

Type-Theoretic Signatures for Algebraic Theories and Inductive Types

András Kovács

2021 September

Contents

1	Introduction	1
1.1	Specification and Semantics for Inductive Types	1
1.2	Overview of the Thesis and Contributions	1
1.3	Notation and Conventions	1
1.3.1	Metatheory	1
1.3.2	Universes	1
2	Simple Inductive Signatures	2
2.1	Theory of Signatures	2
2.2	Semantics	4
2.2.1	Algebras	5
2.2.2	Morphisms	6
2.2.3	Displayed Algebras	8
2.2.4	Sections	10
2.3	Term Algebras	11
2.3.1	Weak Initiality	12
2.3.2	Induction	13
2.4	Discussion	14
2.4.1	Comparison to F-algebras	14
2.4.2	Generic Programming	16
3	Semantics in Two-Level Type Theory	21
3.1	Motivation	21
3.2	Models of Type Theories	23
3.2.1	The Algebraic View	23
3.2.2	Categories With Families	24
3.2.3	Type formers	26

3.3	Two-Level Type Theory	32
3.3.1	Models	32
3.3.2	Internal Syntax and Notation	34
3.4	Standard Semantics of 2LTT	35
3.4.1	Presheaf Model of the Outer Layer	35
3.4.2	Modeling The Inner Layer	41
3.4.3	Functions with Inner Domains	42
3.5	Simple Inductive Signatures in 2LTT	43
4	Finitary Quotient Inductive-Inductive Types	44
4.1	Theory of Signatures	44
4.1.1	Models	44
4.1.2	Examples	44
4.2	Semantics	44
4.2.1	Finite Limit Cwfs	44
4.2.2	Equivalence of Initiality and Induction	44
4.2.3	Model of the Theory of Signatures	44
4.3	Term Algebras	44
4.3.1	Generic Term Algebras	44
4.3.2	Induction for Term Algebras	44
4.3.3	Church Encoding	44
4.3.4	Awodey-Frey-Speight Encoding	44
4.4	Left Adjoints of Signature Morphisms	44
5	Infinitary Quotient Inductive-Inductive Types	45
5.1	Theory of Signatures	45
5.2	Semantics	45
5.3	Term Algebras	45
6	Levitation, Bootstrapping and Universe Levels	46
6.1	Levitation for Closed QIITs	46
6.2	Levitation for Infinitary QIITs	46
7	Higher Inductive-Inductive Types	47
7.1	Theory of Signatures	47
7.2	Semantics	47

<i>CONTENTS</i>	iii
8 Reductions	48
8.1 Finitary Inductive Types	48
8.2 Finitary Inductive-Inductive Types	48
8.3 Closed Quotient Inductive-Inductive Types	48
9 Conclusion	49

CHAPTER 1

Introduction

1.1 Specification and Semantics for Inductive Types

1.2 Overview of the Thesis and Contributions

1.3 Notation and Conventions

1.3.1 Metatheory

1.3.2 Universes

CHAPTER 2

Simple Inductive Signatures

In this chapter, we take a look at a very simple notion of inductive signature. The motivation for doing so is to present the basic ideas of this thesis in the easiest possible setting, with explicit definitions. The later chapters are greatly generalized and expanded compared to the current one, and are not feasible (and probably not that useful) to present in full formal detail. We also include a complete Agda formalization of the contents of this chapter, in less than 200 lines.

potentially in intro

The mantra throughout this dissertation is the following: inductive types are specified by typing contexts in certain *theories of signatures*. For each class of inductive types, there is a corresponding theory of signatures, which is viewed as a proper type theory and comes equipped with an algebraic model theory. *Semantics* of signatures is given by interpreting them in certain models of the theory of signatures. Semantics should at least provide a notion of induction principle for each signature; in this chapter we provide a bit more than that, and substantially more in Chapters 4 and 5.

2.1 Theory of Signatures

Generally, more expressive theories of signatures can describe a larger classes of inductive types. As we are aiming at minimalism right now, the current theory of signatures is as follows:

Definition 1. The **theory of signatures**, or ToS for short, is a simple type theory equipped with the following features:

- An empty base type ι .
- A *first-order function type* $\iota \rightarrow -$; this is a function whose domain is fixed to be ι . Moreover, first-order functions only have neutral terms: there is application, but no λ -abstraction.

We can specify the full syntax using the following Agda-like inductive definitions.

$$\begin{array}{ll}
\text{Ty} & : \text{Set} \\
\iota & : \text{Ty} \\
\iota \rightarrow - & : \text{Ty} \rightarrow \text{Ty} \\
\\
\text{Var} & : \text{Con} \rightarrow \text{Ty} \rightarrow \text{Set} \\
\text{vz} & : \text{Var } (\Gamma \triangleright A) A \\
\text{vs} & : \text{Var } \Gamma A \rightarrow \text{Var } (\Gamma \triangleright B) A \\
\\
\text{Con} & : \text{Set} \\
\bullet & : \text{Con} \\
- \triangleright - & : \text{Con} \rightarrow \text{Ty} \rightarrow \text{Con} \\
\\
\text{Tm} & : \text{Con} \rightarrow \text{Ty} \rightarrow \text{Set} \\
\text{var} & : \text{Var } \Gamma A \rightarrow \text{Tm } \Gamma A \\
\text{app} & : \text{Tm } \Gamma (\iota \rightarrow A) \rightarrow \text{Tm } \Gamma \iota \rightarrow \text{Tm } \Gamma A
\end{array}$$

Here, Con contexts are lists of types, and Var specifies well-typed De Bruijn indices, where vz represents the zero index, and vs takes the successor of an index.

Notation 1. We use capital Greek letters starting from Γ to refer to contexts, A, B, C to refer to types, and t, u, v to refer to terms. In examples, we may use a nameful notation instead of De Bruijn indices. For example, we may write $x : \text{Tm } (\bullet \triangleright (x : \iota) \triangleright (y : \iota)) \iota$ instead of $\text{var } (\text{vs vz}) : \text{Tm } (\bullet \triangleright \iota \triangleright \iota) \iota$. Additionally, we may write $t u$ instead of $\text{app } t u$ for t and u terms.

Definition 2. Parallel substitutions map variables to terms.

$$\begin{array}{l}
\text{Sub} : \text{Con} \rightarrow \text{Con} \rightarrow \text{Set} \\
\text{Sub } \Gamma \Delta \equiv \{A\} \rightarrow \text{Var } \Delta A \rightarrow \text{Tm } \Gamma A
\end{array}$$

We use σ and δ to refer to substitutions. We also recursively define the action of substitution on terms:

$$\begin{array}{l}
-[-] : \text{Tm } \Delta A \rightarrow \text{Sub } \Gamma \Delta \rightarrow \text{Tm } \Gamma A \\
(\text{var } x) \ [-\sigma] \equiv \sigma x \\
(\text{app } t u) [-\sigma] \equiv \text{app } (t[-\sigma]) (u[-\sigma])
\end{array}$$

The identity substitution id is defined simply as var . It is easy to see that $t[\text{id}] = t$ for all t . Substitution composition is as follows.

$$\begin{aligned} - \circ - &: \text{Sub } \Delta \Xi \rightarrow \text{Sub } \Gamma \Delta \rightarrow \text{Sub } \Gamma \Xi \\ (\sigma \circ \delta) x &::= (\sigma x)[\delta] \end{aligned}$$

Example 1. We may write signatures for natural numbers and binary trees respectively as follows.

$$\begin{aligned} \text{NatSig} &::= \bullet \triangleright (\text{zero} : \iota) \triangleright (\text{suc} : \iota \rightarrow \iota) \\ \text{TreeSig} &::= \bullet \triangleright (\text{leaf} : \iota) \triangleright (\text{node} : \iota \rightarrow \iota \rightarrow \iota) \end{aligned}$$

In short, the current ToS allows inductive types which are

- *Single-sorted*: this means that we have a single type constructor, corresponding to ι .
- *Closed*: signatures cannot refer to any externally existing type. For example, we cannot write a signature for “lists of natural number” in a direct fashion, since there is no way to refer to the type of natural numbers.
- *Finitary*: inductive types corresponding to signatures are always finitely branching trees. Being closed implies being finitary, since an infinitely branching node would require some external type to index subtrees with. For example, $\text{node} : (\mathbb{N} \rightarrow \iota) \rightarrow \iota$ would specify an infinite branching (if such type was allowed in ToS).

Remark. We omit λ -expressions from ToS for the sake of simplicity: this causes terms to be always in normal form (neutral, to be precise), and thus we can skip talking about conversion rules. Later, starting from Chapter 4 we include proper $\beta\eta$ -rules in signature theories.

2.2 Semantics

For each signature, we need to know what it means for a type theory to support the corresponding inductive type. For this, we need at least a notion of *algebras*, which can be viewed as a bundle of all type and value constructors, and what it means for an algebra to support an *induction principle*. Additionally, we may

want to know what it means to support a *recursion principle*, which can be viewed as a non-dependent variant of induction. In the following, we define these notions by induction on ToS syntax.

Remark. We use “algebra” and “model” synonymously throughout this thesis.

2.2.1 Algebras

First, we calculate types of algebras. This is simply a standard interpretation into the **Set** universe. We define the following operations by induction; the $-^A$ name is overloaded for **Con**, **Ty** and **Tm**.

$$\begin{aligned} -^A &: \mathbf{Ty} \rightarrow \mathbf{Set} \rightarrow \mathbf{Set} \\ \iota^A & \quad X := X \\ (\iota \rightarrow A)^A X &:= X \rightarrow A^A X \end{aligned}$$

$$\begin{aligned} -^A &: \mathbf{Con} \rightarrow \mathbf{Set} \rightarrow \mathbf{Set} \\ \Gamma^A X &:= \{A : \mathbf{Ty}\} \rightarrow \mathbf{Var} \Gamma A \rightarrow A^A X \end{aligned}$$

$$\begin{aligned} -^A &: \mathbf{Tm} \Gamma A \rightarrow \{X : \mathbf{Set}\} \rightarrow \Gamma^A X \rightarrow A^A X \\ (\mathbf{var} \ x)^A \gamma &:= \gamma \ x \\ (\mathbf{app} \ t \ u)^A \gamma &:= t^A \gamma (u^A \gamma) \end{aligned}$$

$$\begin{aligned} -^A &: \mathbf{Sub} \Gamma \Delta \rightarrow \{X : \mathbf{Set}\} \rightarrow \Gamma^A X \rightarrow \Delta^A X \\ \sigma^A \gamma \ x &:= (\sigma \ x)^A \gamma \end{aligned}$$

Here, types and contexts depend on some $X : \mathbf{Set}$, which serves as the interpretation of ι . We define Γ^A as a product: for each variable in the context, we get a semantic type. This trick, along with the definition of **Sub**, makes formalization a bit more compact. Terms and substitutions are interpreted as natural maps. Substitutions are interpreted by pointwise interpreting the contained terms.

Notation 2. We may write values of Γ^A using notation for Σ -types. For example, we may write $(\mathbf{zero} : X) \times (\mathbf{suc} : X \rightarrow X)$ for the result of computing $\mathbf{NatSig}^A X$.

Definition 3. We define **algebras** as follows.

$$\begin{aligned} \text{Alg} &: \text{Con} \rightarrow \text{Set}_1 \\ \text{Alg } \Gamma &\equiv (X : \text{Set}) \times \Gamma^A X \end{aligned}$$

Example 2. Alg NatSig is computed to $(X : \text{Set}) \times (\text{zero} : X) \times (\text{suc} : X \rightarrow X)$.

2.2.2 Morphisms

Now, we compute notions of morphisms of algebras. In this case, morphisms are functions between underlying sets which preserve all specified structure. The interpretation for calculating morphisms is a *logical relation interpretation* [HRR14] over the $-^A$ interpretation. The key part is the interpretation of types:

$$\begin{aligned} -^M &: (A : \text{Ty})\{X_0 X_1 : \text{Set}\}(X^M : X_0 \rightarrow X_1) \rightarrow A^A X_0 \rightarrow A^A X_1 \rightarrow \text{Set} \\ \iota^M & \quad X^M \alpha_0 \alpha_1 \equiv X^M \alpha_0 = \alpha_1 \\ (\iota \rightarrow A)^M & X^M \alpha_0 \alpha_1 \equiv (x : X_0) \rightarrow A^M X^M (\alpha_0 x) (\alpha_1 (X^M x)) \end{aligned}$$

We again assume an interpretation for the base type ι , as X_0 , X_1 and $X^M : X_0 \rightarrow X_1$. X^M is function between underlying sets of algebras, and A^M computes what it means that X^M preserves an operation with type A . At the base type, preservation is simply equality. At the first-order function type, preservation is a quantified statement over X_0 . We define morphisms for **Con** pointwise:

$$\begin{aligned} -^M &: (\Gamma : \text{Con})\{X_0 X_1 : \text{Set}\} \rightarrow (X_0 \rightarrow X_1) \rightarrow \Gamma^A X_0 \rightarrow \Gamma^A X_1 \rightarrow \text{Set} \\ \Gamma^M X^M \gamma_0 \gamma_1 &\equiv \{A : \text{Ty}\}(x : \text{Var } \Gamma A) \rightarrow A^M X^M (\gamma_0 x) (\gamma_1 x) \end{aligned}$$

For terms and substitutions, we get preservation statements, which are sometimes called *fundamental lemmas* in discussions of logical relations [HRR14].

$$\begin{aligned} -^M &: (t : \text{Tm } \Gamma A) \rightarrow \Gamma^M X^M \gamma_0 \gamma_1 \rightarrow A^M X^M (t^A \gamma_0) (t^A \gamma_1) \\ (\text{var } x)^M & \gamma^M \equiv \gamma^M x \\ (\text{app } t u)^M & \gamma^M \equiv t^M \gamma^M (u^A \gamma_0) \\ -^M &: (\sigma : \text{Sub } \Gamma \Delta) \rightarrow \Gamma^M X^M \gamma_0 \gamma_1 \rightarrow \Delta^M X^M (\sigma^A \gamma_0) (\sigma^A \gamma_1) \\ \sigma^M & \gamma^M x \equiv (\sigma x)^M \gamma^M \end{aligned}$$

The definition of $(\text{app } t u)^M$ is well-typed by the induction hypothesis $u^M \gamma^M : X^M (u^A \gamma_0) = u^A \gamma_1$.

