Models of Intensional Dependent Type Theory

Luis Scoccola

Contents

Introduction **1**
Contextual categories 2
Logical structure 5

The coherence problem **6**
Comprehension categories 7
Strictification theorems 10

Univalence **15**
Univalence holds in the simplicial model 17
Univalence in homotopy theory 18

Tribes and Type theoretic fibration categories **19**
Homotopy canonicity 21

Appendix A – Locally cartesian closed categories **24**

Introduction

The goal of these (incomplete) notes is to give an introduction to categorical models of intensional dependent type theory. They are based on a reading course on models of dependent type theory taken at the University of Western Ontario, with Professor Kapulkin.

The article [Hof97] gives a general introduction to the theory of categorical models of (extensional) dependent type theory. It is a good starting point to learn the basics of categorical semantics of type theory.

Let us start by following the presentation given in [Str91, Chapter 1]. First we define a category that captures the essential components of a given (dependent) type theory, and then we abstract some of its structure to get the definition of *contextual category*. In a sense contextual categories abstract just the right amount of structure: the contextual category induced by a type theory (with certain type constructurs) is initial among contextual categories (with structure representing the type constructors), and this justifies calling contextual categories models of type theory.

Contextual categories

Suppose given a type theory $\mathbb{T}$ with *contexts*, which are (possibly empty) finite lists of *type judgements*. A typical context Γ will have the form $(x_0 : A_0, \ldots, x_n : A_n)$. The types of the theory are allowed to depend on previous declared variables, when we want to make this explicit we will write $(x_0 : A_0, x_1 : A_1(x_0), \ldots, x_n : A_n(x_0, \ldots, x_{n-1}))$.

Throughout the rest of this introduction we fix two contexts

$$\Gamma \equiv (x_0 : A_0, x_1 : A_1(x_0), \ldots, x_n : A_n(x_0, \ldots, x_{n-1}))$$
$$\Delta \equiv (y_0 : B_0, y_1 : B_1(y_0), \ldots, y_m : B_n(y_0, \ldots, y_{m-1})).$$

Definition 1. A *morphism of contexts* $f : \Gamma \to \Delta$ consists of a sequence of terms $t_0, \ldots, t_m$ such that

$$\begin{aligned} &\Gamma \vdash t_0 : B_0 \\ &\Gamma \vdash t_1 : B_1(t_0) \\ &\ \vdots \\ &\Gamma \vdash t_m : B_m(t_0, \ldots, t_{m-1}). \end{aligned}$$

We represent such a morphism by $\langle t_0, t_1, \ldots, t_m \rangle$.

A morphism $f : \Gamma \to \Delta$ can be interpreted as a *substitution* as follows. Suppose given a type $\Delta \vdash X$; we can write $X \equiv X(y_0, y_1, \ldots, y_n)$ since X potentially depends on the variables declared in Δ. Using the morphism f we get an induced type $\Gamma \vdash f^*X$ defined by $f^*X :\equiv X(t_0, t_1, \ldots, t_n)$.

Definition 2. Given a type theory $\mathbb{T}$ its *syntactic category* $\mathrm{Cl}(\mathbb{T})$ has as objects the (well formed) contexts in the type theory up to definitional equality, and as morphisms the context morphisms up to definitional equality.

The syntactic category of a type theory is also called the *category of contexts* or the *classifying category*.

We have the following straighforward lemma.

Lemma 3. *The empty context is a terminal object in this category.* ∎

There are some particularly important morphisms in the classifying category of a type theory, which correspond to context extensions (or dependent types). Let us recall the notion of context extension. If we have $\Gamma \vdash X$ we can extend Γ by adding a judgement $x : X$. The extended context has the form $(x_0 : A_0, x_1 : A_1(x_0), \ldots, x_n : A_n(x_0, \ldots, x_{n-1}), x : X)$. We denote the extended context by $\Gamma.X$.

Definition 4. We have a natural context morphism $p_X : \Gamma.X \to \Gamma$ given by

$$\begin{aligned} &\Gamma \vdash x_0 : X_0 \\ &\Gamma \vdash x_1 : X_1(x_0) \\ &\ \vdots \\ &\Gamma \vdash x_n : X_n(x_0, \ldots, x_{n-1}). \end{aligned}$$

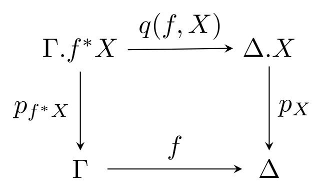
or $\langle x_0, x_1, \ldots, x_n \rangle$. We call these morphisms *display maps*.

One reason why display maps are important is the fact that they classify terms as follows.

Lemma 5. *Given $\Gamma \vdash X$ there is a bijection between terms $\Gamma \vdash x : X$, and sections of the display map $p_X : \Gamma.X \to \Gamma$. The bijection is given by sending a term $\Gamma \vdash x : X$ to the section $\langle x_0, \ldots, x_n, x \rangle$, and a section $\langle x_0, \ldots, x_n, x \rangle$ to the term $\Gamma \vdash x : X$.*

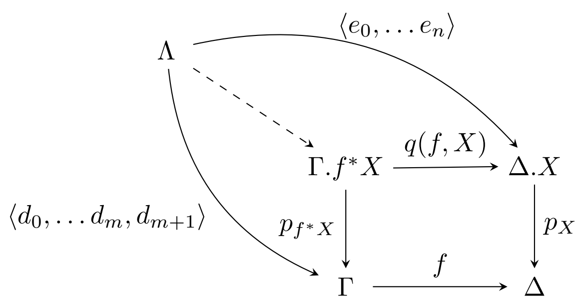
Proof. A section of p_X has the first n terms fixed, so it has the form $\langle x_0, \ldots, x_n, x \rangle$. Moreover, by definition of context morphism, we must have $\Gamma \vdash x : X$, so the correspondence in the statement is well defined. ■

As argued above, context morphisms act as substitutions, so that given an extended context $\Delta.X$ and a context morphism $\langle t_0, \ldots, t_m \rangle : \Gamma \to \Delta$ we can form the extended context $\Gamma.f^*X \equiv (x_0 : A_0, x_1 : A_1, \ldots, x_n : A_n, x : f^*X)$. Moreover, we have a contex morphism $q(f, X) : \Gamma.f^*X \to \Delta.X$ given by $\langle t_0, \ldots, t_m, x \rangle$, since $\Delta \vdash x : X(t_1, \ldots, t_n)$. This gives us a commutative square:



