15-819 Homotopy Type Theory Lecture Notes

Henry DeYoung and Stephanie Balzer September 9 and 11, 2013

1 Contents

These notes summarize the lectures on homotopy type theory (HoTT) given by Professor Robert Harper on September 9 and 11, 2013, at CMU. They start by providing a introduction to HoTT, capturing its main ideas and its connection to other related type theories. Then they present intuitionistic propositional logic (IPL), giving both a proof-theoretic formulation as well as an order-theoretic formulation.

2 Introduction to homotopy type theory

Homotopy type theory (HoTT) is the subject of a very active research community that gathered at the Institute for Advanced Study (IAS) in 2012 to participate in the Univalent Foundations Program. The results of the program have been recently published in the HoTT Book [1].

2.1 HoTT in a nutshell

HoTT is based on Per Martin-Löf's intuitionistic type theory, which provides a foundation for *intuitionistic mathematics* and which is an extension of *Brouwer's program*. Brouwer viewed mathematical reasoning as a human activity and mathematics as a language for communicating mathematical concepts. As a result, Brouwer perceived the ability of executing a step-by-step procedure or algorithm for performing a *construction* as a fundamental human faculty.

Adopting Brouwer's constructive viewpoint, intuitionistic theories view proofs as the fundamental forms of construction. The notion of *proof relevance* is thus a

characteristic feature of an intuitionistic (or constructive¹) approach. In the context of HoTT, proof relevance means that proofs become mathematical objects [3]. To fully understand this standpoint, it is necessary to draw a distinction between the notion of a proof and the one of a formal proof [3, 2]. A formal proof is a proof given in a fixed formal system, such as the axiomatic theory of sets, and arises from the application of the inductively defined rules in that system. Whereas every formal proof is also a proof (assuming soundness of the system) the converse is not true. This fact immediately follows from Gödel's Incompleteness Theorem, which precisely states that there exist true propositions (with a proof), but for which there cannot be given a formal proof, using the rules of the formal system. Unlike conventional formally defined systems, HoTT does not surmise that all possible proofs can be fully circumscribed by its rules, but accepts proofs that cannot be formalized in HoTT. These are exactly the proofs that are considered to be relevant and, being treated as mathematical objects, they can be formulated internally as objects of the type theory.

Being based on intuitionistic type theory, HoTT facilitates some form of axiomatic freedom. This means in particular that there exist fewer assumptions that apply globally. For instance, a typical such assumption that is missing in a intuitionistic interpretation is the law of the excluded middle. As put forth by Brouwer, an assumption like the law of the excluded middle does not need to be ruled out entirely, but can be introduced locally, in a proof, if needed. Whether a particular local assumption is needed or not is mainly determined by the actual proof (i.e., construction). A sparing use of global assumptions results in proofs that are based on less assumptions and thus in stronger results overall.

Another distinguishing characteristics of HoTT is that it adopts a *synthetic*, rather than an *analytic* reasoning approach. The differentiation goes back to Lawvere and is best explained by an example. Euclidean geometry, for instance, represents a synthetic approach to geometry as it treats geometric figures, like triangles, lines, and circles, as "things" in themselves rather than sets of points. In an analytical approach based on the Cartesian coordinate system, on the other hand, geometric figures are treated as sets of points in the plane and thus are based on the real numbers. Whereas traditional formulations of homotopy theory are analytic, HoTT is synthetic. The differentiation between a synthetic and an analytic reasoning approach is mainly relevant with regard to the approach's amenability to mechanized reasoning. It turns out that synthetic approaches are easier to mechanize than analytical approaches. This holds true in particular for HoTT: since proofs of equality in HoTT correspond to paths in a space, they are cleaner, shorter, and completely mechanizable.

¹In this course, intuitionism and constructivism are used interchangeably.

2.2 HoTT in type theory context

HoTT unites homotopy theory with type theory, by embodying Brouwer's intuitionism and drawing from Gentzen's proof theory (see Section 3). It is based on the observation that *types* classify the admissible forms of constructions and thus are programmatically sufficient to encompass all known mathematical constructions. This section briefly sketches how HoTT relates to other existing type theories.

2.2.1 Intensional type theory

Intensional type theory (ITT) is a intuitionistic type theory that serves as the core theory for other type theories. Other type theories are merely extensions of ITT.