Definition 4. To get notions of **algebra morphisms**, we again pack up Γ^M with the interpretation of ι .

$$\begin{aligned} \text{Mor} &: \{\Gamma : \text{Con}\} \rightarrow \text{Alg } \Gamma \rightarrow \text{Alg } \Gamma \rightarrow \text{Set} \\ \text{Mor } \{\Gamma\} (X_0, \gamma_0) (X_1, \gamma_1) &\equiv (X^M : X_0 \rightarrow X_1) \times \Gamma^M X^M \gamma_0 \gamma_1 \end{aligned}$$

Example 3. We have the following computation:

$$\begin{aligned} \text{Mor } \{\text{NatSig}\} (X_0, \text{zero}_0, \text{suc}_0) (X_0, \text{zero}_1, \text{suc}_1) &\equiv \\ (X^M : X_0 \rightarrow X_1) & \\ \times (X^M \text{zero}_0 = \text{zero}_1) & \\ \times ((x : X_0) \rightarrow X^M (\text{suc}_0 x) = \text{suc}_1 (X^M x)) & \end{aligned}$$

Definition 5. We state **initiality** as a predicate on algebras:

$$\begin{aligned} \text{Initial} &: \{\Gamma : \text{Con}\} \rightarrow \text{Alg } \Gamma \rightarrow \text{Set} \\ \text{Initial } \{\Gamma\} \gamma &\equiv (\gamma' : \text{Alg } \Gamma) \rightarrow \text{isContr } (\text{Mor } \gamma \gamma') \end{aligned}$$

Here `isContr` refers to unique existence [Uni13, Section 3.11]. If we drop `isContr` from the definition, we get the notion of weak initiality, which corresponds to the recursion principle for Γ . Although we call this predicate **Initial**, in this chapter we do not yet show that algebras form a category. We provide the extended semantics in Chapter 4. The computed algebras and morphism there remain the same as in the current chapter.

Morphisms vs. logical relations. The $-^M$ interpretation can be viewed as a special case of logical relations over the $-^A$ model: every morphism is a *functional* logical relation, where the chosen relation between the underlying sets happens to be a function. Consider now a more general relational interpretation for types:

$$\begin{aligned} -^R &: (A : \text{Ty}) \{X_0 X_1 : \text{Set}\} (X^R : X_0 \rightarrow X_1 \rightarrow \text{Set}) \rightarrow A^A X_0 \rightarrow A^A X_1 \rightarrow \text{Set} \\ \iota^R \quad X^R \alpha_0 \alpha_1 &\equiv X^R \alpha_0 \alpha_1 \\ (\iota \rightarrow A)^R X^R \alpha_0 \alpha_1 &\equiv (x_0 : X_0)(x_1 : X_1) \rightarrow X^R x_0 x_1 \rightarrow A^R X^R (\alpha_0 x_0) (\alpha_1 x_1) \end{aligned}$$

Here, functions are related if they map related inputs to related outputs. If we know that $X^M \alpha_0 \alpha_1 \equiv (f \alpha_0 = \alpha_1)$ for some f function, we get

$$(x_0 : X_0)(x_1 : X_1) \rightarrow f x_0 = x_1 \rightarrow A^R X^R (\alpha_0 x_0) (\alpha_1 x_1)$$

Now, we can simply substitute along the input equality proof in the above type, to get the previous definition for $(\iota \rightarrow A)^M$:

$$(x_0 : X_0) \rightarrow A^R X^R (\alpha_0 x_0) (\alpha_1 (f x_0))$$

This substitution along the equation is called “singleton contraction” in the jargon of homotopy type theory [Uni13]. The ability to perform contraction here is at the heart of the *strict positivity restriction* for inductive signatures. Strict positivity in our setting corresponds to only having first-order function types in signatures. If we allowed function domains to be arbitrary types, in the definition of $(A \rightarrow B)^M$ we would only have a black-box $A^M X^M : A^A X_0 \rightarrow A^A X_1 \rightarrow \mathbf{Set}$ relation, which is not known to be given as an equality.

In Chapter 4 we expand on this. As a preliminary summary: although higher-order functions have relational interpretation, such relations do not generally compose. What we eventually aim to have is a *category* of algebras and algebra morphisms, where morphisms do compose. We need a *directed* model of the theory of signatures, where every signature becomes a category of algebras. The way to achieve this, is to prohibit higher-order functions, thereby avoiding the polarity issues that prevent a directed interpretation for general function types.

2.2.3 Displayed Algebras

At this point we do not yet have specification for induction principles. We use the term *displayed algebra* to refer to “dependent” algebras, where every displayed algebra component lies over corresponding components in the base algebra. For the purpose of specifying induction, displayed algebras can be viewed as bundles of induction motives and methods.

Displayed algebras over some $\gamma : \mathbf{Alg} \Gamma$ are equivalent to slices over γ in the category of Γ -algebras; we show this in Chapter 4. A slice $f : \mathbf{Sub} \Gamma \gamma' \gamma$ maps elements of γ ’s underlying set to elements in the base algebra. Why do we need displayed algebras, then? The main reason is that if we are to eventually implement inductive types in a dependently typed language, we need to compute induction principles exactly, not merely up to isomorphisms.

For more illustration of using some displayed algebras in a type-theoretic setting, see [AL19]. We adapt the term “displayed algebra” from *ibid.* as a generalization of displayed categories (and functors, natural transformations) to other algebraic structures.

The displayed algebra interpretation is a *logical predicate* interpretation, defined as follows.

$$\begin{aligned} -^D &: (A : \mathbf{Ty})\{X\} \rightarrow (X \rightarrow \mathbf{Set}) \rightarrow A^A X \rightarrow \mathbf{Set} \\ \iota^D \quad X^D \alpha &::= X^D \alpha \\ (\iota \rightarrow A)^D X^D \alpha &::= (x : X)(x^D : X^D x) \rightarrow A^D X^D (\alpha x) \end{aligned}$$

$$\begin{aligned} -^D &: (\Gamma : \mathbf{Con})\{X\} \rightarrow (X \rightarrow \mathbf{Set}) \rightarrow \Gamma^A X \rightarrow \mathbf{Set} \\ \Gamma^D X^D \gamma &::= \{A : \mathbf{Ty}\}(x : \mathbf{Var} \Gamma A) \rightarrow A^D X^D (\gamma x) \end{aligned}$$

$$\begin{aligned} -^D &: (t : \mathbf{Tm} \Gamma A) \rightarrow \Gamma^D X^D \gamma \rightarrow A^D X^D (t^A \gamma) \\ (\mathbf{var} x)^D \gamma^D &::= \gamma^D x \\ (\mathbf{app} t u)^D \gamma^D &::= t^D \gamma^D (u^A \gamma) (u^D \gamma^D) \end{aligned}$$

$$\begin{aligned} -^D &: (\sigma : \mathbf{Sub} \Gamma \Delta) \rightarrow \Gamma^D X^D \gamma \rightarrow \Delta^D X^D (\sigma^A \gamma) \\ \sigma^D \gamma^D x &::= (\sigma x)^D \gamma^D \end{aligned}$$

Analogously to before, everything depends on a predicate interpretation $X^D : X \rightarrow \mathbf{Set}$ for ι . For types, a predicate holds for a function if the function preserves predicates. The interpretation of terms is again a fundamental lemma, and we again have pointwise definitions for contexts and substitutions.

Definition 6 (displayed algebras).

$$\begin{aligned} \mathbf{DispAlg} &: \{\Gamma : \mathbf{Con}\} \rightarrow \mathbf{Alg} \Gamma \rightarrow \mathbf{Set}_1 \\ \mathbf{DispAlg} \{\Gamma\} (X, \gamma) &::= (X^D : X \rightarrow \mathbf{Set}) \times \Gamma^D X^D \gamma \end{aligned}$$

Example 4. We have the following computation.

$$\begin{aligned} \mathbf{DispAlg} \{\mathbf{NatSig}\} (X, \mathit{zero}, \mathit{suc}) &::= \\ & (X^D : X \rightarrow \mathbf{Set}) \\ & \times (\mathit{zero}^D : X^D \mathit{zero}) \\ & \times (\mathit{suc}^D : (n : X) \rightarrow X^D n \rightarrow X^D (\mathit{suc} n)) \end{aligned}$$

2.2.4 Sections

Sections of displayed algebras are “dependent” analogues of algebra morphisms, where the codomain is displayed over the domain.

$$\begin{aligned} -^S &: (A : \mathbf{Ty})\{X X^D\}(X^S : (x : X) \rightarrow X^D) \rightarrow (\alpha : A^A X) \rightarrow A^D X^D \alpha \rightarrow \mathbf{Set} \\ \iota^S & \quad X^S \alpha \alpha^D \equiv X^S \alpha = \alpha^D \\ (\iota \rightarrow A)^S X^S \alpha \alpha^D & \equiv (x : X) \rightarrow A^S X^S (\alpha x) (\alpha^D (X^S x)) \end{aligned}$$

$$\begin{aligned} \mathbf{Con}^S &: (\Gamma : \mathbf{Con})\{X X^D\}(X^S : (x : X) \rightarrow X^D) \rightarrow (\gamma : \Gamma^A X) \rightarrow \Gamma^D X^D \gamma \rightarrow \mathbf{Set} \\ \Gamma^S X^S \gamma_0 \gamma_1 & \equiv \{A : \mathbf{Ty}\}(x : \mathbf{Var} \Gamma A) \rightarrow A^S X^S (\gamma_0 x) (\gamma_1 x) \end{aligned}$$

$$\begin{aligned} -^S &: (t : \mathbf{Tm} \Gamma A) \rightarrow \Gamma^S X^S \gamma \gamma^D \rightarrow A^S X^S (t^A \gamma) (t^D \gamma^D) \\ (\mathbf{var} x)^S \gamma^S & \equiv \gamma^S x \\ (\mathbf{app} t u)^S \gamma^S & \equiv t^S \gamma^S (u^A \gamma) \end{aligned}$$

$$\begin{aligned} -^S &: (\sigma : \mathbf{Sub} \Gamma \Delta) \rightarrow \Gamma^S X^S \gamma \gamma^D \rightarrow \Delta^S X^S (\sigma^A \gamma) (\sigma^A \gamma^D) \\ \sigma^S \gamma^S x & = (\sigma x)^S \gamma^S \end{aligned}$$

Definition 7 (Displayed algebra sections (“sections” in short)).

$$\begin{aligned} \mathbf{Section} &: \{\Gamma : \mathbf{Con}\} \rightarrow (\gamma : \mathbf{Alg} \Gamma) \rightarrow \mathbf{DispAlg} \gamma \rightarrow \mathbf{Set} \\ \mathbf{Section} (X, \gamma) (X^D \gamma^D) & \equiv (X^S : (x : X) \rightarrow X^D x) \times \Gamma^S X^S \gamma \gamma^D \end{aligned}$$

Example 5. We have the following computation.

$$\begin{aligned} \mathbf{Section} \{\mathbf{NatSig}\} (X, \mathbf{zero}, \mathbf{suc}) (X^D, \mathbf{zero}^D, \mathbf{suc}^D) & \equiv \\ (X^S : (x : X) \rightarrow X^D x) & \\ \times (\mathbf{zero}^S : X^S \mathbf{zero} = \mathbf{zero}^D) & \\ \times (\mathbf{suc}^S : (n : X) \rightarrow X^S (\mathbf{suc} n) = \mathbf{suc}^D n (X^S n)) & \end{aligned}$$

Definition 8 (Induction). We define a predicate which holds if an algebra supports induction.

$$\begin{aligned} \mathbf{Inductive} &: \{\Gamma : \mathbf{Con}\} \rightarrow \mathbf{Alg} \Gamma \rightarrow \mathbf{Set}_1 \\ \mathbf{Inductive} \{\Gamma\} \gamma & \equiv (\gamma^D : \mathbf{DispAlg} \gamma) \rightarrow \mathbf{Section} \gamma \gamma^D \end{aligned}$$

We can observe that $\text{Inductive } \{\text{NatSig}\} (X, \text{zero}, \text{suc})$ computes exactly to the usual induction principle for natural numbers. The input DispAlg is a bundle of the induction motive and the methods, and the output Section contains the X^S eliminator function together with its β computation rules.

2.3 Term Algebras

In this section we show that if a type theory supports the inductive types comprising the theory of signatures, it also supports every inductive type which is described by the signatures.

Note that we specified Tm and Sub , but did not need either of them when specifying signatures, or when computing induction principles. That signatures do not depend on terms, is a property specific to simple signatures; this will not be the case in Chapter 4 when we move to more general signatures. However, terms and substitutions are already useful here in the construction of term algebras.

The idea is that terms in contexts comprise initial algebras. For example, $\text{Tm NatSig } \iota$ is the set of natural numbers (up to isomorphism). Informally, this is because the only way to construct terms is by applying the suc variable (given by var vz) finitely many times to the zero variable (given by var (vs vz)).

Definition 9 (Term algebras). Fix an $\Omega : \text{Con}$. We abbreviate $\text{Tm } \Omega \iota$ as T ; this will serve as the carrier set of the term algebra. We additionally define the following.

$$\begin{aligned} -^T &: (A : \text{Ty}) \rightarrow \text{Tm } \Omega A \rightarrow A^A \text{T} \\ \iota^T & \quad t : \equiv t \\ (\iota \rightarrow A)^T t &: \equiv \lambda u. A^T (\text{app } t u) \end{aligned}$$

$$\begin{aligned} -^T &: (\Gamma : \text{Con}) \rightarrow \text{Sub } \Omega \Gamma \rightarrow \Gamma^A \Gamma \text{T} \\ \Gamma^T \nu x &: \equiv A^T (\nu x) \end{aligned}$$

$$\begin{aligned} -^T &: (t : \text{Tm } \Gamma A)(\nu : \text{Sub } \Omega \Gamma) \rightarrow A^T (t[\nu]) = t^A (\Gamma^T \nu) \\ (\text{var } x)^T & \quad \nu \text{ holds by refl} \\ (\text{app } t u)^T & \quad \nu \text{ holds by } t^T \nu \text{ and } u^T \nu \end{aligned}$$

$$\begin{aligned}
& -^T : (\sigma : \mathbf{Sub} \Gamma \Delta)(\nu : \mathbf{Sub} \Omega \Gamma)\{A\}(x : \mathbf{Var} \Delta A) \\
& \quad \rightarrow \Delta^T (\sigma \circ \nu) x = \sigma^A (\Gamma^T \nu) x \\
& \sigma^T \nu x \equiv (\sigma x)^T \nu
\end{aligned}$$

Now we can define the term algebra for Ω itself:

$$\begin{aligned}
& \mathbf{TmAlg}_\Omega : \mathbf{Alg} \Omega \\
& \mathbf{TmAlg}_\Omega \equiv \Omega^T \Omega \mathbf{id}
\end{aligned}$$

In the interpretation for contexts, it is important that Ω is fixed, and we do induction on all Γ contexts such that there is a $\mathbf{Sub} \Omega \Gamma$. It would not work to try to compute term algebras by direct induction on contexts, because we need to refer to the same \mathbf{T} set in the interpretation of every type in a signature.