Lemma 6. *The square defined above is a pullback.*

Proof. Given a context Λ and morphisms $\langle e_0, \ldots e_n \rangle : \Lambda \to \Gamma$ and $\langle d_0, \ldots d_m, d_{m+1} \rangle : \Lambda \to \Delta.X$ making the evident square commute we can define a morphism $\Lambda \to \Gamma.f^*X$ by $\langle e_0, \ldots, e_n, d_{m+1} \rangle$, the dashed arrow in the following commutative diagram:



This is the only way we can define the dashed arrow since it must make the diagram commute. ■

This tells us that pullbacks of display maps always exist but moreover that we have canonical representatives for such pullbacks in such a way that the representatives are functorial with respect to the morphisms we are pulling back along.

Let us conclude our analysis by noting the following structure present in the classifying category. We have an grading of the objects given by the length of a context, with the following properties:

- display maps decrease the grading by one;
- the only object with degree zero is the empty context;
- every object with positive degree $n+1$ has a father with degree n, given by forgetting the last judgement in the context;
- for every object with positive degree we have a unique display map between the object and its father.

The definition of contextual category abstracts the structure that we have been discussing.

Definition 7. A *contextual category* is given by:

1. a category $\mathcal{C}$;
2. a grading of the objects $\mathrm{Ob}\,\mathcal{C} = \coprod_{n \in \mathbb{N}} \mathrm{Ob}_n\,\mathcal{C}$;
3. a distinguished object $1 \in \mathrm{Ob}_0\,\mathcal{C}$;
4. maps $ft_n : \mathrm{Ob}_{n+1}\,\mathcal{C} \to \mathrm{Ob}_n\,\mathcal{C}$;
5. for each $X \in \mathrm{Ob}_{n+1}$, a map $p_X : X \to ft_n(X)$ (called the *canonical projection from* X);
6. for each $X \in \mathrm{Ob}_{n+1}$ and $f : Y \to ft_n(X)$, an object $f^*(X)$ together with a map $q(f, X) : f^*(X) \to X$;

such that:

1. 1 is the only object in $\mathrm{Ob}_0\,\mathcal{C}$;
2. 1 is a terminal object of $\mathcal{C}$;
3. for every $n > 0$, $X \in \mathrm{Ob}_n\,\mathcal{C}$, and $f : Y \to ft_n(X)$, we have $ft_n(f^*X) = Y$ and the following square is a pullback:

$$\begin{array}{ccc} f^*X & \xrightarrow{q(f,X)} & X \\ {\scriptstyle p_{f^*X}}\downarrow & & \downarrow{\scriptstyle p_X} \\ Y & \xrightarrow{f} & ft(X) \end{array}$$

called the *canonical pullback* of X along f;

4. the canonical pullbacks are strictly functorial in the sense that:

$$\begin{aligned} 1^*_{ft(X)}(X) &= X \\ q(1_{ft(X)}, X) &= 1_X \\ (fg)^*(X) &= g^*(f^*(X)) \\ q(fg, X) &= q(f, X) \circ q(g, f^*X). \end{aligned}$$

A *morphism between contextual categories* is a functor respecting all the structure on the nose. The category of contextual categories will be denoted by CxlCat.

Contextual categories form an essentially algebraic theory and thus the notions of morphism and equivalence between contextual categories come for free. Notice that this implies that equivalent contextual categories have *isomorphic* underlying categories.

We have the prototypical example:

Example 8. Given a type theory $\mathbb{T}$, its category of contexts has a natural contextual category structure. As argued above, the grading is given by the length of the contexts, the distinguished terminal object is the empty context, the father maps forget the last judgment of the context and the canonical projections are given by the trivial substitution. Pullbacks are given by syntactic substitution.

A less obvious example is given by the following contextual category where types can be understood as sets and dependent types as set-indexed families of sets.

Example 9. We start by defining the objects of the category $\mathcal{C}$ (with their grading) and a map $S : \mathrm{Ob}(\mathcal{C}) \to \mathsf{Set}$ needed for the construction, by induction:

- The unique object of degree zero is the empty set $\emptyset$ and $S(\emptyset) := \{\emptyset\}$.
- An object of degree $n+1$ is a pair (Δ, F) with Δ an object of degree n and $F : S(\Delta) \to \mathsf{Set}$ a family of sets indexed by $S(\Delta)$. We define $S(\Delta, F) := \{(x, y) \mid x \in S(\Delta), y \in F(x)\}$.

Morphisms between Γ and Δ are exactly functions between the sets $S(\Gamma) \to S(\Delta)$. Fathers are given by projecting the first component $(\Delta, F) \mapsto \Delta$, and display maps are also given by projecting the first component $(x, y) \mapsto x$. Given an object (Δ, F) and a morphism $f : \Gamma \to \Delta$ we define $f^*(\Delta, F) := (\Gamma, F \circ f)$ and $q(f, (\Delta, F))(x, y) := (f(x), y)$. The fact that this gives a functorial choice of pullbacks is a straightforward check.

As mentioned at the very beginning, contextual categories are models of type theory in the following precise sense. For a proof see [Car78].

Theorem 10. *If $\mathbb{T}$ is the dependent type theory given only by structural rules, the category $\mathrm{Cl}(\mathbb{T})$ is initial among contextual categories.* ∎

This theorem is a correctness theorem: contextual categories correcly interpret the syntax of type theory (at least the syntax given by the structural rules). On the other hand we have an obvious completeness theorem: if every model of type theory has some property, then the type theory itself must have it, since the classifying category of the theory encondes its syntax (up to definitional equality).

Logical structure

Useful type theories usually have some logical structure on top of the structural rules. Some examples are Σ-types, Π-types, Id-types, W-types, etc. This logical structure is given by extra rules, which usually consist of a formation rule, an introduction rule, an elimination rule and a computation rule. For example in the case of Σ-types we have.

