CICM 08
Conferences on Intelligent Computer Mathematics

Notes by JHD

August 31, 2008

Part I

Digital Mathematical Libraries 2008

Chapter 1

27 July 2008: Morning

1.1 Petr Sojka — Introduction

He saw the objectives as producing a report on the following questions.

- What technologies, standards etc. should be used, and what metadata should be shared?
- What business models are suitab,e for publishers, authors and funders?
- Is there a sustainable model?

He drew attention to Thierry Bouche's talk at MKM on Monday morning, as well as many of the talks in today's workshop.

1.2 Thierry Bouche: CEDRICS: When CEDRAM meets tralics

1.2.1 The CEDRAM Project

A project to strengthen academic journals. Principally aimed at supporting independent or society publishers. Although centred on tose supported by CNRS, it is open to European partnews. The aim is "to be DML-ready". `http://www.cedram.org`. CEDRAM has the following principles.

- Published articles are original, scientifically checked, research papers.
- Once published, content is unchangeable
- Navigation is free, and metadata is shared openly
- Access to the published articles is subject to a reasonable moving wall (MW). Currently CEDRAM supports 3 open access journals (MW=0), and the maximum MW for the current journals is five years.

Various tools have been produced.

ROUCHE An easychair-like editorial tool.

...

CEDRICS The production tool.

1.2.2 The CEDRAM Production Environment

1. An accepted article is prepared by the journal staff/subcontractors.
2. When an issue is ready, the journal transfers all the required elements to CEDRAM.
3. CEDRAM returns "electronic proofs" to the journal.
4. When agreed, these are exploited via the journla website and published.
5. Annual archive volumes are prepared.

Challenges: jurnals with low resources, hence use of simple standards such as LaTeX. Even compulsory BiBTeX has proved too difficult. But, we want metatdata compatibility with NUMDAM, OAI-PMH etc.

Existing journals had their own house style, so CEDRAM must not trample on these. *But* ther should be a homogeneous interface to users: an interesting compromise.

The journals CEDRAM typically deals with have *one* secretary, who is the interface. CEDRAM outputs PDF proofs, and versions for print and screen, but also metadata as LaTeX snippets. TRALICS converst this into text (UNICODE) and into XML. There is a problem with those medatata that do not appear on the proofs, and so we need a 'proof' websote on which this appears, and authors/journals have to chekc this as well as the page proofs.

The CEdRAM process typically takes half-an-hour: there are always glitches somewhere. It is driven by the following principles.

1. Any medatata is input once into the system.
2. Anything that is not determined by the relevant file's content should stay away from that file, e.g. volume number.
3. Anything that *can* be computed, *should* be computed.
4. Do not reinvent the wheel.

The model is a three-layer one.

$$\text{journal} \ni \text{physical volume} \ni \text{article}.$$

"physical volume" is really more like "issue".

`cedram.cls` is an extension of `amsart.cls`. Hence `cedram-JOUR.cls` The volume file defines numbers (year, month,volume, issue), opening page number, ordered list of articules, and any extras. An article is a directory containing all the data called in to one master file. This declares the *article-level* metadata, with higher-levle metadata being imported.

1.2.3 Layout Versatility

He noted that the *whole* volume is generated: cover pages, table fo contenst etc., and showed the same issue re-formatted for a different journal to demonstrate the versatility.

A page number is a computed artefact.

Annales de l'Institut Fourier, for example, has titles and abstracts in both Englist and French.

1.2.4 Tralics

This is sofware which reads LaTeX and outputs XML.

- A full TeX interpreter.
- A versatile translator from TeX.
- Is used to pruduce the HTML for NUMDAM.

This means that metadata can be input by authors/journal staff in a convenient format. He demostrated an *Annales de l'Institut Fourier* from 1992 (i.e. old LaTeX re-interpreted) with displayed mathematics in the metadata (abstract). This technique has been used to produce the 50^{th} anniversary items for *Annales de l'Institut Fourier*.

1.2.5 Q and A

Q. Why not TeX-for-HT?

A. I tried, but failed. I did look at the concept of using a full TeX engine, and that's why Tralics has a *full* TeX-interpreter. However, there are other converters which I have not tried.

Q. Is it just the metadata that goes into XML?

A. At the moment, yes. Converting a full paper into *production-quality* XML is a major exercise.

Q. What about external references?

A. *If* you know the link, then you can link to it by standard techniques.

Q. What are the licencing conditions for the *sources*? I could convert these *if* I had them.

A. The MW policy applies to the PDF, not the sources. The sources are (apparently: JHD wasn't clear here) owned by the journals.

1.3 J. Rákosník — From pixels and minds to the mathematical knowledge in a digital library

1.3.1 Motivation

- Digital initiatives are world-wide, but following different paths with little cooperation.
- Are we building a new tower of Babel?
- The major problem is money — several failed EU-level attempts, but the CZ Academy of Sciences has an Information Society programme. Hence the DML-CZ programme.

The fundamental target is a user-friendly digital approach to all mathematics literature published in Czech lands (Czechoslovakia, Austro-Hungarian empire etc.). This is about 200,000 pages (90,000 are currently displayed, with more in process).

The project is largely conducted by students, which is both a strength (manpower) and a weakness (continuity). Collaboration with NUMDAM, CEDRAM, Göttingen and CEIC.

The project is run by the Institute of Mathematics of AS CR, with J. Rákosník as co-ordinator. Also Brno (Computer Sciences, Informatics), Charles University and Library of AS.

Primary aim is research journals, but there is an intention to include journals for teachers, society bulletins, and maybe also proceedings series (open-ended topics) and even some monographs.

Journals are *relatively* simple to obtain, the quality is generally good. The journals are *very* multi-lingual: 7–8 languages[1] are common. There are four types of source.

1. paper digitization (600 dpi, 4-bit depth, which helps with transformation of images via BookRestorer).
2. bitmaps from Göttingen (400 dpi and bitonal). We noted that the error rate in OCR with 400 dpi is significantly higher.
3. born digital.

The Czech Copyright Act says that a digital copy of an article is a *new* original. We negotiate with publishers (who may have to deal with authors in the case of monographs).

[1] In answer to a private question from JHD, JR said that the OCR of the cyrillic had not been a major problem, and in fact the fraktur had been worse. Titles are hand-translated (cross-checked with Zbl. or MR), but references are not.

1.3.2 Q and A

Q.—Kohlhase Licencing terms?

A. We print a notice at the start of each article.

Q.—Kohlhase You should really use 'Creative Commons', in view of the weight of lawyers behind it. I am pretty sure there is a Czech version, and I know there is a German one.

1.4 Hlaváč — On-line formula recognition via 2D grammars

Sarting point: limits of 2D context-free grammars. See Schlesinger & Hlaváč *Ten lectures on statistical and structural pattern recognition* (Kluwer, 2002). Mathematical formulae are a case-sudy for us. ICDAR 2007 handled the off-line case: we now talk about the on-line version(e.g. the user is drawing in real-time on a tablet). Problem: 2D sketching ⇒ parse tree.

Most previous work splits into two stages: symbol recognition and grammar. We allow *ambiguity* in the symbol recognition. Schlesinger's group has also worked on electrical circuits and sheet music.

off-line recognition works by bounding boxes, whereas on-line can work with strokes. An important simplifying assumption is that a symbol is composed of at most four strokes. Example $\sqrt{\overline{3}}$ (but handwritten): was the initial part actually a 'v' etc.?

Each production rule has a 'penalty' associated, and we want the minimum-penalty solution. We assume that regions are rectangles, which gives a dramatic reduction in computational complexity. It is easy to express geometric relations amongst non-termninals. Standard 2D grammars, with horizontal and vertical concatenation, is no sufficient, e.g. 5^3 or fractions. In the handwritten case, th regions are not precisiely aligned, so some form of elascicity is needed.

