Type theory and its meaning explanations

Jonathan Sterling

Thanks to Bob Harper, Peter Dybjer, Bengt Nordström, Carlo Angiuli and Darin Morrison for invaluable conversations which helped sharpen my view of the meaning explanations for computational and intensional type theories. All mistakes are mine.

Abstract. At the heart of intuitionistic type theory lies an intuitive semantics called the "meaning explanations"; crucially, when meaning explanations are taken as definitive for type theory, the core notion is no longer "proof" but "verification". We'll explore how type theories of this sort arise naturally as enrichments of logical theories with further judgements, and contrast this with modern proof-theoretic type theories which interpret the judgements and proofs of logics, not their propositions and verifications.

Contents

CHAPTER 1

Logical Theories

To start, we will consider the notion of a *logical theory*; in my mind, it begins with a species of judgements that can be proposed, asserted, and (if they are evident) known.

1. Judgements of a logical theory

The basic forms of judgement for a logical theory will be $\boxed{P\ \mathit{prop}}$ and $\boxed{P\ \mathit{true}}$; and what is P? It is a member of the species of terms, which are made meaningful in the course of making the judgement $P\ \mathit{prop}$ ("P is a proposition") evident for a proposition P.

The forms of judgement may be construed as containing *inputs* and *outputs*; an *input* is something which is inspected in the course of knowing a judgement, whereas an *output* is something which is synthesized (or created) in the course of knowing a judgement. The positions of *inputs* and *outputs* in a judgement constitute what is called its *mode*, and we color-code it in this presentation for clarity.[1]

To each judgement is assigned a *meaning explanation*, which explicates the knowledge-theoretic content of the judgement. For a judgement $\mathcal{J}$, a meaning explanation should be in the form:

> To know $\mathcal{J}$ is to know...

The meaning of the judgement $P\ \mathit{prop}$ is, then, as follows:

> To know $P\ \mathit{prop}$ is to know that P is a proposition, which is to know what would count as a direct verification of P.

So if a symbol P is taken to denote a proposition, we must know *what sort of thing* is to be taken as a direct verification of P, and this is understood as part of the definition of P. A "direct verification" is understood in constrast with an "indirect verification", which is to be thought of as a means or plan for verifying the proposition; these distinctions will be explained in more detail later on. Now, the judgement $P\ \mathit{true}$ ("P is true") is only meaningful in case we know $P\ \mathit{prop}$ (this is called a presupposition). Then the meaning of $P\ \mathit{true}$ is as follows:

[1] We will not see any judgements with *outputs* at first, but it will become necessary as soon as we consider judgements about computation, where the reduction of a term is synthesized from the redex. Modes may be used to construe a judgement as having algorithmic content.

To know P *true* is to have a verification of P.

From the (implicit) presupposition P *prop*, we already know what counts as a verification, so the meaning explanation is well-defined. Note that having a means or plan for verifying P is equivalent to having a (direct) verification; this follows from the fact that one may put into action a plan for verifying P and achieve such a verification, and likewise, it is possible to propound a plan of verification by appeal to an existing verification.

2. Higher-order judgements

The judgements we have described so far are "categorical" in the sense that they are made without assumption or generality.

2.1. Hypothetical judgement. We will need to define a further form of judgement, which is called "hypothetical", and this is the judgement under hypothesis $\boxed{\mathcal{J}\ (\mathcal{J}')}$, pronounced "$\mathcal{J}$ under the assumption $\mathcal{J}'$".Its meaning explanation is as follows:

> To know the judgement $\mathcal{J}\ (\mathcal{J}')$ is to know the categorical judgement $\mathcal{J}$ assuming you know the judgement $\mathcal{J}'$.

Hypothetical judgement may be iterated, and $\mathcal{J}\ (\mathcal{J}_1, \mathcal{J}_2)$ will be used as notation for for $\mathcal{J}\ (\mathcal{J}_2)\ (\mathcal{J}_1)$.

2.2. General judgement. Another kind of higher order judgement is "general judgement", which is judgement with respect to a variable, $\boxed{|_x\, \mathcal{J}}$, pronounced "for an arbitrary x, $\mathcal{J}$". The meaning explanation for this new judgement is as follows:

> To know the judgement $|_x\, \mathcal{J}$ is, to know $[E/x]\mathcal{J}$ (i.e. the substitution of E for x in the expression $\mathcal{J}$) for any arbitrary expression E, [2]

As far as notation is concerned, the bar symbol binds the least tightly of all the other notations we have considered. Likewise, general judgement may be iterated, and the notation $|_{x,y}\, \mathcal{J}$ will be used as notation for $|_x\, |_y\, \mathcal{J}$.

2.3. Hypothetico-general judgement. When hypothetical judgement is used inside general judgement, as in $|_x\, A(x)$ *true* $(B(x)$ *true*$)$, we term the whole thing a "hypothetico-general" judgement. One thing bears clarifying, which is, Why do we write P *true* $(P$ *true*$)$ rather than $|_P\, P$ *true* $(P$ *true*$)$?

The former is really not a single judgement, but rather a *scheme* for judgements, where P is intended to be replaced with a concrete expression by the person

[2]Technically, E is qualified as being of the same valence as x, but because we have not developed a formal theory of expressions in this presentation, I choose to ignore this issue.

asserting the judgements. On the other hand, the latter is itself a single judgement which may be asserted all on its own.

3. Propositions and verifications

Now that we have propounded and explained the minimal system of judgements for a logical theory, let us populate it with propositions. First, we have falsity $\bot$, and we wish to make $\bot$ *prop* evident; to do this, we simply state what counts as a direct verification of $\bot$: there is no direct verification of $\bot$.

The next basic proposition is trivial truth $\top$, and to make $\top$ *prop* evident, we state that a direct verification of $\top$ is trivial. The definition of $\top$ thus validates the judgement $\top$ *true* (i.e. that we have a verification of $\top$; this is immediate).

Next, let us define conjunction; in doing so, we will make evident the hypothetical judgement $P \wedge Q$ *prop* (P *prop*, Q *prop*); equivalently, we can display this as a rule of inference:[3]

$$\frac{P \text{ prop} \quad Q \text{ prop}}{P \wedge Q \text{ prop}}$$

A direct verification of $P \wedge Q$ consists in a verification of P and a verification of Q; this validates the assertion of the judgement $P \wedge Q$ *true* (P *true*, Q *true*). Because it is a valid inference, we can write it as an inference rule:

$$\frac{P \text{ true} \quad Q \text{ true}}{P \wedge Q \text{ true}}$$

A direct verification of $P \vee Q$ may be got either from a verification of P or one of Q. From this definition we know $P \vee Q$ *prop* (P *prop*, Q *prop*), or

$$\frac{P \text{ prop} \quad Q \text{ prop}}{P \vee Q \text{ prop}}$$

The verification conditions of disjunction give rise to two evident judgements $P \vee Q$ *true* (P *true*) and $P \vee Q$ *true* (Q *true*), which we can write as inference rules:

$$\frac{P \text{ true}}{P \vee Q \text{ true}} \qquad \frac{Q \text{ true}}{P \vee Q \text{ true}}$$

Finally, we must define the circumstances under which $P \supset Q$ is a proposition (i.e. when $P \supset Q$ *prop* is evident). And we intend this to be under the circumstances that P is a proposition, and also that Q is a proposition assuming that P is true.

---

[3]Evident hypothetical judgements are often written as rules, i.e.

$$\frac{\textit{premise}}{\textit{conclusion}}$$

rather than *conclusion* (*premise*). It must be stressed that only *evident*/*known* judgements may be written in this way.

