Theorem Proving and the Real Numbers

Applications and Challenges

Lawrence C Paulson

UNIVERSITY OF CAMBRIDGE

1. Interactive Theorem Proving

A partial, biased history

AUTOMATH

L. S. van Benthem Jutting,
Checking Landau's "Grundlagen" in the AUTOMATH system (1977)

- constructing the reals from first principles
- the first formalised mathematics textbook
- the first major case study with type theory

but *not at all* about verification

The Hiatus, 1977–92

When everybody studied lists,
natural numbers, Booleans, …

John Harrison (using HOL)

- a formalisation of the reals including limits of series and the elementary functions (1992)
- quantifier elimination for the reals; integrating HOL with a computer algebra system (with L. Théry) (1993)

PENTIUM FDIV BUG (1994, $475 MILLION)

- *floating point verification* of algorithms for the functions sqrt, ln (1995) and exp (1997)

Jacques Fleuriot (Isabelle)

- another formalisation of the reals, and the functions sin, cos, …
- *nonstandard analysis*: a construction of the *hyperreals* using ultrafilters
- development of a proof calculus for *infinitesimal geometry* (1998)
- application: checking the original proofs in Newton's *Principia*

Assia Mahboubi (Coq)

- a formalisation of *real closed fields*
- real algebraic numbers
- *nonlinear* arithmetic decision procedures
- quantifier elimination based on pseudo-remainder sequences
- theory underlying efficient computer algebra algorithms

(2002–07, later with Cyril Cohen)

PVS (1992–present)

- Created for verification (as opposed to foundations)
- Many early proofs involving the reals
- Reasoning methods for the reals (C. Muñoz et al.)
  - interval arithmetic (for numerical inequalities)
  - Bernstein polynomials (for optimisation)
  - Sturm's theorem (for polynomial inequalities)

And Many Many More…

Probability & Measure theory

Real Algebraic Geometry

Multivariate analysis

Complex analysis

by researchers at Concordia, INRIA, Intel, NASA, TU Munich, etc.

2. Automatic Theorem Proving

MetiTarski

MetiTarski =

Resolution Theorem Proving

+ Real-Valued *Special Functions*

A Few Easy Problems

$$0 < t \wedge 0 < v_f \implies ((1.565 + .313v_f)\cos(1.16t) + (.01340 + .00268v_f)\sin(1.16t))e^{-1.34t} - (6.55 + 1.31v_f)e^{-.318t} + v_f + 10 \geqslant 0$$

$$0 \leqslant x \wedge x \leqslant 289 \wedge s^2 + c^2 = 1 \implies 1.51 - .023e^{-.019x} - (2.35c + .42s)e^{.00024x} > -2$$