We have an implementaion in Java, which has been trained for both printing and hand-written formulae[2]. Currently goes as far as fractions, powers, substripts, sums, integrals etc. Precision (based on tests with 400 formulae) is 88%. If we ignore errors from the (external, public-domain) OCR tool, it raises to 97%. It is fast, averaging 80ms/formula.

An example was shown (using pre-handwritten formulae with stroke timing stored). It seemed quite good on surds in the denominator, where the surd overbar sometimes ran into the fraction delimiter.

1.4.1 Q and A

Q. Penalties?

[2]No mention of training for different authors.

A. Partly manual, but also they could be developed from the training set (work in progress). This is a major part of the whole process.

Q. Most common non-OCR error?

A. Example demonstrated, said to be common, is 'accidental crossing'. However, stroke order does make this problem lesser in the on-line case.

Q. Output format?

A. Currently a home-brewed parse tree — admittedly not very useful.

Q.—Watt There are "1 and a bit"D grammars.

A. We haven't evaluated these.

Q.—Watt What about hypergeometrics, with presubscripts?

A. We should talk.

Main message: 2D grammars exist, and are useful.

1.5 egoMath

What should be a "Mathematical search engine". Currently, the documents available are semantically poor, e.g. PDFs. The WWW is very heterogeneous. Needs to be able to search for text *and* formulae.

Based on Egothor #2. The indexer accepts MathML, using the Infty project converting PDF to presentation MathML. There is an *extensible augmentation algorithm*, which in many ways is the key to the project.

They key object for us is the tokeniser. This is trivial for text searchers: one searches for a word, but in mathematics one wants to be able to search for sub-formulae. We linearise into a LaTeX-like format. There is no 'canonical form'. All constants are replaced by "constant", and so on (JHD was not clear whether variables were flattened).

We hope it will be publically available soon. It *does* work on semantically-poor documents, but the tokeniser is a very important element, which still needs tuning.

1.5.1 Q and A

Q. Abdou Yussef's work?

A. Not recently.

Q. How do variables work?

A. It depends on the tokeniser. They are converted into `id1`, `id2` etc.

Q. If you interchange a and b, is it the same formula?

A. [Not really clear, but]"different variables names are not a problem".

1.6 Authors not present

1.7 Watt — Mathematical Document Classification via Symbol Frequency Analysis

This is a byproduct of work on handwritten formula recognition. Humans can recognise the classification of a piece of mathematical very rapidly — how? The following were the goals.

- Retro-classification (classifications change!).
- Aid in document understanding. $G > H$ might be "H is less than" or "H is a subgroup"; $(4, 6)$ can be all sorts of things; F^* might be the nonzeros of a field or the dual of the Faraday tensor; etc.
- Aid in pen-based mathematical interfaces. In most people's handwriting w and ω are pretty indistinguishable.
- Cool problem — how do we do it?

We would like to look at individual symbols and at n-grams, but since mathematics is not linear, we need to define a traversal order for n-grams. We need to consider identifiers and operators separately.

The data corpora were: arXiV (20,000 with MSC[3] and TEX sources); 2nd year engineering mathematics: author (Greenberg — 13% market share), publisher (O'Neil), and 1500 pages[4] scanned with Infty and hand-corrected (72% market share!).

To get from TEX we used ORCCA TEX→MathML. n is the most frequent in arXiV, but x (followed by y) in engineering texts.

Operator frequencies follow Zipf laws pretty closely, and in the egnineering maths (not done for arXiV), then n-grams are also Zipf. Looking at symbol frequency, each MSC is unique after the top few (often 2, always ≤ 6) symbols.

Do we need to know the formulae? Apparently not: symbol frequency seems to do the the job! Note that this is true for both arXiV and textbooks.

1.7.1 Q and A

Q. What happens if we follow citation cliques, or the mathematical genealogy tree? Do people inherit notation from their supervisors?

A. "A bit like deciding who wrote Shakespeare?" Why don't you do this.

Q. Inter-author distribution?

[3]Only used to top-level two-digits of MSC.

[4]Not ethat this *was* in TEX: justthat the TEX was not available to the project.

A. In general, the distinction between different authors in the same area is less than that between the same author in different authors. One exception: one author was "hat-crazy", an their were circumflex accents all over his texts.

Q. More about n-grams? Maybe the choice of traversal order *does* matter, as in $\sum_{\ldots}^{\infty}$, which might dominate in some documents, but would only show up if the scanning order was $\sum_{2}^{1}$, not $\sum_{1}^{2}$.

A. True, but not researched.

Chapter 2

27 July 2008: Afternoon

2.1 Bartošek — DML-CZ Metadata Editor

Concnetrates on the "integration" step: between scanning/OCR and metadata harvesting (e.g. MR/Zbl.). It is a web-based system with a suite of scripts. Note that this *includes* reference processing and verification. The final output is in Dspace: see next talk. As mentioned earlier (1.3), this has to cope with scanned, retro-digital and born digital.

Part of the challenge is "building the article" from scanned pages (via thumbnails), metadata etc. There can be problems (manually resolved) in recognising inter-article breaks, e.g. if last page of article 1 is nearly full. Pages may also need to be cloned if two articles share a page, etc.

Where possible, authors should be linked to an authoritative database, but the form in the original article should be preserved.

Reference processing is semi-automated. The OCR process picks up the block of references, and breaks them up into individual items. Verification is done manually.

2.1.1 Q and A

Q. How much metadata?

A. For about 150,000 articles. The tools were very helpful

Q. (JHD) What about embedded references as in some older journals.

A. This can be a problem, and there is manual work to be done, but it *has* to be done somehow.

A. (JR) Conversely, there is a journal that started in 1872, which used the 'modern' format of a reference block from the beginning.

A. Page breaks within a reference even in the 'modern' format are a greater problem.

2.2 Automated processing of TEX-typeset articles

Big publishers tend to have an XML-based workflow. We use the CEDRAM model. We distinguish three periods.

1. retro-digital.
2. retro-born-digital, i.e. reconstruction from a `.pdf` or `.ps`.
3. born-digital, i.e. not just the pixels, but the whole workflow.

For our journal (*Archivum Mathematicum*, 1992–2007), about 50% of the "retro-born-digital" material was AMSTEX, the rest other forms. We attempted re-compilation, but this was often unsuccessful, at least to generate the paper. **Therefore** used the postscript files for the page images (having processed them to improve the fonts during the postscript⇒PDF conversion, which uses `PStill`, but much depends on the version of `dvips`, and hence on the information in the comments inserted in the PostScript), but run the minimal LATEX sources through Tralics to extract metadata.

For the "born-digital", we used CEDRAM with some local additions. JabRef is used as a GUI for handling the references. Otehr important looks are `make` and `\write18`. See `http://dml.cz` and `http://project.dml.cz`.

2.3 Krejčiř — DML-CZ in DSpace System

For delivery of DML-CZ, we have the choice between developing a new system, or customizing an existing one.

We decided to use the DSpace core (considered to be the institutional repository of choice) with Manakin. Can expect much development, as the world has bought into DSpace. We also had much past experience with DSpace. As far as Manakin was concerned, we wanted to give it a production trial.

The reviewers said that DSpace was designed for an institutional repository, which does not seem to be what you are doing. True, but DSpace is more powerful than one thinks. Manakin uses the DSpace API, but is logically independent of DSpace.

2.3.1 Q and A

Q. Export format?

A. Currently OAI, with basic Dublin Core. This is a problem, as `identifier` could be either MR or Zbl.

2.4 Poster Introductions

2.4.1 Current status of mathematical publication in Japan

2.4.2 Volker Sorge for his student

About inferring information from a PDF document. This is a by-product of the student's thesis. PDF doesn't actually say where the character is, just where the "font box" should be.