In other words, $P \supset Q\ prop\ (P\ prop, Q\ prop\ (P\ true))$, or

$$\frac{P\ prop \quad Q\ prop\ (P\ true)}{P \supset Q\ prop}$$

Now, to validate this judgement will be a bit more complicated than the previous ones. But by unfolding the meaning explanations for hypothetical judgement, proposition-hood and truth of a proposition, we arrive at the following explanation:

> To know $P \supset Q\ prop\ (P\ prop, Q\ prop\ (P\ true))$ is to know what counts as a direct verification of $P \supset Q$ when one knows what counts as a direct verification of P, and, when one has such a verification, what counts as a direct verification of Q.[4]

If the judgement $P \supset Q\ prop\ (P\ prop, Q\ prop\ (P\ true))$ is going to be made evident, then we must come up with what should count as a direct verification of $P \supset Q$ under the assumptions described above.

And so to have a direct verification of $P \supset Q$ is to have a verification of Q assuming that one has one of P; this is the meaning of implication, and it validates the judgement $P \supset Q\ true\ (Q\ true\ (P\ true))$, and may be written as an inference rule as follows:

$$\frac{Q\ true\ (P\ true)}{P \supset Q\ true}$$

4. Judgements for verifications

So far, we have given judgements which define what it means to be a proposition, namely $P\ prop$, and thence for each proposition, we have by definition a notion of what should count as a verification of that proposition. And we have a judgement $P\ true$, which in its assertion means that one has (a way to obtain) such a verification of P, but we have not considered any judgements which actually refer to the verifications themselves symbolically.

It is a hallmark of Martin-Löf's program to resolve the contradiction between syntax and semantics not by choosing symbols over meanings or meanings over symbols, but by endowing symbols with meaning in the course of knowing the evident judgements. **As such, P is a symbol, but when we assert $P\ prop$ we are saying that we know what proposition P denotes.**

A similar thing can be done with verifications themselves, by representing them with symbols in the same way we have done for the propositions. And then, we can consider a judgement such as "M is a verification of P", and in knowing that judgement, we know what verification M is meant to denote. In practice, this

[4]Note that unless $P\ true$, it need not be evident that $Q\ prop$; in other words, Q only has to be a proposition if P is true. It would also be acceptable to give a stronger definition to implication, but this is the one accepted by Martin-Löf.

judgement has been written in several ways:

Notation	Pronunciation
$M \in P$	M is an element of P
$M \Vdash P$	M realizes P
$P \llcorner \mathsf{ext}\ M \lrcorner$	P is witnessed by M

But they all mean the same thing, so we will choose the notation $\boxed{M \in P}$ and pronounce it "M verifies P". Tentatively, the following defective meaning explanation could be given:

> * To know $M \in P$ is to know that M is a verification of P.

But now that we have started to assign expressions to verifications, we must be more careful about differentiating *direct verifications* (which we will call "canonical") from *indirect verifications* (which we will call "non-canonical"). So the domain of expressions must itself be accorded with a notion of reduction to canonical form, and this corresponds with putting into action a plan of verification in order to get a direct (canonical) verification; reduction to canonical form will be represented by a judgement $\boxed{M \Rightarrow M'}$, pronounced "M evaluates to M'".

> To know $M \Rightarrow M'$ is to know that M is an expression which reduces to a canonical form M'.

An example of an evident reduction judgement in elementary mathematics would be $3+4 \Rightarrow 7$; note that $3+4 \Rightarrow 1+6$ is, on the other hand, not evident, since this judgement describes reduction to *canonical* form, whereas $1+6$ is not a canonical number.

Now, we can correct the previous meaning explanation as follows:

> To know $M \in P$ is to know an M' such that $M \Rightarrow M'$ and M' is a canonical (direct) verification of P.

If it is not yet clear why it would have been a mistake to fail to use the notion of reduction to canonical form in the above meaning explanation, consider that each time a proposition is defined, it should be possible to do so without knowing what other propositions exist in the theory. But if we consider non-canonical forms (as would be necessary if we omitted the $M \Rightarrow M'$ premise), then we would have to fix in advance all the possible non-canonical forms in the computation system in the course of defining each proposition. As such, the open-ended nature of the logic would be destroyed; in a later chapter, the seriousness of this problem will be made even more clear.

The meaning explanation for P *prop* must be accordingly modified to take into account the computational behavior of expressions:

> To know P *prop* is to know a P' such that $P \Rightarrow P'$ and P' is a canonical proposition, which is to say, that one knows what counts as a canonical verification for P'.

In practice, when it is clear that P is canonical, then we will simply say, "To know P *prop* is to know what counts as a canonical verification of P". As an example, then, we will update the evidence of the following assertion:

$$P \supset Q \text{ } prop \text{ } (P \text{ } prop, Q \text{ } prop \text{ } (P \text{ } prop))$$

The meaning of this, expanded into spoken language, is as follows:

> To know $P \supset Q$ *prop* $(P$ *prop*$, Q$ *prop* $(P$ *prop*$))$ is to know what counts as a canonical (direct) verification of $P \supset Q$ under the circumstances that $P \Rightarrow P'$, such that one knows what counts as a canonical verification P', and, if one has such a verification, $Q \Rightarrow Q'$ such that one knows what counts as a canonical verification of Q'.

And the above judgement is evident, since we will say that a canonical verification of $P \supset Q$ is an expression $\lambda x.E$ such that we know the hypothetico-general judgement $|_x E \in Q \ (x \in P)$. This validates the assertion $\lambda x.E \in P \supset Q \ (|_x E \in Q \ (x \in P))$, or, written as an inference rule:

$$\frac{|_x E \in Q \ (x \in P)}{\lambda x.E \in P \supset Q}$$

By the addition of this judgement, we have graduated from a logical theory to a type theory, in the sense of *Constructive Mathematics and Computer Programming* (Martin-Löf, 1979). In fact, we may dispense with the original P *true* form of judgement by *defining* it in terms of the new $M \in P$ judgement as follows:

$$\frac{M \in P}{P \text{ } true}$$

CHAPTER 2

Computational Type Theories

As alluded to at the end of the previous chapter, we may add a judgement $M \in A$ which deals directly with the objects M which verify the propositions (or types) P. We will develop a full theory of dependent types in the sense of Martin-Löf.[1]

1. The categorical judgements

Because we will need to consider the introduction of types which do not have a trivial (intensional) equality relation, we must first amend the meaning explanations for some of our judgements, and add a few new forms of judgement. First, we will refer to *types* rather than *propositions* in order to emphasize the generality of the theory; in some presentations, the word *set* is used instead.

The meaning of hypothetical and general judgement are the same as in the previous chapter, and so we will not reproduce them here. The first form of judgement is $\boxed{A \ type}$, and its meaning explanation is as follows:

> To know A *type* is to know an A' such that $A \Rightarrow A'$, and you know what counts as a canonical verification of A' and when two such verifications are equal.

The next form of judgement is $\boxed{M \in A}$, which remains the same as before:

> To know $M \in A$ (presupposing A *type*) is to know an M' such that $M \Rightarrow M'$ and M' is a canonical verification of A.

We'll need to add new judgements for equality (equality of verifications, and equality of types respectively). First, equality of verifications is written $\boxed{M = N \in A}$, and means the following:

> To know $M = N \in A$ (presupposing $M \in A$ and $N \in A$, and thence A *type*) is to know that the values of M and N are equal as canonical verifications of A.

[1] Please note that the judgements given here, and their meaning explanations, are *not* the same as those used in Constable et al's "Computational Type Theory" and Nuprl. In this chapter, we use the term "computational type theory" in a general sense to characterize a family of type theories which have their origin in Martin-Löf's 1979 paper *Constructive Mathematics and Computer Programming*.

The fact that M and N reduce to canonical values which verify A is known from the presuppositions of the judgement; and what it means for them to be equal as such is known from the evidence of the presupposition $A\ type$ which is obtained from the other presuppositions.