Definition 11. A type theory with Σ-*types* has the following rules.

$$\frac{\Gamma \vdash A \qquad \Gamma, x : A \vdash B(x)}{\Gamma \vdash \Sigma_{x:A} B(x)} \ \Sigma\text{−FORM}$$

$$\frac{\Gamma \vdash A \qquad \Gamma, x : A \vdash B(x)}{\Gamma, x : A, y : B(x) \vdash \langle x, y \rangle_{A,B} : \Sigma_{x:A} B(x)} \ \Sigma\text{−INTRO}$$

$$\frac{\Gamma, z : \Sigma_{x:A} B(x) \vdash C(z) \qquad \Gamma, x : A, y : B(x) \vdash m(x,y) : C(\langle x, y \rangle_{A,B})}{\Gamma, z : \Sigma_{x:A} B(x) \vdash \mathsf{ind}_m(z) : C(z)} \ \Sigma\text{−ELIM}$$

$$\frac{\Gamma, z : \Sigma_{x:A} B(x) \vdash C(z) \qquad \Gamma, x : A, y : B(x) \vdash m(x,y) : C(\langle x, y \rangle_{A,B})}{\Gamma, x : A, y : B(x) \vdash \mathsf{ind}_m(\langle x, y \rangle_{A,B}) \equiv m(x,y) : C(\langle x, y \rangle_{A,B})} \ \Sigma\text{−COMP}$$

When finding models for such a type theory we will want to interpret Σ-types. By inspecting the structure that this type former induces on the classifying category of the type theory one abstracts the following definition.

Definition 12. A Σ-*type structure* on a contextual category $\mathcal{C}$ consists of:

1. for each $(\Gamma, A, B) \in \mathrm{Ob}_{n+2}\,\mathcal{C}$ an object $(\Gamma, \Sigma(A, B)) \in \mathrm{Ob}_{n+1}\,\mathcal{C}$ (the formation rule);
2. for each (Γ, A, B), a morphism $\langle , \rangle_{A,B} : (\Gamma, A, B) \to (\Gamma, \Sigma(A, B))$ that commutes with the canonical projections to Γ (the introduction rule);
3. for each (Γ, A, B), each $(\Gamma, \Sigma(A, B), X)$ and each morphism $m : (\Gamma, A, B) \to (\Gamma, \Sigma(A, B), X)$ such that $p_X \circ m = \langle , \rangle_{A,B}$, a morphism $\mathsf{ind}_m : (\Gamma, \Sigma(A, B)) \to (\Gamma, \Sigma(A, B), X)$ (the elimination rule) such that $\mathsf{ind}_m \circ \langle , \rangle_{A,B} = m$ (the computation rule);

such that for every morphism $f : \Delta \to \Gamma$ we have

$$f^*(\Gamma, \Sigma(A, B)) = (\Delta, \Sigma(f^*A, f^*B))$$
$$f^*\langle , \rangle_{A,B} = \langle , \rangle_{f^*A, f^*B}$$
$$f^*\mathsf{ind}_m = \mathsf{ind}_{f^*m}.$$

Again we have a correctness result (see [Str91] for a proof a particular case):

Theorem 13. *Given $\mathbb{T}$ a dependent type theory with the usual structure rules plus some combination of the logical rules for Σ, Π, W and Id, then its classifying category $\mathrm{Cl}(\mathbb{T})$ is initial among contextual categories with the corresponding structure.* ■

The coherence problem

The first problem one encounters when modelling type theory is that usual categories that one wishes to use as models do not have a contextual category structure. The solution is to consider different kinds of structured categories that are more likely to “appear in nature”, and then find general approaches that let us construct contextual categories out of this more natural structured categories.

Why not work directly with these more natural categories? One problem is that the canonical pullbacks have to be striclty functorial, so that the mere existence of pullbacks is not enough to obtain canonical pullbacks. Even more problematic is the logical structure needed to interpret logical constructors. Concretelly, given a type former, say the coproduct $+$, and a substitution f^*, in the syntactic category of a type theory we have $f^*(A+B) = f^*(A) + f^*(B)$. That is to say, substitution is strictly functorial with respect to the logical structure. Since this is a property of the syntax it has to be a property of every model. On the other hand, in categorical semantics, type formers are usually interpreted using universal properties (or even weaker definitions), which guarantee substitution to be functorial (at best) up to canonical isomorphism. Models satisfying only this weak stability can be thought as weak models. The problem of strictifying weak models is known as the *coherence problem*.

In our case weak models will be *comprehension categories*. The process of strictifying a weak model will take the form of a functor from the category of comprehension categories to the category of contextual categories. To understand the constructions involved it is useful to introduce a third kind of model, living in between: *categories with attributes*. The situation looks like this:

$$\text{Contextual Categories} \longleftarrow \text{Categories with attributes} \longleftarrow \text{Comprehension Categories}$$

$$\text{strict stability} \leftarrow \text{- -} \rightarrow \text{weak stability}$$

References for this section are:

- [Jac93] for an introduction to comprehension categories as models of dependent type theory.
- [KL12] for an introduction to contextual categories and the first general solution to the coherence problem for models of intensional dependent type theory (due to Voevodsky).
- [LW15] for a generalization of Voevodsky's universes approach for solving the coherence problem. This is done in the framework of comprehension categories.
- [AHR99] for an introduction to essentially algebraic theories and their relation to locally presentable categories.

Comprehension categories

Let us start by defining comprehension categories and then motivating the definition.

Definition 14. A *cleaving* for a (Grothendieck) fibration $p : \mathcal{T} \to \mathcal{C}$ consists of a choice function that given $t \in \mathcal{T}$ and $u : b \to p(t)$ in $\mathcal{C}$, it returns a P-cartesian lift $f : t' \to t$ of u.

A cleaving is *normal* if it lifts identities to identities. A cleaving is *split* if it is normal and the lifts of compositions are the compositions of the lifts, this means that the lifts are functorial.

