Homotopy Type Theory in Lean

Floris van Doorn

Department of Philosophy Carnegie Mellon University http://leanprover.github.io

25 June 2016

- The Lean Theorem Prover
- Lean's kernel
- Lean's elaborator
- Demo
- The HoTT library

Lean is a new interactive theorem prover, developed principally by Leonardo de Moura at Microsoft Research, Redmond.

It was "announced" in the summer of 2015.

It is open source, released under a permissive license, Apache 2.0.

The goal is to make it a community project, like Clang.

The aim is to bring interactive and automated reasoning together, and build

- an interactive theorem prover with powerful automation
- an automated reasoning tool that
 - produces (detailed) proofs,
 - has a rich language,
 - can be used interactively, and
 - is built on a verified mathematical library.

Lean is a designed to be a mature system, rather than an experimental one.

- Take advantage of existing theory.
- Build on strengths of existing interactive and automated theorem provers.
- Craft clean but pragmatic solutions.

We have drawn ideas and inspiration from Coq, SSReflect, Isabelle, Agda, and Nuprl, among others.

The Lean Theorem Prover

Notable features:

- based on a powerful dependent type theory
- written in C++, with multi-core support
- small, trusted kernel with an independent type checker
- standard and HoTT instantiations
- powerful elaborator
- can use proof terms or tactics
- Emacs mode with proof-checking on the fly
- browser version runs in javascript
- already has a respectable library
- automation is now the main focus

Currently working on the code base: Leonardo de Moura, Daniel Selsam, Lev Nachman, Soonho Kong

Currently working the standard library: Jeremy Avigad, Rob Lewis, Jacob Gross

Currently working on the HoTT library: Floris van Doorn, Ulrik Buchholtz, Jakob von Raumer

Contributors: Cody Roux, Joe Hendrix, Parikshit Khanna, Sebastian Ullrich, Haitao Zhang, Andrew Zipperer, and many others.

Lean's kernel

Lean's kernel implements dependent type theory with

- A hierarchy of (non-cumulative) universes: Type.{0} : Type.{1} : Type.{2} : Type.{3} : ...
- universe polymorphism: definition id.{u} {A : Type.{u}} : A \rightarrow A := λ a, a
- dependent products: Πx : A, B
- inductive types (à la Dybjer, constructors and recursors)

The kernel is smaller and simpler than those of Coq and Agda.

Daniel Selsam has written an independent type checker in Haskell, which is less than 2,000 lines long.

The kernel can be instantiated into two modes, a standard mode and the HoTT mode.

Lean's kernel

Definitions like these are compiled down to recursors:

```
definition tail {A : Type} :
  \Pi{n}, vector A (succ n) \rightarrow vector A n
| tail (h :: t) := t
definition zip {A B : Type} :
  \Pi{n}, vector A n \rightarrow vector B n \rightarrow vector (A \times B) n
| zip nil nil
                         := nil
| zip (a::va) (b::vb) := (a, b) :: zip va vb
definition diag : \Pi{n}, vector (vector A n) n \rightarrow vector A n
| diag nil
                          := nil
| diag ((a :: v) :: M) := a :: diag (map tail M)
```

Standard mode

Specific to the standard mode:

• Type. {0} (aka Prop) is impredicative and proof irrelevant.

quotient types (lift f H (quot.mk x) = f x definitionally).
 We use additional axioms for classical reasoning: propositional extensionality, Hilbert choice.

Lean keeps track of which definitions are computable.

noncomputable definition inv_fun (f : X \rightarrow Y) (s : set X) (dflt : X) (y : Y) : X := if H : $\exists_0 x \in s$, f x = y then some H else dflt

definition add (x y : \mathbb{R}) : \mathbb{R} := ...

```
noncomputable definition div (x y : \mathbb{R}) : \mathbb{R} := x * y^{-1}
```

In the HoTT mode:

- There is no Prop.
- There are two primitive HITs in Lean: the *n*-truncation and *quotients*. Given A : U and R : A → A → U the quotient is: HIT *quotient_A(R)* :=
 - $q: A \rightarrow quotient_A(R)$
 - $\Pi(x, y : A), R(x, y) \rightarrow q(x) = q(y)$

Many HITs can be constructed from these two.

The Univalence Axiom is assumed.

Fully detailed expressions in dependent type theory are long.

Systems of dependent type theory allow users to leave a lot of information implicit.

Such systems therefore:

- Parse user input.
- Fill in the implicit information.

The last step is known as "elaboration."