For equality of types $\boxed{A = B\ type}$, there are a number of possible meaning explanations, but we'll use the one that Martin-Löf used starting in 1979:

> To know $A = B\ type$ (presupposing $A\ type$ and $B\ type$) is to know $|_M\, M \in B$ $(M \in A)$ and $|_M\, M \in A$ $(M \in B)$, and moreover $|_{M,N}\, M = N \in B$ $(M = N \in A)$ and $|_{M,N}\, M = N \in A$ $(M = N \in A)$.

In other words, two types are equal when they have the same canonical verifications, and moreover, the same equality relation over their canonical verifications. Note that there are other possible explanations for type equality, including more intensional ones that appeal to the syntactic structure of type expressions, and these turn out to be more useful in proof assistants for practical reasons. However, the extensional equality that we have expounded is the easiest and most obvious one to formulate, so we will use it here.

2. The functional sequent judgement

Now, because we are allowing the definition of types with arbitrary equivalence relations, we cannot use plain hypothetico-general judgement in the course of defining our types. For instance, if we were going to try and define the function type $A \supset B$ in the same way as we did in the previous chapter, we would permit "functions" which are not in fact functional, i.e. they do not take equal inputs to equal outputs. As such, we will need to bake functionality (also called extensionality) into the definition of functions, and since we will need this in many other places, we elect to simplify our definitions by baking it into a single judgement which is meant to be used instead of plain hypothetico-general judgement.

The judgement which expresses simultaneously generality, hypothesis and functionality has been written in multiple ways. Martin-Löf has always written it as $\mathcal{J}\ (\Gamma)$, but this is a confusing notation because it appears as though it is merely a hypothetical judgement (but it is much more, as will be seen). Very frequently, it is written with a turnstile, $\Gamma \vdash \mathcal{J}$, and in the early literature surrounding Constable's Computational Type Theory and Nuprl, it was written $\boxed{\Gamma \gg \mathcal{J}}$; we choose this last option to avoid confusion with a similar judgement form which appears in proof-theoretic, intensional type theories; we'll call the judgement form a "(functional) sequent".

First, we will add the syntax of contexts into the computation system:

(Canonical) $$\cdot \Rightarrow \cdot$$
$$(\Gamma, x : A) \Rightarrow (\Gamma, x : A)$$

The sequent judgements will be defined simultaneously with two other judgements, $\boxed{\Gamma\ ctx}$ ("Γ is a context") and $\boxed{x \# \Gamma}$ ("x is fresh in Γ"). Their meaning explanations are as follows:

> To know $\Gamma\ ctx$ is to know that $\Gamma \Rightarrow \cdot$, or it is to know a variable x and expressions Δ, A such that $\Gamma \Rightarrow \Delta, x : A$ and $\Delta \gg A\ type$, and $x \# \Delta$.
>
> To know $x \# \Gamma$ is to know that $\Gamma \Rightarrow \cdot$, or, if $\Gamma \Rightarrow \Delta, y : A$ such that x is not y and $x \# \Delta$.

In other words, the well-formed contexts are inductively generated by the following grammar:

$$\frac{}{\cdot\ ctx} \qquad \frac{\Gamma\ ctx \quad \Gamma \gg A\ type \quad x \# \Gamma}{\Gamma, x : A\ ctx}$$

$$\frac{}{x \# \cdot} \qquad \frac{x \# \Gamma}{x \# \Gamma, y : A}$$

We will say that $\Gamma \gg \mathcal{J}$ is only a judgement under the presuppositions that $\Gamma\ ctx$ and that $\mathcal{J}$ is a categorical judgement of the form $A\ type$ or $M \in A$. Its meaning explanation must be given separately for each kind of conclusion.

Now we may begin giving the meaning explanations for $\Gamma \gg \mathcal{J}$, starting with typehood with respect to a context:

> To know $\Gamma \gg A\ type$ is to know, if $\Gamma \Rightarrow \cdot$, then $A\ type$; otherwise, if $\Gamma \Rightarrow \Delta, x : B$, that $|_M\, \Delta \gg [M/x]A\ type$ $(\Delta \gg M : B)$ and moreover $|_{M,N}\, \Delta \gg [M/x]A = [N/x]A\ type$ $(\Delta \gg M = N \in B)$.

In this way, we have defined the sequent judgement by structural recursion on the context. We can explain type equality sequents in a similar way:

> To know $\Gamma \gg A = B\ type$ is to know, if $\Gamma \Rightarrow \cdot$, that $A = B\ type$; otherwise, if $\Gamma \Rightarrow \Delta, x : C$, it is to know
>
> $$|_M\, \Delta \gg [M/x]A = [M/x]B\ type\ (\Delta \gg M \in C)$$
>
> and moreover, to know
>
> $$|_{M,N}\, \Delta \gg [M/x]A = [N/x]B\ type\ (\Delta \gg M = N \in C)$$

Next, the meaning of membership sequents is explained:

To know $\Gamma \gg M \in A$ is to know, if $\Gamma \Rightarrow \cdot$, that $M \in A$; otherwise, if $\Gamma \Rightarrow \Delta, x : B$, it is to know

$$|_N\, \Delta \gg [N/x]M \in [N/x]A \ (\Delta \gg N \in B)$$

and moreover, to know

$$|_{L,N}\, \Delta \gg [L/x]M = [N/x]M \in A \ (\Delta \gg L = N \in B)$$

Finally, member equality sequents have an analogous explanation:

To know $\Gamma \gg M = M' \in A$ is to know, if $\Gamma \Rightarrow \cdot$, that $M = M' \in A$; otherwise, if $\Gamma \Rightarrow \Delta, x : B$, it is to know

$$|_N\, \Delta \gg [N/x]M = [N/x]M' \in A \ (\Delta \gg N \in B)$$

and moreover, to know

$$|_{L,N}\, \Delta \gg [L/x]M = [N/x]M' \in A \ (\Delta \gg L = N \in B)$$

The simultaneous definition of multiple judgements may seem at first concerning, but it can be shown to be non-circular by induction on the length of the context Γ.

3. The definitions of types

We will now define the types of a simple computational type theory without universes. In the course of doing so, opportunities will arise for further clarifying the position of the judgements, meaning explanations and proofs on the one hand, and the propositions, definitions and verifications on the other hand.

3.1. The unit type. First, we introduce two canonical forms with trivial reduction rules:

(Canonical) $\qquad \mathsf{unit} \Rightarrow \mathsf{unit} \qquad \bullet \Rightarrow \bullet$

Next, we intend to make the judgement unit *type* evident; and this is done by defining what counts as a canonical verification of unit and when two such verifications are equal. To this end, we say that $\bullet$ is a canonical verification of unit, and that it is equal to itself. I wish to emphasize that this is the entire definition of the type: we have introduced syntax, and we have defined the canonical forms, and there is nothing more to be done.

In the presentations of type theory which are currently in vogue, a type is "defined" by writing out a bunch of inference rules, but in type theory, the definitions that we have given above are prior to the rules, which are justified in respect of the definitions and the meaning explanations of the judgements. For instance, based on the meaning of the various forms of sequent judgement, the following rule schemes

are justified:

$$\frac{}{\Gamma \gg \mathsf{unit}\ \mathit{type}} \qquad \frac{}{\Gamma \gg \mathsf{unit} = \mathsf{unit}\ \mathit{type}}$$

$$\frac{}{\Gamma \gg \bullet \in \mathsf{unit}} \qquad \frac{}{\Gamma \gg \bullet = \bullet \in \mathsf{unit}}$$

Each of the assertions above has evidence of a certain kind; since the justification of these rules with respect to the definitions of the logical constants and the meaning explanations of the judgements is largely self-evident, we omit it in nearly all cases. It is just important to remember that it is not the rules which define the types; a type A is defined in the course of causing the judgement $A\ \mathit{type}$ to become evident. These rules merely codify standard patterns of use, nothing more, and they must each be justified.