The interpretation of types embeds terms as A -algebras. At the base type ι , this embedding is simply the identity function, since $\iota^A \mathbf{T} \equiv \mathbf{T} \equiv \mathbf{Tm} \Omega \iota$. At function types we recursively proceed under a semantic λ . The interpretation of contexts is pointwise.

The interpretations of terms and substitutions are coherence properties, which relate the term algebra construction to term evaluation in the $-^A$ model. For terms, if we pick $\nu \equiv \mathbf{id}$, we get $A^T t = t^A \mathbf{TmAlg}_\Omega$. The left side embeds t in the term model via $-^T$, while the right hand side evaluates t in the term model.

A way to view the term algebra construction, is that we are working in a *slice model* over the fixed Ω , and every $\nu : \mathbf{Sub} \Omega \Gamma$ can be viewed as an internal Γ -algebra in this model. The term algebra construction demonstrates that every such internal algebra yields an external element of Γ^A . We will see in Section 4.3 that we can construct term algebras from *any* model of a ToS, not just the ToS syntax; but while term algebras constructed from ToS syntax are themselves initial algebras, in other cases they may not be initial.

2.3.1 Weak Initiality

We show that \mathbf{TmAlg}_Ω supports a recursion principle, i.e. it is weakly initial.

Definition 10 (Recursor construction). We assume $(X, \omega) : \mathbf{Alg} \Omega$; recall that $X : \mathbf{Set}$ and $\omega : \Omega^A X$. We define $\mathbf{R} : \mathbf{T} \rightarrow X$ as $\mathbf{R} t \equiv t^A \omega$. We additionally

define the following.

$$\begin{aligned}
& -^R : (A : \mathbf{Ty})(t : \mathbf{Tm} \, \Omega \, A) \rightarrow A^M \mathbf{R} (A^T t) (t^A \omega) \\
& \iota^R \quad t \equiv (\text{refl} : t^A \omega = t^A \omega) \\
& (\iota \rightarrow A)^R t \equiv \lambda u. A^R (\text{app } t u) \\
\\
& -^R : (\Gamma : \mathbf{Con})(\nu : \mathbf{Sub} \, \Omega \, \Gamma) \rightarrow \Gamma^M \mathbf{R} (\Gamma^T \nu) (\nu^A \omega) \\
& \Gamma^R \nu x \equiv A^R (\nu x)
\end{aligned}$$

We define the recursors for Ω as

$$\begin{aligned}
& \text{Rec}_\Omega : (alg : \mathbf{Alg} \, \Omega) \rightarrow \mathbf{Mor} \, \mathbf{TmAlg}_\Omega \, alg \\
& \text{Rec}_\Omega (X, \omega) \equiv (\mathbf{R}, \Omega^R \Omega \text{id})
\end{aligned}$$

In short, the way we get recursion is by evaluating terms in arbitrary (X, ω) algebras using $-^A$. The $-^R$ operation for types and contexts confirms that \mathbf{R} preserves structure appropriately, so that \mathbf{R} indeed yields algebra morphisms.

We skip interpreting terms and substitutions by $-^R$. It is necessary to do so with more general signatures, but not in the current chapter.

2.3.2 Induction

We take the idea of the previous section a bit further. We have seen that recursion for term algebras is given by evaluation in the “standard” model $-^A$. Now, we show that induction for term algebras corresponds to evaluation in the logical predicate model $-^D$.

Definition 11 (Eliminator construction). We assume $(X^D, \omega^D) : \mathbf{DispAlg} \, \mathbf{TmAlg}_\Omega$. Recall that $X^D : \mathbf{T} \rightarrow \mathbf{Set}$ and $\omega^D : \Omega^D X^D (\Omega^T \Omega \text{id})$. Like before, we first interpret the underlying set:

$$\begin{aligned}
& \mathbf{E} : (t : \mathbf{T}) \rightarrow X^D t \\
& \mathbf{E} t \equiv t^D \omega^D
\end{aligned}$$

However, this definition is not immediately well-typed, since $t^D \omega^D$ has type $X^D (t^A (\Omega^T \Omega \text{id}))$, so we have to show that $t^A (\Omega^T \Omega \text{id}) = t$. This equation says that nothing happens if we evaluate a term with type ι in the term model. We

get it from the $-^T$ interpretation of terms: $t^T \text{id} : t[\text{id}] = t^A (\Omega^T \Omega \text{id})$, and we also know that $t[\text{id}] = t$. We interpret types and contexts as well:

$$\begin{aligned} -^E : (A : \mathbf{Ty})(t : \mathbf{Tm} \Omega A) &\rightarrow A^S \mathbf{E} (t^A (\Omega^T \Omega \text{id})) (t^D \omega^D) \\ \iota^E \quad t : (t^A (\Omega^T \Omega \text{id}))^D \omega^D &= t^D \omega^D \\ (\iota \rightarrow A)^E t &\equiv \lambda u. A^E (\text{app } t u) \end{aligned}$$

$$\begin{aligned} -^E : (\Gamma : \mathbf{Con})(\nu : \mathbf{Sub} \Omega \Gamma) &\rightarrow \Gamma^S \mathbf{E} (\nu^A (\Omega^T \Omega \text{id})) (\nu^D \omega^D) \\ \Gamma^E \nu x &\equiv A^E (\nu x) \end{aligned}$$

In ι^E we use the same equation as in the definition of \mathbf{E} . In $(\iota \rightarrow A)^E$ the definition is well-typed because of the same equation, but instantiated for the abstracted u term this time. All of this amounts to some additional path induction and transport fiddling in the (intensional) Agda formalization. We get induction for Ω as below.

$$\begin{aligned} \text{Ind}_\Omega : (alg : \mathbf{DispAlg} \mathbf{TmAlg}_\Omega) &\rightarrow \mathbf{Section} \mathbf{TmAlg}_\Omega \text{ alg} \\ \text{Ind}_\Omega (X^D, \omega^D) &\equiv (E, \Omega^E \Omega \text{id}) \end{aligned}$$

2.4 Discussion

2.4.1 Comparison to F-algebras

A well-known alternative way for treating inductive types is to use certain cocontinuous endofunctors as a more semantic notion of signatures.

For example, single-sorted inductive types can be presented as endofunctors which preserve colimits of some ordinal-indexed chains. For instance, if we have an κ -cocontinuous $F : \mathbb{C} \rightarrow \mathbb{C}$, then algebras are given as $(X : |\mathbb{C}|) \times (\mathbb{C}(F X, X))$, morphisms as commuting squares, and Adámek's theorem [AK79] establishes the existence of initial algebras.

An advantage of this approach is that we can describe different classes of signatures by choosing different \mathbb{C} categories:

- If \mathbb{C} is **Set**, we get simple inductive types.
- If \mathbb{C} is **Set** ^{I} for some set I , we get indexed inductive types.

- If \mathbb{C} is \mathbf{Set}/I , we get inductive-recursive types.

Another advantage of F -algebras is that signatures are a fairly semantic notion: they make sense even if we have no syntactic presentation at hand. That said, often we do need syntactic signatures, for use in proof assistants, or just to have a convenient notation for a class of cocontinuous functors.

An elegant way of carving out a large class of such functors is to consider polynomials as signatures. For example, when working in \mathbf{Set} , a signature is an element of $(S : \mathbf{Set}) \times (P : S \rightarrow \mathbf{Set})$, and (S, P) is interpreted as a functor as $X \mapsto (s : S) \times (P s \rightarrow X)$. The initial algebra is the W-type specified by S shapes and P positions. This yields infinitary inductive types as well.

However, it is not known how to get *inductive-inductive* signatures by picking the right \mathbb{C} category and a functor. In an inductive-inductive signature, there may be multiple sorts, which can be indexed over previously declared sorts. For example, in the signature for categories we have $\mathbf{Obj} : \mathbf{Set}$ and $\mathbf{Mor} : \mathbf{Obj} \rightarrow \mathbf{Obj} \rightarrow \mathbf{Set}$, indexed twice over \mathbf{Obj} . Some extensions are required to the idea of F -algebras:

- For inductive-inductive definitions with two sorts, Forsberg gives a specification with two functors, and a considerably more complex notion of algebras, involving dialgebras [NF13].
- For an arbitrary number of sorts, Altenkirch et al. [ACD⁺18] use a “list” of functors, specified mutually with categories of algebras: each functor has as domain the semantic category of all previous sorts.

The functors-as-signatures approach gets significantly less convenient as we consider more general specifications. The approach of this thesis is to skip the middle ground between syntactic signatures and semantic categories of algebras: we treat syntactic signatures as a key component, and give direct semantic interpretation for them. Although we lose the semantic nature of F -algebras, our approach scales extremely well, all the way up to infinitary quotient-inductive-inductive types in Chapter 5, and to some extent to higher inductive-inductive types as well in Chapter 7.

If we look back at $-^A : \mathbf{Con} \rightarrow \mathbf{Set} \rightarrow \mathbf{Set}$, we may note that Γ^A yields a functor, in fact the same functor (up to isomorphism) that we would get from an F -algebra presentation. However, this is a coincidence in the single-sorted

case. With the F -algebra presentation, we can view $(X : |\mathbb{C}|) \times (\mathbb{C}(F X, X))$ as specifying the category of algebras as the total category of a displayed category (by viewing the Σ -type here as taking total categories; a Σ in **Cat**). In our approach, we aim to get the displayed categories directly, without talking about functors.

2.4.2 Generic Programming

Let's consider now our signatures and term algebras in the context of generic programming. This is largely future work, and we don't elaborate it much. But we can draw some preliminary conclusions and make some comparisons.

If a language can formalize inductive signatures and their semantics, that can be viewed as an implementation of generic programming over the described types. Compared to a purely mathematical motivation for this formalization, the requirements for practical generic programming are a bit more stringent.

- *Encoding overhead*: there should be an acceptable overhead in program size and performance when using generic representations. Size blowup can be an issue when writing proofs as well, when types and expressions become too large to mentally parse.
- *Strictness properties*: generic representations should compute as much as possible, ideally in exactly the same way as their non-generic counterparts.

Fixpoints of functors. There is a sizable literature of using fixpoints of functors in generic programming, mainly in Haskell [Swi08, BH11, YHLJ09] and Agda [LM11, AAC⁺20]. We give a minimal example below for an Agda-like implementation.

We have an inductive syntax for some strictly positive functors, covering essentially the same signatures as **Con**.

```

Sig    : Set
Id     : Sig
KT     : Sig
-⊗-    : Sig → Sig → Sig
-⊕-    : Sig → Sig → Sig

```

`Id` codes the identity functor, and `K \top` codes the functor which is constantly \top . $-\otimes-$ and $-\oplus-$ are pointwise products and coproducts respectively. So we have the evident interpretation functions:

$$\begin{aligned} \llbracket - \rrbracket &: \text{Sig} \rightarrow \text{Set} \rightarrow \text{Set} \\ \text{map} &: (F : \text{Sig}) \rightarrow (X_0 \rightarrow X_1) \rightarrow \llbracket F \rrbracket X_0 \rightarrow \llbracket F \rrbracket X_1 \end{aligned}$$

In Haskell and Agda, the easiest way to get initial algebras is to directly define the inductive fixpoint for each assumed $F : \text{Sig}$:

$$\begin{aligned} \text{Fix}_F &: \text{Set} \\ \text{con}_F &: \llbracket F \rrbracket \text{Fix}_F \rightarrow \text{Fix}_F \end{aligned}$$

In Haskell, this definition is valid for arbitrary *semantic* F functor, because there is no termination checking or positivity checking. In Agda, the above definition is valid because the positivity checker is willing to look inside the definition of $\llbracket - \rrbracket$. However, Coq and Lean reject this definition. Next we establish weak initiality, by defining the recursor:

$$\begin{aligned} \text{rec} &: (\llbracket F \rrbracket X \rightarrow X) \rightarrow \text{Fix}_F \rightarrow X \\ \text{rec } f \text{ (con } x) &::= f (\text{map } (\text{rec } f) x) \end{aligned}$$

This is again fine in Haskell, but it unfortunately does not pass Agda's termination checker. The most conservative solution is to inline the recursive call into `map`, so that the definition becomes transparent to termination checking.

$$\begin{aligned} \text{rec} &: (\llbracket F \rrbracket X \rightarrow X) \rightarrow \text{Fix}_F \rightarrow X \\ \text{rec } f \text{ (con } x) &::= f (\text{maprec } f x) \end{aligned}$$

$$\begin{aligned} \text{maprec} &: \{G\} \rightarrow (\llbracket F \rrbracket X \rightarrow X) \rightarrow \llbracket G \rrbracket \text{Fix}_F \rightarrow \llbracket G \rrbracket X \\ \text{maprec } \{\text{Id}\} \quad f x &::= \text{rec } f x \\ \text{maprec } \{\text{K}\top\} \quad f x &::= x \\ \text{maprec } \{G \otimes H\} f (x, y) &::= (\text{maprec } f x, \text{maprec } f y) \\ \text{maprec } \{G \oplus H\} f (\text{inj}_1 x) &::= \text{maprec } f x \\ \text{maprec } \{G \oplus H\} f (\text{inj}_2 x) &::= \text{maprec } f x \end{aligned}$$

Alternatively, we might use *sized types* [AVW17], as in [AAC⁺20]. The drawback is dependence on an additional language feature which is only supported in Agda, and which has had several soundness issues since its introduction [?].

There is yet another possible approach: defining initial algebras as sequential colimits, using Adámek’s theorem. This approach was taken by Ahrens, Matthes and Mörtberg in [AMM19]. However, the encoding overhead is excessive, and it is practically unusable for generic programming. Another drawback is that defining colimits requires quotient types, which are often not available natively.

W-types. Given a polynomial $(S, P) : (S : \mathbf{Set}) \times (S \rightarrow \mathbf{Set})$, the corresponding W-type is inductively specified as below:

$$\begin{aligned} \mathbf{W}_{S,P} &: \mathbf{Set} \\ \sup &: (s : S) \rightarrow (P\ s \rightarrow \mathbf{W}_{S,P}) \rightarrow \mathbf{W}_{S,P} \end{aligned}$$

If we assume \top , \perp , **Bool**, Π , Σ , W-types, universes and an intensional identity type, a large class of inductive types can be derived, including infinitary and finitary indexed inductive families; this was shown by Hugunin [Hug21]. The encoding also yields definitional β -rules for recursion and elimination. However, there is also significant encoding overhead here.