Lean has a powerful elaborator that handles:

- implicit universe levels
- higher-order unification
- computational reductions
- ad-hoc overloading
- coercions
- type class inference
- tactic proofs

variables {A : Type} {B : A \rightarrow Type} {C : Π a, B a \rightarrow Type} definition sigma_transport {a₁ a₂ : A} (p : a₁ = a₂) (x : Σ (b : B a₁), C a₁ b) : p \triangleright x = \langle p \triangleright x.1, p \triangleright D x.2 \rangle := by induction p; induction x; reflexivity

 \triangleright is transport. .1 and .2 are projections.

The first transport is along the family $\lambda(a : A)$, $\Sigma(b : B a)$, C a b The second transport is along B The last transport is a "dependent transport" which is a map

 $C a_1 x.1 \rightarrow C a_2 (p \triangleright x.1)$

Structures and type class inference have been optimized to handle the algebraic hierarchy.

The algebraic hierarchy consist of:

- order structures (including lattices, complete lattices)
- additive and multiplicative semigroups, monoids, groups,
- rings, fields, ordered rings, ordered fields,
- modules over arbitrary rings, vector spaces, normed spaces, ...
- homomorphisms preserving appropriate parts of structures

```
structure has_mul [class] (A : Type) := (mul : A \rightarrow A \rightarrow A)
```

```
structure semigroup [class] (A : Type) extends has_mul A :=
  (is_set_carrier : is_set A)
  (mul_assoc : \delta b c, mul (mul a b) c = mul a (mul b c))
```

```
structure monoid [class] (A : Type) extends semigroup A, has_one A :=
  (one_mul : ∀a, mul one a = a) (mul_one : ∀a, mul a one = a)
```

```
variables {A : Type} [monoid A]
definition pow (a : A) : \mathbb{N} \to A
| 0 := 1
| (n+1) := pow n * a
```

```
theorem pow_add (a : A) (m : N) : ∀n, a^(m + n) = a^m * a^n
| 0 := by rewrite [nat.add_zero, pow_zero, mul_one]
| (n+1) := by rewrite [add_succ, *pow_succ, pow_add, mul.assoc]
```

```
definition int.linear_ordered_comm_ring [instance] :
    linear_ordered_comm_ring int := ...
```

```
Floris van Doorn (CMU)
```

Type classes are also used in the standard library:

- to infer straightforward facts (finite s, is_subgroup G)
- to simulate classical reasoning constructively (decidable p)

They are used in the HoTT library:

- for category theory
- to infer truncatedness (is_trunc n A)
- to infer half-adjoint equivalence (is_equiv f)

Calculational proofs

$$\sum_{(a,b):\Sigma_{a:A}B(a)} C(a) \simeq \sum_{(a,b):\Sigma_{a:A}C(a)} B(a)$$

definition sigma_assoc_comm_equiv {A : Type} (B C : A \rightarrow Type) : (Σ (v : Σ a, B a), C v.1) \simeq (Σ (u : Σ a, C a), B u.1) := calc (Σ (u : Σ a, B a), C v.1)

$$\begin{array}{l} (\Sigma(\texttt{v}: \Sigma\texttt{a}, \texttt{B}\texttt{a}), \texttt{C}\texttt{v}.\texttt{I}) \\ \simeq (\Sigma\texttt{a} (\texttt{b}: \texttt{B}\texttt{a}), \texttt{C}\texttt{a}) & : \texttt{sigma_assoc_equiv} \\ \dots \simeq (\Sigma\texttt{a} (\texttt{c}: \texttt{C}\texttt{a}), \texttt{B}\texttt{a}) & : \texttt{sigma_equiv_sigma_right} \\ & (\lambda\texttt{a}, \texttt{!comm_equiv_nondep}) \\ \dots \simeq (\Sigma(\texttt{u}: \Sigma\texttt{a}, \texttt{C}\texttt{a}), \texttt{B}\texttt{u}.\texttt{1}) & : \texttt{sigma_assoc_equiv} \end{array}$$

In Lean, we can present a proof as a term, as in Agda, with nice syntactic sugar.

We can also use tactics.

The two modes can be mixed freely:

- anywhere a term is expected, begin ... end or by enter tactic mode.
- within tactic mode, have ..., from ... or show ..., from ... or exact specify terms.