2.2.2 Extensional type theory

Extensional type theory (ETT) extends ITT with equality of reflection (ER) and uniqueness of identity proofs (UIP):

$$ETT = ITT + ER + UIP$$

Since types are perceived as sets in ETT, ETT gives rise to a intuitionistic theory of sets.

2.2.3 Homotopy type theory

HoTT extends ITT with higher inductive types (HIT) and the univalence axiom (UA):

$$HoTT = ITT + HIT + UA$$

Since types are perceived as abstract spaces in HoTT, HoTT gives rise to a intuitionistic theory of $weak \propto -groupoids$.

3 Intuitionistic propositional logic

What is meant by *intuitionistic* logic? It is a *proof-relevant* logic. One might say its slogan is "logic as if people matter", alluding to Brouwer's principle that mathematics is a social process in which proofs are crucial for communication. Whenever a claim of truth of a proposition is made, it must be accompanied by a proof.

As advanced by Per Martin-Löf, a modern presentation of intuitionistic propositional logic (IPL) distinguishes the notions of *judgment* and *proposition*. A judgment is something that may be known, whereas a proposition is something that sensibly may be the subject of a judgment. For instance, the statement "Every natural number larger than 1 is either prime or can be uniquely factored into a product of primes." is a proposition because it sensibly may be subject to judgment. That the statement is in fact true is a judgment. Only with a proof, however, is it evident that the judgment indeed holds.

Thus, in IPL, the two most basic judgments are A prop and A true:

A prop A is a well-formed propositionA true Proposition A is intuitionistically true, i.e., has a proof.

The inference rules for the **prop** judgment are called formation rules. The inference rules for the **true** judgment are divided into classes: *introduction rules* and *elimination rules*.

Following Martin-Löf, the meaning of a proposition A is given by the introduction rules for the judgment A true. The elimination rules are dual and describe what may be deduced from a proof of A true.

The principle of internal coherence, also known as Gentzen's principle of inversion, is that the introduction and elimination rules for a proposition A fit together properly. The elimination rules should be strong enough to deduce all information that was used to introduce A (local completeness), but not so strong as to deduce information that might not have been used to introduce A (local soundness). In a later lecture, we will discuss internal coherence more precisely, but we can already give an informal treatment.

3.1 Negative fragment of IPL

3.1.1 Conjunction

One familiar group of propositions are the conjunctions. If A and B are well-formed propositions, then so is their conjunction, which we write as $A \wedge B$. This is the content of the formation rule for conjunction: it serves as evidence of the judgment $A \wedge B$ prop, provided that there is evidence of the judgments A prop and B prop.

$$\frac{A \operatorname{prop} \quad B \operatorname{prop}}{A \wedge B \operatorname{prop}} \ \wedge F$$

We have yet to give meaning to conjunction, however; to do so, we must say how to introduce the judgment that $A \wedge B$ is *true*. As the following rule shows, a verification of $A \wedge B$ consists of a proof of A true paired with a proof of B true.

$$\frac{A \text{ true } B \text{ true}}{A \wedge B \text{ true}} \wedge I$$

What may we deduce from the knowledge that $A \wedge B$ is true? Because every proof of $A \wedge B$ true ultimately introduces that judgment from a pair of proofs of A true and B true, we are justified in deducing A true and B true from any proof of $A \wedge B$ true. This leads to the elimination rules for conjunction.

$$\frac{A \wedge B \text{ true}}{A \text{ true}} \wedge E_1 \qquad \qquad \frac{A \wedge B \text{ true}}{B \text{ true}} \wedge E_2$$

Internal coherence. As previously mentioned, the principle of internal coherence says that the introduction and elimination rules fit together properly: the elimination rules are strong enough, but not too strong.

If we mistakenly omitted the $\wedge E_2$ elimination rule, then there would be no way to extract the proof of B true that was used in introducing $A \wedge B$ true—the elimination rules would be too weak.

On the other hand, if we mistakenly wrote the $\wedge I$ introduction rule as

$$\frac{A \text{ true}}{A \wedge B \text{ true}}$$
,

then there would be no proof of B true present to justify deducing B true with the $\wedge E_2$ rule—the elimination rules would be too strong.