3.2. The empty type. The empty type is similarly easy to define. First, we introduce a constant:

(Canonical) $$\mathsf{void} \Rightarrow \mathsf{void}$$

To make the judgement $\mathsf{void}\ \mathit{type}$ evident, we will say that there are no canonical verifications of void, and be done with it. This definition validates some further rules schemes:

$$\frac{}{\Gamma \gg \mathsf{void}\ \mathit{type}} \qquad \frac{}{\Gamma \gg \mathsf{void} = \mathsf{void}\ \mathit{type}}$$

$$\frac{\Gamma \gg M \in \mathsf{void}}{\Gamma \gg \mathcal{J}}$$

The last rule simply says that if we have a verification of void, then we may conclude any judgement whatsoever. Remember that the inference rules are just notation for an *evident* hypothetical judgement, e.g. $\Gamma \gg \mathcal{J}$ ($\Gamma \gg M \in \mathsf{void}$).

Note that we did not introduce any special constant into the computation system to represent the elimination of a verification of void (in proof-theoretic type theories, this non-canonical form is usually called $\mathsf{abort}(R)$). This is because, computationally speaking, there is never any chance that we should ever have use for such a term, since we need only consider the evaluation of closed terms (which is guaranteed by the meaning explanations), and by its very definition, there can never be a closed verification of void.

3.3. The cartesian product of a family of types. This will be our first dependent type, and it will likewise be the first example of a non-trivial addition to the computation system. First, let us add our canonical and non-canonical forms

and their reduction rules:

$$\text{(Canonical)} \qquad \mathsf{\Pi}(A; x.B) \Rightarrow \mathsf{\Pi}(A; x.B) \qquad \lambda(x.E) \Rightarrow \lambda(x.E)$$

$$\text{(Non-canonical)} \qquad \frac{[M/x]E \Rightarrow N}{\mathsf{ap}(\lambda(x.E); M) \Rightarrow N}$$

This is the first time in this chapter that we have introduced a term former with binding structure; it should be noted that the intensional equality of expressions is up to alpha equivalence, and we will not pay attention to issues of variable renaming in our presentation.

We will make evident the following judgement scheme:

$$\frac{A\ \mathit{type} \quad x : A \gg B\ \mathit{type}}{\mathsf{\Pi}(A; x.B)\ \mathit{type}}$$

Or, written as a hypothetical judgement:

$$\mathsf{\Pi}(A; x.B)\ \mathit{type}\ (A\ \mathit{type},\ x : A \gg B\ \mathit{type})$$

This is to say, under the stated assumptions, we know what counts as a canonical verification of $\mathsf{\Pi}(A; x.B)$ and when two such verifications are equal. We will say that $\lambda(x.E)$ is a canonical verification of $\mathsf{\Pi}(A; x.B)$ just when we know $x : A \gg E \in B$; moreover, that two verifications $\lambda(x.E)$ and $\lambda(y.E')$ are equal just when $z \in A \gg [z/x]E = [z/y]E' \in [z/x]B$.

By the meaning of the sequent judgement, this is to say that a lambda expression must be functional with respect to its domain (i.e. it must take equals to equals). We did not need to hypothesize directly two elements of the domain and their equality because this is part of the meaning explanation for the sequent judgement already. Likewise, two lambda expressions are equal when equal inputs yield equal results in both.

The familiar inference rules, which codify the standard mode of use for the family cartesian product, are justified by this definition:

$$\frac{\Gamma \gg A\ \mathit{type} \quad \Gamma, x : A \gg B\ \mathit{type}}{\Gamma \gg \mathsf{\Pi}(A; x.B)\ \mathit{type}} \qquad \frac{\Gamma \gg A = A'\ \mathit{type} \quad \Gamma, z : A \gg [z/x]B = [z/y]B'\ \mathit{type}}{\Gamma \gg \mathsf{\Pi}(A; x.B) = \mathsf{\Pi}(A'; y.B')\ \mathit{type}}$$

$$\frac{\Gamma, x : A \gg E \in B}{\Gamma \gg \lambda(x.E) \in \mathsf{\Pi}(A; x.B)} \qquad \frac{\Gamma, z : A \gg [z/x]E = [z/y]E' \in [z/x]B}{\Gamma \gg \lambda(x.E) = \lambda(y.E') \in \mathsf{\Pi}(A; x.B)}$$

$$\frac{\Gamma \gg M \in \mathsf{\Pi}(A; x.B) \quad \Gamma \gg N \in A}{\Gamma \gg \mathsf{ap}(M; N) \in [N/x]B} \qquad \frac{\Gamma \gg M = M' \in \mathsf{\Pi}(A; x.B) \quad \Gamma \gg N = N' \in A'}{\Gamma \gg \mathsf{ap}(M; N) = \mathsf{ap}(M'; N') \in [N/x]B}$$

Note that the type equality rule scheme that we gave above is structural; it is validated by the meaning explanations, but it is by no means the full totality of

possible equalities between family cartesian product types, which is extensional in this theory.

It will be instructive to explicitly justify the application rule above with respect to the meaning explanations, since I have claimed that such rules are posterior to the definitions we expounded prior to giving these rules.

PROOF. It suffices to consider the case that $\Gamma \Rightarrow \cdot$, because hypotheses may always be added to the context (this is called weakening). And so, by the meaning of the sequent judgement at the empty context, the rule amounts to the assertion

$$\mathsf{ap}(M; N) \in [N/x]B \ (M \in \Pi(A; x.B),\ N \in A)$$

By the meaning explanation for hypothetical judgement, and the definition of the family cartesian product type, we know that $M \Rightarrow \lambda(x.E)$ for some E such that we know $x : A \gg E \in B$; from the meaning of the sequent judgement, we can conclude $|_L\, [L/x]E \in [L/x]B$ $(L \in A)$. On the other hand, from the computation rules, we know that if for some particular E', $[N/x]E \Rightarrow E'$, then we know $\mathsf{ap}(\lambda(x.E); N) \Rightarrow E'$; to demonstrate the evidence of the premise, we may instantiate L at N to know $[N/x]E \in [N/x]B$, whence by the meaning of membership, we know that there exists some canonical E' such that $[N/x] \Rightarrow E'$. □

CHAPTER 3

Proof Theories and Logical Frameworks

1. Proof-theoretic type theory

1.1. Analytic and synthetic Judgement. A synthetic judgement is one for which the experience of *coming to know it* necessarily entails some knowledge which is not implicit in the statement of the judgement; on the other hand, to know an *analytic* judgement is to know it purely on the basis of the information contained inside it. So analytic judgements are decidable, since if they may become evident, it will be purely on the basis of their own content; whereas synthetic judgements become evident to someone when they have obtained some particular evidence for them.

A logical theory has, then, both analytic and synthetic judgements; the judgement P *prop* is analytic, since its evidence follows from the definition of P, whereas the assertion of P *true* entails the knowledge of some extra information, namely a verification of P. When we have extended the logical theory to a type theory in the manner of the previous chapter, the judgement $M \in P$ is also synthetic, since $M \in P$ is not self-evident in general.

But why is it not enough to assert that M verifies P to know whether $M \in P$? It suffices to define a P such that one cannot decide in general whether some term is a verification of it. Let us define the propositional symbol P, and we intend to know the judgement P *prop*, whose meaning is to be expanded as follows:

> To know P *prop* is to know counts as a canonical verification of P.

We will say, then, that $\bullet$ is a canonical verification of P just when Goldbach's conjecture is true. Then it comes immediately that the judgement $M \in \mathsf{P}$ may not be known or refuted on its own basis, nor even the judgement $\bullet \in \mathsf{P}$, since they depend on a proposition whose truth is not known:

> To know $M \in \mathsf{P}$ is to know that $M \Rightarrow M'$ to a canonical verification of P.
>
> $\rightsquigarrow$ To know $M \in \mathsf{P}$ is to know that $M \Rightarrow \bullet$ such that Goldbach's conjecture is true.

