Mario Carneiro
July 1, 2014

Natural Deduction in the Metamath Proof Language

What is Metamath?

- A computer language for writing mathematical proofs
- A program to verify proofs in the Metamath language
- A library of completed proofs
  - Almost 20000 proofs exist in set.mm, the main collection of proofs based primarily on ZFC set theory
  - Covers introductory material in set theory, category theory, real analysis, number theory, algebra, topology, linear algebra, lattice theory, etc.

How does it work?

- Consider a logician's "formal proof"
- Formulas look something like $(v_1 \in v_2 \rightarrow \forall v_0\, v_1 \in v_2)$, with individual variables and no metavariables
- There are an infinite number of axioms, because there are no schemes (although axiomhood is decidable)
- Using schemes, each axiom is a substitution instance of just a few axiom schemes like $(\varphi \rightarrow (\psi \rightarrow \varphi))$
- In metamath, the "scheme" concept is extended to theorems

How does it work?

- Each step of a proof uses metavariables
- The result of the proof is a theorem scheme, which can be substituted in later theorems
- 1-1 correspondence of proof steps to logician's "formal proof"

Assertion

Ref	Expression
idl	$\vdash (\varphi \to \varphi)$

Proof of Theorem idl

Step	Hyp	Ref	Expression
1		ax-1 5	.2 $\vdash (\varphi \to (\varphi \to \varphi))$
2		ax-1 5	..3 $\vdash (\varphi \to ((\varphi \to \varphi) \to \varphi))$
3		ax-2 6	..3 $\vdash ((\varphi \to ((\varphi \to \varphi) \to \varphi)) \to ((\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi)))$
4	2, 3	ax-mp 8	.2 $\vdash ((\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi))$
5	1, 4	ax-mp 8	1 $\vdash (\varphi \to \varphi)$

Colors of variables: wff set class

Syntax hints: $\to$ wi 4

This theorem is referenced by: pm4.24 676 fz0n 13802 ldilval 19735 dib0 20846 dih1 20952 dihglblem5apre 20954

This theorem was proved from axioms: ax-1 5 ax-2 6 ax-mp 8

Advantages

- Conceptually simple foundations
- Core verifier is very small (one independent verifier is ≈300 lines of python)
- Fast proof verification (≈6 sec to verify ≈20000 proofs)
- Axioms are user-specified, so it is not tied to any particular logical foundation
  - Each proof in the Proof Explorer lists the axioms that were used to prove it, so it is possible to, say, track AC usage in a proof

Comparison to Mizar

- Proofs are in the form of formulas, not natural language
- Steps are much smaller in scope
  - Similar to C versus assembly
  - Possible target for “compilation” from higher level languages
- Simple open source verifier, public domain proofs
  - Follows QED philosophy: open source means independent verification
- No concept of “exported theorems”
  - All theorems have globally unique labels and are accessible by any later proof
- Hilbert-style proof system (every step of a proof is a theorem)

Some important theorems

- The following theorems have been formalized in set.mm:
  - Russell's paradox
  - Cantor's theorem
  - Schröder-Bernstein Theorem
  - Zorn's lemma
  - Irrationality of $\sqrt{2}$
  - Countability of $\mathbb{Q}$
  - Euler's thm. & Fermat's little thm.
  - Uncountability of $\mathbb{R}$
  - Bezout's theorem
  - Heine-Borel theorem
  - Bolzano-Weierstrass theorem
  - Infinitude of the primes
  - Fundamental Theorem of Arithmetic
  - Bertrand's postulate
  - Fundamental group of topology
  - Sum of k-th powers
  - Formula for Pythagorean triples
  - Cauchy-Schwarz inequality
  - Descargues's theorem
  - Baire category theorem
  - Riesz representation theorem

Some important theorems

- The following theorems have been formalized in set.mm:
  - Russell's paradox
  - Cantor's theorem
  - Schröder-Bernstein Theorem
  - Zorn's lemma
  - Irrationality of $\sqrt{2}$
  - Countability of $\mathbb{Q}$
  - Euler's thm. & Fermat's little thm.
  - Uncountability of $\mathbb{R}$
  - Bezout's theorem
  - Heine-Borel theorem
  - Bolzano-Weierstrass theorem
  - Infinitude of the primes
  - Fundamental Theorem of Arithmetic
  - Bertrand's postulate
  - Fundamental group of topology
  - Sum of k-th powers
  - Formula for Pythagorean triples
  - Cauchy-Schwarz inequality
  - Descargues's theorem
  - Baire category theorem
  - Riesz representation theorem

Some important theorems

- The following theorems have been formalized in set.mm:
  - Russell's paradox
  - Cantor's theorem
  - Schröder-Bernstein Theorem
  - Zorn's lemma
  - Irrationality of $\sqrt{2}$
  - Countability of $\mathbb{Q}$
  - Euler's thm. & Fermat's little thm.
  - Uncountability of $\mathbb{R}$
  - Bezout's theorem
  - Heine-Borel theorem
  - Bolzano-Weierstrass theorem
  - Infinitude of the primes
  - Fundamental Theorem of Arithmetic
  - Bertrand's postulate
  - Fundamental group of topology
  - Sum of k-th powers
  - Formula for Pythagorean triples
  - Cauchy-Schwarz inequality
  - Descargues's theorem
  - Baire category theorem
  - Riesz representation theorem

Deduction proofs

Theorem ovolicc 13569

Description: The measure of a closed interval. (Contributed by Mario Carneiro, 14-Jun-2014.)