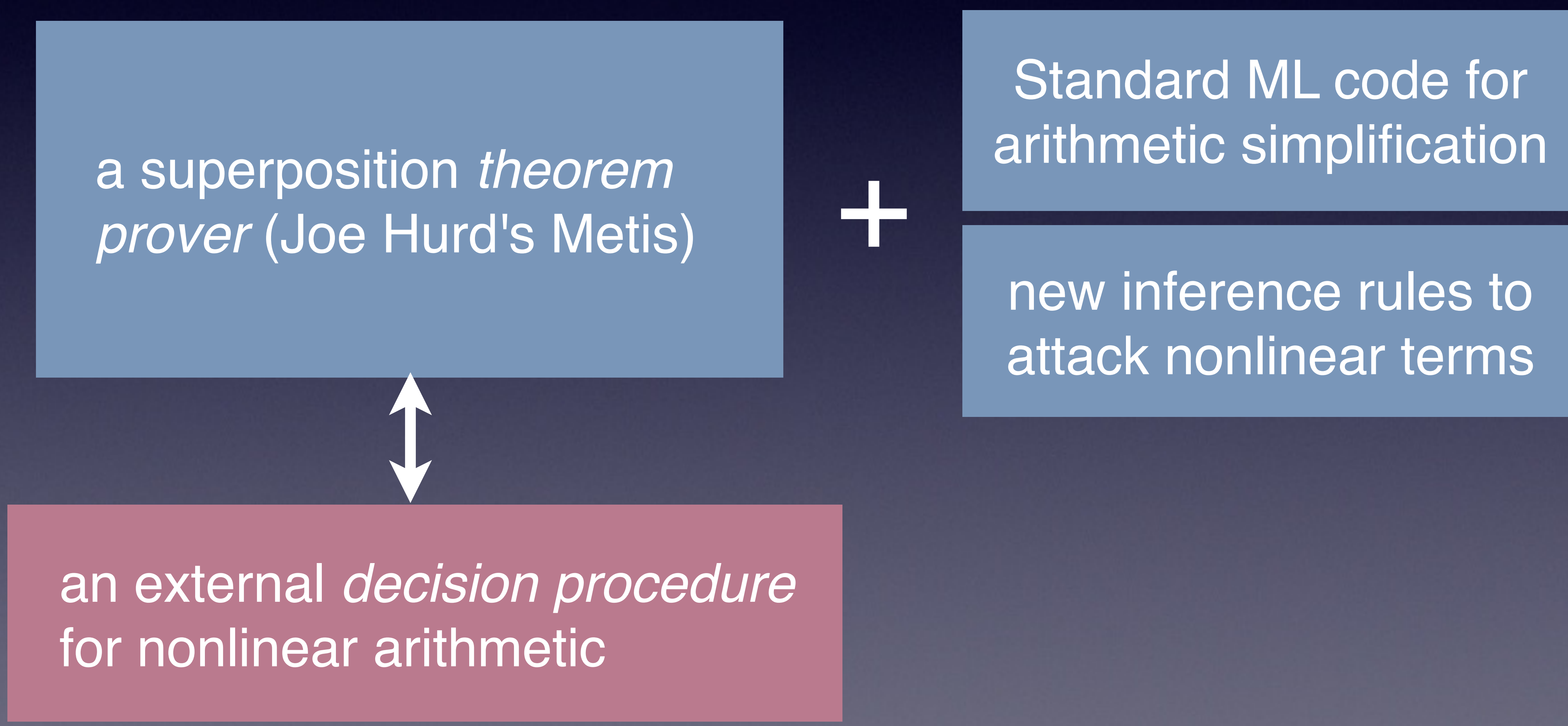
$$0 < x \wedge 0 < y \implies y\tanh(x) \leqslant \sinh(yx)$$

All proved in a few seconds!

How Does It Work?

- It's just *resolution*, augmented with
  - **axioms** giving upper/lower bounds for those functions (as polynomials or rational functions)
  - **heuristics** to isolate and remove occurrences of those functions
  - **decision procedures** to solve the resulting polynomial inequalities

Architecture



Some Upper/Lower Bounds

$$\exp(x) \geqslant 1 + x + \cdots + x^n/n! \qquad (n \text{ odd})$$

$$\exp(x) \leqslant 1 + x + \cdots + x^n/n! \qquad (n \text{ even}, \ x \leqslant 0)$$

$$\exp(x) \leqslant 1/(1 - x + x^2/2! - x^3/3!) \qquad (x < 1.596)$$

Taylor series, …

continued fractions, …

$$\frac{x-1}{x} \leqslant \ln x \leqslant x - 1$$

$$\frac{(1+5x)(x-1)}{2x(2+x)} \leqslant \ln x \leqslant \frac{(x+5)(x-1)}{2(2x+1)}$$

Analysing A Simple Problem

split on signs of expressions

split on sign of x

$$|\exp x - (1 + x/2)^2| \leqslant |\exp(|x|) - (1 + |x|/2)^2|$$

- *isolate* occurrences of functions
- … replace them by their *bounds*
- replace *division* by multiplication
- call some external *decision procedure*

A tweaked resolution loop does all this automatically!

The Decision Procedures

QEPCAD (Hoon Hong, C. W. Brown et al.)

Mathematica (Wolfram research)

Z3 (de Moura et al., Microsoft Research)
[now with nonlinear reasoning!]

A Key Heuristic: Algebraic Literal Deletion

- Resolution works with disjunctions of *literals*.
- We **delete** any literal inconsistent with known facts, according to the decision procedure.
- It’s a fine-grained integration between resolution and a decision procedure.

A Few Applications

- Abstracting non-polynomial dynamical systems (Denman)
- KeYmaera linkup: nonlinear hybrid systems (Sogokon et al.)
- PVS linkup: NASA collision-avoidance projects (Muñoz & Denman)

MetiTarski + PVS

- *Trusted* interface (MetiTarski as an oracle)
- Complementing the PVS support of branch-and-bound methods for polynomial estimation
- It's being tried within NASA's ACCoRD project.
- MetiTarski has been effective in early experiments
- … but there's much more to do.

3. Is MetiTarski Sound?

What Must We Trust?