2.4.3 Doob — A small-scale digitization experiment

Canadian Mathematical Society has an agreement with Google to retrodigitize the Canadian Mathematical Journal. The TIFFs would be at 800 dpi, but Google would not release them. Hence I could do the TIFFs myself (55,000 pages, using nothing more sophisticated than an HP 8390 dsktop scanner), at 600 dpi. From a demonstration, his 600 dpi were actually more legible than Google's[1]. The CMS will make these publically available.

2.4.4 A language engineering architecture for processing mathematics

This came out of the DIALOG project at Saarbrücken. So the linguistic task is the input interpretation to feed the domain reasoner. Problems with phrases such as *vice versa*, pronominal references etc. For example "I'm rewriting the inner parenthesis" does not change the parentheses, but their content, and so on.

2.4.5 Polish Virtual Library of Science

`http://matwbn.icm.edu.pl`. Currently based on a MySQL database, but now trying to move to YaddaWeb. `http://yadda.icm.edu.pl`.

2.5 Digitization of Mathematical Editions in Serbia

Started 15 years ago, at a time when there was no fast internet. First task is old books, being placed in a virtual library.

[1] JHD enquired subsequently. There clearly is degradation at the source level. MD doesn't think this is for technical reasons: more likely constraints imposed by the publishers (not CMS!).

2.6 The RusDML project; Sperber for Wegner

Zentralblatt expects to get 100,000 papers this year. JSTOR, the publishing house archives, the Jahrbuch project etc.

For RusDML, the metadata came from Zentralblatt, and were merged with the digitised images and additional metadata. The core metadata are available in English, Russian (cyrillic) and Russian (transliterated).

There have been political problems and issues with the licencing of the scanned images.

2.6.1 Q and A

Q. What transllliteration? I have never seen some of these before.

A. ISO-Norm and the German D-Norm.

2.7 Panel: How to get to DML?

Sir John Ball ex-president IMU.

Thierry Bouche

Žarko Majajlovic

Masakazu Suzuki

Jiři Rákosník

Petr Sojka

Enrique Macias-Virgos

Wolfram Sperper Zentralblatt, also representing Bernd Wegner.

Katarzyna Zamlynska

Topics might include

1. Technologies, standards etc.
2. Business models, especially what advice we can give publishers and authors.
3. Is there a model of sustainable interoperable working?
4. The three phases.
    - retro-digital.
    - retro-born-digital.
    - born-digital.

JB What should CEIC/IMU do? We don't have money, bu we do have influence. We advocated MW model, but not all organisations (e.g. LMS), and commercial publishers are seeing the value in their back runs.

PS Google Scholar generates 75% of the traffic to SpringerLink, so Springer give their content to Google for indexing.

JR Is this passion for the past just a fad, created by the fact that it is currently feasible? We should worry in case the publishers cease to maintain the past archives.

WS Ask TB.

TB We (?who) should develop a library: in some sense the fact that it is digital is secondary. Archiving is not a publisher's task. Currently digitizing *does* create a product. NUMDAM has passed through the critical mass barrier. The last publisher left is now talking to NUMDAM. It is worth noting that this is all done by "agreements' rather than "contracts", so is fragile.

EM-V We are having problems with Springer as well: suddenly two statistics journals made agreements with Springer, and would not join us. This seems to be naïvety on the part of the journals, hence the importance of point 2 above.

MK Digitization to me means far more than creating bitmaps and getting them elsewhere cheaper. We should try to get *more* out of what we have, e.g. semantic search, proof checking. What you are doing is therefore misguided. Claim: only 100,000 pages of the past are relevant, and we can get out graduate students to retype them (and they will learn someting in the process!).

MS The APS has a trial whereby all new papers (in *Physical Letters*) are *accesssible* (e.g. for the blind). This therefore means accessible to computers. 'Accessible' here means DAISY, and therefore MathML. This is very important for mathematics, since Braille differs from country to country.

This is a very strong argument politically.

PS When there is a robust MathML creator, then we can run our input through it, so we have not ignored this problem.

MS PDF is not the answer.

PS Agreed — XML is the answer, as seen by the fact that every biomedical journal has agreed the Medline DTD. Medline has 750,000 papers/year.

TB But we currently *don't do* full text in XML. He often sees TEX-like code in CDATA.

ZM The semantics of a paper should be defined by metadata. (some dissent from the audience). The first requirement is availability. Retping used to cost $1000/article, whereas scanning these days cost $10.

PS MK is President of OpenMath, which has high ideals. But commercial publsihers do not see the point.

MK Why go *via* images? The author knows the meaning, so should use an editor which captures it. If semantic searching generates more citations, then publishers will use it.

SMW "Who did the owner of the first fax machine send one to?" More seriously, MathML *can* have more semantics than TEX, since every node can be annotated.

JB Getting a graduate student to retype does not necessarily solve the copyright problem.

JHD While good mathematics is cited in perpetuity, we know that mathematics is a small proportion of big publishers deals, and there is the risk that, to put in crudely, when old medicine is thrown away, old mathematics will get thrown away as well.

Barbara Hamilton California[2] Digital Library has "doomsday scenario" digital archives with embargoed material from the publishers.

ZM There are a large number of digitized books (largely textbooks) in Russian — what can we do about this?

?' Legacy is one issue. Why do we need publishers? Everyone else[3] is in this room.

VS To convince Vice-Chancellors and Deans that we should get promoted.

?" And to provide a *definitive* version.

MK But most published papers are wrong — we should have living documents.

AS Assuming that the quality of a living document is monotone increasing. The publisher model may be bad, but that doesn't mean it isn't the best.

MK This "rubber-stamping" is an artefact of a historical process involving dead trees. We should therefore switch from pre-publication review to post-publication review. "I would rather change the future than document the past".

[2]Subsequently, she wasn't sure of the name, but it is part of the University of California system. This has gone to a "one paper copy only" policy (e.g. if it's at UCB, it's not at UCLA) and is therefore heavily reliant on digital versions.

[3]BM disagreed: we are the people who write about writing mathematics, not those who write mathematics.

Part II

Mathematical Knowledge Management 2008

Chapter 3

28 July Morning

3.1 Thierry Bouche — Invited Talk: Digital mathematical libraries: the good, the bad and the ugly

- Mathematical literature never becomes obsolete. He quoted a current article citing Euler.
- It's valid only as a *whole*, building a network of references.
- It's useful to other sciences in a *asynchronous* fashion. He illustrated this with an example of a second paper withdrawing the first.
- About 50% of the citations in today's bibliographies are more than ten years old.
- About 25% of the citations in today's bibliographies are more than twenty years old.
- 96 of the top 100 cited items in MR are books (88 for NUMDAM).

Some milestones of publication.

1665 First scientific journals.

1810 First mathematics journal: *Annales de mathématiques pures et appliquées*.

1978–86 TEX.

1995 JSTOR

2008 10^7 digitised pages. 65% of core journals digitised.

Some milestones of meta-publication.

1868 *Jahrbuch.*

1931 *Zentralblatt.*

1940 *Mathematical Reviews.*

1995 MathSciNet.

2000 links to original articles.

The IMU's (and community's) wishes.

- Free metadata and navigation
- Eventual open access (moving wall).
- No long-term economic, legal, technical barriers.
- No dependence on the viability of any economic agent.

The true situation, e.g. JSTOR, is much more complicated (and very English-biased).