1.2. Proof of a judgement vs. verification of a proposition. Because the judgement $M \in P$ is synthetic, we cannot say that it gives rise to a proof theory for the logic, since the core judgement of a proof theory $M : A$ must be analytic, in order to avoid the infinite regress of a proof theory requiring a proof theory requiring a proof theory, and so on.

The notion of verification of a proposition could never be the same as proof anyway, except in the most trivial circumstances, since a verification is meant to be an effective operation which realizes the truth of a proposition, and no constraints whatsoever (termination, totality, etc.) are placed on these operations except those which come from the meaning of the judgements (see [**3**], [**16**], [**17**]).

So a proof theory is necessarily intensional, and its judgements are to be analytic/decidable. What is it, then, that we have considered so far which corresponds with a proof M such that $M : P$ in a proof theory? As discussed above, M is not merely a term such that $M \in P$, since this is not in general enough information to know whether M is a proof. In fact, M must comprise all the logical inferences which led to the knowledge that P is true, and so a meaning explanation for the judgement $M : P$ in a proof theory immediately suggests itself:

> To know $M : P$ is to know that M is evidence (demonstration, proof, derivation) of the judgement P *true*.

And so the term domain of the proof theory is not the same as the one that we have considered so far; it must consist in terms which represent traces of the inferences made in the course of knowing the judgements of a logical theory. There is a sense in which one can consider the types of a proof theory to interpret the judgements of the logical theory, and this methodology is called "judgements as types" (and this implies "derivations as terms").

What I am calling a "proof-theoretic type theory" is a type theory of the sort used in the proof assistants Agda, Coq and Idris, whereas the kind of type theory that I have described in the previous sections, the one based on meaning explanations, underlies the proof assistant Nuprl.

The proof-theoretic type theories on the one hand are often called "intensional" and the computational type theories on the other hand are usually "extensional"; these characterizations are certainly true, though they are not *essential*; moreover, I fear that comparing one of the former with one of the latter is not quite fair, since there is not any clear analogy to be had. That is to say, the judgement $M \in P$ is a judgement which is added to a logical theory and its meaning is (briefly) "M evaluates to a canonical verification of P", whereas $M : P$ cannot be construed as a judgement added to a logical theory. Instead, it must be understood as part of a (proof) theory which is overlayed atop an existing logical theory; it is possible to

understand the theory which contains the judgement $M : P$ to be a metatheory, or logical framework, for the theory which contains the judgement P *true*, which can be construed as the "object language".

In short, the judgements $M \in P$ and $M : P$ are unrelated to each other in two respects: firstly, that they have different meanings, and secondly that the one is at the same level as the judgements of a logical theory, whereas the latter is a judgement in a theory which is defined over a logical theory.

2. Martin-Löf's equational logical framework

To make this more concrete, let us expound a proof theoretic type theory called **MLLF**, which stands for "Martin-Löf's (equational) logical framework";[1] in the course of introducing each type, we will specify which judgement of the underlying logical theory it is meant to interpret.

We start with four categorical judgements:

Judgement Form	Pronunciation
$\alpha : \mathit{type}$	α is a type
$\alpha = \beta : \mathit{type}$	α and β are equal types
$M : \alpha$	M is of type α
$M = N : \alpha$	M and N are equal at type α

But we have not defined the meaning of the judgements; let us do so below:

> To know $\alpha : \mathit{type}$ is to know what counts as an object of type α, and when two such objects are equal.

For now, we'll leave the question of what is an "object" as abstract; in many cases, types will represent judgements of a logical theory, and the objects will be the derivations (demonstrations, proofs) of those judgements.

> To know $\alpha = \beta : \mathit{type}$ is to know that any object of type α is also an object of type β, and two equal objects of type α are equal as objects of type β (necessarily presupposing $\alpha : \mathit{type}$ and $\beta : \mathit{type}$).
>
> To know $M : \alpha$ is to know that M is an object of type α (necessarily presupposing $\alpha : \mathit{type}$).
>
> To know $M = N : \alpha$ is to know that M and N are equal objects of type α (necessarily presupposing $M : \alpha$ and $N : \alpha$).

Hypothetical judgement $\mathcal{J}$ ($\mathcal{J}'$) and general judgement $|_x \mathcal{J}$ have the same meaning in the logical framework as before. In addition to these forms of judgement, we will need contexts (with their wellformedness judgement Γ *ctx*) and an

[1] For a detailed overview of Martin-Löf's equational logical framework, see [**15**].

intensional sequent judgement $\boxed{\Gamma \vdash \mathcal{J}}$; their meanings here will similar to their meanings in the computational type theory, modulo the fact that the intensional sequent judgement is defined over the categorical judgements of the logical framework ($\alpha : \mathit{type}$, $\alpha = \beta : \mathit{type}$, $M : \alpha$ and $M = N : \alpha$), and that we have no notion of evaluation internally.

At this point, we may begin adding types to the logical framework. In practice, most types which we will introduce in the logical framework will be defined in terms of a judgement of the logical theory which lies below it. For instance, hypothetical judgement in the logical theory is represented by a function type in the logical framework, $(x : \alpha)\beta$, whose typehood is meant to be evident under the following circumstances

$$\frac{\alpha : \mathit{type} \quad x : \alpha \vdash \beta : \mathit{type}}{(x : \alpha)\beta : \mathit{type}}$$

Or as a hypothetical judgement, $(x : \alpha)\beta : \mathit{type}$ $(\alpha : \mathit{type}, x : \alpha \vdash \beta : \mathit{type})$.

Now, to know this judgement is to know that under the circumstances we know what is an object of type α and when two such objects are equal, and that if we have such an object x of type α, we know what an object of type β is, and when two such objects are equal—then we know what an object of type $(x : \alpha)\beta$ is, and moreover, for any two objects y, z of type α, that $[y/x]\beta$ and $[z/x]\beta$ are equal as types. To make this evident, then, we will say that under those circumstances an object of type $(x : \alpha)\beta$ is an object $[x]M$ such that one knows $x : \alpha \vdash M : \beta$ and $|_{x,y}\, [y/x]M = [z/x]M : [y/x]\beta$ $(y = z : \alpha)$; furthermore, two such objects are equal just when they yield equal outputs for equal inputs.

Then, for each atomic proposition P, we can easily define a type $\mathsf{Prf}(P)$, as follows. Under the circumstances that P *prop* in the logical theory, then $\mathsf{Prf}(P)$: *type* in the logical framework, since we will define an object of type $\mathsf{Prf}(P)$ to be a derivation of P *true*; beyond reflexivity, further definitional equalities can be added to reflect the harmony of introduction and elimination rules.

Now, the definitions we have given for the types above are "intuitively" correct, but they actually fail to satisfy the meaning explanation that we have given for α : *type*, because they do not take into account neutral terms. In the following sections, we will investigate this problem in more detail and propose a solution.

2.1. What is an "object"? It is time to revisit what it means to be an "object" of a type in the proof-theoretic type theory; we must note how this will necessarily differ from what it meant to be a "verification" of a proposition in the previous sections. Namely, a verification of a proposition is either a *canonical verification* of that proposition (and what sort of thing this might be is known from the presupposition P *prop*), or it is a means of getting such a canonical verification (i.e. a term which evaluates to a canonical verification).

On the other hand, what we have called an "object" of type P is quite different, since in addition to the possibility that it is a canonical proof of the judgement P *true*, it may also be *neutral* (i.e. blocked by a variable); we will call this "normal" rather than "canonical". Why does this happen?

In order to keep the judgement $M : A$ analytic (decidable), its meaning explanation can no longer be based on the idea of the computation of closed terms to canonical form; instead, we will consider the computation of open terms (i.e. terms with free variables) to *normal* form. The desire for $M : A$ to be analytic follows from our intention that it characterize a *proof theory*: we must be able to recognize a proof when we see one. But why are closed-term-based meaning explanations incompatible with this goal? Consider briefly the following judgement:

$$|_n\, M(n) \in P \ (n \in \mathbb{N})$$

To know this judgement is to know that $M(n)$ computes to a canonical verification of P whenever n is a natural number; when P's use of n is not trivial, this amounts to testing an infinite domain (all of the natural numbers), probably by means of mathematical induction. The judgement is then clearly synthetic: to know it is, briefly, to have come up with an (inductive) argument that $M(N)$ computes to a canonical verification of P at each natural number n.

