# Notes on Type Theory

[DRAFT: April 14, 2025]

Steve Awodey

with contributions from Andrej Bauer

# Contents

4	Hor	notopy Type Theory	5	
	4.1	Identity types	6	
		4.1.1 The naive interpretation	8	
	4.2	The groupoid model	9	
	4.3	Weak factorization systems	13	
	4.4	Natural models	14	
		4.4.1 Modeling the type formers	18	
		4.4.2 Strictification	21	
	4.5	Universes	21	
	4.6 Univalence			
		4.6.1 Function extensionality	21	
		4.6.2 The h-levels	21	
	Bib	liography	23	

4 CONTENTS

## Chapter 4

### Homotopy Type Theory

The extensional dependent type theory of the previous chapter is in some ways a very natural system that admits an intuitively clear model in the locally cartesian closed category of sets and related categories. But for computational purposes, and specifically for the important application of type theory to proof checking in a computer proof assistant such as Agda or Lean, it has some serious defects: the equality relation between terms (or types) is not decidable: there is no algorithm that will determine whether two closed terms of a given type s,t:A are (judgementally) equal  $s \equiv t:A$ . Indeed, there is no normalization procedure for reducing terms to normal forms—otherwise we could use it to decide whether two terms were equal by normalizing them and then comparing their normal forms. Relatedly, one cannot effectively decide whether a given type (e.g an equality type such as  $Eq_A(s,t)$ ) is inhabited (which would be a decision procedure for the provability of  $s =_A t$ ), even given a candidate "proof term"  $p: Eq_A(s,t)$  (which would be a decision procedure for being a proof).

For this reason, the extensional system is often replaced in applications by a weaker one, called *intensional type theory*, which enjoys better computational behavior, such as decidability of equality and type-checking, and normalization. A good discussion of these and several related issues, such as canonicity and consistency can be found in Chapter 3 of the book [AG].

However, this is only one side of the story. The intensional theory was mainly a technical device for specialists in computational type theory (and a conceptual challenge from the semantic point of view) until around 2006, when it was discovered that this theory admitted a homotopical (and higher-categorical) interpretation, which led to the discovery of Homotopy type theory (HoTT) [Awo12]. This interpretation not only helped to clarify the intensional theory, and prove useful in investigating its computational properties, but also opened up a wide range of applications outside of the conventional areas of type theory, vis. computational and constructive mathematics. For, quite independently of such applications, the homotopical interpretation permits the use of intensional type theory as a powerful and expressive internal language for formal reasoning in homotopy theory and higher category theory, both highly abstract areas of mathematics, for which new and rigorous tools for calculation and proof are quite welcome. Moreover, the fortuitous fact that

this system also has the good computational behavior that it does has led to the use of computational proof assistants in homotopy theory and higher category theory, even ahead of some more down-to-earth branches of mathematics, where such exotic semantics were not needed.

The homotopical interpretation was already anticipated by a 2-dimensional one in the category of groupoids, a special case of a higher categorical model that already suffices to make some of the essential features of such models clear. Thus we shall briefly review this model below, after introducing the intensional theory, and before considering the general homotopical semantics using weak factorization systems. Such "weak" interpretations also bring to a head the coherence issues that we deferred in the previous chapter, and we conclude with one approach to strictifying such interpretations using natural models, aka, categories with families.

### 4.1 Identity types

We begin by recalling from Section ?? the rules for *equality* types in the *extensional* system: The formation, introduction, elimination, and computation rules for equality types were as follows:

$$\frac{s:A}{s=_A t} \text{ type} \qquad \frac{a:A}{\text{refl}(a):(a=_A a)}$$
 
$$\frac{p:s=_A t}{s\equiv t:A} \qquad \frac{p:s=_A t}{p\equiv \text{refl}(s):(s=_A s)}$$

The *Identity types* in the intensional theory, also written  $x =_A y$ , or sometimes  $Id_A(x, y)$ , have the same formation and introduction rules as the Equality types, but the *elimination* rule of "equality reflection" is replaced by the following elimination rule:

$$\frac{x:A,y:A,z:\mathtt{Id}_A(x,y)\vdash C(x,y,z)\ \mathtt{type},\qquad x:A\vdash c(x):C(x,x,\mathtt{refl}(x))}{x:A,y:A,z:\mathtt{Id}_A(x,y)\vdash \mathtt{J}(x,y,z,c):C}$$

in which the variable x is bound in the occurance of c within the eliminator J. The associated *computation rule* then becomes:

$$x: A \vdash J(x, x, refl(x), c) \equiv c(x) : C(x, x, refl(x))$$

In HoTT, the elimination rule is called *path induction*, for reasons that will become clear. To see how the elimination rule works, let us derive the basic laws of identity, namely reflexivity, symmetry, and transitivity, as well as Leibniz's Law the *indicernibility of identicals*, also known as the substitution of equals for equals.

• Reflexivity: states that  $x =_A y$  is a reflexive relation, but this is just the Id-formation and intro rules:

$$x:A,y:A\vdash x=_A y$$
 type,  $x:A\vdash \mathtt{refl}(x):x=_A x$ 

• Symmetry: can be stated as  $x:A,y:A,u:x=_Ay\vdash ?:y=_Ax$ , which can be proved with an Id-elim as follows:

$$\frac{x:A \vdash \mathtt{refl}(x): x =_A x}{x:A,y:A,u:x =_A y \vdash \mathtt{J}(x,y,u,\mathtt{refl}): y =_A x}$$

