Mathematical Search Engine for Real-World Document Sets — Mišutka

This is a part of my doctoral thesis.

In general, similarty can be either metric-based (Minkowski, Jaccard, Angular etc.) or non-metric based (generally preferred by domain experts). Consider man↔centaur↔horse which is more plausible than man↔horse. VSM (based on cosine similarity) is in fact not a metric.

Looked at [MHC10], but didn't necessarily agree with it. They extracted a feature vector, based on each `<mo>`, and cosine similarity.

Tried `wikipedia.org`, which has many more formulae, and lots of similarities. 271,103 unique features with the original algorthm, but these replaces variables and constants, and gives bad results. Without this 863,242 features. Used hand ground-truthing. Note that these huge vectors have a performance implication. `wikipedia.org` has 500,000 formulae (300,000 distinct).

Based on EgoMath parse trees, depending on how much "compression" (removing constants, variables, embedded * etc.) is done. Wants to get down to 10,000 features, as studies show that this gives efficient but effective retrieval.

3.3 The Declaratron, semantic specification for scientific computation using MathML — Murray-Rust

Claims $3 \cdot 10^{11}\$$ goes into STEM research, producing 3M articles with no semantics. Shows a line of input for a common Chemistry program. Units are often omitted (see Mars Lander [Nat99]). "The new generation of Open Access publishers are getting interested in moving beyond PDF".

Shows an example of forces in crystals, which they were able to translate into ChemML and process automatically to generate the same answers.

Further example of inter-atomic force fields:

$$\sum_{bonds} K_b(b-b_0)^2 + \sum_{angles} K_\theta(\theta-\theta_0)^2 + \sum_{dihedrals} \ldots + \sum_{electrostatics} \ldots,$$

but the equation in the manual is unimplementable except by a chemist — what are "dihedrals" etc.? Items such as "bond length" are determined by legacy code.

CML (molecules + parameter files), Dictionaries, a black-box library (JUMBO) as existing components, plus a MathML evaluation engine and DeXML our Declaratron ML. Based on Xpath, whch "makes it fantastically powerful to change your mind".

ScMathML is a Scala execution engine for MathML, basically evaluating expressions in a content. Does $\sum_{k=0}^{n} \frac{(-1)^k}{2k+1} \approx \frac{\pi}{4}$. Can say `<assert value="units.joule"`

etc. Uses algorithms like BOBYQA [Pow09]

3.4 Lurch: A Word Processor that Can Grade Students' Proofs — Carter

See also OpenMath sesson tomorrow (5.2). Aimed at introductory university "introduction to proof" courses.

- as indistinguishable from ordinary activities as possible
- supposed to validate mathematics (ths is the whole goal of this conference, but we are talking "tiny math", not 1.1. Note, we don't have a proof assistant: that's the student's job.

Educatonal research tells us (as if we didn't know, but here it is anyway).

1. Frequent immediate feedback is good for learning
2. students work until they are judged correct
3. Professors find the outputs easier to grade
4. Professors spend the time saved with their students.

Rather like "bold" buttons, there's a "make meaningful" button, whose results show up when mouse-over. Traffc-light system — amber means "is that meant to be a premise". Note that Lurch can't even check the work until the student has indicated the premises, reasons etc., which is itself beneficial.

Tried in two courses 2008 (old version) and 2013. This one started with formal proofs "or introducton from line 2", then move to semi-formal proofs.

1. Did students value the feedback — yes?
2. Did they learn the mathematics the instructor put into Lurch — yes from student opinion
3. Did students spend more time — apparently.
4. Is it too helpful "is it possble to do a proof in Lurch by typing and clicking" — 2008 agreed, but 2013 did not.
5. Does it take time away — instructor reckons Lurch actually took only 1 hour of teaching time.
6. Do they want it — when asked to do LaTeX, they used Lurch for validation first.

3.4.1 Plans

- Typeset mathematics
- Efficiency
- Add a type system — please help.

Q How big?

A We ship 50 documents, which are standard "rules". I don't know how much homework has been done.

***** At the moment, it's all an OpenMath document. Want to move to OmDoc.

Q Examples?

A See demo.

Chapter 4

9 July 2013 — PLMMS

4.1 Problems as Types — McBride

James McKinna and I built a programming language on top of a theorem prover (OLEG, as Lego threatened to sue).

Today we are demonstrating in Agda. Say that two elements of different sets are equal ("John Major Equality") if translating elements from one set to another maes them equal [JHD didn't get the exact words, whch clearly matter].

Parameters in a single set of braces are to be solved by unification, in double braces are to be inferred from the context (and must be unique).

Rearrange parameters of `NatElim` to make it easier to see what is being eliminated. The hgh-level program is in `[...]`, there is also some low-level theorem-proving stuff. Two basic operators: Substitution and NoConfusion.

Associative Law of addition (descrbed by Alan Bundy as the E. Coli of theorem proving). The problem statements are not intelligible as they refer to the proof, not the proposition itself. I have a similar problem with Coq. Note also that Agda makes a complete mess of open terms in "this set" (JHD couldn't work out which set). Strictness is a proof is a mstake, except in a proof of absurdity.

"It's just rude to call proofs by their names".

4.2 Sequential decision problems; dependently typed solutions — Brady

This came out of research at the Potsdam Institute of Climate Research (other two authors). They were interested in Idris, but ased whether it did floating-point arithmetic. We need to be able to express invariants about values, e.g. optimality; express relationships between these. See Ionescu's talk at CICM 2012.

Simple dcision problem — the cylinder problem [BirdDeMoor1997]. Most thngs we want to say about floating-point are lies.

4.3 CICM Business Meeting

4.3.1 Business

Kohlhase chaired the meeting, and JHD ended up taking the notes.

4.3.2 Trustees Report

CICM is an organisation of organisations (Calculemus, AISC, DML and MKM). See `http://trac/mathweb.org/CICM/wiki/Charter` adopted in Bertinoro, with minor changes, including a publicity member (Serge Autexier). CICM 2014 will be in Coimbra, with Stephen Watt as General Chair. Manfred Kerber has agreed to be General Chair in 2015, which *may* be at George Washington University. For 2014, James Davenport is the Calculemus track chair for 2014. The other tracks have their business meetings later, and AISC has not yet decided for 2014. The Trustees still need to nominate a S&P Track Chair. Serge Autexier is (permanent) workshop chair. There was a debate about making this also Work-in-progress, but it was felt that the refereeing should stay with the current tracks.

Date for 2014 — SMW to consult.

4.3.3 Treasurer's Report

WMF couldn't attend but sent data, presented by MK. CICM is a charity via IFCOLOG. There is €6380 in the back, and CICM 2012's surplus is estimated to be €10,000, due largely to on-site registrations.