Definition 15. A *comprehension category* is given by:

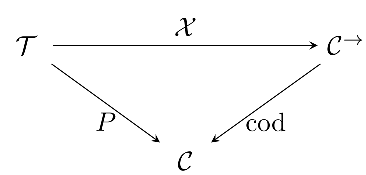
1. a category $\mathcal{C}$;
2. a functor $\mathcal{X} : \mathcal{T} \to \mathcal{C}^{\to}$ (the *comprehension*);

such that:

1. $P := \mathrm{cod} \circ \mathcal{X} : \mathcal{T} \to \mathcal{C}$ is a Grothendieck fibration;
2. for every P-cartesian morphism f, the square $\mathcal{X}f$ is a pullback in $\mathcal{C}$.

A comprehension category is *full* if $\mathcal{X}$ is fully faithful, and *cloven* if P is. *We will assume that all our comprehension categories are full and cloven.*

Notation 16. When working with a comprehension category $(\mathcal{C}, \mathcal{X})$ we will usually denote the composition $\mathrm{cod} \circ \mathcal{X} : \mathcal{T} \to \mathcal{C}$ by P, so that a comprehension category can be pictured as a commutative diagram:



As we did in the previous section we will interpret contexts by objects in a category $\mathcal{C}$ and substitutions by morphisms. Types in contexts are interpreted using the fibration $P : \mathcal{T} \to \mathcal{C}$: a type X in context Γ is an object $X \in \mathcal{T}$ such that $P(X) = \Gamma$. The comprehension $\mathcal{X} : \mathcal{T} \to \mathcal{C}^{\to}$ interprets the display maps as follows.

Definition 17. Given a comprehension category $(\mathcal{C}, \mathcal{X})$ and an object $\Gamma \in \mathcal{C}$ we write $\mathcal{T}_\Gamma$ for its fiber $P^{-1}(\Gamma)$. Given an object $\sigma \in \mathcal{T}_\Gamma$ the object $\mathcal{X}(\sigma) \in \mathcal{C}^{\to}$ is a map in $\mathcal{C}$ that we denote by $p_\sigma : \Gamma.\sigma \to \Gamma$. These maps are called *display maps*.

Substitution is interpreted by the cleaving: the action of a substitution $f : \Gamma \to \Delta$ on a type X in context Δ is given by the source of the P-cartesian lift of f given by the cleaving.

Definition 18. A *category with attributes* is a full split comprehension category. Categories with attributes form an essentially algebraic theory, morphism between comprehension categories are functors between the underlying categories together with a functor between the total spaces commuting with the cleaving. The category of categories with attributes will be denoted by CwA.

We will assume that our categories with attributes come with a distinguished terminal object.

Definition 19. Non necessarily split comprehension categories also form a category $\mathsf{CompCat}$. In this case the morphisms are a little bit more subtle but we won't need the precise definition.

For completeness we mention that in the literature one can find category with attributes defined as a (non necessarily full) comprehension category, such that the fibration p is a discrete fibration. The two definitions are equivalent: given a full split comprehension category forget the arrows in the fibers to obtain a discrete fibration; given a discrete fibration construct a full comprehension category by adding all the arrows present in the target category.

Notice that a split cleaving provides strictly associative substitution. In that sense categories with attributes are strict models. We will now construct adjoint functors $\mathcal{F} : \mathsf{CxlCat} \rightleftarrows \mathsf{CwA} : \mathsf{core}$ in such a way that $\mathcal{F}$ applied to the classifying category of a type theory is still initial.

Definition 20. We construct a functor $\mathcal{F} : \mathsf{CxlCat} \to \mathsf{CwA}$. The underlying category and the distinguished terminal object are the same. The total space $\mathcal{T}$ has as objects pairs (Γ, σ) such that $ft(\sigma) = \Gamma$. Morphisms $(\Delta, \delta) \to (\Gamma, \sigma)$ are exactly morphisms $\Delta.\delta \to \Gamma.\sigma$ in $\mathcal{C}$. The functor $\mathcal{X}$ sends (Γ, σ) to p_σ. The split cleaving is given by the canonical pullbacks of $\mathcal{C}$ and this also implies that $\mathcal{X}$ sends cartesian morphisms to pullbacks.

Given a category with attributes we can ask if there is some grading that makes it into a contextual category, and moreover, if there can be many such gradings.

Proposition 21. *The functor $\mathcal{F}$ is fully faithful, i.e. having a grading is a property of a category with attributes, not extra structure.*

Proof. Having a grading means that for each $\Gamma \in \mathcal{C}$ either Γ is a terminal object or we have a choice of an object $\Delta \in \mathcal{C}$ and a $\sigma \in \mathcal{T}_\Delta$ such that $\Delta.\sigma = \Gamma$ and such that every $\Gamma \in \mathcal{C}$ has finite length. We can prove that having a grading is a property by induction on the length. The base case is immediate. For the inductive case, if for a given Γ we have two choices of $\Delta \in \mathcal{C}, \sigma \in \mathcal{T}_\Delta$ and $\Delta' \in \mathcal{C}, \sigma \in \mathcal{T}_{\Delta'}$ then $\Delta.\sigma = \Gamma = \Delta'.\sigma'$ and thus $\Delta = p(\Delta.\sigma) = p(\gamma) = p(\Delta'.\sigma') = \Delta$. And by inductive hypothesis the grading of Δ is unique. ■

The following construction provides a right adjoint to the functor $\mathcal{F}$.

Definition 22. We now construct the functor $\mathsf{core} : \mathsf{CwA} \to \mathsf{CxlCat}$. Suppose given a category with attributes, consisting of a category $\mathcal{C}$ a functor $\mathcal{X} : \mathcal{T} \to \mathcal{C}^{\to}$ and a distinguished terminal object 1. The underlying category of $\mathsf{core}(\mathcal{C})$ has as objects finite lists $(A_0, A_1, \cdots, A_n)$ such that $A_0 \in \mathcal{T}_1$, with 1 the distinguished terminal object of $\mathcal{C}$, and $A_{k+1} \in \mathcal{T}_{1.A_0.\cdots.A_k}$. The morphisms $(A_0, \cdots, A_n) \to (B_0, \cdots, B_m)$ are exactly the morphisms $1.A_0.A_1.\cdots.A_n \to 1.B_0.B_1.\cdots.B_m$ in $\mathcal{C}$. The terminal object is the empty list.