• Transitivity: we wish to show

$$x: A, y: A, z: A, u: x =_A y, v: y =_A z \vdash ?: x =_A z$$

regarding z:A as a fixed parameter, which we can move to the front of the context, we want to apply an Id-elim with respect to the assumption  $u:x=_A y$ , so we can set x to y, and look for a premiss of the form:

$$z : A, y : A, v : y =_A z \vdash ? : y =_A z$$

We cannot simply take v, however, since the order of the types in the context is still wrong for Id-elim, but we can move the assumption  $v: y =_A z$  to the right with a  $\lambda$ -abstraction to obtain

$$z: A, y: A \vdash \lambda v. v: y =_A z \rightarrow y =_A z$$
,

and now we can apply the planned Id-elim with respect to  $u: x=_A y$  with the "motive" being  $y=_A z \to x=_A z$  to obtain

$$z: A, y: A, x: A, u: x =_A y \vdash J(x, y, u, \lambda v.v): y =_A z \to x =_A z$$

from which follows the desired

$$x: A, y: A, z: A, u: x =_A y, v: y =_A z \vdash J(x, y, u, \lambda v.v) v: x =_A z.$$

• Substitution: to show

$$\frac{x:A \vdash C(x) \text{ type}}{x:A,y:A,u:x=_A y \vdash ?:C(x) \to C(y)}$$

it suffices to have a premiss of the form

$$x: A \vdash c(x): C(x) \rightarrow C(x)$$

for this, we can take  $c(x) = \lambda z : C(x) \cdot z : C(x) \to C(x)$  to obtain

$$x : A, y : A, u : x = A y \vdash J(x, y, u, x, \lambda z : C(x), z) : C(x) \rightarrow C(y)$$
.

Note that the variable x is bound in the J term.

Many more properties of Id-types and their associated J-terms are shown in the introductory texts [Uni13, Rij25]. One key fact is that the higher identity types  $Id_{Id_A(a,b)}(p,q)$  are no longer degenerate, but themselves may have terms that are non-identical, i.e. not propositionally equal, leading to so-called *higher types*. This "failure of UIP" (uniqueness of identity proofs) in the intensional system was first shown using the groupoid model, which sheds considerable light on the intensional system.

**Exercise 4.1.1.** Show that given any a, b, c: A and  $p: a =_A b$  and  $q: b =_A c$ , one can define a composite  $p \cdot q: a =_A c$  (using the transitivity of  $=_A$ ). Then show that, for any  $p: a =_A b$ , the symmetry term  $\sigma(p): b =_A a$  satisfies the (propositional) equation  $\sigma(p) \cdot p = \text{refl}$ . Is either of  $\sigma(p) \cdot p = \text{refl}$  or  $\text{refl} \cdot \sigma(p) = \text{refl}$  judgemental? What about associativity of  $p \cdot q$ 

**Exercise 4.1.2.** Show that  $p \cdot q$  from the previous exercise is (propositionally) associative.

**Exercise 4.1.3.** Show that any term  $f: A \to B$  acts on identities  $p: a =_A b$ , in the sense that there is a term  $ap(f)(p): fa =_B fb$ . Is ap(f) "functorial" (in the evident sense)?

Exercise 4.1.4. Observe that the Substitution property means that the assignment

$$(a:A)\mapsto C(a)$$
 type

is functorial (in some sense). Is it strictly functorial?

#### 4.1.1 The naive interpretation

We can try to interpret the (intensional) identity types in the naive way, as we did for extensional dependent type theory. This would give the formation and introduction rules as a type family  $\mathrm{Id}_A \to A \times A$  with a partial section over the diagonal substitution  $\delta_A : (x : A) \to (x : A, y : A)$ .

$$\begin{array}{c}
\operatorname{Id}_{A} \\
\downarrow^{p} \\
A \xrightarrow{\delta_{A}} A \times A
\end{array}$$

where we are writing  $Id_A$  for the extended context  $(A, A, Id_A)$  and p for the dependent family  $(x : A, y : A \vdash Id_A(x, y))$ . The elimination rule then takes the form:

$$\begin{array}{ccc}
A & \xrightarrow{c} & C \\
 & \downarrow & \downarrow \\
 & \downarrow & \downarrow$$

for any type family  $C \to Id_A$ , with the computation rule asserting that the top triangle commutes (the bottom triangle commutes by the assumption tht J is a section of  $C \to Id_A$ ).

But now recall that in extensional type theory, any map  $f: B \to A$  can be regarded as a type family over A, namely by taking the graph factorization

$$B \cong \Sigma_{a:A}\Sigma_{b:B}\mathsf{Eq}_A(a,fb) \longrightarrow B \times A.$$

So we can take the family C in the elimination to be refl:  $A \to Id_A$ , to obtain:

We therefore get an iso  $A \cong Id_A$ , making the identity type isomorphic to the extensional equality type  $Eq_A = A \to A \times A$ .

**Exercise 4.1.5.** Prove that in the extensional theory, the graph factorization does indeed make any map  $f: B \to A$  isomorphic to a family of types over its codomain. *Hint:* Consider the following two-pullback diagram.

$$\begin{array}{cccc} \Sigma_{a:A}\Sigma_{b:B}\mathrm{Eq}_{A}(a,fb) & \longrightarrow & A \\ & \downarrow & & \downarrow & \\ & A\times B & \xrightarrow{A\times f} & A\times A \\ & p_{2} \downarrow & & \downarrow p_{2} \\ & B & \xrightarrow{f} & A \end{array}$$

### 4.2 The groupoid model

Exercises 4.1.1 - 4.1.4 from the last section suggest an interpretation of the intensional version of dependent type theory, namely with types as groupoids and type families as functors. Such an interpretation was first given by [HS98] in order to show that the principle of Uniqueness of Identity Proofs (UIP) – which holds in the extensional theory – indeed fails in the intensional one. We shall briefly sketch this result here.