Example term proof

```
theorem sqrt_two_irrational {a b : \mathbb{N}} (co : coprime a b) :
  a^2 \neq 2 * b^2 :=
assume H : a^2 = 2 * b^2,
have even (a<sup>2</sup>).
  from even_of_exists (exists.intro _ H),
have even a.
  from even_of_even_pow this,
obtain (c : \mathbb{N}) (aeq : a = 2 * c),
  from exists of even this.
have 2 * (2 * c^2) = 2 * b^2,
  by rewrite [-H, aeq, *pow_two, mul.assoc, mul.left_comm c],
have 2 * c^2 = b^2.
  from eq_of_mul_eq_mul_left dec_trivial this,
have even (b^2).
  from even_of_exists (exists.intro _ (eq.symm this)),
have even b.
  from even_of_even_pow this,
assert 2 | gcd a b,
  from dvd_gcd (dvd_of_even 'even a') (dvd_of_even 'even b'),
have 2 \mid 1,
  by rewrite [gcd_eq_one_of_coprime co at this]; exact this,
show false.
  from absurd '2 | 1' dec_trivial
```

Floris van Doorn (CMU)

Homotopy Type Theory in Lean

Example tactic proof

```
variable (P : S^1 \rightarrow Type)
definition circle.rec_unc (v : \Sigma(p : P base), p =[loop] p)
  : \Pi(x : S^1), P x :=
begin
  intro x, induction v with p q, induction x,
  { exact p},
  { exact q}
end
definition circle_pi_equiv
  : (\Pi(x : S^1), P x) \simeq \Sigma(p : P base), p = [loop] p :=
begin
  fapply equiv.MK,
  { intro f, exact (f base, apd f loop)},
  { exact circle.rec_unc P},
  { intro v, induction v with p q, fapply sigma_eq,
    { reflexivity},
    { esimp, apply pathover_idp_of_eq, apply rec_loop}},
  { intro f, apply eq_of_homotopy, intro x, induction x,
    { reflexivity},
    { apply eq_pathover_dep, apply hdeg_squareover, esimp, apply rec_loop}}
end
```

Leo is now doing a major rewrite of the system.

- The elaborator will be slightly less general, but much more stable and predictable.
- "Let" definitions will be added to the kernel.
- There will be a better foundation for automation (e.g. goals with indexed hypotheses).
- There will be a byte-code compiler and fast evaluator for Lean. This makes it possible to use Lean as an interpreted programming language.
- Monadic interfaces will make it possible to write custom tactics from within Lean.

Plans for automation:

- A general theorem prover and term simplifier, with
 - awareness of type classes
 - powerful unification and e-matching
- Custom methods for arithmetic reasoning, linear and nonlinear inequalities.
- A flexible framework for adding domain specific tools.

Already has:

- datatypes: booleans, lists, tuples, finsets, sets
- number systems: nat, int, rat, real, complex
- the algebraic hierarchy, through ordered fields
- "big operations": finite sums and products, etc.
- elementary number theory (e.g. primes, gcd's, unique factorization, etc.)
- elementary set theory
- elementary group theory
- beginnings of analysis: topological spaces, limits, continuity, the intermediate value theorem

Demo

Browser version available at: https://leanprover.github.io/tutorial/?live&hott

HoTT library

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1					+	+	+	+	+		+	+			
2	+	+	+	+		+	+	+	+	+	+	+	+	+	+
3	+	+	+	+	$\frac{1}{2}$	+	+	+	+		+				
4	-	+	+	+		+	+	+	+						
5	-		$\frac{1}{2}$	-	-			$\frac{1}{2}$							
6		+	+	+	+	+	+	+	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{3}{4}$	+			
7	+	+	+	-	$\frac{3}{4}$	-	-								
8	+	+	+				+	+	+	$\frac{1}{4}$					
9	$\frac{3}{4}$	+	+	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{2}$	-	-	-						
10	$\frac{1}{4}$	-	-	-	-										
11	-	-	-	-	-	-									

HoTT-Lean: ~28k LOC HoTT-Coq: ~31k LOC (in theories/ folder) HoTT-Agda: ~18k LOC (excluding old/ folder) UniMath: ~52k LOC

(excluding blank lines and single-line comments)

There are two primitive HITs in Lean: the *n*-truncation and quotients.

We define all other HITs in terms of these. We can define

- colimits;
- pushouts, hence also suspensions, spheres, joins, ...
- the propositional truncation;
- HITs with 2-constructors, such as the torus and Eilenberg-MacLane spaces K(G, 1);
- [WIP] *n*-truncations and certain (ω -compact) localizations (Egbert Rijke, vD).