The grading is given by the length of the list and the father maps forget the last element. The canonical projections are the projections of Definition 17. The canonical pullbacks are given by the cleaving and the fact that $\mathcal{X}$ sends cartesian morphisms to pullback squares. Since the cleaving is split the canonical pullbacks are functorial.

As a remark observe that while the functor $\mathcal{F}$ preserves the underlying category, the functor core may, in general, add "the same object many times". To avoid this we have to make sure that every object either is the distinguished terminal object, or it has a unique way of reaching the terminal object by writing it as a composite $1.\cdots.\sigma$. If the condition holds, then $|\mathsf{core}(\mathcal{C})| \simeq \mathcal{C}$. But it turns out that this condition is not essentially algebraic so the essential image of $\mathcal{F}$ does not behave as nicely as CwA.

Proposition 23. *The functor* core *is right adjoint to the functor $\mathcal{F}$.*

Proof. Fix a contextual category $\mathcal{C}$ and a category with attributes $(\mathcal{A}, \mathcal{X} : \mathcal{T} \to \mathcal{A}^{\to})$.

Given a morphism of categories with attributes $m : \mathcal{F}(\mathcal{C}) \to \mathcal{A}$ represented by a functor G between the underlying categories and a functor T between the total spaces, we define the contextual functor $\hat{m} : \mathcal{C} \to \mathsf{core}(\mathcal{A})$ first on objects, by induction on the degree of $\Gamma \in \mathcal{C}$. If Γ is the unique object of degree 0 we send it to the empty list in $\mathsf{core}(\mathcal{A})$. If Γ has positive degree $n+1$ we consider the display map $\Gamma \to ft(\Gamma)$ in $\mathcal{C}$. On one hand, by applying G to the display map we obtain a map $G(\Gamma) \to G(ft(\Gamma))$. and by inductive hypothesis we have an object $\hat{m}(ft(\Gamma))$ represented by a list $(A_0, \dots, A_n)$. On the other hand, by definition of the functor $\mathcal{F}$, the display map is represented by an object t in the total space of $\mathcal{F}(\mathcal{C})$. Applying the functor T to t we get an object A_{n+1} over $G(ft(c))$. We then define $\hat{m}(\Gamma) := (A_0, \dots, A_n, A_{n+1})$. Since the morphisms between $(A_0, \dots, A_n)$ and $(B_0, \dots, B_l)$ in $\mathsf{core}(\mathcal{A})$ correspond to morphisms between $1.A_0.\cdots.A_n$ and $1.B_0.\cdots.B_l$ in $\mathcal{A}$ the action of $\hat{m}$ on morphisms is clear.

Now assume given $\hat{m}$ a contextual morphism between $\mathcal{C}$ and $\mathsf{core}(\mathcal{A})$. We first define functor G from the underlying category of $\mathcal{F}(\mathcal{C})$ (which is the underlying category of $\mathcal{C}$) to the underlying category of $\mathcal{A}$. Given $\Gamma \in \mathcal{C}$, consider $\hat{m}(c) = (A_0, \dots, A_n)$ and define $G(\Gamma) := A_0.\cdots.A_n$. The action of G on morphisms is given by the fact that morphisms between $(A_0, \dots, A_n)$ and $(B_0, \dots, B_l)$ in $\mathsf{core}(\mathcal{A})$ correspond to morphisms between $1.A_0.\cdots.A_n$ and $1.B_0.\cdots.B_l$ in $\mathcal{A}$. Next we must define a functor T between the total spaces. Recall that the total space of $\mathcal{F}(\mathcal{C})$ has as objects pairs (Γ, σ) such that $\Gamma, \sigma \in \mathcal{C}$ and $ft(\sigma) = \Gamma$. Consider the morphism $\hat{m}(p_\sigma)$ in $\mathsf{core}(\mathcal{A})$. Since it is a display map it has the form $\hat{m}(p_\sigma) : (A_0, \dots, A_n, A_{n+1}) \to (A_0, \dots, A_n)$. Then define $T(\Gamma, \sigma) := A_{n+1} \in \mathcal{T}$. For a morphisms $f : (\Delta, \delta) \to (\Gamma, \sigma)$ in the total space of $\mathcal{F}(\mathcal{C})$ recall that this corresponds to a morphisms $f : \Delta.\delta \to \Gamma.\sigma$ in $\mathcal{C}$. Applying $\hat{m}$ to it and using the definition of morphisms in $\mathsf{core}(\mathcal{A})$ we get a morphism $1.A_0.\cdots.A_n \to 1.B_0.\cdots.B_l$. Using the fact that $\mathcal{X}$ is fully faithful we get a unique morphism $A_n \to B_l$ in $\mathcal{T}$. ■

Since $\mathcal{F}$ is a left adjoint it preserves initial objects, which implies that if $\mathbb{T}$ is the dependent type theory with just the structural rules, the category with attributes $\mathcal{F}(\mathrm{Cl}(\mathbb{T}))$ is initial. This tells us that categories with attributes model type theory in the same sense that contextual categories do. But moreover categories with attributes are a little bit more natural, since we got rid of the grading.

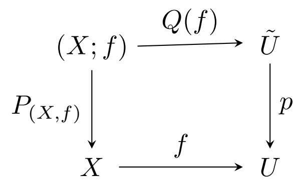
Once again, we are not interested just in comprehension categories, but we wish to have some logical structure on them. As an example we describe the case of dependent sums.

Strictification theorems

The first general strictification theorem is due to Hofmann and provides a functor $(-)_* :$ $\mathsf{CompCat} \to \mathsf{CwA}$ as described in [Hof95] and [CGH14]. The construction behaves well when modeling *extensional* type theory, but it doesn't leave space for non trivial identity types. So for modelling *intensional* type theory one needs to find another strictification. This problem was solved by Lumsdaine and Warren ([LW15]). They define a new strictification functor $(-)_! :$ $\mathsf{CompCat} \to \mathsf{CwA}$ that behaves nicely with respect to identity types. The idea is inspired by Voevodsky's universes method which we now describe.

Given any category with a universe the construction gives back a contextual category. We start with the definition of universe.