Should the rôle of a library change simply because of a format change? One respect probably will: we want more *metadata*.[1]

Cellule MathDoc is a small CNRS-UJF unit (since 1995). In 2000 it started the digitisation of main Fench serials (NUMDAM), and in 2005 started CEDRAM. NUMDAM has expanded, e.g. journals from Amsterdam, Pisa, seminars (Poincaré, Bourbaki etc.), some theses. Tries for internal links, and metadata links to the reviewing journals. About 20% of the items in NUMDAM bibliographies are now linked to the full source (elsewhere in NUMDAM or from publishers sources).

CEDRAM: "NUMDAM for new journals", both the traditional publishing, but also better support of metadata[2]. 1268 articles, 31,400 pages published *de novo* via CEDRAM. *Annales de l'Institut Fourier* is around 50% of the whole.

MathML support, but depends on the browser (the demo (IE) backfired here!, but worked fine on PL's machine with Mozilla). NUMDAM's database now links to CEDRAM, and to the Science Direct metadata for the current issues of the three NUMDAM issues they publish.

He noted that the *forward* links out of NUMDAM encourage people to read new articles, so it could be argued that this is *good* for publishers.

There isn't that much 'M' in DML/NUMDAM: most of what there is is MSC (but these are textual, and also changing). Most matching is fuzzy textual.

[1] JHD: isn't this analogous to a paper library holding copies of other libraries printed catalogues?

[2] But not perfect. Were unable to insist on tagged bibliographies.

3.1.1 Q and A

Q. Could there be an equivalent of an ISBN for mathematical articles.

A. In short, no. For not-so-old articles, the MR number is a close approximation.

Q. You mentioned the absence of Mathware mining. What about MSC?

A. MSC only cover current articles, and aren't always included in the formal metadata.

Q. It should be possible to assign DOIs, e.g. from crossref.

A. JSTOR also have such a mechanism, but the mechanism has changed recently.

Q. You said that there was no 'Math-aware' mining. What about text mining and graph analysis methods, as in Google?

A. There is a certain use of page-rank in NUMDAM to rank returned results.

3.2 Automated classifictaion and categorisation of MK

The goal is to offer 'similar' articles to the reader. Ths could be based on MSC, when they exist, or on full-text. But ful-text has to be robust in the presence of OCR errors. LSA is a purely statisticl methd of topic extraction, LSA seems to be language-independent. We used DML-CZ, 4532 English, 595 Russian, 4?? German etc. articles, but homogeneous as regards source and level of metadata. We checked our similarity coefficients aginst similarity as measured by MSC. A graph of our similarities $S_{i,j}$ where the articles are sorted by MSC, shows a diagonally-dominant matrix, but there are outliers. Looking in detail at 20, we see that `20.30` and `20.Lxx` have strong cross-correlation.

For automated classificatiun, we use a machine learning model. There are many possible algorithms in this area. Depending on whether we insist on #classification $\geq$ 30,40, 60, we have 31, 37, 20 MSC available n our corpus. See the paper for details of which classification system wins.

3.3 Stamerjohanns – transforming the arXiV to XML

It would be nice to have a large corpus of "natural" MathML documents. LaTeX is what the documents *are* in, and the printed quality is still be best. But it mixes form and contant, and has no semantics. Macros, which *can* offer some form of semantics, are not standardised.

arXiV offers $> 500,000$ serious documents, with LaTeX available.

In LaTeX, the author shares implicit context with the reader. Consider $y = c_i x^i$, which *might* be $y = \sum_{i=1}^{3} c_i x^i$. However, $y = c_\alpha x^\alpha$ would equally be $y = \sum_{\alpha=0}^{3} c_\alpha x^\alpha$.

HERMES and TEX4HT rely on the TeXparser. We used Bruce Miller's LaTeXML, which is open source, and under active development, This is a two-pass process. The 400,000 documents we have from arXiV use over 6000 packages (each of which would need a LaTeXML binding file).

We know that we get no errors, but this doesn't prove that the result is *correct*.

3.3.1 Q and A

Q. Is the LaTeX clean?

A. It will have TeXed once, to be in the archive at all.

Q. You should surely be doing more than merely *reproducing* the work of the LaTeX. What about indexing etc.?

A.–MK The paper that did more wasn't accepted!

Q. What about graphics etc.?

A. We have a plan to generate SVG from some common packages like `pstricks`.

Q.–TB How does this give you intuition into MathML given that the MathML is generated by one tool?

A.–MK We have learnt a lot about *mathematics*, e.g. the use of 'bar' and'ket' constructions by the physicists. This converter is also available as a web service (assuming your class file is supported by us/Bruce).

3.4 Wiedijk — Statistics on Digital Libraries of Mathematics

Slides at `http://www.cs.ru.nl/~freek/talks/stats.pdf`; paper (rejected for proceedings) at `http://www.cs.ru.nl/~freek/notes/stats.pdf`. I am interested in *formal* libraries, e.g. mathematical tables, which is now replaced by software. Work with proof assistants is much closer to software.

Out of a list of "100 famous theorems", 80 have been formalised. For example, there are ≥ 17 proofs of the irrationality of $\sqrt{2}$: see the Siekmann festschrift. Four proofs of Fundamental Theorem of Algebra (e.g. FW's thesis). Harrison has a formalisation of the analytic proof of the Prime Number Theorem. There are four proofs of Quadratic Reciprocity. Impossibility of Doubling the Cube has been done: there is an algebraic component and a geometric component.

The Euler polyhedron theorem has just been done, which is interesting, since Lakatos claimed it to be unformalisable.

Why NuPRL isn't included: it turns out that there are only 10 of *these* theorems proved in it. Five systems accounted for a significant number of theorems (surprisingly, not PVS — 15 theorems).

HOL Light 62

ProofPower 42

Isabelle 40

Coq 39

Mizar 45 — by far the biggest in terms of lines of code (around 2,000,000).

He has a collection of pie-charts showing the relative distribution of the kinds of statements in the systems: for examples Mizar has very little automation lines, and 80+% is proofs.

Conclusions. The systems are very different in terms of foundations, but look more similar when you step away from the details. Proof automation is important.

3.4.1 Q and A

Q. What about axioms?

A. There are almost none, e.g. 5 lines in the whole of Coq.

Q. Understandable if you do large theories, but not if you do small theories, which may be more interesting.

Chapter 4

28 July Afternoon

4.1 On correctness of mathematical texts from a logical and practical point of view

Named after the Russian for a 'garden': we want to let things grow.

We are set theory (rather than type theory), because this is the practical mathematica; traditional. Proofs as in textbooks, declarative, rather than imperative. *But* we want large steps, unlike Mizar, so we need a strong prover.

"If you define multiplication for integers, you will not use it for sets. But you might use it for reals, if you extended it."

He showed a proof of Newmann's lemma, which fitted onto one screen, but not at a legible size (to JHD at the back of the room!). We currently do not support the *verified* introduction of well-foundedness: we trust the user.

Theorems proved include irrationality of $\sqrt{p}$, Cauchy-Buniakowsky-Schwartz inequality.

4.1.1 Q and A

Q. You are essentially using controlled English. What happens if you run your parser on other peoples's "english", e.g. Mizar.

A. We are not that close to natural English, and Mizar isn't really English.

Q. How extensible?

A. Part of the grammar is hard-coded, but it is possible to add new [at this point it got rather confused].

4.2 Ωmega

We use TEXmacs. We want to support definitions of notations, and verification of partial proofs. We want to do (some) completion.

This involves detecting the start and end of theories, detecting sentences (a predefined LALR parser was used at MKM07). We now use an earley parser, which seems faster: 0.8s rather than 6s on one example.

Partial proofs are problematic, e.g. "case $x < 0$", is the rest "$x \geq 0$", or are there two: "$X > 0$" and "$X = 0$"? The basic steps are 'inferences', which can be generated from theories.

4.3 Wenzel — Logic-free reasoning in Isabelle/Isar

The aim is not to get first-order/higher-order logic between the user and the application. The Isar is a river, but the concept is the same as the garden.