- First, there is a translation from more convenient signatures to (S, P) polynomials.
- Then, we take the $\mathbf{W}_{S,P}$ type, but we need to additionally restrict it to the *canonical* elements by a predicate, as in $(x : \mathbf{W}_{S,P}) \times \mathbf{canonical}\ x$. This is required because the only way to represent inductive branching is by functions, but functions sometimes contain too many elements up to definitional equality. For example, $\perp \rightarrow A$ has infinitely many definitionally distinct inhabitants.

There is also a performance overhead imposed by the mandatory higher-order constructors. W-types are a great way to have a small basis in a formal setting, both in intensional and extensional type theories, but they are a bit too heavy for practical purposes.

Term algebras. The main advantage of the term algebras described in this chapter is that they can be defined using plain inductive families, which are available in every major dependently typed language. Even in GHC Haskell, the support for generalized ADTs is sufficient for implementing term algebras. In contrast, as we have seen, the direct inductive definition of fixpoints is only

possible in Agda, while the colimit definition requires quotients and is rather heavyweight. Even in Agda, the drawback of the direct fixpoint definition is that many generic operations only pass termination checking after manual inlining. With term algebras, generic operations can be defined by induction on $\text{Tm } \Gamma A$ in an obviously terminating way.

For practical usage it makes sense to slightly modify terms. We switch to a *spine neutral* definition. We mutually inductively define **Spine** and **Tm**:

$$\begin{aligned}
\text{Spine} &: (\Gamma : \text{Con}) \rightarrow \text{Ty} \rightarrow \text{Ty} \rightarrow \text{Set} \\
\epsilon &: \{A\} \rightarrow \text{Spine } \Gamma A A \\
-, - &: \{B C\} \rightarrow \text{Tm } \Gamma \iota \rightarrow \text{Spine } \Gamma B C \rightarrow \text{Spine } \Gamma (\iota \rightarrow B) C \\
\text{Tm} &: \text{Con} \rightarrow \text{Ty} \rightarrow \text{Set} \\
-\$- &: \{A B\} \rightarrow \text{Var } \Gamma A \rightarrow \text{Spine } \Gamma A B \rightarrow \text{Tm } \Gamma B
\end{aligned}$$

With this representation, a term is always a variable applied to a list of arguments. This can be useful, because pattern matching implementations in metalanguages (e.g. Agda or Idris) are more likely to know about which constructors are possible in patterns. Using this, we give here an ad-hoc definition of natural numbers in pseudo-Agda.

$$\begin{aligned}
\text{Nat} &: \text{Set} & \text{zero} &:: \text{vz } \$ \epsilon \\
\text{Nat} &:: \text{Tm } (\bullet \triangleright \iota \rightarrow \iota \triangleright \iota) \iota & \text{suc } n &:: \text{vs vz } \$ (n, \epsilon)
\end{aligned}$$

$$\begin{aligned}
\text{NatElim} &: (P : \text{Nat} \rightarrow \text{Set}) \rightarrow P \text{ zero} \rightarrow ((n : \text{Nat}) \rightarrow P n \rightarrow P (\text{suc } n)) \\
&\rightarrow (n : \text{Nat}) \rightarrow P n \\
\text{NatElim } P \text{ pz ps } (\text{vz } \$ \epsilon) &:: \text{pz} \\
\text{NatElim } P \text{ pz ps } (\text{vs vz } \$ (n, \epsilon)) &:: \text{ps } n (\text{NatElim } P \text{ pz ps } n)
\end{aligned}$$

The actual Agda definition can be found in the supplementary formalization, and it is pretty much the same as above. We recover the exact same behavior with respect to pattern matching as with native inductive definitions.

Rec_Ω and Ind_Ω can be adapted to spine neutral terms with minor adjustments. But what about the β -rules which they produce as part of their output, are they definitional (i.e. proven by **refl**)? In this chapter we do not have a rigorous way of reasoning about definitional equalities; in the next chapter we develop such

reasoning and show that Rec_Ω enjoys definitional β -rules (with or without the spine neutral definition).

However, Ind_Ω only supports propositional β -rules. The issue is the extra transporting in the definition of $\mathbf{E} : (t : \mathbf{Tm} \, \Omega \, \iota) \rightarrow X^D \, t$: transports don't strictly commute with constructors. This appears to be a point of advantage for the direct fixpoint definition in Agda, as it allows generic elimination with strict computation rules [?].

The term algebra presentation can be easily extended to indexed families. In that case, signatures and terms are still definable with basic inductive families, without requiring quotients or complicated encodings; see Kaposi and von Raumer [KvR20].

A related existing work is *sum-of-products* generics by De Vries and Löh [dVL14]. There, signatures for functors are in a normal form: we cannot freely take products and coproducts, instead a signature looks very much like a \mathbf{Con} in this chapter (except in an indexed form). The authors observe that several generic programming patterns are easier to express with normalized signatures. However, they still use explicit fixpoints as the way to get initial algebras.

CHAPTER 3

Semantics in Two-Level Type Theory

In this chapter we describe how two-level type theory is used as a metatheoretic setting in the rest of this thesis. First, we provide motivation and overview. Second, we describe models of type theories in general, and models of two-level type theories as extensions. Third, we describe presheaf models of two-level type theories. Finally, we generalize the semantics and the term algebra construction from Chapter 2 in two-level type theory, as a way to illustrate the applications.

3.1 Motivation

We note two shortcomings of the semantics presented in the previous chapter.

First, we were not able to reason about definitional equalities, only propositional ones. We have a formalization of signatures and semantics in intensional Agda, where the two notions differ¹, but only propositional equality is subject to internal reasoning. For instance, we would like to show that term algebras support recursion with strict β -rules, and for this we need to reason about strict equality.

Second, the semantics that we provided was not as general as it could be. We used the internal `Set` universe to specify algebras, but algebras make sense in many different categories. A crude way to generalize semantics is to simply say that our formalization, which was carried out in the syntax (i.e. initial model) of some intensional type theory, can be interpreted in any model of the type theory. However, this is wasteful: for simple inductive signatures, it is enough to assume a category with finite products as semantic setting. We don't need all the extra

¹As opposed to in extensional type theory, where they're the same.

baggage that comes with a model of a type theory.

Notation 3. We use \bullet for the terminal object in a \mathbb{C} category, with $\epsilon : \mathbb{C}(A, \bullet)$ for the unique morphism. For products, we use $- \otimes -$ with $(-, -) : \mathbb{C}(A, B) \rightarrow \mathbb{C}(A, C) \rightarrow \mathbb{C}(A, B \otimes C)$ and \mathbf{p} and \mathbf{q} for first and second projections respectively.

Example 6. Assuming \mathbb{C} has finite products, natural number algebras and binary tree algebras are specified as follows.

$$\begin{aligned} \mathbf{Alg}_{\mathbf{NatSig}} &::= (X : |\mathbb{C}|) \times \mathbb{C}(\bullet, X) \times \mathbb{C}(X, X) \\ \mathbf{Alg}_{\mathbf{TreeSig}} &::= (X : |\mathbb{C}|) \times \mathbb{C}(\bullet, X) \times \mathbb{C}(X \otimes X, X) \end{aligned}$$

Here, $\mathbf{Alg}_{\mathbf{NatSig}}$ and $\mathbf{Alg}_{\mathbf{TreeSig}}$ are both sets in some metatheory, and the \times in the definitions refer to the metatheoretic Σ . Algebras can be viewed as diagrams which preserve finite products, and algebra morphisms are natural transformations.

How should we adjust \mathbf{Alg} to compute algebras in \mathbb{C} , and \mathbf{Mor} to compute their morphisms? While it is possible to do this in a direct fashion, working directly with objects and morphisms of \mathbb{C} is rather unwieldy. \mathbb{C} is missing many convenience features of type theories.

- There are no variables or binders. We are forced to work in a point-free style or chase diagrams; both become difficult to write and read after a certain point of complexity.
- There no functions, universes or inductive types.
- Substitution (with weakening as a special case) has to be handled explicitly and manually. Substitutions are certain morphisms, while “terms” are also morphisms, and we have to use composition to substitute terms. In contrast, if we are working internally in a type theory, terms and substitutions are distinct, and we only have to explicitly deal with terms, and substitutions are automated and implicit.

The above overlaps with motivations for working in *internal languages* [nc21] of structured categories: they aid calculation and compact formalization by hiding bureaucratic structural details.

A finite product category \mathbb{C} does not have much of an internal language, it is too bare-bones. But we can work instead in the internal language of $\hat{\mathbb{C}}$, the

category of presheaves over \mathbb{C} . This allows faithful reasoning about \mathbb{C} , while also including all convenience features of extensional type theory.

Two-level type theories [ACKS19], or 2LTT in short, are type theories such that they have “standard” interpretations in presheaf categories. A 2LTT has an inner layer, where types and terms arise by embedding \mathbb{C} in $\hat{\mathbb{C}}$, and an outer layer, where constructions are inherited from $\hat{\mathbb{C}}$. The exact details of the syntax may vary depending on what structures \mathbb{C} supports, and which type formers we assume in the outer layer. Although it is possible to add assumptions to a 2LTT which preclude standard presheaf semantics [ACKS19, Section 2.4.], we stick to basic 2LTT in this thesis. By using 2LTT, we are able to use a type-theoretic syntax in the rest of the thesis which differs only modestly from the style of definitions that we have seen so far.

From a programming perspective, basic 2LTT provides a convenient syntax for writing metaprograms. This can be viewed as *two-stage compilation*: if we have a 2LTT program with an inner type, we can run it, and it returns another program, which lives purely in the inner theory.

3.2 Models of Type Theories

Before elaborating on 2LTT-specific features, we review models of type theories in general. Variants of 2LTT will be obtained by adding extra features on the top of more conventional TTs.

It is also worth to take a more general look at models at this point, because the notions presented in this subsection (categories with families, type formers) will be reused several times in this thesis, when specifying theories of signatures.

3.2.1 The Algebraic View

We take an algebraic view [?] of models and syntaxes of type theories throughout this thesis. Models of type theories are algebraic structures: they are categories with certain extra structure. The syntax of a type theory is understood to be its initial model. In initial models, the underlying category is the category of typing contexts and parallel substitutions, while the extra structure corresponds to type and term formers, and equations quotient the syntax by definitional equality.

Type theories can be described with quotient inductive-inductive (QII) signatures, and their initial models are quotient inductive-inductive types (QIITs).

Hence, 2LTT is also a QII theory. We will first talk about QIITs in Chapter 4. Until then, we shall make do with an informal understanding of categorical semantics for type theories, without using anything in particular from the metatheory of QIITs. There is some annoying circularity here, that we talk about QIITs in this thesis, but we employ QIITs when talking about them. However, this is only an annoyance in exposition and not a fundamental issue: Chapter 6 describes how to eliminate circularity by a form of bootstrapping.

The algebraic view lets us dispense with all kinds of “raw” syntactic objects. We only ever talk about well-typed and well-formed objects, moreover, every construction must respect definitional equalities. For terms in the algebraic syntax, definitional equality coincides with metatheoretic equality. This mirrors equality of morphisms in 1-category theory, where we usually reuse metatheoretic equality.

In the following we specify notions of models for type theories. We split this in two parts: categories with families and type formers.

3.2.2 Categories With Families

Definition 12. A **category with family** (cwf) [Dyb95] is a way to specify the basic structural rules for contexts, substitutions, types and terms. It yields a dependently typed explicit substitution calculus [?]. A cwf consists of the following.

- A category with a terminal object. We denote the set of objects as $\mathbf{Con} : \mathbf{Set}$ and use capital Greek letters starting from Γ to refer to objects. The set of morphisms is $\mathbf{Sub} : \mathbf{Con} \rightarrow \mathbf{Con} \rightarrow \mathbf{Set}$, and we use σ, δ and so on to refer to morphisms. We write id for the identity morphism and $- \circ -$ for composition. The terminal object is \bullet with unique morphism $\epsilon : \mathbf{Sub} \Gamma \bullet$. In initial models (that is, syntaxes) of type theories, objects correspond to typing contexts, morphisms to parallel substitutions and the terminal object to the empty context; this informs the naming scheme.
- A *family structure*, containing $\mathbf{Tm} : (\Gamma : \mathbf{Con}) \rightarrow \mathbf{Tm} \Gamma \rightarrow \mathbf{Set}$. We use A, B, C to refer to types and t, u, v to refer to terms. \mathbf{Tm} is a presheaf over the category of contexts and \mathbf{Tm} is a displayed presheaf

over \mathbf{Ty} . This means that types and terms can be substituted:

$$\begin{aligned} -[-] : \mathbf{Ty} \Delta \rightarrow \mathbf{Sub} \Gamma \Delta \rightarrow \mathbf{Ty} \Gamma \\ -[-] : \mathbf{Tm} \Delta A \rightarrow (\sigma : \mathbf{Sub} \Gamma \Delta) \rightarrow \mathbf{Tm} \Gamma (A[\sigma]) \end{aligned}$$

Substitution is functorial: we have $A[\text{id}] = A$ and $A[\sigma \circ \delta] = A[\sigma][\delta]$, and likewise for terms.

A family structure is additionally equipped with *context comprehension* which consists of a context extension operation $- \triangleright - : (\Gamma : \mathbf{Con}) \rightarrow \mathbf{Ty} \Gamma \rightarrow \mathbf{Con}$ together with an isomorphism $\mathbf{Sub} \Gamma (\Delta \triangleright A) \simeq ((\sigma : \mathbf{Sub} \Gamma \Delta) \times \mathbf{Tm} \Gamma (A[\sigma]))$ which is natural in Γ .

The following notions are derivable from the comprehension structure:

- By going right-to-left along the isomorphism, we recover *substitution extension* $-, - : (\sigma : \mathbf{Sub} \Gamma \Delta) \rightarrow \mathbf{Tm} \Gamma (A[\sigma]) \rightarrow \mathbf{Sub} \Gamma (\Delta \triangleright A)$. This means that starting from ϵ or the identity substitution id , we can iterate $-, -$ to build substitutions as lists of terms.
- By going left-to-right, and starting from $\text{id} : \mathbf{Sub} (\Gamma \triangleright A) (\Gamma \triangleright A)$, we recover the *weakening substitution* $\mathbf{p} : \mathbf{Sub} (\Gamma \triangleright A) \Gamma$ and the *zero variable* $\mathbf{q} : \mathbf{Tm} (\Gamma \triangleright A) (A[\mathbf{p}])$.
- By weakening \mathbf{q} , we recover a notion of variables as De Bruijn indices. In general, the n -th De Bruijn index is defined as $\mathbf{q}[\mathbf{p}^n]$, where \mathbf{p}^n denotes n -fold composition.