3.1.2 Truth

Another familiar and simple proposition is *truth*, which we write as \top . Its formation rule serves as immediate evidence of the judgment \top prop, that \top is indeed a well-formed proposition.

$$\overline{\top \operatorname{prop}} \ \top F$$

Once again, to give meaning to truth we must say how to introduce the judgment that \top is true. \top is a trivially true proposition, and so its introduction rule makes the judgment \top true immediately evident.

$$\frac{}{\top \text{ true}} \ ^{\top I}$$

We should also consider an elimination rule: from a proof of \top true, what can we deduce? Since \top is trivially true, any such elimination rule would not increase our knowledge—we put in no information when we introduced \top true, so, by the principle of conservation of proof, we should get no information out. For this reason, there is no elimination rule for \top , and we can see that its absence is coherent with the introduction rule.

The nullary conjunction. An observation about \top that often proves useful is that \top behaves as a nullary conjunction—a conjunction over the empty set of conjuncts, rather than over a set of two conjuncts.

This observation is reflected in the inference rules. Just as the introduction rule for binary conjunction has two premises (one for each of the two conjuncts), the introduction rule for truth has no premises (one for each of the no conjuncts):

$$\frac{A \text{ true } B \text{ true}}{A \wedge B \text{ true}} \wedge I \qquad \qquad \frac{}{\top \text{ true}} \top I$$

Likewise, just as there are two elimination rules for binary conjunction, there are no elimination rules for truth:

$$\frac{A \wedge B \text{ true}}{A \text{ true}} \wedge E_1 \qquad \frac{A \wedge B \text{ true}}{B \text{ true}} \wedge E_2 \qquad \text{(no } \top E \text{ rule)}$$

3.1.3 Entailment

The last form of proposition in the negative fragment of IPL is implication. However, to define implication, a different form of judgment is required: entailment (also known as logical consequence or a hypothetical judgment). Entailment is written as

$$\underbrace{A_1 \text{ true}, \ldots, A_n \text{ true}}_{n>0} \vdash A \text{ true},$$

and expresses the idea that the judgment A true follows from A_1 true, ..., A_n true. You can think of A_1 true, ..., A_n true as being assumptions from which the conclusion A true may be deduced. The metavariable Γ is typically used to stand for such a context of assumptions.

We should note that, thus far, the inference rules have been presented in a *local* form in which the context of assumptions was left implicit. It would also be possible to make this context explicit. For example, the introduction rule for conjunction in the two different forms is:

$$\frac{A \text{ true } B \text{ true}}{A \wedge B \text{ true}} \wedge I \qquad \qquad \frac{\Gamma \vdash A \text{ true } \Gamma \vdash B \text{ true}}{\Gamma \vdash A \wedge B \text{ true}} \wedge I$$

In the remainder of these notes, we will write inference rules in local form whenever possible.

As an entailment, the judgment form satisfies several *structural properties*: reflexivity, transitivity, weakening, contraction, and permutation.

Reflexivity. If the entailment judgment is to express logical consequence, that is, that a conclusion follows from some assumptions, then you must accept a principle of reflexivity,

$$\overline{A \text{ true} \vdash A \text{ true}} \ \mathsf{R}$$

that an assumption is enough to conclude the same judgment. If you tried to deny this principle, the meaning of an assumption would be unclear.

Transitivity. Dual to reflexivity is a transitivity principle that states that a proof of a conclusion satisfies an assumption of the same judgment.

$$\frac{A \text{ true } \vdash C \text{ true}}{C \text{ true}} \top$$

Transitivity is a lemma rule. If you prove a lemma $(A \operatorname{\mathsf{true}})$, then you are justified in using it to prove a theorem that explicitly depends $(A \operatorname{\mathsf{true}} \vdash C \operatorname{\mathsf{true}})$; taken together, they are viewed as a direct proof of the theorem $(C \operatorname{\mathsf{true}})$.

Transitivity can also be thought of as proof inlining. Rather than pairing the lemma with the theorem that depends upon it, we could inline the lemma's proof at every point at which the theorem refers to the lemma. The result is a truly direct proof of the theorem.

Weakening. Reflexivity and transitivity are undeniable properties of entailment because they give meaning to assumptions—assumptions are strong enough to prove conclusions (reflexivity), but are only as strong as the proofs they stand for (transitivity). But there are also structural properties that can be denied: weakening, contraction, and permutation. Logics that deny any of these properties are called substructural logics.