Hypotheses

Ref	Expression
ovolicc.1	$\vdash (\varphi \to A \in \mathbb{R})$
ovolicc.2	$\vdash (\varphi \to B \in \mathbb{R})$
ovolicc.3	$\vdash (\varphi \to A \le B)$

Assertion

Ref	Expression
ovolicc	$\vdash (\varphi \to (\mathrm{vol}^*`(A[,]B)) = (B - A))$

- Developed to allow natural deduction in Metamath
- Each hypothesis and the conclusion start with "$\varphi \to$"
- Substitutions of $(\varphi \land \cdots)$ for φ take the place of assumption discharge
- Relies on Metamath's wff metavariables in an essential way

Deduction proofs

12	6, 10, 11	sylanc	$(B - A) \wedge (B - A) \leq (\mathrm{vol}^*`(A[,]B)))))$
13	1	adantr 490	....5 $\vdash ((\varphi \wedge x \in \mathbb{R}^+) \to A \in \mathbb{R})$
14	2	adantr 490	....5 $\vdash ((\varphi \wedge x \in \mathbb{R}^+) \to B \in \mathbb{R})$
15		ovolicc.3	.....6 $\vdash (\varphi \to A \leq B)$
16	15	adantr 490	....5 $\vdash ((\varphi \wedge x \in \mathbb{R}^+) \to A \leq B)$
17		simpr 485	....5 $\vdash ((\varphi \wedge x \in \mathbb{R}^+) \to x \in \mathbb{R}^+)$
18	13, 14, 16, 17	ovolicc1 13562	...4 $\vdash ((\varphi \wedge x \in \mathbb{R}^+) \to (\mathrm{vol}^*`(A[,]B)) \leq ((B - A) + x))$
19	18	ralrimiva 2317	..3 $\vdash (\varphi \to \forall x \in \mathbb{R}^+ (\mathrm{vol}^*`(A[,]B)) \leq ((B - A) + x))$
			...4 $\vdash (((\mathrm{vol}^*`(A[,]B)) \in \mathbb{R}^* \wedge (B - A) \in \mathbb{R}) \to ((\mathrm{vol}^*`(A$

Theorem ovolicc 13569

Description: The measure of a closed interval. (Contributed by Mario Carneiro, 14-Jun-2014.)

Hypotheses

Ref	Expression
ovolicc.1	$\vdash (\varphi \to A \in \mathbb{R})$
ovolicc.2	$\vdash (\varphi \to B \in \mathbb{R})$
ovolicc.3	$\vdash (\varphi \to A \leq B)$

Assertion

Ref	Expression
ovolicc	$\vdash (\varphi \to (\mathrm{vol}^*`(A[,]B)) = (B - A))$

Theorem ovolicc1 13562

Description: The measure of a closed interval is lower bounded by its length. (Contributed by Mario Carneiro, 13-Jun-2014.)

Hypotheses

Ref	Expression
ovolicc.1	$\vdash (\varphi \to A \in \mathbb{R})$
ovolicc.2	$\vdash (\varphi \to B \in \mathbb{R})$
ovolicc.3	$\vdash (\varphi \to A \leq B)$
ovolicc1.4	$\vdash (\varphi \to C \in \mathbb{R}^+)$

Assertion

Ref	Expression
ovolicc1	$\vdash (\varphi \to (\mathrm{vol}^*`(A[,]B)) \leq ((B - A) + C))$

Deduction Theorem

- A theorem of classical logic which justifies natural deduction proof methods
- If $\Gamma \cup \{\varphi\} \vdash \psi$, then $\Gamma \vdash (\varphi \to \psi)$
- The proof operates by modifying each proof step in a proof of $\Gamma \cup \{\varphi\} \vdash \psi$ to produce an equivalent proof of $\Gamma \vdash (\varphi \to \psi)$
- We can’t use this “theorem” directly in Metamath because it is at the meta-proof level
  - We work with actual proofs – we need to actually show a proof, not just prove that a proof exists nonconstructively
- But we can “implement” the theorem’s reductions to produce an actual proof
  - Efficiency matters!