Comprehension can be characterized either by taking $-, -, \mathbf{p}$ and \mathbf{q} as primitive, or the natural isomorphism. The two are equivalent, and we may switch between them, depending on which is more convenient.

There are other ways for presenting the basic categorical structure of models, which are nonetheless equivalent to cwfs, including natural models [Awo18] and categories with attributes [Car78]. We use the cwf presentation for its immediately algebraic character and closeness to conventional explicit substitution syntax.

Notation 4. As De Bruijn indices are hard to read, we will mostly use nameful notation for binders. For example, assuming $\mathbf{Nat} : \{\Gamma : \mathbf{Con}\} \rightarrow \mathbf{Ty} \Gamma$ and $\text{ld} : \{\Gamma : \mathbf{Con}\}(A : \mathbf{Ty} \Gamma) \rightarrow \mathbf{Tm} \Gamma A \rightarrow \mathbf{Tm} \Gamma A \rightarrow \mathbf{Ty} \Gamma$, we may write $\bullet \triangleright n : \mathbf{Nat} \triangleright$

$p : \text{ld Nat } n \ n$ for a typing context, instead of using numbered variables or cwf combinators as in $\bullet \triangleright \text{Nat} \triangleright \text{ld Nat } q \ q$.

Notation 5. In the following, we will denote family structures by (Ty, Tm) pairs and overload context extension $- \triangleright -$ for different families.

Definition 13. The following derivable operations are commonly used.

- *Single substitution* can be derived from parallel substitution as follows. Assume $t : \text{Tm } (\Gamma \triangleright A) \ B$, and $u : \text{Tm } \Gamma \ A$. t is a term which may depend on the last variable in the context, which has A type. We can substitute that variable with the u term as $t[\text{id}, u] : \text{Tm } \Gamma \ (A[\text{id}, u])$. Note that term substitution causes the type to be substituted as well. $(\text{id}, u) : \text{Sub } \Gamma \ (\Gamma \triangleright A)$ is well-typed because $u : \text{Tm } \Gamma \ A$ hence also $u : \text{Tm } \Gamma \ (A[\text{id}])$.
- We can *lift substitutions* over binders as follows. Assuming $\sigma : \text{Sub } \Gamma \ \Delta$ and $A : \text{Ty } \Delta$, we construct a lifting of σ which maps an additional A -variable to itself: $(\sigma \circ \mathbf{p}, \mathbf{q}) : \text{Sub } (\Gamma \triangleright A[\sigma]) \ (\Delta \triangleright A)$. Let us see why this is well-typed. We have $\mathbf{p} : \text{Sub } (\Gamma \triangleright A[\sigma]) \ \Gamma$ and $\sigma : \text{Sub } \Gamma \ \Delta$, so $\sigma \circ \mathbf{p} : \text{Sub } (\Gamma \triangleright A[\sigma]) \ \Delta$. Also, $\mathbf{q} : \text{Tm } (\Gamma \triangleright A[\sigma]) \ (A[\sigma][\mathbf{p}])$, hence $\mathbf{q} : \text{Tm } (\Gamma \triangleright A[\sigma]) \ (A[\sigma \circ \mathbf{p}])$, thus $(\sigma \circ \mathbf{p}, \mathbf{q})$ typechecks.

Notation 6. As a nameful notation for substitutions, we may write $t[x \mapsto u]$, for a single substitution, or $t[x \mapsto u_1, y \mapsto u_2]$ and so on.

In nameful notation we leave all weakening implicit, including substitution lifting. Formally, if we have $t : \text{Tm } \Gamma \ A$, we can only mention t in Γ . If we need to mention it in $\Gamma \triangleright B$, we need to use $t[\mathbf{p}]$ instead. In the nameful notation, $t : \text{Tm } (\Gamma \triangleright x : B) \ A$ may be used.²

3.2.3 Type formers

A family structure in a cwf may be closed under certain type formers, such as functions, Σ -types, universes or inductive types. We give some examples here for their specification. First, we look at common negative type formers, which can be specified using isomorphisms. Then, we consider positive type formers, and finally universes.

²Moreover, when working in the internal syntax of a theory, we just write Agda-like type-theoretic notation, without noting contexts and substitutions in any way.

Negative types

Definition 14. A $(\mathsf{Ty}, \mathsf{Tm})$ family supports **Π -types** if it supports the following.

$$\begin{aligned}
\Pi & : (A : \mathsf{Ty} \Gamma) \rightarrow \mathsf{Ty} (\Gamma \triangleright A) \rightarrow \mathsf{Ty} \Gamma \\
\Pi[] & : (\Pi A B)[\sigma] = \Pi (A[\sigma]) (B[\sigma \circ \mathbf{p}, \mathbf{q}]) \\
\mathsf{app} & : \mathsf{Tm} \Gamma (\Pi A B) \rightarrow \mathsf{Tm} (\Gamma \triangleright A) B \\
\mathsf{lam} & : \mathsf{Tm} (\Gamma \triangleright A) B \rightarrow \mathsf{Tm} \Gamma (\Pi A B) \\
\Pi\beta & : \mathsf{app} (\mathsf{lam} t) = t \\
\Pi\eta & : \mathsf{lam} (\mathsf{app} t) = t \\
\mathsf{lam}[] & : (\mathsf{lam} t)[\sigma] = \mathsf{lam} (\sigma \circ \mathbf{p}, \mathbf{q})
\end{aligned}$$

Here, Π is the type formation rule. $\Pi[]$ is the type substitution rule, expressing that substituting Π proceeds structurally on constituent types. Note $B[\sigma \circ \mathbf{p}, \mathbf{q}]$, where we lift σ over the additional binder.

The rest of the rules specify a natural isomorphism $\mathsf{Tm} \Gamma (\Pi A B) \simeq \mathsf{Tm} (\Gamma \triangleright A) B$. We only need a substitution rule (i.e. a naturality rule) for one direction of the isomorphism, since the naturality of the other map is derivable.

This way of specifying Π -types is very convenient if we have explicit substitutions. The usual “pointful” specification is equivalent to this. For example, we have the following derivation of pointful application:

$$\begin{aligned}
\mathsf{app}' & : \mathsf{Tm} \Gamma (\Pi A B) \rightarrow (u : \mathsf{Tm} \Gamma A) \rightarrow \mathsf{Tm} \Gamma (B[\mathbf{id}, u]) \\
\mathsf{app}' t u & \equiv (\mathsf{app} t)[\mathbf{id}, u]
\end{aligned}$$

Remark on naturality. The above specification for Π can be written more compactly if we assume that everything is natural with respect to substitution.

$$\begin{aligned}
\Pi & : (A : \mathsf{Ty} \Gamma) \rightarrow \mathsf{Ty} (\Gamma \triangleright A) \rightarrow \mathsf{Ty} \Gamma \\
(\mathsf{app}, \mathsf{lam}) & : \mathsf{Tm} \Gamma (\Pi A B) \simeq \mathsf{Tm} (\Gamma \triangleright A) B
\end{aligned}$$

This is a reasonable assumption; in the rest of the thesis we only ever define structures on cwf's which are natural in this way.

Notation 7. From now on, when specifying type formers in family structures, we assume that everything is natural, and thus omit substitution equations.

There are ways to make this idea more precise, and take it a step further by working in languages where only natural constructions are possible. The term

higher-order abstract syntax (HOAS) is sometimes used for this style. It lets us also omit contexts, so we would only need to write

$$\begin{aligned} \Pi & : (A : \mathbf{Ty}) \rightarrow (\mathbf{Tm} A \rightarrow \mathbf{Ty}) \rightarrow \mathbf{Ty} \\ (\mathbf{app}, \mathbf{lam}) & : \mathbf{Tm} (\Pi A B) \simeq ((a : \mathbf{Tm} A) \rightarrow \mathbf{Tm} (B a)) \end{aligned}$$

Recently several promising works emerged in this area [?]. Although this technology is likely to be the preferred future direction in the metatheory of type theories, this thesis does not make use of it. The field is rather fresh, with several different approaches and limited amount of pedagogical exposition, and the new techniques would also raise the level of abstraction in this thesis, all contributing to making it less accessible. One more reason, and perhaps the most important one, for skipping higher-order abstract syntax, is that this author has not yet acquired the proficiency to comfortably use the new techniques.

Definition 15. A family structure supports **constant families** if we have the following.

$$\begin{aligned} \mathbf{K} & : \mathbf{Con} \rightarrow \{\Gamma : \mathbf{Con}\} \rightarrow \mathbf{Ty} \Gamma \\ (\mathbf{app}_\mathbf{K}, \mathbf{lam}_\mathbf{K}) & : \mathbf{Tm} \Gamma (\mathbf{K} \Delta) \simeq \mathbf{Sub} \Gamma \Delta \end{aligned}$$

Constant families express that every context can be viewed as a non-dependent type in any context. Having constant families is equivalent to the *democracy* property for a cwf [CD14, NF13]. Constant families are convenient when building models, because they let us model non-dependent types as semantic contexts, which are often simpler structures than semantic types. From a programming perspective, constant families specify closed record types, where $\mathbf{K} \Delta$ has Δ -many fields.

Definition 16. A family structure supports **Σ -types** if we have

$$\begin{aligned} \Sigma & : (A : \mathbf{Ty} \Gamma) \rightarrow \mathbf{Ty} (\Gamma \triangleright A) \rightarrow \mathbf{Ty} \Gamma \\ (\mathbf{proj}, (-, -)) & : \mathbf{Tm} \Gamma (\Sigma A B) \simeq ((t : \mathbf{Tm} \Gamma A) \times \mathbf{Tm} \Gamma (B[\mathbf{id}, t])) \end{aligned}$$

We use the shorter specification above, where everything is assumed to be natural. We may write \mathbf{proj}_1 and \mathbf{proj}_2 for composing the metatheoretic first and second projections with \mathbf{proj} .

Definition 17. A family structure supports the **unit type** if we have $\top : \mathbf{Ty} \Gamma$ such that $\mathbf{Tm} \Gamma \top \simeq \top$, where the \top on the right is the metatheoretic unit type, and we overload \top for the internal unit type. From this, we get the internal $\mathbf{tt} : \mathbf{Tm} \Gamma \top$, which is definitionally unique.

Definition 18. A family structure supports **extensional identity** types if there is $\mathbf{ld} : \mathbf{Tm} \Gamma A \rightarrow \mathbf{Tm} \Gamma A \rightarrow \mathbf{Ty} \Gamma$ such that $(\mathbf{reflect}, \mathbf{refl}) : \mathbf{Tm} \Gamma (\mathbf{ld} \, t \, u) \simeq (t = u)$.

It is also possible to give a positive definition for identity types, in which case we get intensional identity. Extensional identity corresponds to a categorical equalizer of terms (a limit), while the Martin-Löf-style intensional identity is characterized as the initial reflexive relation on a type (a colimit).

This choice between negative and positive specification generally exists for type formers with a single term construction rule. For example, Σ can be defined as a positive type, with an elimination rule that behaves like pattern matching. Positive Σ is equivalent to negative Σ , although it only supports propositional η -rules. In contrast, positive identity is usually *not* equivalent to negative identity.

$\mathbf{refl} : t = u \rightarrow \mathbf{Tm} \Gamma (\mathbf{ld} \, t \, u)$ expresses reflexivity of identity: definitionally equal terms are provably equal. $\mathbf{reflect}$, which goes the other way around, is called *equality reflection* [?]: provably equal terms are identified in the metatheory.

Uniqueness of identity proofs (UIP) is often ascribed to the extensional identity type [?]. UIP means that $\mathbf{Tm} \Gamma (\mathbf{ld} \, t \, u)$ has at most a single inhabitant up to \mathbf{ld} . However, UIP is not something which is inherent in the negative specification, instead it's inherited from the metatheory. If \mathbf{Tm} forms a homotopy set in the metatheory, then internal equality proofs inherit uniqueness through the defining isomorphism.

Positive types

We do not dwell much on positive types here, as elsewhere in this thesis we talk a lot about specifying such types anyway. We provide here some background and a small example.

The motivation is to specify initial internal algebras in a cwf. However, specifying the uniqueness of recursors using definitional equality is problematic, if we are to have decidable and efficient conversion checking for a type theory. Consider

the specification of **Bool** together with its recursor.

$$\begin{aligned}
\mathbf{Bool} & : \mathbf{Ty} \Gamma \\
\mathbf{true} & : \mathbf{Tm} \Gamma \mathbf{Bool} \\
\mathbf{false} & : \mathbf{Tm} \Gamma \mathbf{Bool} \\
\mathbf{BoolRec} & : (B : \mathbf{Ty} \Gamma) \rightarrow \mathbf{Tm} \Gamma B \rightarrow \mathbf{Tm} \Gamma B \rightarrow \mathbf{Tm} \Gamma \mathbf{Bool} \rightarrow \mathbf{Tm} \Gamma B \\
\mathbf{true}\beta & : \mathbf{BoolRec} B t f \mathbf{true} = t \\
\mathbf{false}\beta & : \mathbf{BoolRec} B t f \mathbf{false} = f
\end{aligned}$$

$\mathbf{BoolRec}$ together with the β -rules specifies an internal **Bool**-algebra morphism. A possible way to specify definitional uniqueness is as follows. Assuming $B : \mathbf{Ty} \Gamma$, $t : \mathbf{Tm} \Gamma B$, $f : \mathbf{Tm} \Gamma B$ and $m : \mathbf{Tm} (\Gamma \triangleright b : \mathbf{Bool}) B$, such that $m[b \mapsto \mathbf{true}] = t$ and $m[b \mapsto \mathbf{false}] = f$, it follows that $\mathbf{BoolRec} B t f b : \mathbf{Tm} (\Gamma \triangleright b : \mathbf{Bool}) B$ is equal to m .

Unfortunately, deciding conversion with this rule entails deciding pointwise equality of arbitrary **Bool** functions, which can be done in exponential time in the number of **Bool** arguments. More generally, Scherer presented a decision algorithm for conversion checking with strong finite sums and products in simple type theory [Sch17], which also takes exponential time. If we move to natural numbers with definitionally unique recursion, conversion checking becomes undecidable.

One solution is to have propositionally unique recursion instead. However, if such equations are postulated, that would break the canonicity property in intensional type theories, since now we would have equality proofs other than \mathbf{refl} in the empty context.