CICM aims to keep €5000 as a seed/backstop, and use the rest for doctoral support. SMW asked about funding for invited speakers: the General Chair, who finally spends the money on the recommendation of track chairs, has to consult with the local chair, who has to raise the money! **SMW/PQ**

CSC noted, and all agreed, that if 2015 is in North America, we would need more funding then for students, and should be careful. **Trustees**
SMW noted that NSF group travel grants might be possible, also Sloan Foundation.

The membership thanked WMF for his report.

4.3.4 Programme Chair

The acceptance rates were as follows.

Calculemus 5/12 (was 7/9)

DML 6/8 (was 23)

MKM 7/178 (was 13/19)

. . .

JC firmly recommended *not* doing as he did, which was having one PC for CICM, rather than one per track.

4.3.5 Local Report

JHD reported, including London Mathematical Society support of £5600. The meeting thanked the LMS for its support.

4.3.6 Floor Remarks

Clash OpenMath/MathUI was regretted. SMW found the Doctoral Programme clashes unfortunate.

Demos CL noted that there is no way for demonstrations attached to non-S&P papers to be given. It would be nice to allow this. **SMW/PQ**

links/ArXiV Not all papers have been uploaded to ArXiV (and therefore linked from the website). Most posters are linked — could the rest be? **All/CL**

Doctoral MK regretted that there had not been many submissions to the doctoral programme. **Supervisors**
DJW commented that, as a student, he had found it helpful.

Chapter 5

10 July 2013 — OpenMath Workshop

CL pointed out that the proceedings would soon be on the CEUR website, as part of the integrated CICM Workshops proceedings.

5.1 Literate Sources for Content Dictionaries — Hellström

Motivation: publishing the mathematics.

- Small results — hand-edit into papers.
- Very large results — may need highly specialised software, an *ad hoc* database etc. Example of colleague swho wrote coefficients base 36 rather than base 10 for compactness as I/O was a major issue.
- Inbetween? This is where OpenMath has a rôle.

But doing research tends to need new symbols. But the learnng curve must not be underestimated. OpenMath — by page 30 you know what to do, and addng symbols is cler, while MathML rather buries this.

How to write an `.ocd`? Just follow the rules. But there's no software support. It doesn't seem important, and the software that geneerates the XHTML is apparently private to the website. he has not been able to make it work after four years. But we could make a `.ocd` which would also be a document. But we really need the `.ocd`, `.sts` and possibly `.ntn`. These shuld really belog in one document, which begins to sound like *Literate Programming*, as the author has experienced in the LaTeX world, especially the `doc` package. This did not look as if .t work well fro CDs, though.

Showed an example of `.etx` files. The source i

Note — there is Unicode text embedded n the formalconent. Getting LaTeX to generate valid XML is nontrivil, but doable, and he has done much of it.

`.ocd` Obviously

`.sts` Does anyone use this — MK does.

`.xslt` as used in `www.openmath.org`. We need a documented API for the utility functions.

`.ntn` Clerly useful, but what format? Is the version in a MathML-3 draft[1] definitive. There seem to be several formats. CL — there is nothing authority.

* It was noted that there is some functional overlap between the `.xslt` and `.ntn` files. The `.xslt` files do not support multiple notations. MK's group has a version of `.ntn` in OmDoc that supports multiple notations (and a thesis in the area), but this is evolving.

There was a debate about client-side/server-side XSLT. Client-side is apparently XSLT-1, but it could be done on JWK's server.

Best practice?

`CDURL` What to do?

`CDBase http://www.openmath.org/CD`?

`status` Should we use `experimental`?

`.ocd` tends to use CamelCase, but LaTeX tends not to. This was not felt to be a major issue.

JWK PopCorn?

A What to do the whole thing in LaTeX.

MK We're done this in STeX.

A This is different — I am not on the formality diagonal in the STeXpaper.

5.2 Lurch — Carter

Lurch is a free word processor whch can check proof at the "introduction to proof" level. Note that there is no proof assistant — the student does this. See 3.4. No 2-D typesetter, but this matters less here than it would do in a calculus course. One kind of 'bubble' is a 'context', e.g. each exercise is a different one. The OpenMath is constructed from the "meaningful expression" bubbles. Each type of ME has its verification algorithm.

[1] Not in final standard.

Declarations Must be not already declared, and not have a reason.

ME+reasons lots of clauses, for example "one reason only", "reason must exist and be citable", etc., but these are *not* rules of logic.

The rules of logic are coded in Lurch documentts, either directly or "imported", via dependencies. Premises can be "explicitly cited" or "recent". In the example shown, "suppose" is just words, but "contradiction" is a keyword in the logic defined in the sheet.

There are "smart rules" (the whole of propositional calculus, possibly computer algebra), but these are only active if imported by a

Q Foolproof?

A In theory a student could create a worksheet with invalid rules, and then get this to "pass the test". In practice this is too hard for students, and can be solvedbyre-running the sheet in the teacher's browser.

5.3 n-ary quantifiers — Hellström

20th-century mathematics has been dominated by First-order logic and set theory. But the 19th-century was more liberal, such as "most continuous functions are nowhere diferentiable":

$$Q_{most} f.[\text{ContinuousFunction}(f), \text{NowhereDifferentiable}(f)]. \tag{5.1}$$

In general, we use as weak a logc as possible, and do the work elsewhere. But we don't need to. Monadic Second order Logic provides "sets of atoms" without $\in$. There are also logics whch can give transitive closures without quantifying over relations.

OpenMath cannot directly express (5.1), and logicians have *proved* that there is no way to express Q_{most} using only single-body quantifiers.

Proposal 1 *Insert* `zeroOrMore` *around the last* `<ref name="omel">` *in the Relax NG for* `OMBIND`*.*

Q Why zero?

A–FR I use it! In fact the code gets simpler.

But Use a dummy combiner `tuple` would suffice.

A Most logics make to with two non-terminals: term and formula. Having `tuple` would add one, without a clear meaning.

NB My personal first use wouldn't be them, but "$\sum$ with conditions" etc., as have been proposed [DK09], would be relevant.

5.4 OM Business Meeting

5.4.1 Formal Business

Chair Kohlhase chaired the meeting (proposed JHD, Seconded BM).

Minutes JHD ended up taking the notes (proposed BM, seconded MK). Minute-checkers: CAR and DPC.

Annual Formal The corporate meeting had taken place. The last business meeeting was at Bremen in 2012.

Finances No transactions.

Membership JHD/MK proposed Lars Hellström. CL/MK proposed Nathan Carter. MK/CAR proposed Petr Sojka. The Executive would discuss these.

Executive Mike Dewar wishes to stand down as Vice-President. MK found it anomalous that JHD was not in the Committee. He would also propose David Carlisle for membership. These two were elected unanimously.

Meeting Attendance at this OpenMath workshop was gratifyingly high, especially in light of the clash with MathUI. The speakers were thanked, also CL for the proceedings — `http://ceur-ws.org/Vol-1010/`.

5.4.2 Developments

MK listed some change suggestions.