Deduction Theorem

Axioms of propositional calculus		
Axiom *Simp*	ax-1	$\vdash (\varphi \to (\psi \to \varphi))$
Axiom *Frege*	ax-2	$\vdash ((\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi)))$
Axiom *Transp*	ax-3	$\vdash ((\neg \varphi \to \neg \psi) \to (\psi \to \varphi))$
Rule of Modus Ponens	ax-mp	$\vdash \varphi$ & $\vdash (\varphi \to \psi) \Rightarrow \vdash \psi$

- The textbook deduction theorem works on a logician's "formal proof": no theorems or metavariables allowed
- Every step is an instance of one of the axioms, or the inference rule ax-mp
- Each step S_i is converted to $S_i' \stackrel{\text{def}}{=} \varphi \to S_i$
- If S_i is an axiom or in Γ, then $\varphi \to S_i$ is proven in two extra steps from S_i (theorem a1i)
- If S_i is ax-mp applied to two previous steps S_j, $(S_j \to S_i)$, then $\varphi \to S_i$ can be proven in three steps from $\varphi \to S_j$, $\varphi \to (S_j \to S_i)$ (theorem mpd)

Deduction Theorem

Axioms of propositional calculus

Axiom *Simp*	ax-1	$\vdash (\varphi \to (\psi \to \varphi))$
Axiom *Frege*	ax-2	$\vdash ((\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi)))$
Axiom *Transp*	ax-3	$\vdash ((\neg \varphi \to \neg \psi) \to (\psi \to \varphi))$
Rule of Modus Ponens	ax-mp	$\vdash \varphi$ & $\vdash (\varphi \to \psi)$ $\Rightarrow$ $\vdash \psi$

- If S_i is φ (the assumption), then $\varphi \to \varphi$ is proven in five steps (theorem id) from no assumptions
- Result: A ≈3x increase in number of steps of a direct-from-axioms proof
- If we allow the usage of theorems a1i, mpd, id, this can be decreased to 1x, but that's not fair since the original proof had no theorem references

Deduction Theorem, v2

- Consider the general case, with a set of theorems all referencing each other
- We count steps as the sum of the steps in each theorem, even if a theorem is used many times with different substitutions
  - exponential gain over direct-from-axioms step count
- Construct a set A of basic theorems that we will need
  - $\mathrm{a1i}$, mpd, id
  - $\mathrm{a1i}$ applied to each axiom
- Our transformation will add "$\varphi \rightarrow$" as a prefix to each hypothesis and the conclusion of every theorem T_i in the collection

Deduction Theorem, v2

- The result is the statement that $\Gamma \vdash \psi$ implies $(\varphi \to \Gamma) \vdash (\varphi \to \psi)$
  - To get the standard version "$\Gamma \cup \{\varphi\} \vdash \psi$ implies $\Gamma \vdash (\varphi \to \psi)$", apply a1i to each hypothesis and prove the redundant hypothesis $(\varphi \to \varphi)$ using id
- Transformation is the same as before, only now we use one step proofs only
- If S_i is an axiom, then $\varphi \to S_i$ is a theorem from A
- If $S_j, (S_j \to S_i) \Rightarrow S_i$ is an application of ax-mp, then $\varphi \to S_j, \varphi \to (S_j \to S_i) \Rightarrow \varphi \to S_i$ is an application of mpd
- If S_i is an application of a previous theorem T_k, then the transformed theorem T_k', which already has "$\varphi \to$" in its hypotheses and conclusion, correctly proves S_i' from the transformed previous steps

Deduction Theorem, v2

- The net result is that the total number of steps increases only by the number of steps in the theorems of A, which is a fixed constant
  - But we had to change every theorem in the collection just to discharge one hypothesis in one theorem!
- Solution: theorems imp, ex: $\big((\varphi \wedge \psi) \rightarrow \chi\big) \Leftrightarrow \big(\varphi \rightarrow (\psi \rightarrow \chi)\big)$
- Call a 1-deduction a theorem where each hypothesis and conclusion is already in the form $(\varphi \rightarrow \cdots)$, a 2-deduction for theorems of the form $\big(\varphi \rightarrow (\psi \rightarrow \cdots)\big)$, etc.
- We want to keep the old versions of each theorem, to minimize the effect on the collection

Multiple application

- Call a 1-deduction a theorem where each hypothesis and conclusion is already in the form $(\varphi \to \cdots)$, a 2-deduction for theorems of the form $\big(\varphi \to (\psi \to \cdots)\big)$, etc.
- The algorithm just described turns a 0-deduction into a 1-deduction, a 1-deduction into a 2-deduction, etc.
- Any usage of a 2-deduction theorem can be converted to the equivalent 1-deduction theorem by using imp on each hypothesis (turns $\big(\varphi \to (\psi \to \cdots)\big)$ into $\big((\varphi \land \psi) \to \cdots\big)$) and ex on the conclusion (goes the other direction)
- Overhead proportional to the number of hypotheses

Conclusion

- All theorems that are already 1-deductions and only reference 1-deductions are left unchanged
  - Algorithm is idempotent on its output
- In a typical application, only the target theorem is modified, and overhead is proportional to the number of hypotheses to the theorem
  - If $\Gamma \cup \{\varphi\} \vdash \psi$ in n steps, then $\Gamma \vdash (\varphi \to \psi)$ in $n + |\Gamma| + O(1)$ steps
- Natural deduction can be implemented in a Hilbert system like Metamath with only a constant overhead, if $|\Gamma|$ is bounded
- Not discussed: predicate calculus & bound variables
  - Empirical evidence that it is rarely an issue

Questions