- *Arithmetic simplification* and normalisation

  should be easy

- *Specialised axioms* giving upper or lower bounds of special functions

  see below!

- The external decision procedure

  not clear...

But we get machine-readable proofs!
(Resolution steps + extensions)

A Machine-Readable Proof

```
SZS output start CNFRefutation for abs-problem-14.tptp
 cnf(lgen_le_neg, axiom, (X <= Y | ~ lgen(0, X, Y))).

cnf(leq_left_divide_mul_pos, axiom

cnf(leq_right_divide_mul_pos

cnf(leq_right_divide_mul_neg

cnf(exp_positive, ax

cnf(ex
(~ -1 <= X | ~ lg

cnf(exp_l

(1 + X / 3 +
1 / 24 * (
```

nearly 200 steps!

```
.
.
.
cnf(refute_0_191, plain, ($false),
    inference(resolve,
        [$cnf(skoX *
          (21743271936 +
            skoX *
            (10871635968 +
              skoX *
              (3623878656 +
                skoX *
                (891813888 +
                  skoX *
                  (169869312 +
                    skoX *
                    (25657344 +
                      skoX *
                      (3096576 +
                        skoX *
                        (297216 +
                          skoX *
                          (22272 +
        skoX * (1248 + skoX * (48 +
```

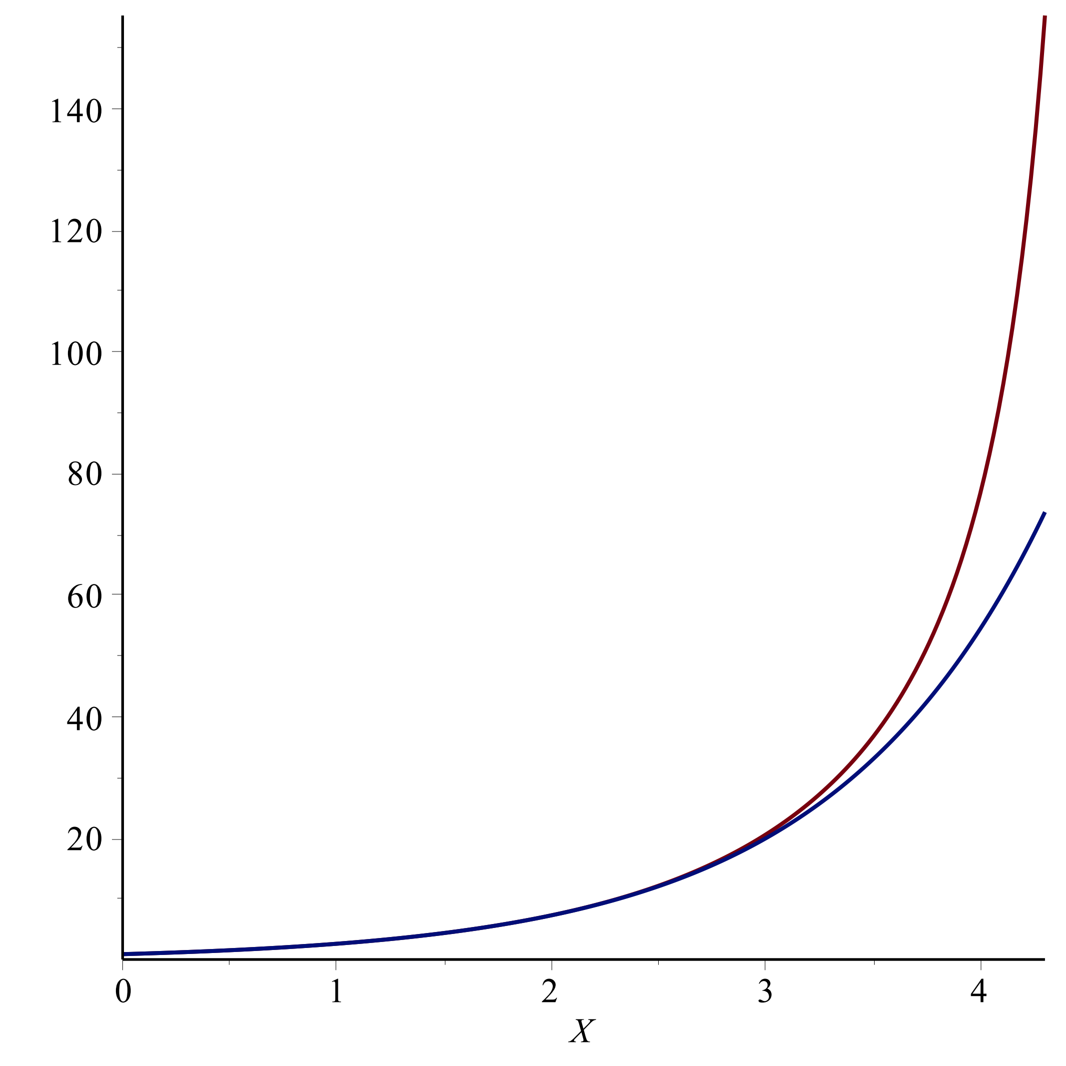
Verifying the Axioms

- *Taylor series expansions* are already verified for the elementary functions (sin, cos, $\tan^{-1}$, exp, ln).
- *Continued fractions* are much more accurate, but rely on advanced theory.
- Many of the axioms have now been verified using Isabelle, PVS, etc.

Bounding exp(x) Above

$$\mathrm{cf3}\,x \triangleq -\frac{x^3 + 12x^2 + 60x + 120}{x^3 - 12x^2 + 60x - 120}$$

- Based on a continued fraction
- **Singularity** around 4.644
- Can it be *proved* to be an upper bound in this range?



$\mathrm{cf3}\,x \geqslant \exp x \quad (0 \leqslant x \leqslant 4.644)$

By monotonicity of ln, enough to show

$$\ln(\mathrm{cf3}\,x) \geqslant x$$

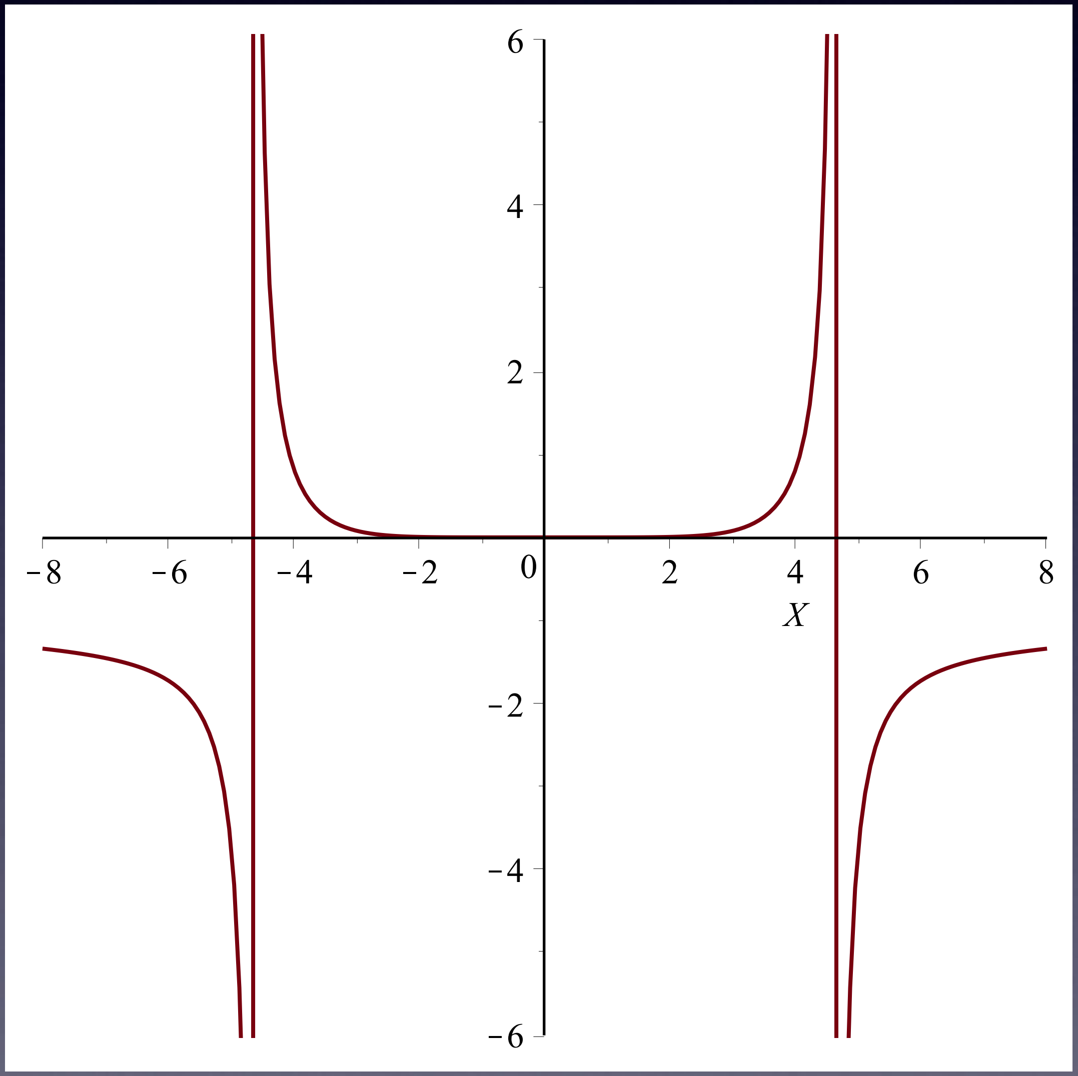
Take the derivative of the difference:

$$\frac{d}{dx}\left[\ln(\mathrm{cf3}\,x) - x\right] =$$

$$-\frac{x^6}{(x^3 - 12x^2 + 60x - 120)(x^3 + 12x^2 + 60x + 120)}$$

Here's that Derivative

- Singularities at ±4.644
- Nonnegative within that interval



That derivative is positive provided

$$x^3 - 12x^2 + 60x - 120 < 0$$

and in particular if $0 < x < 4.644$.

The result follows because also $\mathrm{cf3}(0) = 1 = \exp 0$

Similar techniques justify a *lower bound* axiom:

$$\mathrm{cf3}\, x \leqslant \exp x \quad (x \leqslant 0)$$

So the axioms are okay. What about the *decision procedures*?

- Nonlinear decision procedures rely on complicated computer algebra techniques …
- and real quantifier elimination is *doubly exponential* in the number of variables.
- Can they justify their answers with **evidence**?

This is a crucial research question!

4. The Way Forward

Goal: to Integrate MetiTarski with Other Tools

Computer algebra proof methods in various ITPs demonstrate the power of integrated tools.

Integration requires a way to validate nonlinear reasoning

The MetiTarski-PVS linkup is promising, but it's an oracle …

… and in turn, a substantial library of formalised mathematics.

Our Disorganised Libraries of Formal Mathematics

- created in bits and pieces by students and postdocs
- spread over many incompatible systems: Coq, HOL4 or HOL Light, Isabelle, Mizar, PVS, …
- based on a great variety of source texts

Goal: to Formalise a Body of Applied Mathematics

- *complex analysis*: the cornerstone of physics, engineering mathematics, etc.
- *real algebraic geometry*: the foundation of many computer algebra algorithms
- *approximation theory*: the foundation of numerical methods

Remember the QED Project?

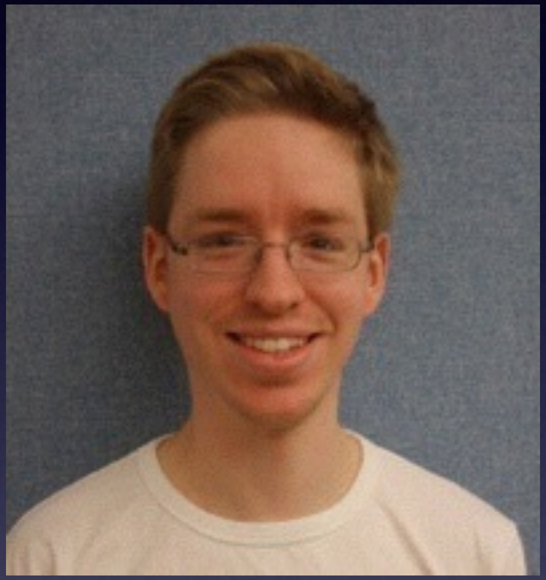
- That 1993 proposal to formalise all mathematics was too ambitious, and unconvincing to funders.
- Let's fix a more modest goal:

to formalise, and organise, the *core developments of applied mathematics*.

Can we do this?

The Cambridge Team

James Bridge



William Denman

Zongyan Huang

(to 2008: Behzad Akbarpour)

Acknowledgements

- *Edinburgh Team*: Paul Jackson, G Passmore, A Sogokon.
- Assistance from J. H. Davenport, J. Hurd, D. Lester, C. Muñoz, E. Navarro-López, etc.
- Supported by the Engineering and Physical Sciences Research Council [grant numbers EP/C013409/1,EP/I011005/1,EP/I010335/1].

MetiTarski (like Isabelle) is coded in **Standard ML**.