The principle of weakening says that we can add assumptions to a proof without invalidating that proof:

$$\frac{A \operatorname{true}}{B \operatorname{true} \vdash A \operatorname{true}} \ \mathsf{W}$$

Of course, the new proof is of a weaker statement, but it is nevertheless a valid proof.

Denying weakening leads, in part, to relevance logic. It is called relevance logic because the proofs may not contain the unnecessary, irrelevant assumptions that weakening allows. In this course, we will always have the principle of weakening, however.

Contraction. The principle of contraction says that we are unconcerned about the number of copies of an assumption A true; one copy is as good as two:

$$\frac{A \operatorname{true}, A \operatorname{true} \vdash C \operatorname{true}}{A \operatorname{true} \vdash C \operatorname{true}} \subset$$

Denying contraction (along with weakening) leads to linear logic, in which we wish to reason about the number of copies of an assumption. This is a powerful way to express consumable resources. In this course, we will always have the principle of contraction, however.

Permutation. The principle of permutation, or exchange, says that the order of assumptions does not matter; we can apply any permutation π to the assumptions and still have a valid proof:

$$\frac{\Gamma \vdash C \text{ true}}{\pi(\Gamma) \vdash C \text{ true}} \mathsf{P}$$

(Note that it is difficult to state the permutation principle in local rule form.)

Denying permutation (along with weakening and contraction) leads to ordered, or noncommutative, logic. It is a powerful way to express ordered structures, like lists or even formal grammars. In this course, we will always have the principle of permutation, however.

3.1.4 Implication

With the entailment judgment in hand, we can give rules for implication.

Like conjunction, if A and B are well-formed propositions, then so is their implication, which we write as $A \supset B$.

$$\frac{A \ \mathsf{prop} \quad B \ \mathsf{prop}}{A \supset B \ \mathsf{prop}} \supset \!\! F$$

Once again, to give meaning to implication, we must say how to introduce the judgment $A \supset B$ true. To prove $A \supset B$ true, we assume A true and prove B true.

$$\frac{A \operatorname{true} \vdash B \operatorname{true}}{A \supset B \operatorname{true}} \supset I$$

In this way, implication internalizes the entailment judgment as a proposition, while we nonetheless maintain the distinction between propositions and judgments. (As an aside for those familiar with category theory, the relationship between entailment and implication is analogous to the relationship between a mapping and a collection of mappings internalized as an object in some categories.)

This introduction rule for implication is the key distinction between Gentzenstyle natural deduction calculus and a Hilbert-style axiomatic calculus. In a Hilbert presentation of IPL, there is no separate notion of entailment, making it difficult to reason hypothetically. Instead, one must contort proofs to make use of several seemingly unmotivated axioms about implication.

Thankfully, we will work with natural deduction and be able to reason hypothetically using the introduction rule. But what does the elimination rule for $A \supset B$ look like? Because every proof of $A \supset B$ true ultimately introduces that judgment from a proof of the entailment A true $\vdash B$ true, we might like to write the elimination rule as

$$\frac{A\supset B \text{ true}}{A \text{ true} \vdash B \text{ true}}.$$

This is a valid principle of reasoning, but it turns out to be useful to instead adopt an uncurried form as the actual elimination rule:

$$\frac{A\supset B \text{ true} \quad A \text{ true}}{B \text{ true}}\supset E \ .$$

This rule is sometimes referred to as modus ponens.

Internal coherence. These introduction and elimination rules are coherent. The elimination rule is strong enough to recover the entailment that any proof of $A \supset B$ true ultimately uses in introduction, as the following derivation shows.

$$\frac{\frac{A\supset B \text{ true}}{A \text{ true} \vdash A\supset B \text{ true}} \text{ W} \quad \frac{}{A \text{ true} \vdash A \text{ true}} \text{ } \frac{\text{R}}{\supset E}$$

On the other hand, the elimination rule is not too strong because it is just an uncurrying of the inverted introduction rule.