The standard solution is to have dependent elimination principles instead: this allows inductive reasoning, canonicity and effectively decidable definitional equality at the same time. For **Bool**, we would have

$$\begin{aligned}
\mathbf{BoolInd} & : (B : \mathbf{Ty} (\Gamma \triangleright b : \mathbf{Bool})) \rightarrow \mathbf{Tm} \Gamma (B[b \mapsto \mathbf{true}]) \\
& \rightarrow \mathbf{Tm} \Gamma (B[b \mapsto \mathbf{false}]) \rightarrow (t : \mathbf{Tm} \Gamma \mathbf{Bool}) \rightarrow \mathbf{Tm} \Gamma (B[b \mapsto t])
\end{aligned}$$

together with $\mathbf{BoolInd} B t f \mathbf{true} = t$ and $\mathbf{BoolInd} B t f \mathbf{false} = f$.

Of course, if we assume extensional identity types, we have undecidable conversion anyway, and definitionally unique recursion is equivalent to induction. But decidable conversion is an essential part of type theory, perhaps its main selling point as a foundation for mechanized mathematics, which makes it possi-

ble to relegate a deluge of boilerplate to computers. Hence, decidable conversion should be kept in mind.

Universes

Universes are types which classify types. There are several different flavors of universes.

Definition 19. A **Tarski-style** universe consists of the following data:

$$U : \text{Ty } \Gamma \quad \text{El} : \text{Tm } \Gamma \, U \rightarrow \text{Ty } \Gamma$$

This is a weak classifier, since not all $\text{Ty } \Gamma$ are necessarily represented as terms of the universe. Rather, this kind of universe can be viewed as an internal subfamily of (Ty, Tm) . Like families, Tarski universes can be closed under type formers. For instance, if U has Nat , we have the following:

$$\text{Nat} : \text{Tm } \Gamma \, U \quad \text{zero} : \text{Tm } \Gamma \, (\text{El } \text{Nat}) \quad \text{suc} : \text{Tm } \Gamma \, (\text{El } \text{Nat}) \rightarrow \text{Tm } \Gamma \, (\text{El } \text{Nat})$$

$$\begin{aligned} \text{NatElim} : & \quad (P : \text{Ty } (\Gamma \triangleright n : \text{El } \text{Nat})) \\ & \rightarrow \text{Tm } \Gamma \, (P[n \mapsto \text{zero}]) \\ & \rightarrow \text{Tm } (\Gamma \triangleright n : \text{El } \text{Nat} \triangleright np : P[n \mapsto n]) \, (P[n \mapsto \text{suc } n]) \\ & \rightarrow (n : \text{Tm } \Gamma \, (\text{El } \text{Nat})) \rightarrow \text{Tm } \Gamma \, (P[n \mapsto n]) \end{aligned}$$

If all type formers in U follow this scheme, U may be called a **weakly Tarski** universe. If we assume that every type former in U is also duplicated in (Ty, Tm) , moreover El preserves all type formers, so that e.g. $\text{El } \text{Nat}$ is definitionally equal to the natural number type in Ty , then U is **strongly Tarski**.

It is often more convenient to have stronger classifiers as universes, so that *all* types in a given family structure are represented.

Definition 20. Ignoring size issues for now, **Coquand universes** [?] are specified as follows:

$$U : \text{Ty } \Gamma \quad (\text{El}, c) : \text{Tm } \Gamma \, U \simeq \text{Ty } \Gamma$$

c maps every type in Ty to a code in U . Now we can ignore El when specifying type formers, as c can be always used to get a code in U for a type.

Unfortunately, the exact specification above yields an inconsistent “type-in-type” system, because \mathbf{U} itself has a code in \mathbf{U} . The standard solution is to have multiple family structures $(\mathbf{T}y_i, \mathbf{T}m_i)$, indexed by universe levels, and have $\mathbf{U}_i : \mathbf{T}y_{i+1} \Gamma$ and $\mathbf{T}m_{i+1} \Gamma \mathbf{U}_i \simeq \mathbf{T}y_i \Gamma$. For a general specification of consistent universe hierarchies, see [Kov21]. As mentioned in Section 1.3.2, we omit universe indices in the following, and implicitly assume “just enough” universes for particular purposes.

Definition 21. **Russell universes** are Coquand universes additionally satisfying $\mathbf{T}m \Gamma \mathbf{U} = \mathbf{T}y \Gamma$ as an equality of sets, and also $\mathbf{E}! t = t$. This justifies omitting $\mathbf{E}!$ and \mathbf{c} from informal notation, implicitly casting between $\mathbf{T}m \Gamma \mathbf{U}$ and $\mathbf{T}y \Gamma$.

Russell-style universes are commonly supported in set-theoretic models. They are also often inherited from meta-type-theories which themselves have Russell-universes. Major implementations of type theories (Coq, Lean, Agda, Idris) are all such.

3.3 Two-Level Type Theory

3.3.1 Models

We describe models of 2LTT in the following. This is not the only possible way to present 2LTT; our approach differs from [ACKS19] in some ways. We will summarize the differences at the end of this section.

Definition 22. A model of a **two-level type theory** is a model of type theory such that

- It supports a Tarski-style universe $\mathbf{T}y_0 : \mathbf{T}y \Gamma$ with decoding $\mathbf{T}m_0 : \mathbf{T}m \Gamma \mathbf{T}y_0 \rightarrow \mathbf{T}y \Gamma$.
- $\mathbf{T}y_0$ may be closed under arbitrary type formers, however, it is only possible to eliminate from $\mathbf{T}y_0$ type formers to types in $\mathbf{T}y_0$.

Types in $\mathbf{T}y_0$ are called *inner* types, while other types are *outer*. Alternatively, we may talk about *object-level* and *meta-level* types.

For example, if we have inner functions, we have the following:

$$\begin{aligned} \Pi_0 & : (A : \mathbf{T}m \Gamma \mathbf{T}y_0) \rightarrow \mathbf{T}m (\Gamma \triangleright \mathbf{T}m_0 A) \rightarrow \mathbf{T}m \Gamma \mathbf{T}y_0 \\ (\mathbf{app}_0, \mathbf{lam}_0) & : \mathbf{T}m \Gamma (\mathbf{T}m_0 (\Pi_0 A B)) \simeq \mathbf{T}m (\Gamma \triangleright \mathbf{T}m_0 A) (\mathbf{T}m_0 B) \end{aligned}$$

If we have inner Booleans, we have the following (with β -rules omitted):

$$\begin{aligned}
\mathbf{Bool}_0 & : \mathbf{Tm} \Gamma \mathbf{T}_{y_0} \\
\mathbf{true}_0 & : \mathbf{Tm} \Gamma (\mathbf{Tm}_0 \mathbf{Bool}_0) \\
\mathbf{false}_0 & : \mathbf{Tm} \Gamma (\mathbf{Tm}_0 \mathbf{Bool}_0) \\
\mathbf{BoolInd}_0 & : (B : \mathbf{Tm} (\Gamma \triangleright b : \mathbf{Tm}_0 \mathbf{Bool}_0) \mathbf{T}_{y_0}) \\
& \rightarrow \mathbf{Tm} \Gamma (\mathbf{Tm}_0 (B[b \mapsto \mathbf{true}_0])) \\
& \rightarrow \mathbf{Tm} \Gamma (\mathbf{Tm}_0 (B[b \mapsto \mathbf{false}_0])) \\
& \rightarrow (t : \mathbf{Tm} \Gamma (\mathbf{Tm}_0 \mathbf{Bool}_0)) \rightarrow \mathbf{Tm} \Gamma (\mathbf{Tm}_0 (B[b \mapsto t]))
\end{aligned}$$

Intuitively, we can view outer types and terms as metatheoretical, while \mathbf{T}_{y_0} represents the set of types in the object theory, and \mathbf{Tm}_0 witnesses that any object type can be mapped to a metatheoretical set of object terms. The restriction on elimination is crucial. If we have a Boolean term in the object language, we can use the object-level elimination principle to construct new object terms. But it makes no sense to eliminate into the metatheory. In fact, an object-level Boolean term is not necessarily **true** or **false**, it can also be just a variable or neutral term in some context, or it can be an arbitrary non-canonical value in a given model.

We review some properties of 2LTT. An important point is the action of \mathbf{Tm}_0 on type formers. In general, \mathbf{Tm}_0 preserves the negative type formers but not others.

For example, we have the isomorphism $\mathbf{Tm}_0 (\Pi_0 A B) \simeq \Pi_1 (\mathbf{Tm}_0 A) (\mathbf{Tm}_0 B)$, where Π_1 denotes outer functions. We move left-to-right by mapping t to $\mathbf{lam}_1 (\mathbf{app}_1 t)$, and the other way by mapping t to $\mathbf{lam}_0 (\mathbf{app}_0 t)$. The preservation of Σ , \top , \mathbf{K} and extensional identity is analogous.

In contrast, we can map from outer positive types to inner ones, but not the other way around. From $b : \mathbf{Tm} \Gamma \mathbf{Bool}_1$, we can use the outer \mathbf{Bool}_1 recursor to return in $\mathbf{Tm}_0 \mathbf{Bool}_0$. In the other direction, we only have constant functions since the \mathbf{Bool}_0 recursor only targets types in \mathbf{T}_{y_0} .

It may be the case that there are universes in the inner layer. For example, disregarding size issues (or just accepting an inconsistent inner theory), there may be an \mathbf{U}_0 in \mathbf{T}_{y_0} such that we have $\mathbf{Tm} \Gamma (\mathbf{Tm}_0 \mathbf{U}_0) = \mathbf{Tm} \Gamma \mathbf{T}_{y_0}$. This amounts to having a Russell-style inner universe with type-in-type. Assume that we have \mathbf{U}_1 as well, as a meta-level Russell universe. Then we can map from $\mathbf{Tm}_0 \mathbf{U}_0$ to \mathbf{U}_1 , by taking A to $\mathbf{Tm}_0 A$, but we cannot map in the other direction.

3.3.2 Internal Syntax and Notation

In the rest of this thesis we will mostly work internally to a 2LTT, i.e. we use 2LTT as metatheory. Thus, we mostly adapt the metatheoretical notations used up until now. We list used features and conventions below.

- We keep previous notation for type formers. For instance, Π -types are written as $(x : A) \rightarrow B$ or as $A \rightarrow B$.
- We assume a Coquand-style universe in the outer layer, named **Set**. As before, we leave the sizing levels implicit; if we were fully precise, we would write Set_i for a hierarchy of outer universes. Despite having a Coquand universe, we shall omit encoding and decoding in the internal syntax, and instead work in Russell-style. In practical implementations, elaborating Russell-style notation to Coquand-style is straightforward to do.
- If the same type formers are supported both in the inner and outer layers, we may distinguish them by $_0$ and $_1$ subscripts, e.g. by having Bool_0 and Bool_1 . We omit some inferable subscripts, e.g. for Π and Σ -types. In these cases, we usually know from the type parameters which type former is meant. For example, $\text{Tm}_0 \text{Bool}_0 \rightarrow \text{Bool}_1$ can only refer to outer functions.
- By having **Set**, we are able to have $\text{Ty}_0 : \text{Set}$ and $\text{Tm}_0 : \text{Ty}_0 \rightarrow \text{Set}$. So we don't have to deal with proper meta-level types, and have a more uniform notation. Notation and specification for inner type formers changes accordingly. For example, for inner Π -types we may write $(x : A) \rightarrow B$ if $A : \text{Ty}_0$ and B depends on $x : \text{Tm}_0 A$. This also enables a higher-order specification: if $B : \text{Tm}_0 A \rightarrow \text{Ty}_0$, then $(x : A) \rightarrow B x : \text{Ty}_0$, and the specifying isomorphism for Π can be written as $\text{Tm}_0((x : A) \rightarrow B x) \simeq ((x : \text{Tm}_0 A) \rightarrow \text{Tm}_0(B x))$.
- We will omit inferable Tm_0 applications. For instance, $\text{Bool}_1 \rightarrow \text{Bool}_0$ can be “elaborated” to $\text{Bool}_1 \rightarrow \text{Tm}_0 \text{Bool}_0$ without ambiguity, since the function codomain must be on the same level as the domain, and the only thing we can do to make sense of this is to lift the codomain by Tm_0 . Sometimes there is some ambiguity: $(\text{Bool}_0 \rightarrow \text{Bool}_0) \rightarrow \text{Bool}_1$ can be elaborated both to $\text{Tm}_0(\text{Bool}_0 \rightarrow \text{Bool}_0) \rightarrow \text{Bool}_1$ and to $(\text{Tm}_0 \text{Bool}_0 \rightarrow \text{Tm}_0 \text{Bool}_0) \rightarrow \text{Bool}_1$. However, in this case the two output types are definitionally isomorphic,

because of the Π -preservation by Tm_0 . Hence, the elaboration choice does not make much difference, so we shall still omit Tm_0 -s in situations like this.

Example 7. Working in the internal syntax of 2LTT, the specification of Bool_0 looks like the following (omitting β again):

$$\begin{aligned} \mathsf{Bool}_0 & : \mathsf{Ty}_0 \\ \mathsf{true}_0 & : \mathsf{Bool}_0 \\ \mathsf{false}_0 & : \mathsf{Bool}_0 \\ \mathsf{BoolInd}_0 & : (B : \mathsf{Bool}_0 \rightarrow \mathsf{Ty}_0) \rightarrow B \mathsf{true}_0 \rightarrow B \mathsf{false}_0 \rightarrow (t : \mathsf{Bool}_0) \rightarrow B t \end{aligned}$$

If we elaborate the type of $\mathsf{BoolInd}_0$, we get the following:

$$\begin{aligned} \mathsf{BoolInd}_0 & : (B : \mathsf{Tm}_0 \mathsf{Bool}_0 \rightarrow \mathsf{Ty}_0) \rightarrow \mathsf{Tm}_0 (B \mathsf{true}_0) \rightarrow \mathsf{Tm}_0 (B \mathsf{false}_0) \\ & \rightarrow (t : \mathsf{Tm}_0 \mathsf{Bool}_0) \rightarrow \mathsf{Tm}_0 (B t) \end{aligned}$$

Here, the type is forced to live in the outer level because of the dependency on Ty_0 . Since Ty_0 is an outer type, $\mathsf{Bool}_0 \rightarrow \mathsf{Ty}_0$ must be lifted, which in turn requires all other types to be lifted as well.

3.4 Standard Semantics of 2LTT

We review the standard semantics of 2LTT which we use in the rest of the thesis. This justifies the metaprogramming view, that 2LTT allows meta-level reasoning about an inner theory.