Definition 24 ([Voe15, Definition 2.1])**.** Given a category $\mathcal{C}$, a *universe* in $\mathcal{C}$ is a morphism $\tilde{U} \to U$ in $\mathcal{C}$ together with, for every morphism $f : X \to U$ a choice of pullback square:



Usually one refers to the universe as U, leaving the rest of the structure implicit.

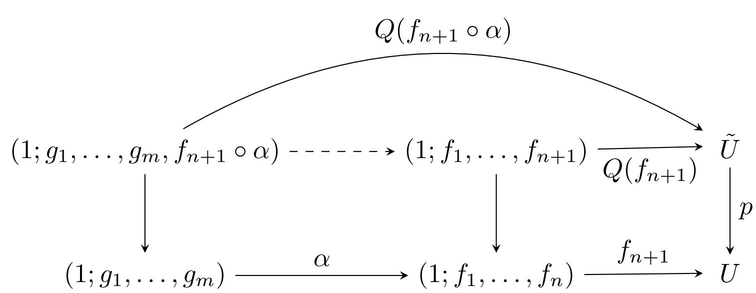
Observe that we can form successive pullbacks inductively: given a two morphisms $f_1 : X \to U$, $f_2 : (X; f_1) \to U$ we can consider the object $((X; f_1); f_2)$.

Notation 25. Given a sequence of n morphisms $f_1, \dots, f_n$ of this form we denote the iterated pullback by $(X; f_1, \dots, f_n)$.

The contextual category is then given by pulling back the objects from U. Concretely:

Definition 26 ([Voe15, Construction 2.12])**.** Given a category $\mathcal{C}$ together with a universe U and a terminal object 1 we define the contextual category $\mathcal{C}_U$ as follows:

- $\mathrm{Ob}_n\, \mathcal{C}_U := \{(f_1, \dots, f_n) \mid f_i : (1; f_1, \dots, f_{i-1}) \to U, 1 \leq i \leq n\}$;
- The morphisms between $(f_1, \dots, f_n)$ and $(g_1, \dots, g_m)$ in $\mathcal{C}_U$ are exactly the morphisms between $(1; f_1, \dots, f_n)$ and $(1; g_1, \dots, g_m)$ in $\mathcal{C}$;
- The distinguished terminal object in $\mathcal{C}_U$ is $()$, the empty sequence;
- Fathers are given by forgetting the last component $ft(f_1, \dots, f_{n+1}) := (f_1, \dots, f_n)$;
- The projection $p_(f_1, \dots, f_{n+1}) : (f_1, \dots, f_{n+1}) \to (f_1, \dots, f_n)$ is the morphism $P_{(1; f_1, \dots, f_{n+1})}$ given by the universe structure on U;
- Given an object $(f_1, \dots, f_{n+1})$ and a morphism $\alpha : (g_1, \dots, g_m) \to (f_1, \dots, f_n)$ in $\mathcal{C}_U$, the canonical pullback $\alpha^*(f_1, \dots, f_{n+1})$ is the object $(g_1, \dots, g_m, f_{n+1} \circ \alpha)$ in the diagram:



and the corresponding projection is the dashed arrow, induced by the pullback on the right.

A straightforward check shows that this defines in fact a contextual category. Moreover the contextual category $\mathcal{C}_U$ depends only on $p : \tilde{U} \to U$ and not on the choice of pullbacks and terminal object, up to canonical isomorphism. The key is that strict functoriality of the data is attained by using the composition in $\mathcal{C}$, which is strictly associative.

Another pleasant property is the following:

Proposition 27 ([Voe15, Construction 5.2]). *Given a small contextual category $\mathcal{C}$ we can consider the universe U in the presheaf category $[\mathcal{C}^{op}, \mathsf{Set}]$ given by:*

$$U = \{Y \mid ft(Y) = X\}$$
$$\tilde{U}(X) = \{(Y, s) \mid ft(Y) = X, s \text{ is a section of } p_Y\}$$

and the evident projection map. Then $[\mathcal{C}^{op}, \mathsf{Set}]_U$ and $\mathcal{C}$ are isomorphic as contextual categories. ■

The really useful aspect of this construction is that structure on the morphism $p : \tilde{U} \to U$ translates to logical structure on the contextual category $\mathcal{C}_U$. For these constructions it is convenient to assume that $\mathcal{C}$ is locally cartesian closed.

Definition 28. A *universe* in a contextual category $\mathcal{C}$ is an object $(\mathsf{U}, \mathsf{El}) \in \mathrm{Ob}_2\, \mathcal{C}$.

We illustrate this definition in Example 36 and Example 37. If the contextual category has some logical structure one usually asks for some compatibility between the logical structure and the universe. In the case of Σ we have:

Definition 29. In a contextual category $\mathcal{C}$ with a Σ-structure we say that a universe $(\mathsf{U}, \mathsf{El})$ is *closed under Σ-types* if for every $a : \Gamma \to \mathsf{U}$ and $b : (\Gamma, a^* \mathsf{El}) \to \mathsf{U}$ we have $\sigma(a, b) : \Gamma \to \mathsf{U}$ (the type constructor) such that $(\Gamma, \sigma(a, b)^* \mathsf{El}) = (\Gamma, \Sigma(a^* \mathsf{El}, b^* \mathsf{El}))$ (compatible with the Σ-structure) and such that for every substitution $f : \Gamma' \to \Gamma$ we have $f^*(\sigma(a, b)) = \sigma(f^* a, f^* b)$ (stable under substitution).

Definition 30. We now describe the functor $(-)_! : \mathsf{CompCat} \to \mathsf{CwA}$ on objects. Given a comprehension category $(\mathcal{C}, \mathcal{T}, \mathcal{X})$ we define a new comprehension category $(\mathcal{C}_!, \mathcal{T}_!, \mathcal{X}_!)$ where $\mathcal{C}_! := \mathcal{C}$. The objects in $\mathcal{T}_!$ over $\Gamma \in \mathcal{C}$ consist of tuples $(V_A, E_A, \{A\})$, where V_A is an object of $\mathcal{C}$, E_A is and object of $\mathcal{T}(V_A)$ and $\{A\}$ is a morphism $\Gamma \to V_A$ in $\mathcal{C}$. We write $[A]$ for the reindexing of E_A along $\{A\}$.