3.2 Positive fragment of IPL

3.2.1 Disjunction

As for conjunction and implication, the disjunction, $A \vee B$, of A and B is a well-formed proposition if both A and B are themselves well-formed propositions.

$$\frac{A \operatorname{prop} \quad B \operatorname{prop}}{A \vee B \operatorname{prop}} \vee F$$

A disjunction $A \vee B$ may be introduced in either of two ways: $A \vee B$ is true if A is true or if B is true.

$$\frac{A \text{ true}}{A \vee B \text{ true}} \vee I_1 \qquad \qquad \frac{B \text{ true}}{A \vee B \text{ true}} \vee I_2$$

To devise the elimination rule, consider what may we deduce from the knowledge that $A \vee B$ is true. For $A \vee B$ to be true, it must have been ultimately introduced using one of the two introduction rules. Therefore, either A or B is true (or possibly both). The elimination rule allows us to reason by these two cases: If C true follows from A true and also follows from B true, then C is true in either case.

$$\frac{A \vee B \text{ true} \quad A \text{ true} \vdash C \text{ true} \quad B \text{ true} \vdash C \text{ true}}{C \text{ true}} \vee E$$

3.2.2 Falsehood

The unit of disjunction is falsehood, the proposition that is trivially never true, which we write as \bot . Its formation rule is immediate evidence that \bot is a well-formed proposition.

$$\overline{\perp \operatorname{prop}} \ \bot F$$

Because \perp should never be true, it has no introduction rule. The elimination rule captures $ex\ falso\ quodlibet$: from a proof of \perp true, we may deduce that any proposition C is true because there is ultimately no way to introduce \perp true.

$$\frac{\perp \text{ true}}{C \text{ true}} \perp E$$

Once again, the rules cohere. The elimination rule is very strong, but remains justified due to the absence of any introduction rule for falsehood.

The nullary disjunction. We previously noted that \top behaves as a nullary conjunction. In the same way, \bot behaves as a nullary disjunction. For a binary disjunction, there are two introduction rules, $\lor I_1$ and $\lor I_2$, one for each of the two disjuncts; for falsehood, there are no introduction rules:

$$\frac{A \text{ true}}{A \vee B \text{ true}} \vee I_1 \qquad \qquad \frac{B \text{ true}}{A \vee B \text{ true}} \vee I_2 \qquad \qquad \text{(no } \perp I \text{ rule)}$$

Likewise, for a binary disjunction, there is one elimination rule with a premise for the disjunction and one premise for each of the disjuncts; for falsehood, there is one elimination rule with just a premise for falsehood:

$$\frac{A \vee B \text{ true } \vdash C \text{ true } \quad B \text{ true} \vdash C \text{ true}}{C \text{ true}} \vee E \qquad \qquad \frac{\bot \text{ true}}{C \text{ true}} \perp E$$

4 Order-theoretic formulation of IPL

It is also possible to give an order-theoretic formulation of IPL because entailment is a preorder (reflexive and transitive). We want $A \leq B$ to hold exactly when $A \text{ true} \vdash B \text{ true}$. We can therefore devise the order-theoretic formulation with these soundness and completeness goals in mind.

4.1 Conjunction as meet

The elimination rules for conjunction (along with reflexivity of entailment) ensure that $A \wedge B$ true $\vdash A$ true and $A \wedge B$ true $\vdash B$ true. To ensure completeness of the order-theoretic formulation, we include the rules

$$\overline{A \wedge B \leq A}$$
 $\overline{A \wedge B \leq B}$,

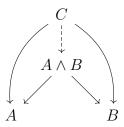
which say that $A \wedge B$ is a lower bound of A and B.

The introduction rule for conjunction ensures that C true $\vdash A \land B$ true if both C true $\vdash A$ true and C true $\vdash B$ true. Order-theoretically, this is expressed as the rule

$$\frac{C \le A \quad C \le B}{C \le A \land B},$$

which says that $A \wedge B$ is as large as any lower bound of A and B. Taken together these rules show that $A \wedge B$ is the greatest lower bound, or meet, of A and B.

Graphically, these order-theoretic rules can be represented with a commuting *product diagram*, where arrows point from smaller to larger elements:



4.2 Truth as greatest element

The introduction rule for \top ensures that C true $\vdash \top$ true. Order-theoretically, we have

$$\overline{C \leq \top}$$
,

which says that \top is the greatest, or final, element.

In the proof-theoretic formulation of IPL, we saw that truth \top is the nullary conjunction. We should expect this analogy to hold in the order-theoretic formulation of IPL as well, and it does—the greatest element is indeed the greatest lower bound of the empty set.