In order to give a model of intensional type theory we should define what it means to be be a model of (intensional) type theory. We will do this in section ?? below – for now we simple describe a single model in the category  $\mathsf{Gpd}$  of groupoids, which will turn out to be an instance of the general notion. For the extensional theory, we defined a model simply to be an interpretation into an LCCC  $\mathcal{E}$ , with contexts  $\Gamma$  interpreted as objects of  $\mathcal{E}$ , substitutions  $\sigma: \Delta \to \Gamma$  as arrows of  $\mathcal{E}$ , type families  $\Gamma \vdash A$  as objects of  $\mathcal{E}/\Gamma$ , and terms  $\Gamma \vdash a: A$  as sections of the associated families. Substitution into families and terms was (weakly) interpreted' as pullback (in the sense that there was an unresolved coherence issue), and the  $\Sigma$  and  $\Pi$  type formers were adjoints to pullback. Finally, the equality type

 $x:A,y:A\vdash \mathsf{Eq}_A(x,y)$  was interpreted as the diagonal  $A\to A\times A$ .

We may simplify the notation for category of contexts by using  $\Sigma$ -types, writing e.g.  $(x:A,y:A,z: \mathsf{Eq}_A(x,y)) = \Sigma_{x:A}\Sigma_{y:A}\mathsf{Eq}_A(x,y)$  or even  $\mathsf{Eq}_A$ , and  $(x:A,y:A) = A \times A$ , etc.

The groupoid interpretation of the intensional theory is based on the idea that the identity type of a type (interpreted as a groupoid) G can be interpreted by the *path groupoid* of G, which we shall write as  $G^I$ .

**Definition 4.2.1.** If the groupoid  $G = G_0 \rightrightarrows G_1$  has objects  $G_0$  and arrows  $G_1$ , the *path groupoid*  $G^I = (|(G^I)^I| \rightrightarrows |G^I|)$  has as objects  $|G^I| = G_1$ , and as arrows  $|(G^I)^I|$  the set of all commutative squares in G, with the obvious source and target maps.

In other words, the path groupoid is the arrow category  $G^{\downarrow}$ . Recall that the category  $\mathsf{Gpd}$  of (small) groupoids is a cartesian closed subcategory of  $\mathsf{Cat}$ , and that there is a walking arrow groupoid I with exactly two objects and two (mutually inverse) non-identity arrows,

$$I = (0 ) 1)$$

The notation  $G^I$  for the path groupoid is then correctly the exponential of G by I in Gpd (and in Cat). Observe that there are functors  $dom, cod: G^I \rightrightarrows G$ , as well as one  $id: G \to G^I$ , making  $G^I \rightrightarrows G$  into an internal groupoid in Gpd, for any object G. This will be our interpretation of the identity type of the type interpreted by G.

More formally, we interpret:

- Contexts Γ: groupoids, i.e. objects of Gpd,
- Substitutions  $\sigma: \Delta \to \Gamma$ : homomorphisms of groupoids, i.e. arrows of Gpd,
- Types  $\Gamma \vdash A$ : functors  $A : \mathsf{G} \to \mathsf{Gpd}$ , where  $\mathsf{G}$  interprets  $\Gamma$ ,
- Terms  $\Gamma \vdash a : A$ : natural transformations  $a : 1 \to A$  between functors, where 1 is the terminal functor in  $\mathsf{Gpd}^\mathsf{G}$ ,
- Context extension  $(\Gamma,A) \to \Gamma$ : the Grothendieck construction  $\int_{\mathsf{G}} A \to \mathsf{G}$ .

In order to model the type formers  $\Sigma$ ,  $\Pi$ , etc. of intensional type theory in  $\mathsf{Gpd}$ , we must deal with the fact that  $\mathsf{Gpd}$  is not locally cartesian closed, although it is cartesian closed. Recall that in order to model extensional type theory in presheaves we used the fact that  $\mathsf{Set}^{\mathbb{C}^{\mathsf{op}}}$  is always a CCC and that for any  $P \in \mathsf{Set}^{\mathbb{C}^{\mathsf{op}}}$  we have an equivalence

$$\mathsf{Set}^{\mathbb{C}^\mathsf{op}}\!/_{\!P} \; \simeq \; \mathsf{Set}^{\int P^\mathsf{op}} \, .$$

and thus every slice is also a CCC. For groupoids, something similar is the case, but instead of the full slice category  $\mathsf{Gpd}/\mathsf{G}$  we use the subcategory of *fibrations*  $\mathsf{Fib}_\mathsf{G} \hookrightarrow \mathsf{Gpd}/\mathsf{G}$ , for which we have an equivalence

$$\mathsf{Fib}_\mathsf{G} \simeq \mathsf{Gpd}^\mathsf{G} \simeq \mathsf{Gpd}(\mathsf{Set}^{\mathsf{Gop}})$$

Since the proof that  $\mathsf{Gpd}$  is a CCC doesn't depend on the classical logic of  $\mathsf{Set}$ , the category of internal groupoids in a topos like  $\mathsf{Set}^{\mathsf{G}^{\mathsf{op}}}$  is also a CCC. Thus we have that  $\mathsf{Fib}_\mathsf{G}$  is a CCC for any groupoid  $\mathsf{G}$ .

**Definition 4.2.2.** A (split op-) fibration of groupoids  $p: A \to G$  is a functor satisfying the condition: for every  $a \in A$  and  $\gamma: pa \to g$  there is given a "lift"  $\ell(a, \gamma): a \to \tilde{g}$  with  $p(\ell(a, \gamma)) = p$ , and moreover,

- 1.  $\ell(a, 1_{pa}) = 1_a : a \to a$ ,
- 2. for  $\gamma': g = p(\tilde{g}) \to h$ , the lift of the composite is the composite of the lifts:

$$\ell(a, \gamma' \circ \gamma) = \ell(\tilde{g}, \gamma') \circ \ell(a, \gamma) : a \to \tilde{h}.$$

**Proposition 4.2.3.** The category  $\mathsf{Fib}_\mathsf{G}$  of fibrations of groupoids and functors  $f:\mathsf{A}\to\mathsf{B}$  over  $\mathsf{G}$  that preserve the lifts is equivalent to the functor category  $\mathsf{Gpd}^\mathsf{G}$ .