1. Better rôle system (MK/FR)
2. n-ary binders (see Hellström's second presentation)
3. first-class sequences (Horozal/Kohlhase)
4. first-class records (Kohlhase)
5. flexForm CDs
6. Notation Definitions
7. DefMPs
8. Document/develop CD writing tools (see Hellström's first presentation)
9. Recognise Content MathML as an encoding.
10. Bug reports

He therefore proposed a Standards Extension Committee, to meet and deliver an opinion *before* the OpenMath Meeting 2014 (at CICM 2014 in Coimbra). This proposal was carried.

DPC noted that MathML was proposed by W3C to be "elevated" to ISO standards, with the process to start in 2013 already. This might impose some time constraints, though it was not totally obvious precisely which these were at this stage.

Membership MK, DPC, JHD, LH, CL, CAR, JWK. the Committee was given the power to co-opt.

Working The Committee should work via an **open** mailing list, and Skype calls.

5.4.3 CD issues

Various issues about the management of CDs had been discussed above.

Statistics JWK noted that this was weak.

Probability CAR noted that we haad none here, and this impacted negatively on statistics.

Q What if a CD depends on OM3 features? Good question —

Roadmap LH suggested a roadmap of CDs - which areas of mathematics (defined by MSC?) were covered, which were weak, etc. This was felt to be a good idea.

5.4.4 Infrastructure

Website Currently at DFKI with PL, but PL is not at DFKI! JWK volunteered Eindhoven for this.

SVN is currently at DFKI. This would also move to Eindhoven, with a backup in Bremen. There was a debate about SVN/Git — it was felt that SVN was adequate, and we could stay with that.

DNS for `openmath.org` — currently is in Helsinki, and will stay there.

TRAC for the Standards Extension Committee. This would be in Eindhoven, linked to the SVN.

Chapter 6

10 July 2013 — Others

6.1 InkChat: A Collaboration Tool for Mathematics — Hu

This is a whiteboard system (common style, these days). Microsoft oneNote, for example, does not allow one to write (talk?) and draw at the same tme. We therefore believe in multi-modal input. We would like record/rollback facilities. See also 6.6.

6.1.1 Prior Art

QuickSet US Navy, voice and pen input.

Classroom 2000 Does not support real-time collaboration.

InkBoard Real-time, but based on Windows XP.

Electronic Chalkboard

Regni/Watt A mixture of C# and Python, depending on platform. Less portable than one would like.

Hu/Mazalov/Watt InkML.

We are based on the last.

6.1.2 Current work

We note that most systems support pen and voice input, and that this multimodality is useful. InkChat's user interface is fairly standard, but attempts are made to minmse pen movement — harder to click a button. Built on the Streaming Digital Framework, so gives Windows(8/7/Vitsa/XP)/Linux/MacOS 10 portable. A Java Application. InkML means there's Microsoft Office compatibility.

For communication currently use Skype and GoogleChat, but extensible. Conference mode is supported if the underlying medium does. Working on Google Hangout implementation.

Q Collaborative efforts — multiple speakers/actors.

A Yes, not easy.

6.2 Semantic Enrichment for Mathematical Expressions and its application to math search problems – Nghiem

Motivations: automatic calculation, semantuc search, automatic recommendation. Three phases to my research: done frist, and third is necessary to make much progress on second.

Relation Extraction between Math and Text

Math Search Expression

Semantic Enrichment

Example: $\delta(n) = 1|n = 0$ has a MathML-P tree. What is the MathM-C? Used a Machine Learnning Approach.

1. Reorder tree to make operator come first
2. SMT — rules with probabilities extracted from SMT training data. `mo:=` $\rightarrow$ `ci:eq` etc. Use a Support Vector Machine to drive disambiguation.

   Two evaluation sets: Wofram Function site (136,629 expressons, seven subclasses), and archives of the ACL Corpus (729 expressions). SVM-1 uses P-MathLL and text, SVM-2 P-MathML only. Also SMT. For each classes, results are pretty close, but SVM-1 wins by a whisker in most cases.

Q Why SVM-1 and SVM-2 almost equal? You said "few textual features".

A Not a very clear answer.

Q Maybe you should design better features for text. EuDML has shown that even statistics of symbol distributions are sufficient to cassify papers

A

Q Reordering? How would this work for non-binary operators such as mod.

A Yes — these are hard.

6.3 Advances in Cylindrical Algebraic Decomposition — Wilson

Define CAD, describe projection/lifting. Note that MetiTarski [AP10] passes hard problems to CAD. Formulating the problem is key [BDEW13]. [McC93, Str00] produce full-dimensional CAD, extended to layered CAD. If we have an equation, and are only interested in that being true **and** others, then use layered CAD to give Manifold CAD.

6.3.1 Adjacency

This is a key problem, and hard to define. Showed 2D CAD for Piano-Movers, and shows that there is a vast amount of "junk".

6.3.2 Goals

Would like the equivalent of Coq tactics,

Q Can you take advantage of symmetry?

A I don't know how do?

6.4 Machine Learning Techniques — Evans

Mathematics is inherently two-dimensional. There is progress in recovering `\sqrt{\frac{n-1}n}` from $\sqrt{\frac{n-1}{n}}$, but this is not enough. But how about $(0, 1)$. What I want is "open interval from 0 to 1" — but it might be Cartesian coordinates[1].

6.4.1 Problems

Compound symbols — multiple connected components, as in $\leq$, or even i and $=$.

6.4.2 Hidden Operators

As in ab or $2x$: often implicit multiplication, but also function application or addition $1\frac{1}{2}$. Can be word buidlng or number building. These combine, e.g. $\sin x$. These can sometimes be disambiguated by font descriptions.

[1] SMW: clearly "the gcd of 0 and 1".

6.4.3 Expression Graphs

Each connected component of the expression is a node, and initially fully connected. $E_{\phi_1,\phi_2,\theta,d}$ where $\theta \in \{1,\ldots 8\} = \{N, NE, \ldots, NW\}$ and $d \in \{1,2\} = \{near, far\}$.

SMW Suppose it's NNE — what do you do?

A This becomes noise in the training data.

MK What is your training data?

A ArXiV — at the moment I don't need the semantics, since I am looking for the intra-relations in the formulae.

Then use graph-cutting to group "common" edges.

6.4.4 Latent Dirchlet Allocation

For grouping documents — originally in the text field. Also note SMW's work [Wat08] on allocation by symbol count.

6.4.5 Language Analysis

Used in machine translation. Can we use this to correlate mathematical expressions and the surrounding text.

Q–MK Look at other peple doing classification — they've tried but haven't succeeded!

A Noted.

Q–APS Note that `$v$` and `$\nu$` often have identical bit pattersn, so there are intrinsic limts to machie learning.

A Noted.

6.5 Offline Segmentation and Recognition of Handwritten Mathematical Notes on Whiteboards — Saroui

offline means without timing or pressure information, hence harder. Trajectory recovery attempts to recover the online information. This has also been used in signature recognition. One problem is "hidden endpoint" — where does a calliraphic 'd' start?