Morphisms $(V_B, E_B, \{B\}) \to (V_A, E_A, \{A\})$ in $\mathcal{T}_!$ over $m : \Delta \to \Gamma$ are maps $[B] \to [A]$ in $\mathcal{T}$ over m.

This already gives the projection $p_! = \mathcal{T}_! \to \mathcal{C}_!$, sending $(V_B, E_B, \{B\})$ to the domain of $\{B\}$.

Given a morphism $m : \Delta \to \Gamma$ in $\mathcal{C}$ and $(V_A, E_A, \{A\})$ an object in $\mathcal{T}_!$ over Γ, we define $A[m] := (V_A, E_A, \{A\} \circ m)$. Then take $A_m : A[m] \to A$ to be the canonical map $[A[m]] \to [A]$ over m in $\mathcal{T}$ given by the fact that $(E_A)_{\{A\}}$ is P-cartesian. This choice makes $p_!$ a split fibration.

Finally we can send an object $(V_A, E_A, \{A\})$ to $\mathcal{X}(\{A\})$ which gives us $\mathcal{X}_! : \mathcal{T}_! \to \mathcal{C}_!^{\to}$. The fact that $\mathcal{X}_!$ sends P-cartesian morphisms to pullback squares follows from the fact that $\mathcal{X}$ has this property.

Using this construction we get a very general method for constructing models of intensional dependent type theory. To state the theorem we need the following definition.

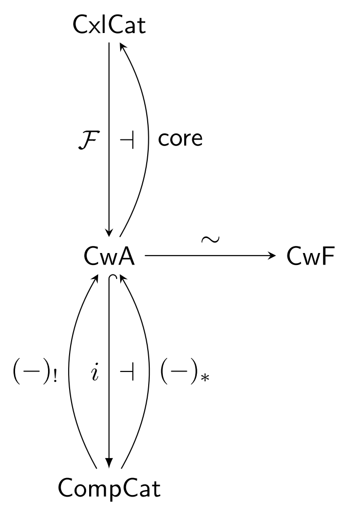
Definition 31. A[1] comprehension category $\mathcal{C}$ satisfies *condition (LF)* if its underlying category has finite products, and given composable morphisms $Z \xrightarrow{g} Y \xrightarrow{f} X$, if f is a display map and g is either a display map or a product projection, then the dependent product[2] $\Pi[f, g]$ exists.

[1] Cite LW in the definition?

[2] They say "dependen exponential", why?

Theorem 32 ([LW15])**.** *Let $\mathcal{C}$ be a comprehension category satisfying condition (LF). Then $\mathcal{C}$ is equivalent to a split comprehension category $\mathcal{C}_!$, and if $\mathcal{C}$ has weakly Σ-types (resp. Π-types,* Id*-types,* W*-types, ...) then $\mathcal{C}_!$ has strictly stable Σ-types (resp. Π-types,* Id*-types,* W*-types, ...).*

The rough diagram given at the beginning of this section can now be refined to the following diagram of categories and functors that relates the different categorical models that we presented.

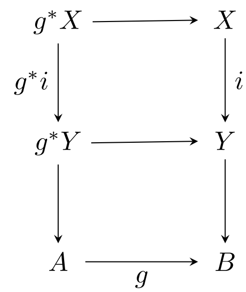


As an application of the theorem we construct a model in the category of simplicial sets. We start by recalling some results from category theory.

Lemma 33. *Given an adjunction between categories with weak factorization systems, then the left adjoint preserves the left class if and only if the right adjoint preserves the right class.* ∎

Proposition 34. *If $g : A \to B$ is a fibration between simplicial sets then the adjunction $g^* :$ sSet $/B \rightleftarrows$ sSet $/A : g_*$ is a Quillen adjunction.*

Proof. We have to prove that g^* preserves cofibrations (equiv. g_* preserves acyclic fibrations) and acyclic cofibrations (equiv. g_* preserves fibrations). The equivalences follow from Lemma 33. Recall that the cofibrations in sSet are the monomorphisms, and that monomorphisms are stable under pullback. So given a cofibration $i : X \to Y$ in sSet $/B$ we can consider the diagram:



given by pulling back X and Y along g and using the universal property of pullbacks to get the map $g^*i : g^*X \to g^*Y$, which is exactly the action of the functor g^* on morphisms. Since both the bottom square and the outer square are pullbacks, the upper square is a pullback, by the two pullback lemma. This implies that g^*i is also a monomorphism, and thus a cofibration.

In our case, to prove that g^* respects trivial cofibrations it is enough to show that it respects weak equivalences. Recall that sSet is a right proper model category, which means that weak equivalences are stable under pullback along fibrations. Considering an analogous diagram to the one above but with a weak equivalence w instead of a cofibration i, we see that the horizontal arrows are fibrations, since they are closed under pullback. But then the map g^*w is a weak equivalence by right properness, since the upper square is a pullback. ■

We are now ready to construct the model.

Example 35. Consider the comprehension category $\mathsf{Kan}^{\twoheadrightarrow} \to \mathsf{Kan}$, where Kan is the full subcategory of sSet spanned by Kan complexes and $\mathsf{Kan}^{\twoheadrightarrow}$ is the full subcategory of $\mathsf{Kan}^{\to}$ consisting of Kan fibrations. To see that this is a (Grothendieck) fibration notice that fibrations are stable under pullback (which happens in any model category). Notice that by definition display maps are exactly the fibrations. To get a model of Π, Σ, Id we want $\mathsf{fib} \to \mathsf{sSet}$ to satisfy condition (LF). So suppose given maps $Z \xrightarrow{g} Y \xrightarrow{f} X$ such that f is a display map and g is either a display map or a product projection. If g is a display map, then it is a fibration, and since it is a product projection of a product between fibrant objects, then it is also a fibration. By the proposition above, the dependent product functor preserves fibrations. Being a right adjoint it also preserves the terminal object, and hence it preserves fibrant objects. This implies that the dependent product $\Pi_f g$ is fibrant, so it lives inside our category Kan, which proves that (LF) holds.

Finally we have to exhibit weak Πs, Σs and Ids. For Π and Σ we use the internal versions, that is, the left and right adjoints to the pullback functors. Since these are characterized by universal property weak stability is guaranteed.