The interpretation of the context extension  $(\Gamma,A)\to \Gamma$  is to be projection  $\int_{\mathsf{G}} A\to \mathsf{G}$  given by the Grothendieck construction, and this is indeed a fibration of groupoids. Indeed, the functor taking  $A:\mathsf{G}\to\mathsf{Gpd}$  to  $\int_{\mathsf{G}} A\to \mathsf{G}$  mediates the equivalence  $\int:\mathsf{Gpd}^{\mathsf{G}}\simeq\mathsf{Fib}_{\mathsf{G}}$ .

For the base change functors along a fibration  $p : A \to G$ , we then have left and right adjoints as follows:

$$\begin{array}{cccc} \mathsf{A} & & \mathsf{Gpd}^\mathsf{A} & \xrightarrow{\int} \mathsf{Fib}_\mathsf{A} \\ p & & p^* \! \uparrow & & \Sigma \! \begin{pmatrix} \uparrow \\ p^* \end{pmatrix} \Pi \\ \mathsf{G} & & \mathsf{Gpd}^\mathsf{G} & \xrightarrow{\int} \mathsf{Fib}_\mathsf{G} \\ \end{array}$$

To show this, it needs to be shown that:

- 1. the pullback of a fibration is a fibration,
- 2. the composite of fibrations is a fibration,
- 3. there is a push-forward fibration of a fibration along a fibration, which is right adjoint to pullback.

The proof uses the CCC structure in the categories  $Fib_A \simeq Gpd^A \simeq Gpd(Set^A)$  and is similar to the proof of the LCCC structure for a category of presheaves.

#### Identity types

To interpret the Id-type of a type (interpreted as, say)  $A = (A_1 \rightrightarrows A_0)$  in Gpd (or indeed in any relative version  $Gpd(Set^G)$ ), we shall use the path groupoid

$$Id_A = A^I \rightarrow A \times A$$
,

which is easily seen to be a fibration, and therefore corresponds to a functor  $Id_A : A \times A \rightarrow Gpd$ , namely that with *discrete groupoids* as its values:

$$\operatorname{Id}_A(a,b) = \{p : a \to b\} \subseteq A_1.$$

Note that for two objects  $a, b \in A$  there may be many different arrows  $f: a \to b$  in  $\mathrm{Id}_A(a,b)$ , but for two such parallel arrows  $f,g:a \rightrightarrows b$  in A, regarded as objects in the path groupoid  $\mathrm{Id}_A = A^{\mathsf{I}}$ , there need be no arrow between them in the (discrete) groupoid  $\mathrm{Id}_A(a,b)$ ; and indeed, there will be one (which is then unique) just if f=g. Thus we will have the desired violation of UIP, once we have shown that this interpretation satisfies the rules for intensional Id-types.

To show that, consider the diagram below, which we have already encountered as (4.3). We take any fibration  $p: C \to Id_A$  and any section  $c: A \to C$  over the insertion of identity arrows into the path groupoid refl:  $A \to A^I = Id_A$ , and we need a diagonal filler J.

$$\begin{array}{c} \mathsf{A} \stackrel{c}{\longrightarrow} \mathsf{C} \\ \text{refl} \downarrow \qquad \qquad \mathsf{J} \downarrow^p \\ \mathsf{Id}_A =\!\!\!\!\!= \mathsf{Id}_A \end{array}$$

Since the diagram commutes by assumption, for any  $a \in A$  we have  $pca = 1_a$ . Let  $\alpha : a \to b$  be any object in  $\mathrm{Id}_A = A^{\mathsf{I}}$  and observe that there is always an arrow  $\chi_\alpha : 1_a \Rightarrow \alpha$  in  $\mathrm{Id}_A = A^{\mathsf{I}}$ , namely  $\chi_\alpha = (1_a, \alpha)$ .

$$\begin{array}{ccc}
a & \xrightarrow{1_a} & a \\
\downarrow^{1_a} & \chi_\alpha & \downarrow^{\alpha} \\
a & \xrightarrow{\alpha} & b
\end{array}$$

Since  $p: C \to Id_A$  is a fibration, there is a lift  $\ell(ca, \chi_\alpha): ca \to \tilde{\alpha}$ . We then set

$$J(\alpha) = \tilde{\alpha}$$

to obtain a functor  $J: \mathtt{Id}_A \to A$  making the two triangles in the diagram commute.

Exercise 4.2.4. Prove this!

**Exercise 4.2.5.** Show that the composition of fibrations  $B \to A$  and  $A \to G$  is a fibration. (This will be used for the interpretation of the type  $\Gamma \vdash \Sigma_A B$ , where  $\Gamma \vdash A$  and  $\Gamma, A \vdash B$ .)

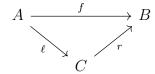
### 4.3 Weak factorization systems

WFSs generalize the groupoid model and include many familiar examples.

We can axiomatize the features of the groupoid model that allowed us to model intensional type theory using the notion of a *weak factorization system*, which is important in modern axiomatic homotopy theory. This is a weakening of the notion of an orthogonal factorization system from Definition ??:

**Definition 4.3.1.** An weak factorization system (wfs) on a category C consists of two classes of arrows (L, R) such that:

- 1. The classes  $L, R \subseteq \mathcal{C}_1$  are closed under retracts in the arrow category.
- 2. Every map  $f: A \to B$  factors  $f = r \circ \ell$  into  $\ell \in L$  followed by  $r \in R$ ,



3. Given any commutative square with an L-map on the left and an R-map on the right,

$$\begin{array}{ccc}
A & \longrightarrow & B \\
\ell \downarrow & & & \downarrow r \\
C & \longrightarrow & D
\end{array}$$

there is a (not necessarily unique) diagonal filler as indicated, making both triangles commute.