Why whteboards? Widely used in academia, and tends to be sloppier than paper. Colour information has been largely ignored, but there's no excuse today.

Connected components are still an issue — see previous talk. I group "small components".

Stroke analysis includes colour analysis of the entire stroke (seems to depend on characteristic of whiteboard pens). ML based on relative width and HSB gives 90+% correct analysis of direction. Various rules for grouping [ChanYeung].

6.6 Pen-based collaboration tools for Mathematics — Hu

Pen-based input is 3 times faster and less error-prone than keyboard [AYK05]. Showed InkChat (see 6.1). For each symbol, we associate "special points" to determine, e.g. baseline. Showed various $P9$, P_9, Pq and P_q. Note problems of handwriting, device resolution etc.

Database of 240 symbols, 382 classes (JHD would say "styles", e.g. ϕ and φ) and 64944 classified samples. Trying to resolve this caused us to split some classes (now 421). Error rate is now 2%. Some samples are far from the 'norm', so used a homotopy method.

[HuWattICFHR2012] — Efficient Proceessing of Digital Ink Data. Can we achieve minimal error for a given number of points, or achieve a given error with fewest points. Good results here, apparently.

6.7 DML Panel

Panelists: Thierry Bouche, Patrick Ion, Wolfram Sperber, moderated by Petr Sojka. The theme was as follows.

> There are continuation efforts ongoing with the dream of WDML in mind: EuDML has finished its initial phase, preparatory project granted by Sloan foundation, and there are other DML projects maturing (RusDML, Euclid, CEDRAM, etc.) Views how to proceed differ. Some decisions were discussed at DML panel last year. It is time now to stop for a while, and ask for feedback from the community.
>
> **aa)** Have running projects as EuDML, Euclid or preparatory WDML (Sloan) met community expectations?
>
> **ab)** What are pros and cons?
>
> **ac)** Which services are good and which are lacking?
>
> **ba)** How DMLs relate to established services as MR and ZMath?
>
> **bb)** Do they complement their service?
>
> **bc)** What are the prospects of their future interactions?

c) What is the current view of WDML services and how should sustainability and curation be secured by the community, if any?

TB thought the main success of EuDML was breaking the "project barrier" whereby no [national] project talked to another. It also tested in the "real world" some MKM technology. No wthat the project is finished, we are trying to build an ongoing community. EuDML have full text (PDF) for 225,000 items, with corresponding metadata.

WS Apologies from Chief Editor of Zbl. There are questions what Zbl can offer as a free service.

PI EuDML is clearly a success, and Sloan showed so.

MK Can you shed more light on the 'one stop shop"? I see it has 148 users[2].

TB The majority of accesses to NumDam come via EuDML, as EuDML is better indexed in Google Scholar.

MK Is this site mashable?

TB No. That's the next project.

JHD This project's description of itself is the worst advertisement the WDML community could possibly have.

MD Without dissenting from JHD's evaluation of the surface presentation, the achievements of the project are wonderful.

TB One slight failure is smooth input of proceedings. I think the project is 60% successful.

APS I disagree. The core aims were 100% achieved. The contents of the site are mashable (the difference is that EuDML can't host the mashup). Some bits, like accessibility, have moved forward but aren't great successes.

MK I note the success of ArXiV: how can we do the same.

PI JSTOR has a different economic model — note that it was set up by economists.

WS Not sure that ArXiV s a model — libraries have to pay.

PI MR has not noted (nor been really looking for) any impact of EuDML, either positive or negative.

WS Zbl similarly.

[2]This is registered people, not IP addresses etc. The main point of being registered seems to be the ability to add one's own MSCs, but the utility of this is questionable.

6.8 DML Business Meeting

6.8.1 PS's Report

DML at CICM 2013 was a success. Number of submissions significantly up.

We need to appoint a DML Track Chair for CICM 2014, notng that SMW is the Overall Chair. JHD proposed and MD seconded, PS for Track Chair.

MK called the process into question — if we are not careful PS will be perpetual Track Chair.

6.8.2 Future Management of DML

It was observed that, unlike MKM and Calculemus the other subject tracks of CICM this year, DML had no formal management structure. Conversely, AISC has no structure either, being run as an absolute diarchy.

SMW proposed that PS continue as Chair until such time as he was tired of this, when it would be up to him to find a new solution. This was carried. **PS**

The meetng conveyed its thanks to PS.

Chapter 7

11 July 2013

7.1 Social Machines and the Production of Research Mathematics — Martin

1.1 described a major research effort, this paper is about a hypothesis. Showed Fermat's original scribble in Diophantus, and the opening page of [Wil95]. Also an interview with Wiles, noting the large amount of "slog". Use of Mathematical Reviews (many other disciplines don't have the equivalent), computation etc., either Maple etc., or OEIS [Slo03]. Lots of informal collaboration tools: Group Pub Forum, mathoverflow etc.

> "There is no ... mathematician so expert ... as to place entire confidence in his proof immedi- ately on his discovery of it... Every time he runs over his proofs his confidence encreases; but still more by the approbation of his friends. — [Hum39]

See also [Lak76, Had54].

Did a survey of Mathoverflow: 36% conjecture, 28% what is this, 22% other, 14% examples. Classics that don't get captured in the literature are "X doesn't work because of Y". 90% of questions asked have "moved forward" because of the answers. 56% of people share references.

Polymath is a Gowers[1]/Tao experiment. A straightforwardblog, but relying on the fact that blogs *can* now do mathematics. Good style rules. Note that it "needs" a choreographer, partly because there is no more sophstication than blogging software. Used as an example discussion of a GCS Olympiad question. 14% is (stuff that will end up in the) proof, 33% examples, 10% concept, 20% conjectures and the rest "other" — largely social glue. Has a time-evolution graph. Brief summary: Lakatos was right.

We don't have much computational support for concept formulation, or dealing with error. The systems described above are "designed" by, and, more

[1] Gowers does talk about "the difficulty of navigation".

importantly, used by, practicing mathematicians. There are attempts to do "mathematics as as e-science". Note that other communities worry much more about "accountablity, provenance and trust" issues than we do. Mathoverflow etc. users can be anonymous. OEIS requires "justification".

7.2 Structural Similarity Search for Mathematics Retrieval — Kamali

Mathematical expressions are objects with complex structures and rather few distinct symbols and terms. $\int x\sqrt{x^2+a^2}dx$ and ... are very different expressions even though they have the same symbols and (MathML) operators. Our starting point is (presentation) MathML. We ignore the text, and just look at the expressions. Result is "top k documents".

A standard approach is "tree edit distance": number of insert/delete/rename with costs of 1/1/2. Normally calculated via dynamic programming, but this is expensive. Hence want to discard "distant" expressions quickly, and also make use of repetitive sub-expressions. There is a trivial upper (? I think he means lower) bound on edit distance, so first compute this and use "top k selection".