The classic document structure is definitions: statements: proofs, with the proofs being the hard part. Gentzen-style natural deduction with $\rightarrow$ and $\forall$.

Isar has blocks and contexts, which lets one handle intermediate assumptions such as $\exists a : B(a)$. It seems to be quite natural to introduce non-recursive inductive definitions. He showed well-foundedness of multiset ordering. He claims that the benefits are similar to currying and pattern-matching in ML.

4.4

Linear Monadic Context-Free Tree Grammars. Fan-out rules define the expansion of constructors such as $\sum$. Cocke-Younger-Kasami $O(n^3)$ parsing algorithm. Weakly equivalent to various mildly context-sensitive grammars common in NLP, such as LInear Indexed Grammars. We have 35 foun-out rules and 170 context-free rules (this latter will grow). The parsing process involves searching for matching fan-oout rules This has been tested on 21967 formulae in INFTY CDB-1, in 10 seconds.

This has found errors in OCR, e.g. $j = 1$ mis-read as $= -1$ which is not a valid initialiser, and therefore not valid as a subscript on a $\sum$. In at least one case, this found errors in the original document. Claim: as well as OCR, this can also be used for parsing MathML-P. We aim to integrate this with InftyReader.

4.5 Christine Müller — Notations for Living Mathematical Documents

Case study: a single human interacting with a knowledge base. The notation is fixed by context (nationality, subject, level etc.).

- decimal point/comma.
- Binomial coefficient: $\binom{n}{k}$, C_n^k, C_k^n. These are German, Russian, French, in some order.
- `\textstyle` versus `\displaystyle`.

One needs to collect a variety of notations into a 'notational context'. There needs to be inference/defaulting if the context is not specified. Need a CSS-like mechanism of 'Cascading Context Files'.

4.6 Sojka — From pixels and minds to the mathematical knowledge in a digital library

Who should do it?

- publishers have money, but no interest.
- societies such as AMS, EMS. Interest, but no money.
- IPR firms, such as Google scholar, have money, *should* have interest but don't (DML-CZ offered in vain).
- Librarians actually do have *some* money. Göttingen has paid Springer a flat fee for open access to work of Göttingen authors.

DML-CZ is an attempt to solve this problems 'in the small'. There are convoluted copyreiht issues over the pages digitised in Göttingen. Must distinquish between

image processing (scanning) — was 33% of the budget, but now down to 10%;

semantic processing — both data and metadata (marriage with Zbl. metadata).

DML-CZ has tried various learning methods for automatic classification via MSC — see paper. No off-the-shelf OCR system was perfect: quite a lot of fine-tuning was needed.

4.6.1 Q and A

Q.–JHD OCR error rates?

A. Less than 1% error rate per character (errors of all types, e.g. *italics* when not, or weight, or size).

Q.–FW That's one error every two lines!

A.–MS It depends drastically on the source: old paper documents give a much higher error rate than modern digital.

Chapter 5

29 July Morning

5.1 Bundy — signature evolution

We are used to dealing with a *fixed* grammar, but I will argue, following Polyá, that representation changes, as part of the problem-solving process. Background: work on 'ontology repair', and the evolution of ontologies.

To address this, we looked at Physics, where the evolution is well-documented.

- "Where's my stuff": splits 'stuff' into visible, invisible (and total).
- inconstancy: adds extra arguments.

Uses λProlog, since the higher-order features are needed to, e.g., define an orbit as a path λ-abstracted over time.

Example: Joseph Blacks's discovery of latent heat: before then, heat and temperature were conflated. More recent example over 'dark matter', postulated because of a contradiction between the newtonian prediction and reality for orbital velocity of stars against radial distance from centre of galaxy. Inthis case we *literally* need 'invisible stuff'. New planets are generally discovered by errors in previous predictions (Uranus, Neptune). There's also an anomaly with Mercury, hence the hypothesis of Vulcan. But the answer is actually *relativity*, not Vulcan. Note that Vulcan *had* to lie on the opposite side of the Sun. there's an alternative to 'dark matter': (MOND) MOdified Newtonian Dynamics — assume G is not constant at low accelerations. This is an 'inconstancy' explanation, rather than a 'where's my stuff' one.

The usual definition of a conservative extension won't work, since we're extending the signature, but it can be fixed.

There's also 'merging', e.g. "morning star" versus "evening star". 'Dropping arguments' is also valid, e.g. Galilean/Aristotlean gravity.

5.1.1 Q and A

Q. Does this work show that Popper is wrong?

A. Yes!

5.2 Joseph Collins — A Mathematical Type for Physical Variables

Fortran code and TeX papers: no connection!

Use SI as a default, with `kg` as mass. Need to allow for different systems, e.g. CGS, Planck. Use $[X]$ to denot ethe domensions of X, which ives a multiplicative Abelian group.

Geometric algebras are Clifford algebras, so we want geometric interpretation of these multivectors. Hasternes has a pedigree of geometric algebra. We need scalar, outer/wedge and Clifford (which has an inverse) products. For 3D, actually have 8 elements: scalar, three vectors, three cross products and the triple product.

Would like to express some of these ideas in CDs. This has a place in SysML, an extensio to UML adopted by the Object Management Group.

5.2.1 Q and A

Q.–MK Is it the case that you have models as specifications, and Fortran as implementations?

A. Essentially, but the problem is that there is no connection. The constraint satisfaction problems may be in TEX, or Mathematica, or (commonly) MatLab.

Q. Can you provide the community with 'small' challenge problems that we can understand the questions?

A. I think I'm ready to start writing CDs.

Q. That's not quite the question — how about a page from a Physics book, or some such.

A. "OK - I'll take the challenge".

5.3 Davenport — Unit Knowledge Management

5.4 Sloman — Kantian Philosophy of Mathematics and Young Robots

Hume's two kinds of knowledge — 'empirical' and 'analytic/definitional'. Kant added 'mathematical' between the two. Can we build a 'robot' to do mathematics, which would "prove Kant right".

Need to end the 'factional' wars inside AI, and cooperate. Claims that we have under-recognised examples such as a child 'learning' that counting objects in different orders gives the same result, i.e. 'empirical' facts *become* mathematical, and possibly even analytic. Noted also recent research on "buggy rules", especially pointed out in tutoring research.

We can't learn something (by imitation or instruction), if you don't have the cognitive resources to do so.

5.4.1 Q and A

Q. How does this cope with 'unlearning', e.g. Alzheimer's disease?

A. It doesn't directly, but in general we can't theorize about what happens when things go wrong unless we understand what happens when they go right. This is often ignored.

A. Furthermore, why don't children see anything wrong with Escher's pictures? What it is that has to develop?

Chapter 6

29 July Afternoon

6.1 Kerber — normalisation issues in mathematical representation

How many clause sets are there? Encode with $s_i \in \{0, 1, \#\}$, depending on whether x_i occurs positively, negatively or not at all in the clause. Therefore there are 3^n clauses. Of course, what we really want is "up to equivalence", or indeed "up to subsumption" .

6.2 Mediated access to comouter algebra systems

The general direction of computer algebra systems is research. Kenzo is very powerful, but it is written in Common Lisp. Hence we use XML-RPC to build an interface. In particular, this allows the user to chain commands.

6.3 Miller — Search in Presentation MathML

Not deep (*ipse dixit*), text search optimisations/user expectations. Problems

- strange symbols.
- case (sometimes) matters.
- symbols don't always matter.
- $a + b$ is (often) the same as $b + a$ etc.

Our context is users with the "google paradigm" and a summary of each document. It turns out that the summary is important and ours had problems.

What should `\BesselJ` match to? `j`, `J`, `BesselJ` ...? Should `sin` match `trigonometric`? Users tend to ask "why on earth did that match my query?"