Given such a wfs on a finitely complete category C, we shall interpret the contexts and substitutions as the objects and arrows of C, the type families as the right maps, and the terms as the sections of the right maps.

**Lemma 4.3.2.** In a wfs (L,R) on a finitely complete category C, the right maps are stable under pullback along all maps. Moreover, both L and R are closed under composition.

$$Proof.$$
 ...

It follows that we can interpret the structural rules and the context extension. The rules for  $\Sigma$  types are also satisfied, since these state that  $\Sigma$  is left adjoint to pullback, which is just composition of right maps. Let us see that we can also interpret the rules for Id-types.

The formation rule for  $\mathrm{Id}_A$  is interpreted by factoring the diagonal substitution  $\delta: A \to A \times A$  into a left map followed by a right map:

$$\begin{array}{c}
\operatorname{Id}_{A} \\
\downarrow^{p} \\
A \xrightarrow{\delta} A \times A
\end{array}$$

This also interprets the introduction rule, using the left map in the factorization as the interpretation of the refl term. For the elimination rule, suppose we have a type family  $p: C \to Id_A$  and a section  $c: A \to C$  over refl:  $A \to Id_A$ ; then we need a diagonal filler J.

$$\begin{array}{ccc}
A & \xrightarrow{c} & C \\
\text{refl} & & \downarrow p \\
Id_A & & & Id_A
\end{array}$$

But since  $ref1: A \to Id_A$  is a left map by the factorization, and  $C \to Id_A$  is a right map by the interpretation of type families as right maps, there is such a filler by the third axiom of wfs's. Thus we have already shown:

**Proposition 4.3.3** ([AW09]). In a fintely complete category  $\mathcal{C}$  with a wfs, the rules of intensional identity types are soundly modeled by interpreting the type families as the right maps and the identity type  $\mathrm{Id}_A$  as a factorization of the diagonal  $\delta: A \to A \times A$  into a left map  $\mathrm{refl}: A \to \mathrm{Id}_A$  followed by a right map  $\mathrm{Id}_A \to A \times A$ .

This kind of interpretation includes many important "naturally occurring" examples involving *Quillen model categories*, which are categories equipped with two interlocking wfs's (see [AW09]). The Π-types can also be interpreted in this way, if the right maps of the wfs pushforward along right maps, as is the case in examples such as right-proper Quillen model categories and Π-tribes in the sense of [?]. Indeed, the groupoid model from the previous section was an instance of such a wfs: as the right maps one can take the isofibrations, and the left maps are the equivalences that are injective on objects [?].

Remark 4.3.4. The coherence issue that we have been postponing is now even more pressing, however, because the factorizations in a wfs need not be stable under pullback, and so the substitution rule for the Id type former will not be soundly modeled, even up to isomorphism. A similar problem occurs with respect to the J-term, which is required to respect certain substitutions in the type theory, but need not do so under this interpretation. One solution to this problem makes use of a further algebraic structure on the factorization system, called an *algebraic weak factorization system*. This approach is discussed in [?]. We shall develop a different solution in the next section.

#### 4.4 Natural models

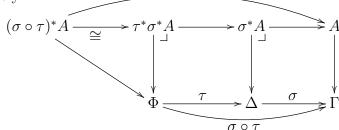
The semantics of DTT in LCCCs described in the previous sections uses the "slice category" hyperdoctrine of an LCCC to interpret the dependent types. Thus the contexts  $\Gamma$  and substitutions  $\sigma: \Delta \to \Gamma$  are interpreted as the objects and arrow of a LCC category  $\mathcal{C}$ , and the dependent types  $\Gamma \vdash A$  and terms  $\Gamma \vdash a: A$  are interpreted as objects  $A \to \Gamma$  in the slice category  $\mathcal{C}/\Gamma$  and their global sections  $a: \Gamma \to A$  (over  $\Gamma$ ). As we mentioned in Remark ??, however, there is a problem with this kind of semantics (as first pointed out by [Hof95]): as a hyperdoctrine, this interpretation is a pseudofunctor  $\mathcal{C}/: \mathcal{C}^{\mathsf{op}} \to \mathsf{Cat}$ ,

4.4 Natural models 15

but the syntax of DTT produces an actual *presheaf* of types in context  $\mathsf{Ty}: \mathcal{C}^{\mathsf{op}} \to \mathsf{Set}$ , since substitution into dependent types is *strictly* functorial with respect to composition of substitutions, in the sense that for a type in context  $\Gamma \vdash A$  and substitutions  $\sigma: \Delta \to \Gamma$  and  $\tau: \Phi \to \Delta$  we have an *equality of types in context*,

$$\Phi \vdash (A[\sigma])[\tau] \equiv A[\sigma \circ \tau] ,$$

rather than the (canonical) isomorphism  $\cong$  fitting into the two-pullbacks diagram of the hyperdoctrine, namely:



A similar problem occurs in the Beck-Chavalley conditions, where the hyperdoctrine structure has only canonical isos, rather than the strict equalities that obtain in the syntax, such as

$$(\Pi_{x:A}B)[\sigma] \equiv (\Pi_{x:A[\sigma]}B[\sigma])$$
.

There are various different solutions to this problem in the literature, some involving "strictifications" of the LCC slice-category hyperdoctrine (including both left- and right-adjoint strictifications [?, ?]), as well as other semantics altogether, such as categories-with-families [Dyb96], categories-with-attributes [?], and comprehension categories [?].

A solution based on the notion of a universe  $U \to U$  was first proposed by Voevodsky [?]; in [Awo16], the universe approach is combined with the notion of a representable natural transformation as follows.