On the other hand, the judgement $n : \mathbb{N} \vdash M(n) : P$ must have a different meaning, one which admits its evidence or refutation purely on syntactic/analytic grounds. In essence, it is to know that $M(n)$ is a proof of P for any *arbitrary* object/expression n such that $n : \mathbb{N}$ (i.e., the only thing we know about n is that it is of type $\mathbb{N}$; we do not necessarily know that it is a numeral).

3. A critique of MLLF

The type theory which we constructed in the previous section is to be considered a proof theory for a logic with the judgements P *prop*, P *true* and $\mathcal{J}$ ($\mathcal{J}'$). There are a few reasons to be dissatisfied with this state of affairs, which I shall enumerate in this section.

3.1. Lack of computational content. Unlike the type theory in the first chapter, there is no built-in computational content. In a computational type theory which is defined by the verificationist meaning explanations, the computational content of terms is understood immediately by means of the $M \Rightarrow M'$ relation; that is, computation is prior to the main judgements because their meaning explanations are defined in terms of evaluation to canonical form.

On the other hand, in the type theory above we did not give a primitive reduction relation; instead, we simply permitted the endowement of proofs with definitional equalities which reflect the harmony of introduction and elimination rules.

That is, if we have known the judgement P *true* by means of an indirect argument (derivation), it must be the case that this derivation corresponds to a direct one; we reflect this in the proof theory by defining the indirect derivation to be definitionally equal to the direct one.

However, this does not amount to computational content being present in terms: only *post facto* may the definitional equality be construed as giving rise to computation, through a metamathematical argument which shows that the definitional equality is confluent and can be used to define a functional normalization relation.

And this is the reason for the peculiarity of the proof-theoretic meaning explanations, namely that they do not include phrases like "evaluates to a canonical...", since evaluation may only be understood after taking the meanings of the judgements ($\alpha : \mathit{type}$, $\alpha = \beta : \mathit{type}$, $M : \alpha$, $M = N : \alpha$) as giving rise to a closed formal system which is susceptible to metamathematical argument: to refer to evaluation in the meaning explanations for the core judgements, then, would be impredicative.

3.2. Modularity of definition. By the same token, the distinction between canonical (direct) and non-canonical (indirect) proof may not be understood as a core notion in the theory, but must be understood separately, secondarily. Why is this a problem? It means that the definition of each type must be made with the full knowledge of the definitions of every other type; in essence, the open-ended nature of type theory is obliterated and one is forced into a fixed formal system; this is in addition to the fact that it causes the epistemic content of $\alpha : \mathit{type}$ for any type α to be extremely complicated.

To illustrate, let us consider as an example a type theory which has four type-formers: trivial truth $\top$, trivial falsity $\bot$, implication $(\alpha)\beta$, and conjunction $\alpha\&\beta$; we will then introduce the following terms to represent proofs: the trivial element $\bullet$, *reductio ad absurdum* $\mathsf{abort}(\alpha; E)$, abstraction $[x : \alpha]E$, application $E(E')$, pairing $\langle E, E' \rangle$, and projections $\mathsf{fst}(E)$, $\mathsf{snd}(E)$.

If we will try to make the judgement $\top : \mathit{type}$ evident, the deficiencies of the formulation will immediately present themselves.

> To know $\top : \mathit{type}$ is to know what counts as an object of type $\top$, and when two such objects are equal. An object of type $\top$, then, is either the expression $\bullet$, or an expression $\mathsf{abort}(\top; E)$ such that we know $E : \bot$, or an expression $E(E')$ such that we know $E : (\alpha)\top$ and $E' : \alpha$, or an expression $\mathsf{fst}(E)$ such that we know $E : \top\&\beta$ for some β, or an expression $\mathsf{snd}(E)$ such that we know $E : \alpha\&\top$ for some α; and we additionally have that $\bullet$ is equal to $\bullet$, and ...

To save space, we elide the rest of the definition of equality for $\top$; what we have seen so far already suffices to bring to light a serious problem: the definition of any type requires knowledge of the entire syntax of the theory. The judgement $\alpha : \mathit{type}$ may never be made evident in isolation, but must be done with full understanding of all the other types and their definitions.

Furthermore, to extend an existing theory with a new type, the definitions of every other type must be rewritten to account for the elimination forms of the new type.

4. A modular logical framework via verifications & uses

We will first try and address the easier of the two critiques by factoring the judgements of the logical framework in such a way as to permit different types to be defined modularly in isolation, and then composed freely into theories. In order to simplify matters, we will restrict our syntax to normal forms, thus eliminating the need for a judgemental equality any more advanced than simple α-equivalence.

The source of the anti-modularity in the previous section's presentation was the fact that the introduction of a type and the elimination of a type were considered within the same judgement. A simple way to resolve this issue is to separate the typing judgement $M : A$ into two separate judgements, a checking judgement $M \Leftarrow A$ and a synthesizing judgement $R \Rightarrow A$; the basic idea is that introduction forms ("verifications") are to be checked, and elimination forms ("uses") are to be synthesized (or inferred). **Note that in this chapter, we have no further use for the big-step reduction judgement $M \Rightarrow M'$; since there will be no ambiguity, so we will re-use its notation for the type synthesis judgement.**

In fact, the technique described above is commonplace in practical implementations of type theory for proof assistants. In practice, checkable terms and inferrable terms are usually put into separate syntactic categories as well; because we need our syntax, judgements and definitions to be as modular as possible, we will find even further granularity to be useful.

4.1. Syntax and judgements. We introduce several syntactic categories (sorts):

sort	names	description
ctx	Γ, Δ	contexts
val	M, N, α, β	values (canonical forms and types)
use	U, V	uses (eliminators)
neu	R, S	neutral terms (variables and applied uses)
exp	E, F, A, B	expressions (normal forms, i.e. canonical forms, neutral terms)

A *value* is an introduction form, like $\langle E_1, E_2 \rangle$ or $[x]E$. A *use* is an unapplied elimination form, like fst, snd, or $\mathsf{ap}(N)$. A *neutral term* is a variable x, or it is an application $U@R$ of a use U to a neutral term R, like $\mathsf{fst}@R$. Lastly, an *expression* is a value $\uparrow M$ or it is a neutral term $\mathsf{neu}(R)$.

Then, we will have the following forms of judgement, each of which will be given a meaning explanation:

judgement	pronunciation
$\Gamma : \mathit{ctx}$	"Γ is a context"
$x \# \Gamma$	"x is fresh in Γ"
$\Gamma(x) \rightsquigarrow A$	"Γ contains $x : A$"
$\Gamma \vdash \alpha \; \epsilon \; \mathit{type}$	"α is a canonical type in context Γ"
$\Gamma \vdash A : \mathit{type}$	"A is a type in context Γ"
$\Gamma \vdash M \; \epsilon \; \alpha$	"M is a canonical verification of A in context Γ"
$\Gamma \vdash U : \alpha > x.\, B(x)$	"U is a use of $x : \alpha$ yielding $B(x)$ in context Γ"
$\Gamma \vdash R \Rightarrow A$	"R synthesizes type A in context Γ"
$\Gamma \vdash E \Leftarrow A$	"E checks at type A in context Γ"
$U \diamond M \rightsquigarrow E$	"U applied to M yields E
$[E/x]M \rightsquigarrow M'$	"the substitution of E for x in M is M'
$[E/x]U \rightsquigarrow U'$	"the substitution of E for x in U is U'
$[E/x]F \rightsquigarrow F'$	"the substitution of E for x in F is F'
$[E/x]R \rightsquigarrow R'$	"the substitution of E for x in R is R'
$[E/x]R \rightsquigarrow E'$	"the substitution of E for x in R is E'

Note that we have made many of the above judgements with respect to a context. It is tempting to first define the each judgement "categorically" (i.e. without a context), and it is often done this way in presentations of type theory, but this is not actually possible for an intensional type theory, where the meaning of the sequent judgement is not compositional. In fact, the sequent judgement for intensional type theories may not even be defined in terms of Martin-Löf's hypothetical judgement, since the former represents derivability, whereas the latter encompasses admissibility as well. In this sense, then, even the sequent judgements above are "categorical".

