The benefits and challenges of teaching proof with Lean

Gihan Marasingha

University of Exeter

Computer-verified proof: 48 hours in Rome, Tor Vergata, January 2024

Outline

Ideas for talk

- Focus on USP: very first introduction to proof and Lean.
- Use of parse trees.
- Radically changing maths to fit with Lean. Example: use of type theory rather than set theory.
- Problems this raises with colleagues.
- Challenges: need more Lean in higher-level modules. Embed Lean throughout programme.
- Lean and GPT.
- Writing legible Lean code.

Teaching with Lean

- Logic and Proof at Carnegie Mellon University, for general audience, by Avigad, Lewis, and van Doorn. Term-based.
- Natural number game at Imperial College, for general audience, by Buzzard. Game engine by Pedramfar. Tactic-based.
- Logique et démonstrations assistées par ordinateur at Université Paris-Saclay, by Patrick Massot. Uses controlled natural language. Based on ideas of convergence of sequences.
- The Mechanics of Proof at Fordham University, by Heather Macbeth. An introduction to proof module.
- Modern Mathematics with Lean at the University of Exeter, by Gihan Marasingha.

See the Lean Community Courses Using Lean web page.

Context

- Large ($\sim$ 260 students) first-year pure maths class: logic, sets, functions, sequences, series, groups, vector spaces, elementary multiplicative number theory.
- Difficulties for new undergraduates in mathematics:
  - the syntax and semantics of mathematical language;
  - expectations concerning the nature of mathematics;
  - expectations concerning the nature of education.

Context

- Large ($\sim$ 260 students) first-year pure maths class: logic, sets, functions, sequences, series, groups, vector spaces, elementary multiplicative number theory.
- Difficulties for new undergraduates in mathematics:
  - the syntax and semantics of mathematical language;
  - expectations concerning the nature of mathematics;
  - expectations concerning the nature of education.
- I encountered Lean in summer 2019 via the Logic and Proof teaching resource of Avigad, Lewis, and van Doorn.
  I have taught my large module using Lean since then.

Context

- Large ($\sim$ 260 students) first-year pure maths class: logic, sets, functions, sequences, series, groups, vector spaces, elementary multiplicative number theory.
- Difficulties for new undergraduates in mathematics:
  - the syntax and semantics of mathematical language;
  - expectations concerning the nature of mathematics;
  - expectations concerning the nature of education.
- I encountered Lean in summer 2019 via the Logic and Proof teaching resource of Avigad, Lewis, and van Doorn.
  I have taught my large module using Lean since then.
- If Lean is the answer, what is the question?

Introduction

Talk overview:

- Identify common student ‘errors’ and challenges.
- Position challenges within a pedagogical framework.
- Discuss how I have used Lean in teaching.
- Criticise my use of Lean in relation to the framework.
- Discuss future developments.

Challenges and common errors

- Errors in propositional logic
- Category mistakes
- Using descriptions rather than definitions
- Quantifier blindness
- Relying on examples rather than definitions
- Poor engagement with feedback

Errors in propositional logic

- Believing $p \vee q$ means 'p or q, but not both p and q'.
- Treating a statement and its converse as being equal.
- Believing p follows from $p \to q$.
- Believing $\neg(p \to q)$ equals $\neg p \to \neg q$.

Errors in propositional logic

- Believing $p \vee q$ means 'p or q, but not both p and q'.
- Treating a statement and its converse as being equal.
- Believing p follows from $p \to q$.
- Believing $\neg(p \to q)$ equals $\neg p \to \neg q$.

Framework: set-befores and met-befores (Tall 2008)

Set-before: innate mental structure that shapes learning: pattern recognition, repetition of actions, language use to describe and refine the way we think about things.
Met-before: 'a current mental facility based on specific prior experiences of the individual'.

Unhelpful met-befores for propositional logic

Framework: unhelpful 'met-before'

Task: Solve the equation $2x + 3 = 7$.

Student's Incorrect Solution: '$\implies$' as punctuation

$$2x + 3 = 7, \qquad 2x = 4, \qquad 2x \implies x = 2$$

Unhelpful met-befores for propositional logic

Framework: unhelpful 'met-before'

Task: Solve the equation $2x + 3 = 7$.