Initial attempts for DLMF were in the LaTeX, but as LaTeXML evolved, we moved to the presentation MathML (note that we have lost some of the semantics by this stage). Choice: index where the semantics are, or where the text the user sees is? Our answer: carry some semantics with the pMML nodes. Example. stick `meaning=Bessel J` on `<mi>J</mi>`.

6.4 Kohlhase

- Substitution instances.
- Homonomy — e.g. binomial coeffieients.

So use content-based representations.

- Mathematical equivalence

But not full logical equivalence, since Fermat's Last Theorem is (now known to be) logically equivalent to 1+1=2.

Binding objects have to allow for α-equivalence: use de Bruijn indices. We have this working.

In general, the problem isn't the search engine, it's finding the corpus to search (usefully) in.

Searching in. say, Wikipedia, requires to know that `Let` binds in the subsequent text. It would also be nice to, at least, warn the user of side-conditions, such as "M should be a metric space".

6.5 MKM Business Meeting

1. MKo presented the agenda, and JHD was chosen as scribe.
2. VS reported on CICM 2008: three conferences, a doctoral programme and five workshops. DML 2008 could well have expanded to two days. Total registration is about 120, with a total of 10 invited speakers. VS has the exciting job of distributing the surplus. Funding from EPSRC, LMS, SSAISB. AS asked who paid for pens/memory sticks. The local team had secured an excellent deal.

   The meeting thanked VS.
3. MS and SA reported as the MKM chairs. 45 abstracts, but only 34 full papers, of which 18 were accepted. He was very pleased with the DML co-presence (but it was noted that it clashed with MathUI).

   AS reported that Springer were rumoured to want a 1-in-3 acceptance rate, which might be a problem in the future. Various members volunteered to submit papers for rejection, if Springer were 'being that stupid'. MKe

reported that this had been discussed by the trustees, and it was noted that some teams had not been present this year.

There was a discussion of a 'best paper'(s) award. AB pointed out that these might seem trite, but were important, at least in the UK scene. JC mentioned the retrospective 'most influential paper' idea.

BM said that the plurality of conferences had confused reviewers and probably authors. MKo asked for a more explicit 'systems descriptions' track, and at other conferences these were often the best part.

AS said that for this sort of meeting, freely-available web pages actually gave greater dissemination. AB added that IJCAI released proceedings on CD-ROM, and back proceedings were freely available. IJCAI proceedings were certainly accepted by promotion committees (not everyone was convinced by this). It was pointed out that Springer could provide CDs rather than printed volumes.

The meeting thanked the programme chairs.

4. WF presented the treasurer's report. The loss on 2006 had eaten up the balance, and 2007 was cost-neutral.

5. MKo presented the Trustees report. There is one bid for the next MKM (initiated by the Trustees, who felt that it was time for MKM to leave Europe).

   JC presented a bid to host CICM, with special reference to MKM, standing in for SMW.

   25–28 June ACA in Montreal.

   early July CICM mainly MKM since it's not an AISC year.

   29–31 July ISSAC Seoul.

   Proposed location is one of two "great lakes" locations near ORCCA.

   Grand Bend (on Lake Huron) Lots of sandy beaches, 'nature' tourism.

   Niagara on the Lake (i.e. lake Ontario) a bustling city, with the Falls for tourists. It was pointed out that Brock University was close to Niagara, and probably had cheap student rooms.

   This meeting was asked for a non-binding preference, since it would depend on the availability. JHD argued that being close to ACA in terms of time was a good idea. The meeting voted *nem. con.* to accept the bid.

   WF said that there had been two MKM conferences in North America (2000, 2002), with substantial interest from students.

   JHD proposed, and the meeting voiced general agreement, that "the meeting was concerned about cost, and recommended that SMW look at the cheaper options".

   MKo said that the Trustees proposed, for co-chairs, SMW and Claudio Sacerdoti Coen. Both would be willing to serve.

6. There was a Trustee vacancy (*ex* Cairns) for a 2008-11 term. The Trustees felt that the meeting with DML had been very successful, and wished to suggest PS as a candidate.

   The election is an e-mail ballot by the Condorcet method. The constituency is the mailing list: attendees at MKM are automatically added (but can opt out of the list). Nominations are needed to MKo or any other Trustee (self-nominations are allowed, but a seconder is needed).

7. Any other business — none.

Chapter 7

30 July Morning

7.1 Libbrecht — Cross-curriculum search for Intergeo

Interoperable interactive geometry. Proven to demonstrtate learning, and part of the core curriculum in Germany. *But* a lot of software with interoperability barriers. To be fair, there are also interoperatility issues between curricula, textbooks etc. The opening survey found 20,000 resources.

MSC is just too broad for us, and we would like to use full-text search, *but* multi-linguality is a major issue. Uses the English curriculum, as it's quite rich in resources, but all the annotations are manual. Also use the U.S. `curriki`.

The current fashion (Pisa) is to demonstrate 'competencies', which means the GeoSkills project links competencies. GeoSkills is an OWL DL ontology, designed in Protégé.

7.1.1 Q and A

Q.–TB It is not reasonable to expect MSC to handle this sort of problem.

A. Possibly, but if not MSC, then what? We found something by `Eurydice`, but it seemed unusable.

7.2 Specifying Strategies for Exercises

Will show how a strategy can interact with diagnosis of errors. Many subjects require the learning of procedurla skills, which are often supported by tutoring tools. Demonstrated some tools, including ActiveMath, and MathXPert.

Many studies have shown that, evne if immediate feedback does not *improve* learning, it certainly *speeds* it. He showed various strategies (e.g. "eliminate negation bottom-up") expressed in his language. This can be used to give

feedback, populate a 'progress bar' etc. Can also be extended with 'buggy rules'.

7.3 Aspinall — A tactics language for Hiproofs

Hiproofs are *system-independent* labelled proof trees, with small-step semantics.

7.4 Herbrand Sequence Extraction

Asking the 'converse' question: can we go from a formal proof to an informal one? One step is the elimination of lemmas, since thes may contain notions that are not in (the signature) of the theorem. Now lemmas correspond to cuts, and we can use cut-elimination. Demonstrated cut-elimination by resolution.

Part III

Calculemus 2008

Chapter 8

30 July Afternoon

8.1 Coquand — Type Theory and Linear Algebra

Ultimate goal: verify and implement Kenzo in type theory. Kenzo is a large (16000 LISP program: Rubio & Sergerart (*Bull. Sci. Math.* 162(2002) pp. 389-412). memoization is a key technique, but is this doen well? Is related work actually relevant. Kenzo can compuet things not computable elsewhere, so cross-checking is impossible. Aransay has proved one lemma correct in Isabelle.

Kenzo is a purely *algebraic* system. However, standard homological algebra is not constructive.

Constriuctive Type Theory by Martin-Löf. Gregoire and LeRroy have effieientc omputations in type theory. See also Thery's proof of the associativity of elliptic-curve addition *inside* type theory. Key ideas in constructive type theory are dependent products and sums.

8.1.1 Q and A

Q. Is 'natural' a problem?

A. Not as such, but we do have issues that, say, ker is only defined up to isomorphism.

8.2 Dominguez — formalising Hidden Algebras in Coq to Flrmalize Computer Algebra

Isabelle and ACL2 have been used to prove parts of Kenzo, but we want an actual Coq version.

An explanation of algebraic specification: sorts, operations and signatures. Kenzo has the normal data structures (lists, numbers), but also algebaric structures (groups etc.), and many hundreds of groups can be create din the course

of a computation. Initial semanics are useful for the first layer, but not the second. The *implementation* of a group can be regarded as a hidden sort for the signature. Example: n is a hidden sort for representations of integers mod n as a group. The coresponding *final* algebra corresponds to the Kenzo/EAT data structure.