Lean-HoTT library

Furthermore, the library includes:

• A library of squares, cubes, pathovers, squareovers, cubeovers (based on the paper by Licata and Brunerie)

Lean-HoTT library

Furthermore, the library includes:

• A library of squares, cubes, pathovers, squareovers, cubeovers (based on the paper by Licata and Brunerie)

• A library of pointed types, pointed maps, pointed homotopies, pointed equivalences

Lean-HoTT library

Furthermore, the library includes:

• A library of squares, cubes, pathovers, squareovers, cubeovers (based on the paper by Licata and Brunerie)

• A library of pointed types, pointed maps, pointed homotopies, pointed equivalences

• Some category theory which is not in the book, e.g. limits, colimits and exponential laws:

definition functor_functor_iso (C D E : Precategory) : (C ^c D) ^c E \cong c C ^c (E \times c D) Loop space of the circle:

definition base_eq_base_equiv : base = base $\simeq \mathbb{Z}$ definition fundamental_group_of_circle : $\pi_1(S^1.) = g\mathbb{Z}$ Connectedness of suspensions (by Ulrik Buchholtz): definition is_conn_susp (n : \mathbb{N}_{-2}) (A : Type) [H : is_conn n A] : is_conn (n .+1) (susp A)

```
The Hopf fibration (by Ulrik Buchholtz):

variables (A : Type) [H : h_space A] [K : is_conn 0 A]

definition hopf : susp A \rightarrow Type :=

susp.elim_type A A

(\lambdaa, equiv.mk (\lambdax, a * x) !is_equiv_mul_left)
```

definition hopf.total (A : Type) [H : h_space A] [K : is_conn 0 A] : sigma (hopf A) \simeq join A A

definition circle_h_space : h_space S¹
definition sphere_three_h_space : h_space (S 3)

Some results are ported from Agda, such as the Freudenthal equivalence

and the associativity of join (by Jakob von Raumer)

```
definition join.assoc (A B C : Type) :
join (join A B) C \simeq join A (join B C)
```

Truncated Whitehead's principle:

Here 'pmap_of_map f a' is the pointed map $(A, a) \xrightarrow{f} (B, f(a))$.

Eilenberg-MacLane spaces (based on the paper by Licata and Finster):

```
definition EM : CommGroup \rightarrow \mathbb{N} \rightarrow \text{Type}^*
variables (G : CommGroup) (n : \mathbb{N})
definition homotopy_group_EM :
\pi g[n+1] (EM G (n+1)) \simeq g G
theorem is_conn_EM : is_conn (n.-1) (EM G n)
theorem is_trunc_EM : is_trunc n (EM G n)
```

```
definition EM_pequiv_1 {X : Type*} (e : \pi_1 X \simeq g G)
[is_conn 0 X] [is_trunc 1 X] : EM G 1 \simeq* X
-- TODO for n > 1
```

Seifert-van Kampen Theorem:

 $\begin{array}{l} \mbox{definition vankampen {S A B C : Type} (k : S \rightarrow C) \\ (f : C \rightarrow A) (g : C \rightarrow B) [is_surjective k] \\ (x y : A + B) : \\ \mbox{@hom (Π_1 (pushout f g)) _ $ (pushout_of_sum f g y) $ \simeq $ (pushout_of_sum f g x) (pushout_of_sum f g y) _ $ x y $ \end{tabular}$

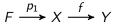
The long exact sequence of homotopy groups.

```
Given a pointed map f: X \to Y.
```



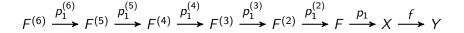
The long exact sequence of homotopy groups.

Given a pointed map $f: X \to Y$.



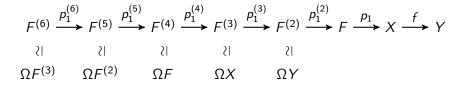
The long exact sequence of homotopy groups.

Given a pointed map $f : X \to Y$.



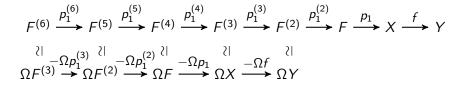
The long exact sequence of homotopy groups.

Given a pointed map $f: X \to Y$.



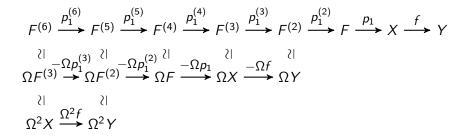
The long exact sequence of homotopy groups.

Given a pointed map $f : X \to Y$.



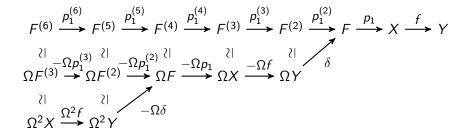
The long exact sequence of homotopy groups.