Student's Incorrect Solution: ' $\implies$ **' as punctuation**

$$2x + 3 = 7, \qquad 2x = 4, \qquad 2x \implies x = 2$$

What $p \to q$ means

To the student $p \to q$ may mean "we **have** p and we get q from p"

Unhelpful met-befores for propositional logic

Framework: unhelpful 'met-before'

Task: Solve the equation $2x + 3 = 7$.

Student's Incorrect Solution: ' $\implies$ ' as punctuation

$$2x + 3 = 7, \qquad 2x = 4, \qquad 2x \implies x = 2$$

What $p \to q$ means

To the student $p \to q$ may mean "we **have** p and we get q from p"

Conversational Implicature (Huddleston and Pullum 2002)

How does a student interpret: 'If you submit your work late, it will be capped at 40%'?

Category mistakes

- Set membership and set inclusion.

Task

Prove or disprove $\emptyset \in \emptyset$.

Incorrect answer

Nothing belongs to the empty set. And $\emptyset$ is nothing, so $\emptyset \in \emptyset$.

- Vectors and expressions involving their components.

Task

Is $U = \{(x, y, z) \in \mathbb{R}^3 : x + yz = 0\}$ a subspace of $\mathbb{R}^3$?

A (partial) student answer

Let $\lambda \in F$ $u \in U$, $\quad \lambda u = \lambda(x + yz) = \lambda x + \lambda yz = 0 \in U$.

Functions are expressions

Confusion between predicates and 'ordinary' functions (or expressions), equally between propositions and terms of other types.

Task

Prove $\sum_{i=0}^{n} i = \frac{1}{2}n(n+i)$ for every natural number n.

Common error

Let $P(n) = \frac{1}{2}n(n+1)$. Then $P(0) = \frac{1}{2} \cdot 0 \cdot (0+1)$ is true. Assume $P(k) = \frac{1}{2}k(k+1)$, then [...]'.

Descriptions vs. definitions

Exam question

What does it mean for a sequence $(a_n)_{n=0}^{\infty}$ of real numbers to converge to a real number a?'

Common errors

It means $a_n \to a$ as $n \to \infty$.
It means a_n gets closer and closer to a, but never reaches a.

Descriptions vs. definitions

Exam question

What does it mean for a sequence $(a_n)_{n=0}^{\infty}$ of real numbers to converge to a real number a?'

Common errors

It means $a_n \to a$ as $n \to \infty$.
It means a_n gets closer and closer to a, but never reaches a.

Framework: stipulative vs. descriptive definitions (see Gupta 2021)

A stipulative definition imparts a meaning to the defined term, and involves no commitment that the assigned meaning agrees with prior uses (if any) of the terms.
Descriptive definitions spell out meaning but also aim to be adequate to existing usage.

Quantifier blindness; scope problems

Exam question

Prove that there exists a set A such that for every set B, $A \subseteq B$.

Common errors

Let $A = B$. Then $A \subseteq B$ as $B \subseteq B$.

Quantifier blindness; scope problems

Exam question

Prove that there exists a set A such that for every set B, $A \subseteq B$.

Common errors

Let $A = B$. Then $A \subseteq B$ as $B \subseteq B$.

The Three Worlds of Mathematics (Tall 2008)

- Conceptual-embodied
- Proceptual-symbolic
- Axiomatic-formal

Relying on examples

Task

Let $\phi : V \to U$ be a linear map. Let $\theta : V \to U$ where $\theta(x) = 2\phi(x)$. Prove that θ is a linear map.

A student dialogue

Student: How do you do this question? I haven't seen anything like this in the lecture notes.
Instructor: What's the definition of linear map?
Student: I don't know.

Effective Engagement with Feedback

- Students often ignore feedback.
- Traditional feedback mechanisms result in delayed feedback.
- Lack of opportunity for incremental improvement.

Potential for Lean

Pros: instant feedback, incremental improvement possible.
Cons: Error messages can be inscrutable. The error may not be on the indicated line.

Proof with context

- Student vocabulary of goals, target, context, hypotheses.
- Repetition of work with stipulative definitions.
- Opportunity to undo incorrect ‘met-befores’ of logic. In Lean, we do not *have* p given $p \to q$.
- Category mistakes are type errors.
- Working explicitly with quantifiers avoids scope problems.
- Reasoning must be explicit. Use simp lemmas to avoid messiness.