So needs a Coq implementation of 'signature', 'algebra' and 'term algebra'. Also 'category', 'functor' and 'initial object'. Then can define 'implements'. As far as I could tell, he only had one 'hhidden' sort, which was identified specially[1].

Conclusion: it can be done, but a lot of technical lemmas are needed in order to simplify the syntactic compbinations.

8.2.1 Q and A

Q.–JC Why 'hidden algebras' rather than higher-order theories?

A. Seems to it better with the OO style of Kenzo.

8.3 Using COQ to prove properties of the cache levle of a video-on-demand server

Software certification is hard and expensive. The video cache server was programmed in Erlang. The cahche is divided into blocks, but has to maintain locality, hence we speak of 'block intervals'. Hence need to model 'sequence' and 'interval', and to model the implementations handling of a canonical model of the block structure.

We can verify part of a *real* operating system, but only the Coq model of it, not the actual production code. The difficulty came in the verification of the actual algorithms (which have to be efficient), rather than a simple abstract structure. 794 lines of Coq, 330K compiled.

8.3.1 Q and A

Q.–LP The interesting part of Erlang is the parallelism.

A. True.

8.4 Davenport

8.5 Sorge — Towards abstract matrix arithmetic

[1] JHD notes that this looks awfully like the 'environment' of Axiom.

Part IV

Mixed Calculemus/AISC

Chapter 9

31 July Morning

9.1 Linton — Symmetry and Search

Claims that *any* classical search problem can have a symmetry issue. Consider n Queens: generally 8-fold symmetry, but not for small n: $n = 6$ has a symmetry group of order $3 \cdot 10^{24}$.

Symmetry breaking is nice. In teh worst case, can need exponentially many constraints (Crawford it et al.), but recent work at St. A. has shwn that 'often' it needs few, and the work is pretty mechanisable.

9.2 Linton — Symbolic Computation Software Composability

SCIENCE — Symbolic Computation Infrastructure in Europe. `http://www.symbolic-computation.org`. SCSCP: OpenMath-enabled RPC communication. GAP, Kant/Kash, Maple, MuPAD are the partner software, but the system should be open, and has in fact linked to others.

We encourage SCSCP to be embedded in the systems, not just wrappers. Two downsides to OpenMath

1. Limited to CDs: have a skeletal CD for 'transient' definitions
2. Encoding is verbose — e.g. matrices over finite fields[1]. Can use `OMBYTES` or `OMFOREIGN` for some of these.

Similarly, sending groups around is a problem.

[1] Are also working on a CD for this which uses `OMBYTES` for the rows.

9.3 Ida — Eos project

Noted that all his students were at RISC, unlike SMW's commuting. Stands for e-Origami system. A fundamental theorem for origami is Morley's theorem, which (bizarrely) was not known to the Greeks. Of course, in general Morley's theorem is not (ruler-and-compass) constructive.

EOS is build on Mathematica, so there's a version WebEOS, which uses WebMathematica on the server.

9.4 Lucas — improving linear and geometric techniques for termination proofs

There are well-known techniques for proving termination in term-rewriting. One technique is via polynomial interpretation, which have integer ceefficients. Real coefficients were mentioned by Dershowitz (IPL '79), but only revived by Lucas in 2005. We have shown (AISC08) that real coefficients are *strictly* more powerful. It *may* be necessary to deal with irrational coefficients (e.g. $\sqrt{2}$), which can be difficult.

From the theoretical point of view, can reduce to a Tarski problem. Can be transformed into a purely existential problem. The Real Algebraic Geometry solutions are costly.

In 1994, Steinback proposed using *linear* programming, in particular the simplex method. Inequalities between sums are replaced by inequalities between products, and then logarithms are used to linearize (!).

Chapter 10

31 July Afternoon

10.1 Paulson — MetiTarski

Combines the theorem prover (metis) with QEPCAD. We eeplace "awkward" functions with algebraic (i.e. rational unction) uppe/lower bounds. The resulting polynomial equations/inequalities are decidable. Taylor series give good approximatoin to, say, exp in good regions. For upper bounds, use. e.g. $\exp(x) \leq 1/(1 - x + x^2/2 - \cdots)$. All variables range over the reals!

In some sense[1], we are only interested in `FALSE`, i.e. $\not\exists$ counterexample. HOL-LIGHT implements Hörmander, but this hands if degree ≥ 6, so now use QEPCAD, which has good behaviour w.r.t. degree (inherently doubly exponential in # variables).

Changes to METIS (but these are more generic)

1. Algebraic literal deletion. Keep a list of all ground algebraic clauses. Any literal inconsistent with these (via QEPCAD) can be deleted. E.g. delete $x^1 + < 0$. Similarly in $x^2 > 3 \vee x > 3$ can delete $x > 3$

2. Algebraic sbsumption. Again use QEPCAD to detect clauses that cannot be false given previous clauses.

3. Formula canonicalization, via Horner form for polynomials. Isolate the leading no-algebraic term of each formula. Simplify to rational function form, *but* need guard clauses.

4. Knuth–Bendix ordering, as implemented in Metis, counts numbers of occurrences, so replacing exp by a bound, which may have manymore occurrences of x, is still regareded as simpler.

5. Explicit division rules: $f(t) \cdot u \leq v$ generates three clauses for f, e.g. $u > 0 and f(t) \leq v/u$.

[1] That of resolution.

6. Omit laws like transitivity, since QEPCAD will tend to pick this up.

7. lgen(R, X, Y): $X < Y$ if $R = 0$, otherwise $X \leq Y$. Simplifies human work.

8. $|X|$ is replaced bu conditions, and $|\cdot|$ has a high weight.

Various examples, including many ODE problem which Maple converts to elementary functions and MetiTarski then solves.

No range reduction, so proofs about sin(3000) fail. Not everything can be proved. π is represented by fixed fractions, which causes problems with, say, $\cos(\pi x) < 1 - 2x$ fails at the edge. The resolution proofs can (in principle!) be checked independently.

10.2 Meikle — Isabelle and QEPCAD-b

Verifying geometric algorithms, e.g. convex hulls. Summary: possible, beneficial, hard[2]. Manually discovering the loop invariants is particularly hard.

Want graphical cues for novice use, but a richer GGUI for advanced use. Built on Eclipse ProofGeneral, a broker-centred architectire. Many of these theorems required manipulaton of reals, which wasn't possible in Isabelle, so had to use QEPCAD. So the UI provides tips to eliminate, say, `left turn` until the result is QEPCAD-able. Coudl sovle a problem involving 8 variables! Currently have to take QEPCAD's results on trust, but are looking at ways of uing QEPCAD to guide a prover. If QEPCAD finds a witness, this can be fed into the prover. For non-theorems, can produce a counter-example, but no real hint about what is missing. When trying to find a missing assumption, the "black art" is knowing which variables to bind and which to leave free.

Can also detect reductant or contrdictory assumptions. Can be used in both interactive and automatic modes. This seems to be a step forward in modular construction. Parts f the prover are re-usable, as is a lot of the "infrastructure". Need heuristics for #variable reduction[3]. We would like to integrate mre tools (Maple alost fimished).

10.3 Chaieb — Parametric liear arithmetic over ordered fields in Isabelle/HOL

Division by zero *is* allowed. v are variables (can be quantified), v_p are3 parameters and may not. The variables may only occur linearly. QE is an axtension of Ferrante–Rackoff, and this produce existential witnesses

Use the HOL in Isabelle. Our formulation should work in any linearly ordered field. Use HOL's code generation to produce Standard ML, OCaml and

[2]Found bugs in Hilbert's book!

[3]Example with 8 variables had 10 before specialisation.