4.3 Disjunction as join

The introduction rules for disjunction (along with reflexivity of entailment) ensure that A true $\vdash A \lor B$ true and B true $\vdash A \lor B$ true. To ensure completeness of the order-theoretic formulation, we include the rules

$$\overline{A \le A \lor B}$$
 $\overline{B \le A \lor B}$,

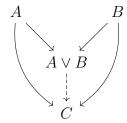
which say that $A \vee B$ is an upper bound of A and B.

The elimination rule for disjunction (along with reflexivity of entailment) ensures that $A \vee B$ true $\vdash C$ true if both A true $\vdash C$ true and B true $\vdash C$ true. Order-theoretically, we have the corresponding rule

$$\frac{A \le C \quad B \le C}{A \lor B \le C},$$

which says that $A \vee B$ is as small as any upper bound of A and B. Taken together these rules show that $A \vee B$ is the least upper bound, or join, of A and B.

Graphically, this is captured by a commuting *coproduct diagram*:



4.4 Falsehood as least element

The elimination rule for falsehood (along with reflexivity of entailment) ensures that \bot true $\vdash C$ true. The order-theoretic counterpart is the rule

$$\overline{\perp \leq C}$$
,

which says that \perp is the least, or initial, element.

Once again, because we saw that falsehood is the nullary disjunction in the proof-theoretic formulation, we should expect this analogy to carry over to the order-theoretic formulation. Indeed, the least element is the least upper bound of the empty set.

4.5 Order-theoretic IPL as lattice

As seen thus far, the order-theoretic formulation of IPL gives rise to a *lattice* as it establishes a preorder with finite meets and joins. The definition of a lattice assumed in this course may deviate from the one typically found in the literature, which usually considers a lattice to be a partial order with finite meets and joins. In this course, we deliberately ignore the property of antisymmetry. If we were to impose the property of antisymmetry on the order defined by entailment, then we would need to introduce equivalence classes of propositions, which requires associativity. As we will see later in this course, the axiom of univalence provides an elegant way of dealing with equivalence of propositions.

4.6 Implication as exponential

The elimination rule for implication (along with reflexivity of entailment) ensures that A true, $A \supset B$ true $\vdash B$ true. For the order-theoretic formulation to be complete, we include the rule

$$\overline{A \wedge (A \supset B) \leq B}$$

The introduction rule for implication ensures that C true $\vdash A \supset B$ true if A true, C true $\vdash B$ true. Once again, so that the order-theoretic formulation is complete, we have

$$\frac{A \land C \le B}{C \le A \supset B},$$

Taken together, these rules show that $A \supset B$ is the exponential of A and B.

As we have seen previously, the order-theoretic formulation of IPL gives rise to a lattice. Now we have just seen that it also supports exponentials. As a result, the order-theoretic formulation of IPL gives rise to a *Heyting algebra*. A Heyting algebra is a lattice with exponentials. As we will see later in this course, the notion of a Heyting algebra is fundamental in proving completeness of IPL. The proof also relies on the notion of a *complement* in a lattice. The complement \overline{A} of A in a lattice is such that

- 1. $\top \leq \overline{A} \vee A$;
- 2. $\overline{A} \wedge A < \bot$.

It follows that a complement, if present, is a suitable notion of negation, but negation, defined via the exponential, is not necessarily a complement.

References

- [1] Institute for Advanced Study. Homotopy Type Theory: Univalent Foundations of Mathematics. The Univalent Foundations Program, 2013. http://homotopytypetheory.org/book/.
- [2] Robert Harper. Extensionality, intensionality, and Brouwer's dictum. http://existentialtype.wordpress.com/2012/08/11/extensionality-intensionality-and-brouwers-dictum/, August 2012.
- [3] Robert Harper. Constructive mathematics is not metamathematics. http://existentialtype.wordpress.com/2013/07/10/constructive-mathematics-is-not-meta-mathematics/, July 2013.
- [4] Frank Pfenning. Lecture notes on harmony. http://www.cs.cmu.edu/~fp/courses/15317-f09/lectures/03-harmony.pdf, September 2009.
- [5] Frank Pfenning. Lecture notes on natural deduction. http://www.cs.cmu.edu/~fp/courses/15317-f09/lectures/02-natded.pdf, August 2009.