7.2.1 Repetition

Use expressions from Wikipedia and DLMF. Expressions of length 1–5 have 325 repetitons, 6–10 have 10.5 and so on. Use a hash function to assign signatures to subtrees, and use these to resolve duplicates. Once we know this we can cache the dit distance between subtrees. In practice don't bother for very small subtrees, or very rare subtrees. In fatc, has a theory of "memoization with Memory Constraint"

7.2.2 Experiments

Two sets of queries: "interview" and "mathematics forum". Three measurements

MRR Mean Reciprocal Rank of first correct answr

NFRR Non-fail retrieval rate.

Time (elapsed)

ExactMatch as 13% NFRR (87% failures) but MFRR=1. SimSearch has NFRR 10% and MFRR 0.74. But unoptimised search takes 50 seconds, whereas best optmised ones 0.2 seconds, or 1000ms with memory constraints.

Q–MK Why not in the search engine evaluation earlier this year?

A Informed late — MK denies this.

Q What about sub-expressions?

A Some of our algorithms do?

Q–PS Your smilarity is a metric space, and there is known technology here, e.g. in image search.

A

7.3 A hybrid approach for semantic enrchment of Mathemetical Expressions — Nghiem

See 6.2. $\delta(n) = 1|n = 0$ is a typical example. Use SMT — Static Machine Translation. First re-order the P-MathML tree to be "operator-first" to agree with C-MathML (but this seems to depend on the P-MathML being "well-behaved" — problems with nested `mrow`s). Our training data is aligned, and only has `mi` issues.

Q–PL Other languages?

A If we had the training data.

7.4 Three Years of DLMF: Web, Math & Search — Miller

Slightly different interpretaton of "Library". Bib;iometrically, over the period 1974–1995 [AS64] was growing more than the total number of citations! [Nat10] came out in 2010. Comments mostly positive: of the negatives

1/2 misunderstood

1/3 technical glitches

1/6 errata.

6 minor updates and 21 errata. [Nat10] has 10% of the total [Nat10]+ [AS64]. Some citations are just to the web site: people don't like citing on-line resources.

Log analysis shows 150M hits, but 1/2–2/3 are from robots. Caching is also a problem (probably less for us than for other sites). daily visits:

595 total visits/day

389 Permalinks

313 Presentation MathML

303 TEX

44 bib

or the 2010 delivery we did browser sniffing XHTML+MathML (1 HTTP request) for Firefox, IE w/MathPlayer, Cookies, and HTML+Images (typically 25 HTTP requests) for everything else. Hence non-MathML browsers look 25× more popular! After allowing for this, over half the visitors (to this site!) use a MathML browser. half the rest have "mostly MathML" — e.g. Safari which is MathML without prescripts, which is pretty useless for special functions.

7.4.1 Search

Text-based. We attempt to decipher unstructured queries. 30 queries/day. Text searches (83%) have 1–10 terms, math searches (17%) more, but the behaviour is pretty similar. Some people (10%) do use wild-card matching. Also some LaTeX searches, including `\hbox` etc. A large number of misspellings of "Whittaker" show up in the logs! About 40% of searches result in a click-through to one of the (top 10) results.

7.4.2 Conclusions

The outlook for scientific publishng on the web is mxed. We'll move to HTML5 with MathML. Probably MathJax. We're looking at DLMF "tables on demand" — 50% of [AS64] was tables. Looking at WebGL.

Q–APS How far are you from Content MathML, and why isn't it on your list.

A The macros are half-way semantic as it is. It's on the long list: I would want to do it.

Q Did you detect the disappearance of MathML from Chrome?

A Not realy — we never actually served MathML to Chrome, since it has limitations.

Q–SMW Is Text/Math ratio a reflection of quality, or of what people want?

A Not sure.

Q–UHM What user investgation have you done?

A Not much. We haven't really looked at 'refers'. We will be getting Google Analytics information soon.

Q–PI Chrome's dropping MathML was a bizarre side-effect. We don't know how much people *want* formula search.

A True.

7.5 Escaping the trap of too precise subjects — Libbrecht

We tend to assume a a definedvocabulary, but this is a trap, and we tend to miss resources. One example is MSC 2010: note that "complexity" occurs in many places.

`i2geo.net` is a tool to share resources. Proposed solution "escape ladders " — suggesting other queries.

Q Is it a mistake to insist that taxonomies be trees?

A Possibly, but the alternatives seem to be worse.

Q Do you use all language versions of Wikipedia

A Yes. The ambiguties users have are visible in the correspondng Wikipedia.

7.6 Using MathML to Represent Units of Measurement for Improved Ontology Alignment — Do

Our group focuses on sensors, especially in energy, e.g. smart grids. COSM is a web site for sensor data. There are various symbols for temperature in it, such as $\circ C$, "Celsius" etc.

Ontologies: `subClass` and `Instance`. OWL. Ontologies are used when there's need for consistent unambiguous usage of terms and concepts. The BBC uses ontologies to organise news items. An alignment is a set of correspondences $C = \langle e_1, e_2, r, c \rangle$ where c is the confidence.

Alignments can be contructing by lexical matching: propcoressing, distance (Levenshtei, Jaccard etc.), structural comparison (typically similarity flooding). These are very general approaches, which suffer from term variation; automoble/motor car, also homonyms, distance measurement, how to do structural comparison.

We would like to do this in the area of units, where we might have equations like $F = ma$ and can solve for the unknowns. Considered QUDT (NASA), SWEET (also NASA!) and OM (Dutch). Generate the MathML to 'improve' the ontologies. Showed `mm/day` in OM, which is defned as a quotient. Preecision is around 0.8, recall is 0.97 and the F-measure is 0.86, which is good by industry standards.

Why false positives? Hertz versus Becquerel (both T^{-1}), also ontology errors.

Q–JHD How dod you deal with relative/absolute temperatures, as in [SD08]?

A Not our problem — the ontologies describe them differently.

7.7 Towards Machine-actionable Modules of a Digital Mathematical Library — Sojka

EuDMLhad these goals:

- To master the technology ...
- concept of moving wall to motivate and engage commercial publishers ...
- To collect data from existing local or publishers databases ....

Aggregation in EuDML is via the standard OAI-PMH protocol. EuDML metadata schema heavily based on NLM (U.S. NIH product), extended to allow Math-awareness (MathML and LaTeX).

Workflow varied: sometimes from paper, sometimes PDF, DVI or TeX, and getting metadata from the paper was difficult. Birmingham's MaxTract detected the mathematics in PDF. Lisbon's REPOX remapped metadata formats to a unique representation. ParsCit will parse citations. Negotiations with Google Scholar have resulted in much better visibility. Did a Latent Dirichlet Allocation for topics.

While there is a lot to be done, we can conclude that automation is feasible, and large-scale processing possible.

Chapter 8

11 July 2013 MKM

8.1 Determining Points on Handwritten Mathematical Symbols — Hu

By a "determining point", we mean one that defines a feature such as basepoint. Consider p, where the determining point is the bottom of the circle. Consider Pq P_q $P9$ and P_9.