**Definition 4.4.1.** For a small category  $\mathbb{C}$ , a natural transformation  $f: Y \to X$  of presheaves on  $\mathbb{C}$  is called *representable* if for every  $C \in \mathbb{C}$  and  $x \in X(C)$ , there is given a  $p: D \to C$  and a  $y \in Y(D)$  such that the following square is a pullback.

$$yD \xrightarrow{y} Y$$

$$yp \downarrow \qquad \downarrow f$$

$$yC \xrightarrow{x} X$$

$$(4.2)$$

We will show that a representable natural transformation is essentially the same thing as a *category with families* in the sense of Dybjer [Dyb96]. Indeed, let us write the objects of  $\mathbb{C}$  as  $\Gamma, \Delta, \ldots$  and the arrows as  $\sigma : \Delta \to \Gamma, \ldots$ , thinking of  $\mathbb{C}$  as a "category of contexts". Let  $t : \dot{T} \to T$  be a representable map of presheaves, and interpret the elements as

$$T(\Gamma) = \{A \mid \Gamma \vdash A\}$$
  
$$\dot{T}(\Gamma) = \{a \mid \Gamma \vdash a : A\},\$$

so that under Yoneda we have:

$$A \in \mathsf{T}(\Gamma)$$
 iff  $\Gamma \vdash A$   
 $a \in \dot{\mathsf{T}}(\Gamma)$  iff  $\Gamma \vdash a : A$ ,

where  $t \circ a = A$ , as indicated in:

$$y\Gamma \xrightarrow{a} T$$

Thus we regard T as the *presheaf of types*, with  $T(\Gamma)$  the set of all types in context  $\Gamma$ , and  $\dot{T}$  as the *presheaf of terms*, with  $\dot{T}(\Gamma)$  the set of all terms in context  $\Gamma$ , while the component  $t_{\Gamma}: \dot{T}(\Gamma) \to T(\Gamma)$  is the typing of the terms in context  $\Gamma$ .

The naturality of  $t : T \to T$  just means that for any substitution  $\sigma : \Delta \to \Gamma$ , we have an action on types and terms:

$$\Gamma \vdash A \quad \mapsto \quad \Delta \vdash A\sigma$$

$$\Gamma \vdash a : A \quad \mapsto \quad \Delta \vdash a\sigma : A\sigma .$$

While, by functoriality, given any further  $\tau: \Phi \to \Delta$ , we have

$$(A\sigma)\tau = A(\sigma \circ \tau)$$
  $(a\sigma)\tau = a(\sigma \circ \tau),$ 

as well as

$$A1 = A$$
  $a1 = a$ 

for the identity substitution  $1:\Gamma\to\Gamma$ .

Finally, the representability of the natural transformation  $p: E \to U$  is exactly the operation of *context extension*: given any  $\Gamma \vdash A$ , by Yoneda we have the corresponding map  $A: y\Gamma \to T$ , and we let  $p_A: \Gamma.A \to \Gamma$  be (the map representing) the pullback of t along A, as in (4.2). We therefore have a pullback square:

$$y\Gamma.A \xrightarrow{q_A} \dot{\mathsf{T}}$$

$$yp_A \downarrow \qquad \qquad \downarrow t$$

$$y\Gamma \xrightarrow{A} \mathsf{T}$$

$$(4.3)$$

where the map  $q_A: \Gamma.A \to \dot{\mathsf{T}}$  now determines a term

$$\Gamma.A \vdash q_A : Ap_A.$$

We may hereafter omit the y for the Yoneda embedding, letting the Greek letters serve to distinguish representable presheaves and their maps.

4.4 Natural models

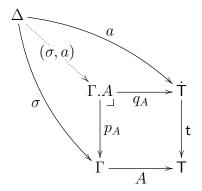
**Exercise 4.4.2.** Show that the fact that (4.3) is a pullback means that given any  $\sigma : \Delta \to \Gamma$  and  $\Delta \vdash a : A\sigma$ , there is a map

$$(\sigma, a): \Delta \to \Gamma.A,$$

and this operation satisfies the equations

$$p_A \circ (\sigma, a) = \sigma$$
  
 $q_A(\sigma, a) = a$ ,

as indicated in the following diagram.



Show moreover that the uniqueness of  $(\sigma, a)$  means that for any  $\tau : \Delta' \to \Delta$  we also have:

$$(\sigma, a) \circ \tau = (\sigma \circ \tau, a\tau)$$
  
 $(p_A, q_A) = 1.$ 

Comparing the foregoing with the definition of a category with families in [Dyb96], we have shown:

**Proposition 4.4.3.** Let  $t: \dot{T} \to T$  be a representable natural transformation of presheaves on a small category  $\mathbb C$  with a terminal object. Then t determines a category with families, with  $\mathbb C$  as the contexts and substitutions,  $T(\Gamma)$  as the types in context  $\Gamma$ , and  $\dot{T}(\Gamma)$  as the terms in context  $\Gamma$ .

**Remark 4.4.4.** A category with families is usually defined in terms of a presheaf

$$\mathsf{Ty}:\mathcal{C}^\mathsf{op}\to\mathsf{Set}$$

of types on the category  $\mathcal{C}$  of contexts, together with a presheaf

$$\mathsf{Tm}': \left( \int_{\mathcal{C}} \mathsf{Ty} \right)^{\mathsf{op}} \to \mathsf{Set}$$

of typed-terms on the category  $\int_{\mathcal{C}} \mathsf{Ty}$  of types-in-context. We are using the equivalence of categories, valid for any category of presheaves  $\mathsf{Set}^{\mathbb{C}^{\mathsf{op}}}$ ,

$$\mathsf{Set}^{\mathbb{C}^{\mathsf{op}}}/_{P} \; \simeq \; \mathsf{Set}^{(\int_{\mathbb{C}} P)^{\mathsf{op}}}$$

between the slice category over a presheaf P and the presheaves on its category of elements  $\int_{\mathbb{C}} P$ , to turn the presheaf  $\mathsf{Tm}': (\int_{\mathcal{C}} \mathsf{Ty})^{\mathsf{op}} \to \mathsf{Set}$  into one  $\mathsf{Tm}: \mathcal{C}^{\mathsf{op}} \to \mathsf{Set}$  together with a map  $\mathsf{Tm} \to \mathsf{Ty}$  in  $\mathcal{C}^{\mathsf{op}} \to \mathsf{Set}$ .