We present it two steps, by assuming progressively more structure in the inner theory. First, we only assume a category. This already lets us present a presheaf semantics for the outer layer. Then, we assume a cwf as the inner theory, which lets us interpret Ty_0 and Tm_0 and also consider inner type formers.

3.4.1 Presheaf Model of the Outer Layer

In this subsection we present a presheaf model for the outer layer of 2LTT, that is, the base category together with the terminal object, the $(\mathsf{Ty}, \mathsf{Tm})$ family and some type formers. This presheaf semantics is well-known in the literature [?]. We give a specification which follows [?] and [?] most closely.

In the following, we work outside 2LTT (since we are defining a model of 2LTT), in a suitable metatheory; an extensional type theory with enough **Set** universes suffices.

We assume a \mathbb{C} category. We write $i, j, k : |\mathbb{C}|$ for objects and $f, g, h : \mathbb{C}(i, j)$ for morphisms. We use a different notation than for cwfs before, in order to disambiguate components in \mathbb{C} from components in the presheaf model of 2LTT. We use $\hat{\mathbb{C}}$ to refer to the model which is being defined. We use the same component names for $\hat{\mathbb{C}}$ as in Section 3.2.

Model of Cwf

Definition 23. $\Gamma : \mathbf{Con}$ is a presheaf over \mathbb{C} . Its components are as follows.

$$\begin{aligned} |\Gamma| & : |\mathbb{C}| \rightarrow \mathbf{Set} \\ -\langle - \rangle & : |\Gamma| j \rightarrow \mathbb{C}(i, j) \rightarrow |\Gamma| i \\ \gamma\langle \text{id} \rangle & = \gamma \\ \gamma\langle f \circ g \rangle & = \gamma\langle f \rangle\langle g \rangle \end{aligned}$$

We flip around the order of arguments in the action of Γ on morphisms. This is more convenient because of the contravariance; we can observe this in the statement of preservation laws already. The action on morphisms is sometimes called *restriction*.

Definition 24. $\sigma : \mathbf{Sub} \Gamma \Delta$ is a natural transformation from Γ to Δ . It has action $|\sigma| : |\Gamma| i \rightarrow |\Delta| i$, such that $|\sigma|(\gamma\langle f \rangle) = (|\sigma|\gamma)$.

Definition 25. $A : \mathbf{Ty} \Gamma$ is a displayed presheaf over Γ . The “displayed” here is used in exactly the same sense as in “displayed algebra” before. As we will see in Chapter 4, presheaves can be specified with a signature, in which case a presheaf is an algebra, and a displayed presheaf is a displayed algebra. The definition here is equivalent to saying that A is a presheaf over the category of elements of Γ , but it is more convenient to use in concrete definitions and calculations. The components of A are as follows.

$$\begin{aligned} |A| & : |\Gamma| i \rightarrow \mathbf{Set} \\ -\langle - \rangle & : |A| \gamma \rightarrow (f : \mathbb{C}(i, j)) \rightarrow |A| (\gamma\langle f \rangle) \\ \alpha\langle \text{id} \rangle & = \alpha \\ \alpha\langle f \circ g \rangle & = \alpha\langle f \rangle\langle g \rangle \end{aligned}$$

Definition 26. $t : \mathsf{Tm} \Gamma A$ is a section of the displayed presheaf A . This is again the same notion of section that we have seen before, instantiated for presheaves.

$$\begin{aligned} |t| : (\gamma : |\Gamma|) i &\rightarrow |A| \gamma \\ |t|(\gamma \langle f \rangle) &= (|t| \gamma) \langle f \rangle \end{aligned}$$

Definition 27. $\Gamma \triangleright A : \mathsf{Con}$ is the total presheaf of the displayed presheaf A . Its action on objects and morphisms is the following.

$$\begin{aligned} |\Gamma \triangleright A| &\equiv (\gamma : |\Gamma|) \times |A| \gamma \\ (\gamma, \alpha) \langle f \rangle &\equiv (\gamma \langle f \rangle, \alpha \langle f \rangle) \end{aligned}$$

The id and $-\circ-$ preservation laws follow immediately.

Definition 28. $A[\sigma] : \mathsf{Ty} \Gamma$ is defined as follows, assuming $A : \mathsf{Ty} \Delta$ and $\sigma : \mathsf{Sub} \Gamma \Delta$.

$$\begin{aligned} |A[\sigma]| \gamma &\equiv |A| (|\sigma| \gamma) \\ \alpha \langle f \rangle &\equiv \alpha \langle f \rangle \end{aligned}$$

In the second component, we use $-\langle - \rangle$ for A on the right hand side. The definition is well-typed since $|A| (|\sigma| (\gamma \langle f \rangle)) = |A| ((|\sigma| \gamma) \langle f \rangle)$ by the naturality of σ . Functoriality follows from functoriality of A .

It's easy to check that the above definitions can be extended to a cwf.

- For the base category, we take the category of presheaves.
- The empty context \bullet is the terminal presheaf, i.e. the constantly \top functor.
- Type substitution is functorial, as it is defined as simple function composition of actions on objects.
- Term substitution is defined as composition of a section and a natural transformation; and also functorial for the same reason.
- Context comprehension structure follows from the Σ -based definition for context extension.

Yoneda Embedding

Before continuing with interpreting type formers in $\hat{\mathbb{C}}$, we review the Yoneda embedding, as it is useful in subsequent definitions.

Definition 29. The **Yoneda embedding**, denoted y , is a functor from \mathbb{C} to the underlying category of $\hat{\mathbb{C}}$, defined as follows.

$$\begin{aligned} y : |\mathbb{C}| &\rightarrow \mathbf{Con} & y : \mathbb{C}(i, j) &\rightarrow \mathbf{Sub}(y i)(y j) \\ y i &\equiv \mathbb{C}(-, i) & |y f| g &\equiv f \circ g \end{aligned}$$

Lemma 1 (Yoneda lemma). We have $\mathbf{Sub}(y i) \Gamma \simeq |\Gamma| i$ as an isomorphism of sets, natural in i [?].

Corollary. If we choose Γ to be $y j$, it follows that $\mathbf{Sub}(y i)(y j) \simeq \mathbb{C}(i, j)$, i.e. that y is bijective on morphisms; hence it is an embedding.

Notation 8. For $\gamma : |\Gamma| i$, we use $\gamma \langle - \rangle : \mathbf{Sub}(y i) \Gamma$ to denote transporting right-to-left along the Yoneda lemma. In the other direction we don't really need a notation, since from $\sigma : \mathbf{Sub}(y i) \Gamma$ we get $\sigma \text{id} : |\Gamma| i$.

Type Formers

Definition 30. Constant families are displayed presheaves which do not depend on their context.

$$\begin{aligned} K & : \mathbf{Con} \rightarrow \{\Gamma : \mathbf{Con}\} \rightarrow \mathbf{Ty} \Gamma \\ |K \Delta| \{i\} \gamma & \equiv |\Delta| i \\ \delta \langle f \rangle & \equiv \delta \langle f \rangle \end{aligned}$$

With this definition, we actually have $\mathbf{Tm} \Gamma (K \Delta) = \mathbf{Sub} \Gamma \Delta$ which is stronger than the required isomorphism in Definition 15.

Notation 9. It is useful to consider any set as a constant presheaf, so given $A : \mathbf{Set}$ we may write $A : \mathbf{Con}$ for the constant presheaf as well.

Definition 31. From any $A : \mathbf{Set}$, we get $K A : \mathbf{Ty} \Gamma$. This can be used to model negative or positive **closed type formers**. For example, natural numbers are modeled as $K \mathbb{N}$, Booleans as $K \mathbf{Bool}$, the unit type as $K \top$, and so on.

Definition 32. **Coquand universes** can be defined as follows. We write $\mathbf{Set}_{\hat{\mathbb{C}}}$ for the outer universe in the model, to distinguish it from the external \mathbf{Set} . Since the $\mathbf{Set}_{\hat{\mathbb{C}}}$ is a non-dependent type, it is helpful to define it as a $\mathbf{Set}_{\hat{\mathbb{C}}} : \mathbf{Con}$ such that $\mathbf{Sub} \Gamma \mathbf{Set}_{\hat{\mathbb{C}}} \simeq \mathbf{Ty} \Gamma$. The usual universe can be derived from this as $\mathbf{K} \mathbf{Set}_{\hat{\mathbb{C}}}$. Again, we ignore size issues; the fully formal definition would involve indexing constructions in $\hat{\mathbb{C}}$ by universe levels.

We can take a hint from the Yoneda lemma. We aim to define $|\mathbf{Set}_{\hat{\mathbb{C}}}| i$, but by the Yoneda lemma it is isomorphic to $\mathbf{Sub} (y i) \mathbf{Set}_{\hat{\mathbb{C}}}$. However, by specification this should be isomorphic to $\mathbf{Ty} (y i)$, so we take this as definition:

$$\begin{aligned} \mathbf{Set}_{\hat{\mathbb{C}}} & : \mathbf{Con} \\ |\mathbf{Set}_{\hat{\mathbb{C}}}| i & \equiv \mathbf{Ty} (y i) \\ A \langle f \rangle & \equiv A[yf] \end{aligned}$$

In the $A \langle f \rangle$ definition, we substitute $A : \mathbf{Ty} (y i)$ with $yf : \mathbf{Sub} (y j) (y i)$ to get an element of $\mathbf{Ty} (y j)$. The required $\mathbf{Sub} \Gamma \mathbf{Set}_{\hat{\mathbb{C}}} \simeq \mathbf{Ty} \Gamma$ is straightforward, so we omit the definition.

We note that *Russell* universes are not supported in the outer layer, as $\mathbf{Sub} \Gamma \mathbf{Set}_{\hat{\mathbb{C}}}$ and $\mathbf{Ty} \Gamma$ are not strictly the same, in particular they have a different number of components as iterated Σ -types. Nevertheless, as we mentioned in 3.3.2, we use Russell-style notation in the internal 2LTT syntax, and assume that encoding/decoding is inserted by elaboration.

Definition 33. Σ -types are defined pointwise. The definitions for pairing and projections follow straightforwardly.

$$\begin{aligned} \Sigma & : (A : \mathbf{Ty} \Gamma) \rightarrow \mathbf{Ty} (\Gamma \triangleright A) \rightarrow \mathbf{Ty} \Gamma \\ |\Sigma A B| \gamma & \equiv (\alpha : |A| \gamma) \times |B| (\gamma, \alpha) \\ (\alpha, \beta) \langle f \rangle & \equiv (\alpha \langle f \rangle, \beta \langle f \rangle) \end{aligned}$$

Definition 34. We define Π -types in the following. This is a bit more complicated, so first we look at the simpler case of presheaf exponentials.

The exponential $\Delta^\Gamma : \mathbf{Con}$ is characterized by the isomorphism $\mathbf{Sub} (\Gamma \otimes \Delta) \Xi \simeq \mathbf{Sub} \Gamma (\Xi^\Delta)$, where we write \otimes for the pointwise product of two presheaves. We can again use the Yoneda lemma. We want to define $|\Delta^\Gamma| i$, but this is isomorphic to $\mathbf{Sub} (y i) (\Delta^\Gamma)$, which should be isomorphic to $\mathbf{Sub} (y i \otimes \Gamma) \Delta$ by the specification

of exponentials. Hence:

$$\begin{aligned} |\Delta^\Gamma| i &\equiv \mathbf{Sub} (y i \otimes \Gamma) \Delta \\ \sigma \langle f \rangle &\equiv \sigma \circ (y f \circ \mathbf{p}, \mathbf{q}) \end{aligned}$$

In the definition of presheaf restriction, we use \mathbf{p}, \mathbf{q} as projections and $-, -$ as pairing for \otimes . In short, $(y f \circ \mathbf{p}, \mathbf{q})$ is the same as the morphism lifting from Definition 13: it weakens $y f : \mathbf{Sub} (y j) (y i)$ to $\mathbf{Sub} (y j \otimes \Gamma) (y i \otimes \Gamma)$.

The dependently typed case follows the same pattern, except that we use \mathbf{Tm} and $-\triangleright-$ instead of \mathbf{Sub} and $-\otimes-$. Additionally, the action on objects depends on $\gamma : |\Gamma| i$, and we make use of $\gamma \langle - \rangle : \mathbf{Sub} (y i) \Gamma$ (introduced in Notation 8).

$$\begin{aligned} \Pi & : (A : \mathbf{T y} \Gamma) \rightarrow \mathbf{T y} (\Gamma \triangleright A) \rightarrow \mathbf{T y} \Gamma \\ |\Pi A B| \{i\} \gamma &\equiv \mathbf{Tm} (y i \triangleright A[\gamma \langle - \rangle]) (B[\gamma \langle - \rangle \circ \mathbf{p}, \mathbf{q}]) \\ t \langle f \rangle & : \equiv t[y f \circ \mathbf{p}, \mathbf{q}] \end{aligned}$$

Let's unfold the above definition a bit. Assuming $t : |\Pi A B| \{i\} \gamma$, we have

$$|t| : \{j : |\mathbb{C}|\} \rightarrow ((f, \alpha) : (f : \mathbb{C}(j, i)) \times |A| (\gamma \langle f \rangle)) \rightarrow |B| (\gamma \langle f \rangle, \alpha)$$

This is a bit clearer if we remove the Σ -type by currying.

$$|t| : \{j : |\mathbb{C}|\} (f : \mathbb{C}(j, i)) (\alpha : |A| (\gamma \langle f \rangle)) \rightarrow |B| (\gamma \langle f \rangle, \alpha)$$

Restriction is functorial since it is defined as \mathbf{Tm} substitution. The definitions for **lam** and **app** are left to the reader.

Definition 35. Extensional identity is defined as pointwise equality of outer terms, i.e. sections of displayed presheaves.

$$\begin{aligned} \mathbf{Id} : \mathbf{Tm} \Gamma A \rightarrow \mathbf{Tm} \Gamma A \rightarrow \mathbf{T y} \Gamma \\ |\mathbf{Id} t u| \gamma &\equiv |t| \gamma = |u| \gamma \end{aligned}$$

For the restriction operation, we have to show that $|t| \gamma = |u| \gamma$ implies $|t| (\gamma \langle f \rangle) = |u| (\gamma \langle f \rangle)$. This follows from congruence by $-\langle f \rangle$ and naturality of t and u . The defining $(\mathbf{reflect}, \mathbf{refl}) : \mathbf{Tm} \Gamma (\mathbf{Id} t u) \simeq (t = u)$ isomorphism is evident, assuming UIP and function extensionality for the metatheoretic $=$ relation (which we do assume).