Mode of delivery

- Lecture notes alternate 'straight maths' and Lean chapters.
- Online Lean via the Modern Mathematics with Lean game.
- Student induction via the MIU game.
- Assessed via GitHub Classroom + GitHub Codespaces.
- Support via computer labs and 'coding clubs'.

Summative assessment

GitHub Classroom:

- Each student has a repository.
- Work in online vs code container via GitHub Codespaces.
- Submit by pushing commit.
- Work is autograded: either perfect score or zero.

Reflections:

- Work needed to reconcile with University mark system.
- Students need to be taught how to use the system.
- Manually marking required if a student submits an incomplete proof.
- Plagiarism is easy.
- GitHub Classroom is poorly maintained.

Reflections on process

- Web Assembly Lean in a browser can be very slow.

Reflections on process

- Web Assembly Lean in a browser can be very slow.
- Use of ‘gapped proofs’. Do the ‘messy’ bits for the students.

Theorem

Let S be the set of non-negative real numbers such that $x^2 < 2$. Then S has a supremum.

```
lemma sqrt2_set_has_sup : has_sup {x : ℝ | 0 ≤ x ∧ x ^ 2 < 2} :=
```

Proof :

```
begin
let S := {x : ℝ | 0 ≤ x ∧ x ^ 2 < 2},
show ∃ (u : ℝ), is_sup {x : ℝ | 0 ≤ x ∧ x ^ 2 < 2} u,
have h₁ : (0 : ℝ) ∈ S,
{ show (0 : ℝ) ≤ 0 ∧ (0 : ℝ)^2 < 2,
  sorry, },
have h₂ : S.nonempty, from nonempty_of_mem h₁,
have h₃ : bounded_above S,
{ sorry, },
show has_sup S, sorry,
```

Reflections on process

- Web Assembly Lean in a browser can be very slow.
- Use of ‘gapped proofs’. Do the ‘messy’ bits for the students.
- Coercions: unnatural from a set-theoretic perspective, but maybe not to students who think $5 \notin \mathbb{C}$.

Reflections on process

- Web Assembly Lean in a browser can be very slow.
- Use of ‘gapped proofs’. Do the ‘messy’ bits for the students.
- Coercions: unnatural from a set-theoretic perspective, but maybe not to students who think $5 \notin \mathbb{C}$.
- *Teach* and *assess* how to deal with error messages.

Reflections on process

- Web Assembly Lean in a browser can be very slow.
- Use of 'gapped proofs'. Do the 'messy' bits for the students.
- Coercions: unnatural from a set-theoretic perspective, but maybe not to students who think $5 \notin \mathbb{C}$.
- *Teach* and *assess* how to deal with error messages.
- Tactic-mode (programming) vs. term-style proof.

```
example (P Q : X → Prop) (h : ∃ (x : X), P(x) ∧ Q(x)) :
∃ (z : X), Q(z) :=
begin
  cases h with y h₂,
  given h₂ : P(y) ∧ Q(y),
  use y,
  show Q(y), from h₂.right,
end
```

Annoyances and difficulties for students

- (Incorrectly) believing they have a correct traditional proof that they cannot implement in Lean.
- The error messages are incomprehensible. Students would prefer natural language errors.
- The error messages do not help the student to fix their code.
- How is Lean used in later level courses.
- Challenges working with inductive and recursive definitions.
- Not understanding what variables and hypothesis are 'available'. For example, given a proof of

$$\exists x, f(x) \leq 5$$

a student make suspect that they have access to a variable x.

Development

Book:

- Starts with creating parse trees for expressions, understanding definitions, types, functions.
- Covers natural numbers, logic, sets, functions, real numbers, sequences.
- Alternates between ‘straight maths’ and Lean.
- Teaches how to deal with feedback.

Conclusion

- Potential of Lean to develop proof skills.
- It's a new mode of teaching: to use it successfully, understand how students think.

Questions?

Thank you for your attention. Questions?

Gupta, Anil (2021). "Definitions". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2021. Metaphysics Research Lab, Stanford University.

Huddleston, Rodney D. and Geoffrey K. Pullum (Apr. 2002). *The Cambridge Grammar of the English Language*. Cambridge University Press.

Tall, David (2008). "The transition to formal thinking in mathematics". In: *Mathematics Education Research Journal* 20.2, pp. 5–24.