4.2. Meaning explanations. For each form of judgement listed in the previous section, a meaning explanation will be given.

> To know $\Gamma : \mathit{ctx}$ is to know that $\Gamma \equiv \cdot$; or it is to know that $\Gamma \equiv \Delta, x : A$ such that $\Delta : \mathit{ctx}$, $x \# \Delta$ and $\Delta \vdash A : \mathit{type}$.

To know $x \# \Gamma$ is to know that $\Gamma \equiv \cdot$, or to know a context Δ and a variable $y \# \Delta$ such that x is not y.

To know $\Gamma(x) \rightsquigarrow A$ (presupposing $\Gamma : \mathit{ctx}$) is to know that $\Gamma \equiv \Delta, x : A$, or to know that $\Gamma \equiv \Delta, y : A$ such that $\Delta(x) \rightsquigarrow A$.

To know $\Gamma \vdash \alpha \ \epsilon \ \mathit{type}$ (presupposing $\Gamma : \mathit{ctx}$) is to know what counts as a canonical verification of α in context Γ, and what is a use of $x : \alpha$ yielding what in context Γ; moreover, for any such use U and any such canonical verification M, to know what expression is yielded by applying U to M. Moreover, for any such verification M and any variable x, to know the value of the substitution of E for x in M for any E; and for any such use U and any variable x, to know the substitution of E for x in U for any E; finally, to know the substitution of E for x in α for any E and x.

To know $\Gamma \vdash A : \mathit{type}$ is to know a type α such that $A \equiv \uparrow \alpha$ and $\Gamma \vdash \alpha \ \epsilon \ \mathit{type}$.

To know $\Gamma \vdash M \ \epsilon \ \alpha$ (presupposing $\Gamma \vdash \alpha \ \epsilon \ \mathit{type}$) is to know that M is a canonical verification of α in context Γ.

To know $\Gamma \vdash U : \alpha > x.B(x)$ (presupposing $\Gamma \vdash \alpha \ \epsilon \ \mathit{type}$, $\Gamma, x :\uparrow \alpha \vdash B(x) : \mathit{type}$ and $x \# \Gamma$) is to know that U is a use of $x : \alpha$ yielding $B(x)$ in context Γ.

To know $\Gamma \vdash R \Rightarrow A$ (presupposing $\Gamma : \mathit{ctx}$) is, if $R \equiv x$, to know that $\Gamma(x) \rightsquigarrow A$; if $R \equiv U@S$, then it is to know that $\Gamma \vdash U : \beta > x.C(x)$, $\Gamma \vdash S \Rightarrow \uparrow \beta$, and $C(U@S) \equiv A$.

To know $\Gamma \vdash E \Leftarrow A$ (presupposing $\Gamma \vdash A : \mathit{type}$) is, if $E \equiv \mathsf{neu}(R)$, to know that $\Gamma \vdash R \Rightarrow A$; if $E \equiv \uparrow M$, then it is to know $A \equiv \uparrow \alpha$ and $\Gamma \vdash M \ \epsilon \ \alpha$.

To know $U \diamond M \rightsquigarrow E$ is to know that U applied to M yields E.

To know $[E/x]M \rightsquigarrow M'$ is to know that the substitution of E for x in M is M'.

To know $[E/x]U \rightsquigarrow U'$ is to know that the substitution of E for x in U is U'.

To know $[E/x]\mathsf{neu}(R) \rightsquigarrow \mathsf{neu}(R')$ is to know that $[E/x]R \rightsquigarrow R'$; to know $[E/x]\mathsf{neu}(R) \rightsquigarrow \uparrow M$ is to know that $[E/x]R \rightsquigarrow \uparrow M$; to know $[E/x]\uparrow M \rightsquigarrow \uparrow M'$ is to know that $[E/x]M \rightsquigarrow M'$.

To know $[E/x]R \rightsquigarrow R$ is to know that x does not occur free in R; to know $[E/x]U@R \rightsquigarrow V@S$ is to know that $[E/x]U \rightsquigarrow V$ and $[E/x]R \rightsquigarrow S$.

To know $[E/x]x \rightsquigarrow E$ is immediate; to know $[E/x]U@R \rightsquigarrow E'$ is to know that $[E/x]U \rightsquigarrow V$, $[E/x]R \rightsquigarrow F$, and $V \diamond F \rightsquigarrow E'$.

Many of the above explanations may be given more clearly as inference rules; it is not some peculiarity of meaning explanations that they must be given in prose rather than mathematical notation, though we have stuck to the former here for uniformity.

4.2.1. *Discussion.* The idea of separating out unapplied elimination forms (*uses* in our parlance) is closely related to the notion of *spines*, which are lists of *uses*, which may be applied all at once to a variable. It would be possible to add further syntax and judgements for spines into our system, but the primary reason for separating out the *uses* in the first place was simply to make the evidence for $\Gamma \vdash \alpha \ \epsilon \ \mathit{type}$ as minimal as possible, so as to make the definitions of the types both economical and modular.

The substitution judgements that we gave over each relevant syntactic category give rise to a system of *hereditary substitution* (see [**18**]); the purpose of hereditary substitutions is to provide a way to substitute a normal form into a term without creating redexes. The hereditary substitutions may be divided into two groups. First, the structural substitutions which must be given separately for the operators of each type, $[E/x]M \rightsquigarrow M'$ and $[E/x]U \rightsquigarrow U'$; these may be implemented mechanically, by induction on the arity of each operator. Second, the substitutions which may be implemented once and for all in terms of the other judgements, namely $[E/x]R \rightsquigarrow M$, $[E/x]R \rightsquigarrow S$ and $[E/x]F \rightsquigarrow F'$.

Lastly, in its full presentation, Martin-Löf's logical framework has a single universe Set such that $\mathsf{El}(A) : \mathit{type}$ $(A : \mathsf{Set})$; in a proper development of our modular logical framework, we would need to add such a universe. In order to do this, the meaning explanation of $\Gamma \vdash A : \mathit{type}$ (which is defined trivially in terms of $\Gamma \vdash \alpha \ \epsilon \ \mathit{type}$) would have to be adjusted, in order to take into account the neutral types which would be introduced by the universe.

4.3. The definitions of types. Now we will proceed with defining some basic types according to the meaning explanations we gave above.

4.3.1. *The unit type.* The following operators are introduced with their sorts:

$$\mathsf{unit} : \mathbf{val} \qquad \bullet : \mathbf{val}$$

Then, we will make the judgement $|_\Gamma\, \Gamma \vdash \mathsf{unit}\ \epsilon\ \mathit{type}$ evident; note that we will use Martin-Löf's general judgement in order to express that the typehood assertion is valid in *any* context. The evidence of this judgement is the following list of definitions:

(1) $\bullet$ is a canonical verification of unit in context Γ.
(2) There are no applicable uses.
(3) For any E and x, the substitution of E for x in $\bullet$ is $\bullet$.
(4) For any E and x, the substitution of E for x in unit is unit.