We think of a representable map of presheaves on an arbitrary category  $\mathbb{C}$  as a "type theory over  $\mathbb{C}$ ", with  $\mathbb{C}$  as the category of contexts and substitutions. We will show in Section 4.4.2 that such a map of presheaves is essentially determined by a class of maps in  $\mathbb{C}$  that is closed under all pullbacks, corresponding to the "incoherent" interpretation of types in context as maps  $A \to \Gamma$ .

**Definition 4.4.5.** A natural model of type theory on a small category  $\mathbb{C}$  is a representable map of presheaves  $t: \dot{T} \to T$ .

**Exercise 4.4.6** (The natural model of syntax). Let  $\mathbb{T}$  be a dependent type theory and  $\mathcal{C}_{\mathbb{T}}$  its category of contexts and substitutions. Define the presheaves  $\mathsf{Ty}: \mathcal{C}_{\mathbb{T}}^{\mathsf{op}} \to \mathsf{Set}$  of types-in-context and  $\mathsf{Tm}: \mathcal{C}_{\mathbb{T}}^{\mathsf{op}} \to \mathsf{Set}$  of terms-in-context, along with a natural transformation

$$\mathsf{tp}:\mathsf{Tm}\to\mathsf{Ty}$$

that takes a term to its type. Show that  $tp: Tm \to Ty$  is a natural model of type theory.

#### 4.4.1 Modeling the type formers

Given a natural model  $t: T \to T$ , we will make extensive use of the associated *polynomial* endofunctor  $P_t: \hat{\mathbb{C}} \longrightarrow \hat{\mathbb{C}}$  (cf. [?]), defined by

$$P_t = T_! \circ t_* \circ \dot{T}^* : \hat{\mathbb{C}} \longrightarrow \hat{\mathbb{C}} \,,$$

The action of  $P_t$  on an object X may be depicted:

$$\begin{array}{cccc}
X & \longleftarrow X \times \dot{\mathsf{T}} & & \mathsf{P}_{\mathsf{t}}(X) \\
\downarrow & & \downarrow \\
\dot{\mathsf{T}} & \longrightarrow \mathsf{T}
\end{array}$$

We call  $t: \dot{T} \to T$  the signature of  $P_t$  and briefly recall the following universal mapping property from [?].

**Lemma 4.4.7.** For any  $p: E \to B$  in a locally cartesian closed category  $\mathcal{E}$ , the polynomial functor  $P_p: \mathcal{E} \to \mathcal{E}$  has the following universal property: for any objects  $X, Z \in \mathcal{E}$ , maps

4.4 Natural models

 $f: Z \to \mathsf{P}_p(X)$  correspond bijectively to pairs of maps  $f_1: Z \to B$  and  $f_2: Z \times_B E \to Z$ , as indicated below.

The correspondence is natural in both X and Z, in the expected sense.

This universal property is also suggested by the conventional type theoretic notation, namely:

$$\mathsf{P}_p(X) = \Sigma_{b:B} X^{E_b}$$

The lemma can be used to determine the signature  $p \cdot q$  for the composite  $P_p \circ P_q$  of two polynomial functors, which is again polynomial, and for which we therefore have

$$\mathsf{P}_{p \cdot q} \cong \mathsf{P}_p \circ \mathsf{P}_q \,. \tag{4.5}$$

Indeed, let  $p: B \to A$  and  $q: D \to C$ , and consider the following diagram resulting from applying the correspondence (4.6) to the identity arrow,

$$\langle a, c \rangle = 1_{\mathsf{P}_p(C)} : \mathsf{P}_p(C) \to \mathsf{P}_p(C) ,$$

and taking Q to be the indicated pullback.

$$D \longleftarrow Q$$

$$q \downarrow \qquad \qquad \downarrow \qquad p \cdot q$$

$$C \longleftarrow \pi^* B \longrightarrow B \qquad \qquad \downarrow p$$

$$\downarrow \qquad \qquad \downarrow p$$

$$P_p(C) \longrightarrow A$$

$$(4.6)$$

The map  $p \cdot q$  is then defined to be the indicated composite,

$$p \cdot q = a^* p \circ c^* q.$$

The condition (4.5) can then be checked using the correspondence (4.6) (also see [?]).

**Definition 4.4.8.** A natural model  $t : \dot{T} \to T$  over  $\mathbb{C}$  will be said to *model* the type formers  $1, \Sigma, \Pi$  if there are pullback squares in  $\hat{\mathbb{C}}$  of the following form,

where  $\mathbf{t} \cdot \mathbf{t} : \dot{\mathsf{T}}_2 \to \mathsf{T}_2$  is determined by  $\mathsf{P}_{\mathbf{t} \cdot \mathbf{t}} \cong \mathsf{P}_{\mathbf{t}} \circ \mathsf{P}_{\mathbf{t}}$  as in (4.5).

The terminology is justified by the following result from [?].

**Theorem 4.4.9** ([Awo16] Theorem XXX). Let  $t : \dot{T} \to T$  be a natural model. The associated category with families satisfies the usual rules for the type-formers  $1, \Sigma, \Pi$  just if  $t : \dot{T} \to T$  models the same in the sense of Definition 4.4.8.

We only sketch the case of  $\Pi$ -types, but the other type formers will be treated in detail in Section ??.