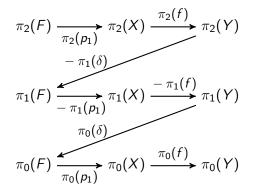
Given a pointed map $f : X \to Y$.

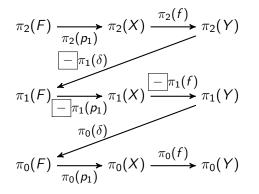


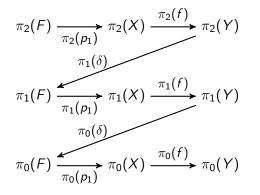
The long exact sequence of homotopy groups.

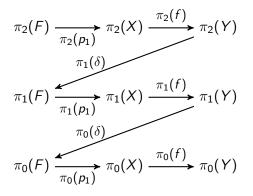
Given a pointed map $f : X \to Y$.











How do we formulate this?

The obvious thing is to have a sequence $Z : \mathbb{N} \to \mathcal{U}$ and maps $f_n : Z_{n+1} \to Z_n$.

Problem: $Z_{3n} = \pi_n(Y)$ doesn't hold definitionally, hence $f_{3n} = \pi_n(f)$ isn't even well-typed.

Better: Take $Z : \mathbb{N} \times 3 \rightarrow \mathcal{U}$, we can define Z by

$$Z_{(n,0)} = \pi_n(Y)$$
 $Z_{(n,1)} = \pi_n(X)$ $Z_{(n,2)} = \pi_n(F).$

Then we can define the maps $f_x : Z_{succ(x)} \to Z_x$, where succ is the successor function for $\mathbb{N} \times 3$.

Better: Take $Z : \mathbb{N} \times 3 \rightarrow \mathcal{U}$, we can define Z by

$$Z_{(n,0)} = \pi_n(Y)$$
 $Z_{(n,1)} = \pi_n(X)$ $Z_{(n,2)} = \pi_n(F).$

Then we can define the maps $f_x : Z_{succ(x)} \to Z_x$, where succ is the successor function for $\mathbb{N} \times 3$.

We define chain complexes over an arbitrary type with a successor operation.

```
definition homotopy_groups : +3\mathbb{N} \rightarrow \text{Set}^*
| (n, fin.mk 0 H) := \pi^* [n] Y
| (n, fin.mk 1 H) := \pi^* [n] X
| (n, fin.mk k H) := \pi^*[n] (pfiber f)
definition homotopy_groups_fun : \Pi(n : +3\mathbb{N}),
  homotopy_groups (S n) \rightarrow^* homotopy_groups n
| (n, fin.mk 0 H) := \pi \rightarrow^* [n] f
| (n, fin.mk 1 H) := \pi \rightarrow^*[n] (ppoint f)
| (n, fin.mk 2 H) := \pi \rightarrow^*[n] boundary_map \circ^*
    pcast (ap (ptrunc 0) (loop_space_succ_eq_in Y n))
(n, fin.mk (k+3) H) := begin exfalso,
    apply lt_le_antisymm H, apply le_add_left end
```

Then we prove:

- These maps form a chain complex
- This chain complex is exact
- homotopy_groups (n + 2, k) are commutative groups.
- homotopy_groups (1, k) are groups.
- homotopy_groups_fun (n + 1, k) are group homomorphisms.

Corollaries:

theorem is_equiv_ $\pi_of_is_connected {A B : Type*} {n k : <math>\mathbb{N}$ } (f : A \rightarrow * B) [H : is_conn_fun n f] (H2 : k \leq n) : is_equiv ($\pi \rightarrow$ [k] f)

Combine with Hopf fibration:

 $\begin{array}{l} \text{definition } \pi 2\text{S2} : \pi g[1+1] \ (\text{S. 2}) \simeq g \ g\mathbb{Z} \\ \text{definition } \pi n\text{S3}_eq_\pi n\text{S2} \ (n : \mathbb{N}) : \\ \pi g[n+2 \ +1] \ (\text{S. 3}) \simeq g \ \pi g[n+2 \ +1] \ (\text{S. 2}) \end{array}$

Combine with Freudenthal Suspension Theorem:

definition πnSn (n : \mathbb{N}) : $\pi g[n+1]$ (S. (succ n)) $\simeq g g\mathbb{Z}$ definition $\pi 3S2$: $\pi g[2+1]$ (S. 2) $\simeq g g\mathbb{Z}$

- Lean is an exciting new proof assistant.
- The Lean-HoTT library is quite big and growing quickly.
- The Lean-HoTT library contains a good basis for serious formalizations.

Thank you