- Manua annotation is not feasible
- Many symbols
- People's handwriting varies
- Digital ink data depends on device resolution (digital ink is a sequence of points)).

Represent a character as a sum of orthogonal polynomials (degree 12). This lets one talk about an "average symbol" from a set of handwritten samples. We mnaually annotate the average symbol. Mislabelled samples and classes that should be split account for some of the error rate — fixing this drops the error rate to under 2%. Even the homotopy method (see 6.1) fails on some badly-written examples.

Q This reminds me of `detexify`.

A–Hu I don't know of that.

A–SMW trying to avoid device dependence, so very different.

Q Neatenng bad handwriting?

A Yes, we have an application for this.

8.2 Understanding branch cuts — England

Now in Maple 17 — thanks Edgardo. Many mathematicians recommend Remann surfaces, but these are not used in computer algebra. Very similar identities have different branch cuts, and the induced topology is different. Maple's Function Advisor has great aims, and will produce branch cuts for the building blocks, and do the rest by reverse engineering.

We move from z to (x, y) and use semi-algebraic sets, and UWO solver. can cope with $f(p(z))$ where f is a polynomial. Can combine functions, but issues with cancelling cuts. Otherwise use denesting to solve $f(r(\ldots))$ where r is radical.

Parametric method as in [DF94]. Often quicker, but the output is less intuitive.

Two sources of spurious cuts:

- Formulation — when the input as posed has cancellation;
- Denesting cuts.

This is a major contribution to the agenda in [BD02]. [BDE$^+$13] for TTICAD helps with the cylindrical algebraic decomposition induced by branch cuts.

8.3 Students' Comparison o ftehir Trigonometric Answers with the Answers of a Computer Algebra system — Tönisson

Macsyma's answer (say) does not have the naunces required in schools. Consider [(ed99].

- First-year university students
- Course "elementary mathematics" — somewhat repetitious course for school mathematics
- 90 minutes, of which 70 were actual working
- students in pairs (26 pairs), with discussions audio-taped
- students first solve by hand (3 equations), then given a CAS (Maxima, Wiris, Wolfram Alpha), then analyse differences.

Note that our goals are different from STACK ([San08]), but that is also interesting. See also [Dri02].

Degrees/radians is one issue, also challenges of presentation. Consider: "solve $\sin x = m$" — in Estonia we say $x = (-1)^k \arcsin m + k\pi$, but many people/cultures have a case statement or set union.

1. "$\sin(2x+2) = \frac{\sqrt{3}}{2}$" with Wolfram Alpha. Estonian School solution $x = -\frac{1}{2} + (-1)^n \frac{\pi}{12} + \frac{n\pi}{4}$. 12 out of 17 pairs used $(-1)^n$ form. The discussions around the $(-1)^n$ form were sometimes helpful.
2. "$\tan^3 x = \tan x$" with Wiiris.
3. "$\cos\left(x - \frac{\pi}{6}\right) = 0.5$" with Maxima. Good discussions.

Q Prior knowledge of being taped?

A Yes, and knew it was anonymous research.

Q How do you think this will influence their use of technology?

A In future more sceptical, their teachers say.

Q Might this extend to peer marking, without needing computer algebra?

A Probably.

8.4 A Qualitative Comparison of Four Theorem Provers for Basic Auction Theory – Lange

Birmingham authors plus four prover experts. The ideal theorem-prover woul dhave libaries like Isabelle/Mizar, textbook-like term syntax as Theorema etc.

Auctions are mechanisms for allocating goods. The challenge is to define the right auction for the aim. European 3G spectrum got between €20 and €600/capita — not clear why. Not all auctions are well-defined (in the sense that the outcome is unique), or at least it is NP-complete to show well-defined.

We need an expressive language to descrbe the auctions.

Theorem 2 (Vickrey1961) *In a second-price auction [sealed bids: winner pays second-highest price], "truth-telling" is a weakly dominant strategy. Furthermore, the auction is efficient.*

First theorem of modern auction theory. Actually have a new (non-constructive) proof, initially wth nine cases, then reduced to four.

Mizar FOL+set theory. Proved.

Isabelle/HOL Proved.

Hets/CASL/TPTP We could prove it, but a restructured proof failed.

Theorema 2.0 Used as FOL+set theory. Couldn't finish but not all libraries translated to 2.0.

De Bruijn factor 1.3–1.7 over fully-detailed LaTeX. Need to handle exceptional case of one bidder!

Mizar represented the bids as a relation $\{bidders\} \times \mathbf{N}$, but deeply unintuitive to auction experts:

```
winnerof R equals the Element of R"{union rng r}
```

Fully-automated provers needed a great deal of hand-holding.

Isabelle was again hard to read:

```
second_price_aucton_winner = bunch of ML-like stuff
```

We did find some bugs. Will write TPTP challenge problems. This is only the first theorem — later ones require integrals etc. Also problems with combinatorial auctions — currently working on formalising well-definedness. We are pursuing Isabelle in future work.

Q–CSC I believe that the theorem is true for infinitely-many bidders?

A The Mizar version did indeed have this generalisation.

A–MKe But there is an arg max, so need to respecify!

8.5 MKM Business Meeting

8.5.1 Business

Sacerdoti Coen chaired the meeting, and JHD ended up taking the notes.

8.5.2 Track Chair Report

We didn't duplicate the MKM call against the CICM call, whch may have been a mistake. There was an issue of scope versus other tracks, particularly DML. Seconded the previous comments about the difficulties of having a unified PC, but it was good that everyone could see all bids. He was asked "who was the MKM PC", and has a retrospective definition: those who refereed MKM papers, therefore 26. 18 submissions, 7 accepted, and 2 send to WiP/S&P. (Last year was 13 out of 19).

BM liked the coherent CICM advertising, and wish it had had the workshops in as well.

JC Workshops not known at time of first call.

8.5.3 Trustees

CICM 2014 in Coimbra (Portugal!) 28 July – 1 August. Aiming for 2015 at George Washington University. General Chairs will be SMW (2014) and MKe (2015). JU will be the MKM Track Chair for 2014. We are trying to decide the Track Chair for 2015, so that they can get Trustee experience.

FR will be continued as CICM delegate, and PI as Treasurer. SA and FR are leaving the Board of Trustees (leaving CSC, MW, WMF and Da, with JU joining). APS and SMW nominated BM. PL and MKe nominated AK. There will therefore be an election.

8.5.4 General

SMW noted that CICM had been together for several years, and it was maybe time to "stop dating and get married". Do we still need several boards of trustees? JHD reminded the meeting of the Calculemus decision.

> The meeting mandated the trustees to investigate abolishing the separate administrative structure of Calculemus, while retaining the academic identity. **Trustees**

MKe suggested one business meeting with speciality "5-minute slots". To continue the analogy, MK pointed out that the conferences already had a joint account. A unified structure still had to have some mechanism for choosing track chairs. This could be done by the current ex-officio trustees.