3.4.2 Modeling The Inner Layer

We assume now that \mathbb{C} is a cwf. We write types as $a, b, c : \mathbf{Ty}_{\mathbb{C}} i$ and terms as $t, u, v : \mathbf{Tm}_{\mathbb{C}} i a$. We reuse \bullet for the terminal object and $- \triangleright -$ for context extension, and likewise reuse notation for substitutions.

Definition 36 ($\mathbf{Ty}_0, \mathbf{Tm}_0$). First, note that $\mathbf{Ty}_{\mathbb{C}}$ is a presheaf over \mathbb{C} , and $\mathbf{Tm}_{\mathbb{C}}$ is a displayed presheaf over $\mathbf{Ty}_{\mathbb{C}}$; this follows from the requirement that they form a family structure over \mathbb{C} . Hence, in the presheaf model $\mathbf{Ty}_{\mathbb{C}}$ is an element of \mathbf{Con} and $\mathbf{Tm}_{\mathbb{C}}$ is an element of $\mathbf{Ty} \mathbf{Ty}_{\mathbb{C}}$. Also recall from Definition 30 that $\mathbf{Tm} \Gamma (\mathbf{K} \Delta) = \mathbf{Sub} \Gamma \Delta$. With this in mind, we give the following definitions:

$$\begin{aligned} \mathbf{Ty}_0 : \mathbf{Ty} \Gamma & & \mathbf{Tm}_0 : \mathbf{Tm} \Gamma \mathbf{Ty}_0 \rightarrow \mathbf{Ty} \Gamma \\ \mathbf{Ty}_0 &\equiv \mathbf{K} \mathbf{Ty}_{\mathbb{C}} & \mathbf{Tm}_0 A &\equiv \mathbf{Tm}_{\mathbb{C}}[A] \end{aligned}$$

$\mathbf{Tm}_{\mathbb{C}}[A]$ is well-typed since $A : \mathbf{Tm} \Gamma (\mathbf{K} \mathbf{Ty}_{\mathbb{C}})$, thus $A : \mathbf{Sub} \Gamma \mathbf{Ty}_{\mathbb{C}}$. In other words, A is a natural transformation from Γ to the presheaf of inner types.

Inner Type Formers

Can type formers in $(\mathbf{Ty}_{\mathbb{C}}, \mathbf{Tm}_{\mathbb{C}})$ be transferred to $(\mathbf{Ty}_0, \mathbf{Tm}_0)$ in the presheaf model of 2LTT? For example, if \mathbb{C} supports **Bool**, we would like to model **Bool**₀ in \mathbf{Ty}_0 as well. The following explanation is adapted from Capriotti [Cap17, Section 2.3].

Generally, a type former in \mathbb{C} transfers to $\hat{\mathbb{C}}$ if it can be specified in the internal language of $\hat{\mathbb{C}}$; if the type former “always has been” in $\hat{\mathbb{C}}$ to begin with. To be describable in $\hat{\mathbb{C}}$, a type former needs to be natural with respect to \mathbb{C} morphisms. This is also a core idea of HOAS: when working in $\hat{\mathbb{C}}$, everything is natural, and we can omit boilerplate related to contexts and substitutions. For example, consider the specification of inner Π -types in the internal syntax of 2LTT:

$$\begin{aligned} \Pi_0 & : (A : \mathbf{Ty}_0) \rightarrow (\mathbf{Tm}_0 A \rightarrow \mathbf{Ty}_0) \rightarrow \mathbf{Ty}_0 \\ (\mathbf{app}_0, \mathbf{lam}_0) & : \Pi_0 A B \simeq ((a : \mathbf{Tm}_0 A) \rightarrow \mathbf{Tm}_0 (B a)) \end{aligned}$$

We can say that this *defines* what it means for \mathbb{C} to support Π . We can recover the usual non-higher-order specification of Π in the following way, up to isomorphism:

- First, we interpret the higher-order specification as a context or closed Σ -type in the standard presheaf model of 2LTT. This yields a presheaf over \mathbb{C} .

- Then, we evaluate the resulting presheaf at the terminal object \bullet . This yields a set which is isomorphic to the conventional specification of Π .

In summary, if by “type formers” we mean extra structure on $(\mathsf{Ty}_0, \mathsf{Tm}_0)$ which is definable in 2LTT, then *by definition* all such type formers transfer from \mathbb{C} to $(\mathsf{Ty}_0, \mathsf{Tm}_0)$. This holds for every type former mentioned in this thesis.

3.4.3 Functions with Inner Domains

There is a useful semantic simplification in the standard presheaf model, in cases where we have functions of the form $\Pi(\mathsf{Tm}_0 A) B$. This will allow us to almost completely eliminate encoding overhead when interpreting inductive signatures in 2LTT; we look at examples in Section 3.5. First we look at the simply-typed case with presheaf exponentials. We need some observations about the Yoneda embedding.

Lemma 2. \mathbf{y} preserves finite products up to isomorphism, i.e. $\mathbf{y}\bullet \simeq \bullet$ and $\mathbf{y}(i \otimes j) \simeq (\mathbf{y}i \otimes \mathbf{y}j)$.

Proof. $\mathbf{y}\bullet$ is $\mathbb{C}(-, \bullet)$ by definition, which is pointwise isomorphic to \top , hence isomorphic to $\bullet \equiv \mathsf{K} \top$. $\mathbf{y}(i \otimes j)$ is $\mathbb{C}(-, i \otimes j)$, which is isomorphic to $\mathbf{y}i \otimes \mathbf{y}j$ by the specification of products. \square

Lemma 3. We have the following isomorphism.

$$\begin{aligned}
 |\Gamma^{\mathbf{y}i}| j &\equiv \\
 \text{Sub}(\mathbf{y}j \otimes \mathbf{y}i) \Gamma &\simeq && \text{by product preservation} \\
 \text{Sub}(\mathbf{y}(j \otimes i)) \Gamma &\simeq && \text{by Yoneda lemma} \\
 |\Gamma|(j \otimes i) &
 \end{aligned}$$

It is possible to rephrase the above derivation for Π -types. For that, we would need to define the action of \mathbf{y} on types and terms, consider the preservation of $- \triangleright -$ by \mathbf{y} , and also specify a “dependent” Yoneda lemma for Tm . For the sake of brevity, we omit this, and present the result directly:

$$|\Pi(\mathsf{Tm}_0 A) B| \{i\} \gamma \simeq |B| \{i \triangleright |A| \gamma\} (\gamma \langle \mathbf{p} \rangle, \mathbf{q})$$

In short, depending on an inner domain is the same as depending on an extended context in \mathbb{C} . We expand a bit on the typing of the right hand side. We have

$\gamma : |\Gamma| i$, moreover

$$\begin{aligned}
|B| & : \{j : \mathbb{C}\} \rightarrow |\Gamma \triangleright \mathbf{Tm}_0 A| j \rightarrow \mathbf{Set} \\
|B| & : \{j : \mathbb{C}\} \rightarrow ((\gamma' : |\Gamma| j) \times \mathbf{Tm}_{\mathbb{C}} j (|A| \gamma')) \rightarrow \mathbf{Set} \\
|B| \{i \triangleright |A| \gamma\} & : ((\gamma' : |\Gamma| (i \triangleright |A| \gamma)) \times \mathbf{Tm}_{\mathbb{C}} (i \triangleright |A| \gamma) (|A| \gamma')) \rightarrow \mathbf{Set} \\
\gamma \langle \mathbf{p} \rangle & : |\Gamma| (i \triangleright |A| \gamma) \\
\mathbf{q} & : \mathbf{Tm}_{\mathbb{C}} (i \triangleright |A| \gamma) ((|A| \gamma) [\mathbf{p}]) \\
\mathbf{q} & : \mathbf{Tm}_{\mathbb{C}} (i \triangleright |A| \gamma) (|A| (\gamma \langle \mathbf{p} \rangle))
\end{aligned}$$

3.5 Simple Inductive Signatures in 2LTT

CHAPTER 4

Finitary Quotient Inductive-Inductive Types

4.1 Theory of Signatures

4.1.1 Models

4.1.2 Examples

4.2 Semantics

4.2.1 Finite Limit Cwfs

4.2.2 Equivalence of Initiality and Induction

4.2.3 Model of the Theory of Signatures

4.3 Term Algebras

4.3.1 Generic Term Algebras

4.3.2 Induction for Term Algebras

4.3.3 Church Encoding

4.3.4 Awodey-Frey-Speight Encoding

4.4 Left Adjoints of Signature Morphisms

CHAPTER 5

Infinitary Quotient Inductive-Inductive Types

5.1 Theory of Signatures

5.2 Semantics

5.3 Term Algebras

CHAPTER 6

Levitation, Bootstrapping and Universe Levels

6.1 Levitation for Closed QIITs

6.2 Levitation for Infinitary QIITs

CHAPTER 7

Higher Inductive-Inductive Types

7.1 Theory of Signatures

7.2 Semantics

CHAPTER 8

Reductions

8.1 Finitary Inductive Types

8.2 Finitary Inductive-Inductive Types

8.3 Closed Quotient Inductive-Inductive Types

CHAPTER 9

Conclusion

Bibliography

- [AAC⁺20] Guillaume Allais, Robert Atkey, James Chapman, Conor McBride, and James McKinna. A type and scope safe universe of syntaxes with binding: Their semantics and proofs. *CoRR*, abs/2001.11001, 2020.
- [ACD⁺18] Thorsten Altenkirch, Paolo Capriotti, Gabe Dijkstra, Nicolai Kraus, and Fredrik Nordvall Forsberg. Quotient inductive-inductive types. In Christel Baier and Ugo Dal Lago, editors, *Foundations of Software Science and Computation Structures - 21st International Conference, FOSSACS 2018, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2018, Thessaloniki, Greece, April 14-20, 2018, Proceedings*, volume 10803 of *Lecture Notes in Computer Science*, pages 293–310. Springer, 2018.
- [ACKS19] Danil Annenkov, Paolo Capriotti, Nicolai Kraus, and Christian Sattler. Two-level type theory and applications. *ArXiv e-prints*, may 2019.
- [AK79] Jirí Adámek and Václav Koubek. Least fixed point of a functor. *J. Comput. Syst. Sci.*, 19(2):163–178, 1979.
- [AL19] Benedikt Ahrens and Peter LeFanu Lumsdaine. Displayed Categories. *Logical Methods in Computer Science*, Volume 15, Issue 1, March 2019.
- [AMM19] Benedikt Ahrens, Ralph Matthes, and Anders Mörtberg. From signatures to monads in unimath. *J. Autom. Reason.*, 63(2):285–318, 2019.
- [AVW17] Andreas Abel, Andrea Vezzosi, and Théo Winterhalter. Normalization by evaluation for sized dependent types. *Proc. ACM Program. Lang.*, 1(ICFP):33:1–33:30, 2017.

- [Awo18] Steve Awodey. Natural models of homotopy type theory. *Math. Struct. Comput. Sci.*, 28(2):241–286, 2018.
- [BH11] Patrick Bahr and Tom Hvitved. Compositional data types. In Jaakko Järvi and Shin-Cheng Mu, editors, *Proceedings of the seventh ACM SIGPLAN workshop on Generic programming, WGP@ICFP 2011, Tokyo, Japan, September 19-21, 2011*, pages 83–94. ACM, 2011.
- [Cap17] Paolo Capriotti. Models of type theory with strict equality. *arXiv preprint arXiv:1702.04912*, 2017.
- [Car78] John Cartmell. *Generalised algebraic theories and contextual categories*. PhD thesis, Oxford University, 1978.
- [CD14] Pierre Clairambault and Peter Dybjer. The biequivalence of locally cartesian closed categories and martin-löf type theories. *Mathematical Structures in Computer Science*, 24(6), 2014.
- [dVL14] Edsko de Vries and Andres Löf. True sums of products. In José Pedro Magalhães and Tiark Rompf, editors, *Proceedings of the 10th ACM SIGPLAN workshop on Generic programming, WGP 2014, Gothenburg, Sweden, August 31, 2014*, pages 83–94. ACM, 2014.
- [Dyb95] Peter Dybjer. Internal type theory. In Stefano Berardi and Mario Coppo, editors, *Types for Proofs and Programs, International Workshop TYPES’95, Torino, Italy, June 5-8, 1995, Selected Papers*, volume 1158 of *Lecture Notes in Computer Science*, pages 120–134. Springer, 1995.
- [HRR14] Claudio Hermida, Uday S. Reddy, and Edmund P. Robinson. Logical relations and parametricity - A reynolds programme for category theory and programming languages. *Electron. Notes Theor. Comput. Sci.*, 303:149–180, 2014.
- [Hug21] Jasper Hugunin. Why not w?, 2021.
- [Kov21] András Kovács. Generalized universe hierarchies and first-class universe levels, 2021.

- [KvR20] Ambrus Kaposi and Jakob von Raumer. A syntax for mutual inductive families. In Zena M. Ariola, editor, *5th International Conference on Formal Structures for Computation and Deduction, FSCD 2020, June 29-July 6, 2020, Paris, France (Virtual Conference)*, volume 167 of *LIPICs*, pages 23:1–23:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.
- [LM11] Andres Löb and José Pedro Magalhães. Generic programming with indexed functors. In Jaakko Järvi and Shin-Cheng Mu, editors, *Proceedings of the seventh ACM SIGPLAN workshop on Generic programming, WGP@ICFP 2011, Tokyo, Japan, September 19-21, 2011*, pages 1–12. ACM, 2011.
- [nc21] nLab contributors. Internal logic, 2021.
- [NF13] Fredrik Nordvall Forsberg. *Inductive-inductive definitions*. PhD thesis, Swansea University, 2013.
- [Sch17] Gabriel Scherer. Deciding equivalence with sums and the empty type. In Giuseppe Castagna and Andrew D. Gordon, editors, *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages, POPL 2017, Paris, France, January 18-20, 2017*, pages 374–386. ACM, 2017.
- [Swi08] Wouter Swierstra. Data types à la carte. *J. Funct. Program.*, 18(4):423–436, 2008.
- [Uni13] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. <https://homotopytypetheory.org/book>, Institute for Advanced Study, 2013.
- [YHLJ09] Alexey Rodriguez Yakushev, Stefan Holdermans, Andres Löb, and Johan Jeuring. Generic programming with fixed points for mutually recursive datatypes. In Graham Hutton and Andrew P. Tolmach, editors, *Proceeding of the 14th ACM SIGPLAN international conference on Functional programming, ICFP 2009, Edinburgh, Scotland, UK, August 31 - September 2, 2009*, pages 233–244. ACM, 2009.