4.3.2. *The empty type.* For the empty type, we introduce the following operators:

$$\mathsf{void} : \mathbf{val} \qquad \mathsf{abort}(A) : \mathbf{use}\ (A : \mathbf{exp})$$

The judgement $|_\Gamma\, \Gamma \vdash \mathsf{void}\ \epsilon\ \mathit{type}$ is made evident by means of the following definitions:

(1) There is no canonical verification of void.
(2) $\mathsf{abort}(A)$ is a use of $x : \mathsf{void}$ yielding A in context Γ, assuming $\Gamma \vdash A : \mathit{type}$.
(3) For any E and x, the substitution of E for x in $\mathsf{abort}(A)$ is $\mathsf{abort}(A')$, assuming $[E/x]A \rightsquigarrow A'$.
(4) For any E and x, the substitution of E for x in void is void.

This is the first time we have seen a significant difference in practice from the extensional type theory explored in the previous chapter; note how we include an elimination form for void in its definition, whereas in extensional type theory this is entirely unnecessary.

4.3.3. *The dependent function type.* For the dependent function type, the following operators are introduced:

$$(x : A)B(x) : \mathbf{val}\ (A : \mathbf{exp}, B(x) : \mathbf{exp}\ (x : \mathbf{neu}))$$
$$[x]E(x) : \mathbf{val}\ (E(x) : \mathbf{exp}\ (x : \mathbf{neu}))$$
$$\mathsf{ap}(E) : \mathbf{use}\ (E : \mathbf{exp})$$

The judgement $|_\Gamma\, \Gamma \vdash (x : A)B(x)\ \epsilon\ \mathit{type}$ $(\Gamma \vdash A : \mathit{type};\ \Gamma, x : A \vdash B(x) : \mathit{type})$ is made evident via the following definitions:

(1) A canonical verification of $(x : A)B(x)$ in context Γ is $[x]E(x)$, such that $\Gamma, x : A \vdash E(x) \Leftarrow B(x)$
(2) $\mathsf{ap}(E)$ is a use of $z : (x : A)B(x)$ yielding B' in context Γ, assuming $\Gamma \vdash E \Leftarrow A$ and $[E/x]B \rightsquigarrow B'$.

(3) Applying $\mathsf{ap}(E)$ to $[x]F(x)$ yields F', assuming $[E/x]F \rightsquigarrow F'$.
(4) The substitution of E for z in $[x]F(x)$ is $[x]F'(x)$, assuming $[E/z]F \rightsquigarrow F'$.
(5) The substitution of E for z in $\mathsf{ap}(F)$ is $\mathsf{ap}(F')$, assuming $[E/z]F \rightsquigarrow F'$.
(6) The substitution of E for z in $(x : A)B(x)$ is $(x : A')B'(x)$, assuming $[E/z]A \rightsquigarrow A'$ and $[E/z]B \rightsquigarrow B'$.

Based on this definition, and the meanings of the judgements, the familiar inference rules are valid (justified):

$$\frac{\Gamma \vdash A : \mathit{type} \quad \Gamma, x : A \vdash B(x) : \mathit{type}}{\Gamma \vdash \uparrow (x : A)B(x) : \mathit{type}}$$

$$\frac{\Gamma, x : A \vdash E(x) \Leftarrow B(x)}{\Gamma \vdash \uparrow [x]E(x) \Leftarrow \uparrow (x : A)B(x)} \qquad \frac{\Gamma \vdash F \Rightarrow \uparrow (x : A)B(x) \quad [E/x]B \rightsquigarrow B'}{\Gamma \vdash \mathsf{ap}(E)@F \Rightarrow B'}$$

I leave the proofs of the validity of these rule schemes as an exercise to the reader.

4.4. Discussion. In this chapter, a logical framework has been developed which, unlike Martin-Löf's logical framework, permits the modular/separate definitions of types. On the other hand, we have *not* addressed the first critique of **MLLF**, which is that it lacks computational content.

To make the framework truly computational, it would be necessary to modify the syntax to include redexes, add judgements for reduction of open terms to normal form (as opposed to reduction of closed terms to canonical form), and adjust the meaning explanations of the judgements $\Gamma \vdash M \Leftarrow A$ and $\Gamma \vdash R \Rightarrow A$ to normalize their subjects.

Bibliography

[1] Stuart F. Allen, Mark Bickford, Robert L. Constable, Richard Eaton, Christoph Kreitz, Lori Lorigo, and E. Moran. Innovations in computational type theory using nuprl. *J. Applied Logic*, 4(4):428–469, 2006.

[2] Robert L. Constable, Stuart F. Allen, Mark Bromley, Rance Cleaveland, J. F. Cremer, R. W. Harper, Douglas J. Howe, Todd B. Knoblock, N. P. Mendler, Prakash Panangaden, James T. Sasaki, and Scott F. Smith. *Implementing mathematics with the Nuprl proof development system.* Prentice Hall, 1986.

[3] Michael Dummett. *Elements of intuitionism*, volume 39 of *Oxford Logic Guides*. The Clarendon Press Oxford University Press, New York, second edition, 2000.

[4] Peter Dybjer. Program testing and the meaning explanations of intuitionistic type theory. In Peter Dybjer, Sten Lindström, Erik Palmgren, and Göran Sundholm, editors, *Epistemology versus Ontology*, volume 27 of *Logic, Epistemology, and the Unity of Science*, pages 215–241. Springer, 2012.

[5] Johan Georg Granström. *Treatise on Intuitionistic Type Theory*. Springer Publishing Company, Incorporated, 2013.

[6] Robert Harper. Constructing type systems over an operational semantics. *J. Symb. Comput.*, 14(1):71–84, 1992.

[7] Robert Harper. *Practical Foundations for Programming Languages*. Cambridge University Press, New York, NY, USA, 2012.

[8] Robert Harper, Furio Honsell, and Gordon Plotkin. A framework for defining logics. *J. ACM*, 40(1):143–184, January 1993.

[9] Robert Harper and Daniel R. Licata. Mechanizing metatheory in a logical framework. *J. Funct. Program.*, 17(4-5):613–673, July 2007.

[10] Per Martin-Löf. Constructive mathematics and computer programming. In L. Jonathan Cohen, Jerzy Łoś, Helmut Pfeiffer, and Klaus-Peter Podewski, editors, *Logic, Methodology and Philosophy of Science VI, Proceedings of the Sixth International Congress of Logic, Methodology and Philosophy of Science, Hannover 1979*, volume 104 of *Studies in Logic and the Foundations of Mathematics*, pages 153–175. North-Holland, 1982.

[11] Per Martin-Löf. *Intuitionistic Type Theory*. Bibliopolis, 1984.

[12] Per Martin-Löf. Truth of a proposition, evidence of a judgement, validity of a proof. *Synthese*, 73(3):407–420, 1987.

[13] Per Martin-Löf. Analytic and synthetic judgements in type theory. In Paolo Parrini, editor, *Kant and Contemporary Epistemology*, volume 54 of *The University of Western Ontario Series in Philosophy of Science*, pages 87–99. Springer Netherlands, 1994.

[14] Per Martin-Löf. On the meanings of the logical constants and the justifications of the logical laws. *Nordic Journal of Philosophical Logic*, 1(1):11–60, 1996.

[15] B. Nordström, K. Petersson, and J.M. Smith. *Programming in Martin Löf's Type Theory*, volume 7 of *Monographs on Computer Science*. Oxford University Press, 1990.

[16] Dag Prawitz. Truth and proof in intuitionism. In Peter Dybjer, Sten Lindström, Erik Palmgren, and Göran Sundholm, editors, *Epistemology versus Ontology*, volume 27 of *Logic, Epistemology, and the Unity of Science*, pages 45–67. Springer, 2012.

[17] A.S. Troelstra and D. van Dalen. *Constructivism in Mathematics*. North-Holland, 1988. 2 volumes.

[18] Kevin Watkins, Iliano Cervesato, Frank Pfenning, and David Walker. A Concurrent Logical Framework: The Propositional Fragment. *Types for Proofs and Programs*, pages 355–377, 2004.