JHD proposed, and SMW seconded

> The meeting mandated the trustees to investigate abolishing the separate administrative structure of MKM, while retaining the academic identity, and bring proposals to a unified business meeting in 2014. **Trustees**

This was carried $n - 1 : 1$ with $n \gg 1$.

8.5.5 A.O.B.

MK noted that the PC of 26 was a two-edged weapon, and a smaller PC would be more coherent and consistent. MKe also noted that people looked at the PC in the CfP in order to judge relevance. JC confirmed that the opening size of the MKM committee, rather than DA's retrospective definition, was roughly half the size.

MK also noted with approval ISSAC's process under SMW's guidance, of "meta-reviews" synthesising the comments, especially in the rebuttal phase. It is quite likely the the 2014 General Chair will follow this guidance.

FR liked the buffet lunch and small meetings rooms. JWK would have liked to have tables as well.

Chapter 9

12 July 2013

9.1 Mathematics and the World-Wide Web — Ion

Subtitle: a report and some suggestions.

"Mathematics is the Queen and Language of Science" — Gauss.

9.1.1 WWW issues

- accessibility
- Internationalisation
- Security and Privacy
- Intellectual Property — if you are Stallman, you reject the concept.

Note that "The Web" has $2.5 \cdot 10^9$ citizens — half immigrants and half born in it. If it were a state, we would say it has state enemies: China, Iran, Bahrein etc. Estimated "GDP" $4.2 \cdot 10^{12}$\$.

9.1.2 Print Issues

- Earliest form of remote learning
- archival, persistent, easy reproduction
- "how will anyone know anything when they don't need to memorize"
- Reduces rôle of the ndividual teacher.
- Took time — 50 years for page numbers, Leibniz invented indices and keywords.
- Earliest use of Gutenberg's technology was apparently Church bureaucracy.

9.1.3 Start of technology

- Leibniz (again).
- Otlet's vision and UDC — being revived under Google, apparently. Note how modern his design is.
- Vannevar Bush (and his intern Shannon)
- Then digital age ...

Mathematical Reviews now occupies the Ann Arbor Brewery. Database over 3M items. TEX-based since 1984.

9.1.4 MathML etc.

XML vocabulary — classes of rooted labeled planar trees. Essentially pasigraphy. Is it HTML-5, generally being adopted, despite hiccups like Chrome. We expect it to be an ISO standard next year: get your corrections in now!

MSC2010 is RDF/XML labeled graph description of triples: `http://www.msc2010.org`. MSC/SKOS is the definitive machine-processable form.

9.1.5 Graph analysis

- Virginia Bioinformations using MR obfuscated data.
- Network statistics over time, using R
- Apparently there was a 1995 breakpoint
- [BD12] tracing the influence of Soviet influx to USA, on mathematics, and particularly on mathematics employment.
- Other bibliometrics.

Note Google, Wikipedia and disruptive search — "the smartphone disrupts discussions as people keep looking up facts". EuDML etc., also WDML ("≈ Whatever Digital Mathematical Library").

9.1.6 Disruptive Changes

- Hoem pages, preprnts, lectures, videos etc.
- Blogs (Tao, Gowers etc.)
- ePub3

Note the scope for bad actors as well. Plagiarism is certainly more visible, and the knock-on effects of this may be more disruptive than the direct effects. Note thee are new developments, such as Summly, a novel summarization algorithm[1].

[1] Apparently developed by a 15-year old Brit. Now bought by Yahoo.

9.2 Automated Reasoning Service for HOL-Liht — Urban

Example is the Flyspeck project Currently 14,000 theorem collaboration, repetitions etc. Aim to produce a LaTeX bok linked to the formal proof. Our research project is to produce TP (TPTP) problems from HOL Light goalstates. Try ATPs on eay probems (those proved by MESON). We then want to try to rpove whole theorems against the whole library, using Bayes-traned methods and possibly others.

The HOL(y) hammer architecture: complex diagram — backends, integration, clients. We believe we need an AI component to make ATP over very large libraries possibe. Ther are several heuristic, learning and semantic methods developed in the last decade. Presmise selectors preprocess (typically train on) the theoream characterizations and their proof dependencies. They work as raners: ran available theorems by relevance to a new conjecture The best premises produce ATP (FOF, TFF1 etc.) problems.

1. Extract feaures from statements (now 30k)
2. Do feature similarity, with 14–16k examples
3. Needs parse, memory-efficient algorithms
4. Want ones with fast update for new problems.
5. Then goal translation to TPTP, removng nstances of the Hilbert operator.

The union of these methods can prove 47% of the training set. Has an Emacs interaction mode.

- Pervasive parallelism
- lots of OCML

Q Does it depend on amount of trraining?

A It's an iterative process, gradualy improves.

9.3 Capturing Hiproofs in HOL Light

The subgoal package is 30-years old (Larry Paulson) used in all the HOL family. Isabelle and Coq have now largely superceded it, but it is in use in Flyspeck, for example. Shows one example — the first three subgoals have identical proofs.

These Hiproofs are a tree diagram notation for conveyng understanding about a proof. Nodes represent proof tactics. Branches represent subgoals. Main constructors are sequencing and tensor (applying different methods to different subgoals). They can be labelled, and inside a label we can have a further hiproof.

9.3.1 Tactician

My project: a means of packaging proof scripts. This is a process that can take a day. Also HOL Light consists of packaged proofs, and to understand it you have to unpackage by hand (until Tactician). It also outputs the proof as a graph for visualisation. Note that Flyspeck is 300KLOC of ML, so Hales intends using this to tidy up. Works by automatically applying wrapper functions (approx 80) which overwrite existing ML bindings.

9.3.2 HipCam

Stephen's tool. Captures the proof in all its detail. Completely automatic for all proof scripts.Works in the opposite way — traps the HOL kernel functions to capture each application. But the user can explicitly promote somethng so that it gets captured atomically (necessary in practice).

9.4 Formal Mathemaics on Display: A Wiki for Flyspeck — Urban

We want to connect formal and informal mathematics va a Wiki. Formal Mathematics looks like programming, and is unreadable except by experts.

Hence we built the Agora platform. Provides HTMLization of formal proofs (also Proviola), and some editing of formal proofs (prototype for Coq), and a documentation language (Creole/Wiki markup, and MathJax). We have formal HOL Light and Flyspeck documents loaded, and presentations. We have communication between Agora and Hol Light server (keeping a checkpointed image to reduce start-up times).

Showed Flyspeck Chapter 5[2] in Wiki view, and zoom into the formal side. Need a proper hyperlinked for HOL Light (nontrivial!). Agora allows dual formal/informal blocks.

Note that tossing a 300-page math-heavy book at MathJax works! However, we did divide chapter 5 into subsections for faster loading.

CodeMirror is a programmable web editor. We send content ofetn: on-change + timeout.

-
-

[2]Fan. 1981 lines, 15 definitions, 31 lemmas.

