Programs and Proofs

Anthony Wang (xy)

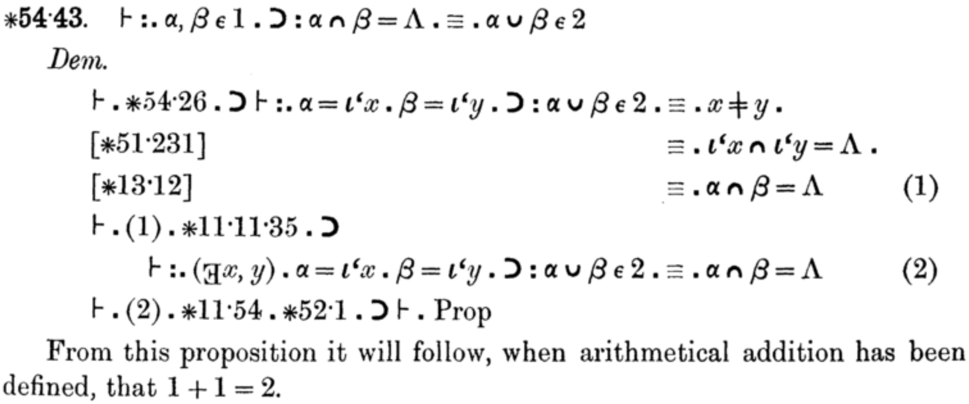
January 7, 2026

Cool Lean projects

- Raytracer
- Video player
- Rupert
- SciLean
- Equational theories
- Webring generator
- Functorio
- HouLean
- Mathlib
- Analysis textbook
- Erdős 707
- LeanTeX

History of formalized math

- 1910: Principia Mathematica



$*54\cdot43.\quad \vdash :. \alpha, \beta \in 1 . \supset : \alpha \cap \beta = \Lambda . \equiv . \alpha \cup \beta \in 2$

Dem.

$$\vdash . *54\cdot26 . \supset \vdash :. \alpha = \iota\text{‘}x . \beta = \iota\text{‘}y . \supset : \alpha \cup \beta \in 2 . \equiv . x \neq y .$$

$$[*51\cdot231] \qquad \equiv . \iota\text{‘}x \cap \iota\text{‘}y = \Lambda .$$

$$[*13\cdot12] \qquad \equiv . \alpha \cap \beta = \Lambda \tag{1}$$

$$\vdash . (1) . *11\cdot11\cdot35 . \supset$$

$$\vdash :. (\exists x, y) . \alpha = \iota\text{‘}x . \beta = \iota\text{‘}y . \supset : \alpha \cup \beta \in 2 . \equiv . \alpha \cap \beta = \Lambda \tag{2}$$

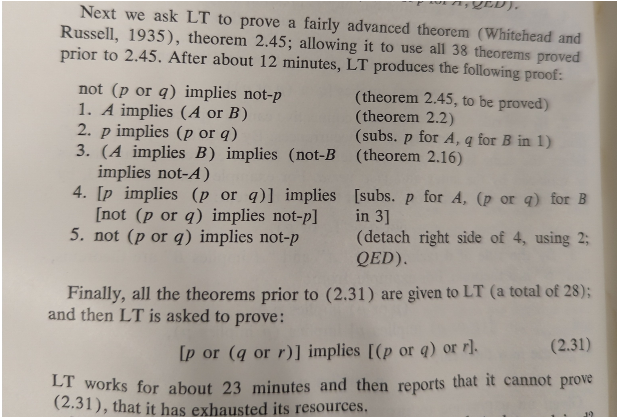
$$\vdash . (2) . *11\cdot54 . *52\cdot1 . \supset \vdash . \text{Prop}$$

From this proposition it will follow, when arithmetical addition has been defined, that $1 + 1 = 2$.

- 1931: Gödel's incompleteness theorems

History (cont.)

- 1936: Entscheidungsproblem proven undecidable
- 1956: Logic Theorist ("first AI program")



Next we ask LT to prove a fairly advanced theorem (Whitehead and Russell, 1935), theorem 2.45; allowing it to use all 38 theorems proved prior to 2.45. After about 12 minutes, LT produces the following proof:

not (*p* or *q*) implies not-*p*	(theorem 2.45, to be proved)
1. *A* implies (*A* or *B*)	(theorem 2.2)
2. *p* implies (*p* or *q*)	(subs. *p* for *A*, *q* for *B* in 1)
3. (*A* implies *B*) implies (not-*B* implies not-*A*)	(theorem 2.16)
4. [*p* implies (*p* or *q*)] implies [not (*p* or *q*) implies not-*p*]	[subs. *p* for *A*, (*p* or *q*) for *B* in 3]
5. not (*p* or *q*) implies not-*p*	(detach right side of 4, using 2; *QED*).

Finally, all the theorems prior to (2.31) are given to LT (a total of 28); and then LT is asked to prove:

[*p* or (*q* or *r*)] implies [(*p* or *q*) or *r*]. (2.31)

LT works for about 23 minutes and then reports that it cannot prove (2.31), that it has exhausted its resources.

History (cont.)

- 1976: Four color theorem proved using brute force (verified in Rocq in 2005)

ITPs vs ATPs

- Two main paradigms
- ITP = Interactive theorem prover, uses tactics, ex: Rocq, Lean
- ATP = Automated ..., uses SMT, ex: Dafny
- ATPs are buggier, more brittle, require learning arcane SMT magic

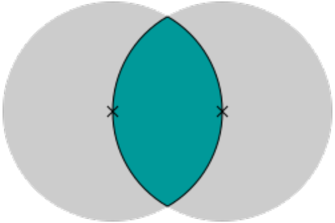
Foundations

- Set theory (Mizar, Metamath)
- Simple type theory (Isabelle/HOL)
- Dependent type theory (Lean, Rocq, Agda, Idris)

Lean bio

- 2013: Created by Leo de Moura at Microsoft, previously created Z3
- 2023: Lean 4 released, rewritten in Lean (except type checker)
- Not named after the drug

Is Lean practical?



- "Invisible math"
- Automated tactics: grind, hammer, canonical
- AI: Harmonic's Aristotle, AlphaProof