**Proposition 4.4.10.** The natural model  $t : \dot{T} \to T$  models  $\Pi$ -types just if there are maps  $\lambda$  and  $\Pi$  making the following a pullback diagram.

$$P_{t}(\dot{T}) \xrightarrow{\lambda} \dot{T}$$

$$P_{t}(t) \downarrow \qquad \qquad \downarrow t$$

$$P_{t}(T) \xrightarrow{\Pi} T$$

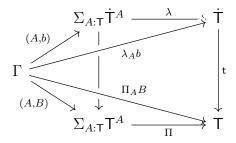
$$(4.8)$$

*Proof.* Unpacking the definitions, we have  $P_t(T) = \Sigma_{A:T} T^A$ , etc., so diagram (4.8) becomes:

$$\begin{array}{ccc} \Sigma_{A:\mathsf{T}} \dot{\mathsf{T}}^A & \xrightarrow{\lambda} & \dot{\mathsf{T}} \\ \Sigma_{A:\mathsf{T}} \mathbf{t}^A \bigg\downarrow & & & \downarrow \mathbf{t} \\ \Sigma_{A:\mathsf{T}} \mathbf{T}^A & \xrightarrow{\quad \Pi \quad} & \mathsf{T} \end{array}$$

For  $\Gamma \in \mathbb{C}$ , maps  $\Gamma \to \Sigma_{A:T}\mathsf{T}^A$  correspond to pairs (A, B) with  $A : \Gamma \to \mathsf{T}$  and  $B : \Gamma, A \to \mathsf{T}$ , and thus to  $\Gamma \vdash A$  and  $\Gamma, A \vdash B$ . Similarly, a map  $\Gamma \to \Sigma_{A:T}\dot{\mathsf{T}}^A$  corresponds to a pair (A, b) with  $\Gamma \vdash A$  and  $\Gamma, A \vdash b : B$ , the typing of b resulting from composing with the map

$$\Sigma_{A:\mathsf{T}}\mathsf{t}^A:\Sigma_{A:\mathsf{T}}\dot{\mathsf{T}}^A\to\Sigma_{A:\mathsf{T}}\mathsf{T}^A$$
.



The composition across the top is then the term  $\Gamma \vdash \lambda_{x:A}b$ , the type of which is determined by composing with t and comparing with the composition across the bottom, namely  $\Gamma \vdash \Pi_{x:A}B$ . In this way, the lower horizontal arrow in the diagram models the  $\Pi$ -formation rule:

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash \Pi_{x:A}B}$$

4.5 Universes 21

and the upper horizontal arrow, along with the commutativity of the diagram, models the  $\Pi$ -introduction rule:

$$\frac{\Gamma, A \vdash b : B}{\Gamma \vdash \lambda_{x:A}b : \Pi_{x:A}B}$$

The square (4.8) is a pullback just if, for every  $(A, B): \Gamma \to \Sigma_{A:T}\mathsf{T}^A$  and every  $t: \Gamma \to \mathsf{T}$  with  $\mathsf{t} \circ t = \Pi_A B$ , there is a unique  $(A, b): \Gamma \to \Sigma_{A:T}\mathsf{T}^A$  with b: B and  $\lambda_A b = t$ . In terms of the interpretation, given  $\Gamma, A \vdash B$  and  $\Gamma \vdash t: \Pi_{x:A}B$ , there is a term  $\Gamma, A \vdash t': B$  with  $\lambda_{x:A}t' = t$ , and t' is unique with this property. This is just what is provided by the  $\Pi$ -elimination rule:

$$\frac{\Gamma, A \vdash B \qquad \Gamma \vdash t : \Pi_{x:A}B \qquad \Gamma \vdash x : A}{\Gamma, A \vdash t \, x : B}$$

in conjunction with the  $\Pi$ -computation rules:

$$\lambda_{x:A}(t x) = t : \Pi_A B$$
$$(\lambda_{x:A} b) x = b : B$$

4.4.2 Strictification

4.5 Universes

4.6 Univalence

4.6.1 Function extensionality

4.6.2 The h-levels

## Bibliography

- [AG] C. Angiuli and D. Gratzer. Principles of dependent type theory. Online at https://carloangiuli.com/courses/b619-sp24/notes.pdf. Version 2024-11-26.
- [AW09] Steve Awodey and Michael A. Warren. Homotopy theoretic models of identity types. *Math. Proc. Cambridge Philos. Soc.*, 146(1):45–55, 2009.
- [Awo12] Steve Awodey. Type theory and homotopy. In Peter Dybjer, Sten Lindström, Erik Palmgren, and Göran Sundholm, editors, Epistemology Versus Ontology: Essays on the Philosophy and Foundations of Mathematics in Honour of Per Martin-Löf, pages 183–201. Springer, 2012. arXiv:1010.1810.
- [Awo16] Steve Awodey. Natural models of homotopy type theory. *Mathematical Structures* in Computer Science, 28:1–46, 11 2016.
- [Dyb96] P. Dybjer. Internal type theory. LNCS, 1158:120–134, 1996.
- [Hof95] Martin Hofmann. Syntax and semantics of dependent types. In *Semantics and logics of computation*, volume 14 of *Publ. Newton Inst.*, pages 79–130. Cambridge University Press, Cambridge, 1995.
- [HS98] Martin Hofmann and Thomas Streicher. The groupoid interpretation of type theory. In *Twenty-five years of constructive type theory (Venice, 1995)*, volume 36 of Oxford Logic Guides, pages 83–111. Oxford Univ. Press, New York, 1998.
- [Rij25] Egbert Rijke. Introduction to Homotopy Type Theory. Cambridge University Press, 2025.
- [Uni13] The Univalent Foundations Program. Homotopy Type Theory: Univalent Foundations of Mathematics. https://homotopytypetheory.org/book, Institute for Advanced Study, 2013.