Bibliography

[AP10] B. Akbarpour and L.C. Paulson. MetiTarski: An Automatic Theorem Prover for Real-Valued Special Functions. *J. Automated Reasoning*, 44:175–205, 2010.

[AS64] M. Abramowitz and I. Stegun. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. *US Government Printing Office*, 1964.

[AYK05] Lisa Anthony, Jie Yang, and Kenneth R. Koedinger. Evaluation of multimodal input for entering mathematical equations on the computer. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1184–1187, New York, NY, USA, 2005. ACM.

[BD02] R.J. Bradford and J.H. Davenport. Towards Better Simplification of Elementary Functions. In T. Mora, editor, *Proceedings ISSAC 2002*, pages 15–22, 2002.

[BD07] C.W. Brown and J.H. Davenport. The Complexity of Quantifier Elimination and Cylindrical Algebraic Decomposition. In C.W. Brown, editor, *Proceedings ISSAC 2007*, pages 54–60, 2007.

[BD12] George J Borjas and Kirk B Doran. The collapse of the soviet union and the productivity of american mathematicians*. *The Quarterly Journal of Economics*, 127(3):1143–1203, 2012.

[BDE+13] R.J. Bradford, J.H. Davenport, M. England, S. McCallum, and J. Wilson. Cylindrical algebraic decompositions for boolean combinations. In *Proceedings ISSAC 2013*, pages 125–132, 2013.

[BDEW13] R.J. Bradford, J.H. Davenport, M. England, and J. Wilson. Optimising Problem Formulation for Cylindrical Algebraic Decomposition. `http://arxiv.org/abs/1304.7222`, 2013.

[BGC94] Helmut Bender, George Glauberman, and Walter Carlip. *Local analysis for the odd order theorem*, volume 188 of *LMS Lecture Notes*. Cambridge University Press, 1994.

[DF94] A. Dingle and R.J. Fateman. Branch cuts in computer algebra. In *Proceedings ISSAC 1994*, pages 250–257, 1994.

[DK09] J.H. Davenport and M. Kohlhase. Quantifiers and Big Operators in OpenMath. *Draft Proceedings 22nd OpenMath Workshop*, pages 119–125, 2009.

[Dri02] Paul Drijvers. Learning mathematics in a computer algebra environment: obstacles are opportunities. *Zentralblatt für Didaktik der Mathematik*, 34(5):221–228, 2002.

[DSS04] A. Dolzmann, A. Seidl, and Th. Sturm. Efficient Projection Orders for CAD. In J. Gutierrez, editor, *Proceedings ISSAC 2004*, pages 111–118, 2004.

[(ed99] Michael J. Wester (ed.). Computer Algebra Systems, A Practical Guide. *John Wiley & Sons Ltd.*, 1999.

[FT63] W. Feit and J.G. Thompson. Solvability of Groups of Odd Order. *Pacific J. Math.*, 13:775–1029, 1963.

[Gon08] G. Gonthier. Formal Proof — The Four-Color Theorem. *Notices A.M.S.*, 55:1382–1393, 2008.

[GT12] G. Gonthier and L. Théry. Formal Proof — The Feit–Thompson Theorem. `http://www.msr-inria.inria.fr/events-news/feit-thompson-proved-in-coq`, 2012.

[Had54] J. Hadamard. An essay on the psychology of invention in the mathematical field. *Dover*, 1954.

[Hal05] T.C. Hales. A proof of the Kepler conjecture. *Ann. Math.*, 162:1065–1185, 2005.

[Hal08] T.C. Hales. Formal Proof. *Notices A.M.S.*, 55:1370–1380, 2008.

[Hal11] T.C. Hales. The Strong Dodecahedral Conjecture and Fejes Toth's Contact Conjecture. `http://arxiv.org/abs/1110.0402`, 2011.

[Hum39] D. Hume. Treatise on human nature. *Edinburgh*, 1739.

[KRW11] M. Kerber, C. Rowat, and W. Windsteiger. Using Theorema in the Formalization of Theoretical Economics. In J.H.Davenport *et al.*, editor, *Proceedings CICM 2011*, pages 58–73, 2011.

[Lak76] I. Lakatos. Proofs and Refutations. *Cambridge University Press*, 1976.

[Lan65] S. Lang. Algebra. *Addison-Wesley*, 1965.

[McC93] S. McCallum. Solving polynomial strict inequalities using cylindrical algebraic decomposition. *Computer J.*, 36:432–438, 1993.

[McC99] S. McCallum. On Projection in CAD-Based Quantifier Elimination with Equational Constraints. In S. Dooley, editor, *Proceedings ISSAC '99*, pages 145–149, 1999.

[MHC10] K. Ma, S.C. Hui, and K. Chang. Feature extraction and clustering-based retrieval for mathematical formulas. In *Proceedings Software Engineering and Data Mining*, pages 372–277, 2010.

[Nat99] National Aeronautical and Space Administration. Mars Climate Orbiter: Phase I Report. `ftp://ftp.hq.nasa.gov/pub/pao/reports/1999/MCO_report.pdf`, 1999.

[Nat10] National Institute for Standards and Technology. The NIST Digital Library of Mathematical Functions. `http://dlmf.nist.gov`, 2010.

[Pet00] Thomas Peterfalvi. *Character theory for the odd order theorem*, volume 272 of *LMS Lecture Notes*. Cambridge University Press, 2000.

[Pow09] M.J.D. Powell. The BOBYQA algorithm for bound constrained optimization without derivatives. Technical Report University of Cambridge DAMTP 2009/NA06, 2009.

[RK13] F. Rabe and M. Kohlhase. A scalable module system. `http://www.sciencedirect.com/science/article/pii/S0890540113000631`, 2013.

[San08] C.J. Sangwin. Assessing Elementary Algebra with STACK. *International Journal of Mathematical Education in Science and Technology*, 38:987–1002, 2008.

[SD08] J.D. Stratford and J.H. Davenport. Unit Knowledge Management. In S. Autexier *et al.*, editor, *Proceedings AISC/Calculemus/MKM 2008*, pages 382–397, 2008.

[Slo03] N.J.A. Sloane. The Online Encyclopedia of Integer Sequences. *Notices A.M.S.*, 50:912–915, 2003.

[Str00] A.W. Strzeboński. Solving systems of strict polynomial inequalities. *J. Symbolic Comp.*, 29:471–480, 2000.

[Wat08] S.M. Watt. Mathematical Document Classification via Symbol Frequency Analysis. In S. Autexier *et al.*, editor, *Proceedings AISC/Calculemus/MKM 2008*, pages 29–40, 2008.

[WBD12] D.J. Wilson, R.J. Bradford, and J.H. Davenport. Speeding up Cylindrical Algebraic Decomposition by Gröbner Bases. In J. Jeuring *et al.*, editor, *Proceedings CICM 2012*, pages 279–293, 2012.

[Wil95] A. Wiles. Modular Elliptic Curves and Fermat's Last Theorem. *Annals Math.*, 141:443–551, 1995.