Haskell. Reflection: prove $(t :: \tau) = t'$ where t is obtained algorithmically from t, and τ is inaccessible from HOL.

10.4 Davenport

10.4.1 Q and A

Q.-SMW If we breathed life into the line-breaking package, then 'your lines get broken sensibly iff you use semantic tags'.

A. Certainly — that's alrger carrot.

10.5 Implementing CAS in a functional language

Haskell and its papallel extension Eden (which has referential transparency). Makes heavy use of higher order functions. This means that 'algorithm skeletons' are possible, e.g. divide-conquer(split,work,combine). We could replace this skeleton with a parallel version transparently from the 'worked' functions.

Prior work: DoCon is a sequential CA in Haskell, making heavy use of algebraic constructs.

10.5.1 Q and A

Q.–JC DoCon gave up because the Haskell type checker could not cope with the complexity. The paper kept being rejected, so you have to look for it on the web site.

10.6 Watt — Towards Real-time Recognition of Handwritten Mathematics Symbols

A huge problem: what can we take from the NLP context. "How do we get the computer to think *while* we're writing" — Yes.

- Mathematics is a mixture of drawing and writing.
- No fixed dictionary.
- Many more similar few-stroke characters: ∂, α, the proportionality symbol, ∞.

Note we are doing *on-line* character recognition, so ew have a trace to start with. and then idntify features and use a classification method.

Chebyshev series can't be computed until we know the length we're integrating over. What we need is a linear weight function (Legendre: weight=1). We are looking at the Hausdorff moment problem (1921), proved to be numerically instable bt Talenti (1987). *However* this instability only shows for high orders:

much higher than we need. Legendre series (n) are worse than Chebyshev series (n) are worse than Legendre series $(n+1)$.

We are looking at a few hundred instructions at pen up.

10.7 Calculemus business meeting

FW opened the meeting. There were three important items.

PC Chair Carette, Sorge and Dixon had expressed an interest

Location options might be

MKM 6–10 July, Ontario

ISSAC end July, Seoul

TPHOLS 17–20 or 24–27 August, Munich

Format There was a low number of submissions this year. Some called for the TYPES format, i.e. post-proceedings. SAL pointed out that this was the mathematics format. The drawback might be that many people would only get funding to attend if a publication was guaranteed.

VS withdrew as a candidate for PC, so the meeting recommended JC and LD (the PC chairs are formally appointed by the trustees). Therefore the conference would co-locate with MKM.

Thre was a shortage of trustees for historical reasons (last year's election never happened). FW proposed extending Rioboo and McCasland by *one* year to stagger terms: the alternative was to leave two vacancies, which was what was decided.

There is then a vacancy for four new trustees. Excluding JC (who would be *ex-officio*), there had been four nominations for the non-existent 2007 elections: Thierry Coquand, James Davenport, Marc Moreno Maza, Makarios Wenzel. These were elected by acclamation.

Chapter 11

1 August Morning

11.1 Cuyt — validated evaluation of special functions

Inspired by W.B. Jones, original plans were for a handbook, now handbook and software. The book is out with Springer since April 2008, in A&S style. Maple lbrary of functions, C++ library of functions. The other components should be avaiable 'early 2009' – the date depends on the cooperation with NIST, which is now pretty close, and this will complement DLMF. Only 10% of our CF representation are in DLMF ∪ Wolfram. Use

$$f(z) = b_0(z) + K_{i=0}^{n} \frac{a_i(z)}{b_i(z)},$$

which abbreviates

$$f(z) = b_0(z) + \cfrac{a_1(z)}{b_1(z) + \cfrac{a_2(z)}{\ddots \\ b_2(z)+\frac{a_n}{b_n}}}.$$

[K='Kettenbruche'.] Much better convergence properties than series.

Note that, as in the case of continued fractions of integers, the fractions converge, but, unlike series, the tails do not converge.

11.1.1 Structure of the book

In three Parts.

Part I Basic Theory (120 pages). In general, a 'historic' reference and an 'accessible ' one.

Part II Numerics (80 pages). Construction, numerical error bounds, evaluation algorithms.

Part III Special Functions (300 pages). Covers about 50% of the special functions in DLMF.

Mathematicians have asymptotic formulae for the tail, but in practice we find it buys at most 2–3 significant digits, and at times actually loses accuracy. How do we *guarantee* accuracy? Currently only implemented for $\mathbf{R}$ — $\mathbf{C}$ coming. We need to bound both truncation error and roundoff error: (apparently) choose each to be $\frac{1}{2}$ of the user's desired error. The key result on the truncation error is an adaptatin of the "Oval sequence theorem", which tells us how the error in the tail affects the whole result.

Can also do series, but the theory here is not new — see also Higham's book. Our underlying library is IEEE 754r (!) compliant.

She showed guaranteed evaluation at a point. In $\mathbf{R}$, the theorems extend to an interval. In $\mathbf{C}$, the theorems extend to discs, but not very well for boxes. They are looking at sectors. In general, the theorems in the lierature are crude bounds on large intervals, whereas we want better bounds, even on shorter interals.

11.2 Dixon — Reasoning about Graphs for Quantum Computations

Following Abramsky *et al/*, we view quantum cmputations as circuit-like graphs. "Not entangled " $\equiv$ "not connected". These graphs ae hard to reason with by hand, hence there's a need for tool support. The formalisation will have to include ellipses notation.

Isabelle-like 'trusted kernel' + derived rules. Compact closed categories provide a useful formalism. Boundary nodes descibe input/output edges. There are two families of nodes, X and Z. Thee is a problem with redundant representations. The semantics of a vraible-nod egraph is the semantics of all the concrete graphs it represents.

11.3

These MAS diagrams *do* describe the behaviour of multi-agent systems, and construsts *minimal* ushout complements.

11.4 A Groebner-based model for multi-modal logics

Based on algebaric models for classic Boolean logic. Although Maple people, we had to use CoCoA because wanted GB over finite fields: Maple has now improved in this area.

Want Rule-Based Expert Systems (RBES) to do this. On the classic case, the first question is consistency, and then maaximality. Use `Groebner` and `Ore_algebra`. However, this isn't a correct model, since Maple thinks we have a field, but we *do* care in the modal case. (JHD not sure why.[1])

[1]Subsequent conversation implied that the problems leave Maple's `Groebner`, the finite field information is forgotten, and then "silly" results like $\frac{1}{2}$ appear.

Chapter 12

1 August Afternoon

12.1 Pfalzgraf – Mechanical Robots

Two types of joints: translational (t of them) and rotational (r of them). Then we have t copies of $\mathbf{R}$ and r copies of the torus as the space to operate it. A major problme is the 'inverse kinematics' problem: given a situation, how do choose the settings of the various joints to achieve it. need to consider how local coordinates interact. Buchberger was the first to produce formal solutions with $t = 0, r = 2$, introducing variables for $\sin\theta\ \cos\theta$ etc. I then considered the case with $t = 1, r = 2$. The algebra was quite difficult.

Then moved to a connectionist approach. Tracing a circle gave ony an approximate result, but, since it was *topologically* correct, modern sensor/feedback based robotics can use it.

Needed to use non-commutative geometry. Built a categorical model of these spaces, e.g. ray spaces or sphere spaces. Having seen limitations of symbolic and conectionist approach, would like a hybrid approach, where some joints are modelled one way and some the other, e.g. an RTRR arm with the first two (RT) symbolic and the last two (RR) connectionist with sensor feedback. JHD asked whether tis is the 'right' way of doing this, but JP replied that this was just the first method he had tried. However, 'symbolic' did seem better for translational joints. Another strategy might be to divide the arms into sub-arms, treating each sub-arm symbolically and combining them is a commectionist manner.

This produced some serious performance problems, so proposed as a family of benchmarks for GB.