Since we have Π-types we don't need to consider *strong* identity types, it suffices to exhibit weak identity types as in [LW15, Definition 2.3.1]. Given a fibration $p : E \to B$ we define a fibration $(s,t) : P_B(E) \to E \times_B E$ together with a factorization $E \xrightarrow{r} P_B(E) \xrightarrow{(s,t)} E \times_B E$ of the diagonal map $E \to E \times_B E$.

The complex $P_B(E)$ is given by the pullback:

$$\begin{array}{ccc} P_B(E) & \xrightarrow{c'} & E^{\Delta^1} \\ \downarrow & & \downarrow p^{\Delta^1} \\ B & \xrightarrow[ct]{} & B^{\Delta^1} \end{array}$$

The fibration $(s,t) : P_B(E) \to E \times_B E$ is given by the fibered product of the maps $P_B(E) \to E$ given by composing c' with the source and target maps $s,t : E^{\Delta^1} \to E$. Applying the universal property of pullbacks to the maps $ct : E \to E^{\Delta^1}$ and $p : E \to B$ we get the fibration $r : E \to P_B(E)$. The fact that $E \xrightarrow{r} P_B(E) \xrightarrow{(s,t)} E \times_B E$ is equal to the diagonal $E \to E \times_B E$ follows immediately from the constructions.

The above construction gives in fact a functorial factorization of the diagonal map as a trivial cofibration followed by a fibration.The existence of j, given d, is then given by lifting property of r.

Finally weak stability is given by the fact that the pullback functors preserve trivial cofibrations.

A similar (and simpler) application of the theorem gives us models in Set and in Gpd. In these it is easy to define internal universes in the sense of Definition 28. We will come back to these examples when discussing univalence.

Example 36. Consider the category Set together with the subobject classifier $\tilde{\mathcal{U}} \to \mathcal{U}$ defined by $\{\emptyset\} \to \{\emptyset, *\}$, where $*$ is some singleton set, say $\{\emptyset\}$. For every map $f : X \to \mathcal{U}$ we can consider the subset $f^{-1}(*)$ and the inclusion $f^{-1}(*) \to X$. This gives us a choice of pullback $f^*(p)$ for every $f : X \to \mathcal{U}$ and thus it defines an internal universe for the category Set. Notice that we can interpret this as the "universe of propositions" of the model Set, since any proposition P about a set X corresponds exactly to a map $[P] : X \to \mathcal{U}$ such that $[P](x) = *$ if and only if $P(x)$ holds.

Example 37. We can perform an analogous construction in the category Gpd. Here we define $\mathcal{U}$ to be the groupoid with small sets as objects and bijections as arrows, and $\tilde{\mathcal{U}}$ to be the groupoid with small pointed sets as objects and point preserving bijections as arrows. The projection $\tilde{\mathcal{U}} \to \mathcal{U}$ forgets the pointing. Again we can choose a pullback of $\tilde{\mathcal{U}}$ for each groupoid morphism $X \to \mathcal{U}$, using for example the category of elements construction. We know that for a given groupoid X we have a bijective correspondence between groupoid morphisms $X \to \mathcal{U}$ and discrete fibrations $Y \to X$ with small fibers, up to canonical isomorphism. Then we can interpret the universe $\tilde{\mathcal{U}} \to \mathcal{U}$ as the "universe of small sets" since, as we argued above, any X-indexed family of small sets corresponds exactly to one groupoid morphism $X \to \mathcal{U}$.

Univalence

As we will see, univalence is a property of fibrations. As many homotopical concepts it is easy to define in a synthetic homotopy theory such as HoTT. Let us start by recalling the definition of equivalence between types. For a discussion about alternative definitions of equivalence see [Pro13, Chapter 4].

Definition 38. Given a map between types $f : A \to B$ an *equivalence structure* for f is given by:

- a map $g^l : B \to A$, together with a homotopy $h^l : g^l \circ f \sim 1_A$, exhibiting g^l as a left inverse of f.
- a map $g^r : B \to A$, together with a homotopy $h^r : f \circ g^r \sim 1_B$, exhibiting g^r as a right inverse of f.

The type of proofs that a map $f : A \to B$ is an equivalence is then defined by:

$$\mathsf{isEquiv}(f) :\equiv \left(\sum_{g^l : B \to A} g^l \circ f \sim 1_A \right) \times \left(\sum_{g_r : B \to A} f \circ g^r \sim 1_B \right)$$

where $\sim$ denotes pointwise equality. The *type of equivalences* between two types A and B is defined by:

$$\mathsf{Equiv}(A, B) :\equiv \sum_{f : A \to B} \mathsf{isEquiv}(f).$$

A key property of this definition is that for any map f the type $\mathsf{isEquiv}(f)$ is a proposition. This means that any two proofs that f is an equivalence are homotopical.

Suppose given a type family $P : A \to \mathcal{U}$. Then, by induction on identity types we can construct a map

$$w_P : \prod_{x,y:A} \mathsf{Id}_A(x, y) \to \mathsf{Equiv}(P(x), P(y))$$

by mapping refl_x to the identity map of $P(x)$, together with the natural proofs that an identity map is an equivalence. Then a fibration is univalent when it classifies the identity types of the base space in the following precise sense.

Definition 39. A type family $P : A \to \mathcal{U}$ is *univalent* if $w_P(x, y)$ is an equivalence for every $x, y : A$. Formally we define:

$$\mathsf{IsUnivalent}(P) :\equiv \prod_{x,y:A} \mathsf{isEquiv}(w_P(x, y)).$$

A fibration $p : B \to A$ is *univalent* if the corresponding type family $p^{-1} : A \to \mathcal{U}$ is univalent.

Now suppose our type theory has a universe type $\mathcal{U}$, and consider the type family $1_{\mathcal{U}} : \mathcal{U} \to \mathcal{U}$. The *Univalence Axiom* says that this type family is univalent. The axiom provides a characterization of the identity types of the universe $\mathcal{U}$, since it implies that the type of equalities between two types $A, B : \mathcal{U}$, seen as inhabitants of the universe $\mathcal{U}$, is equivalent to the type of equivalences between them $\mathsf{Equiv}(A, B)$.

Notice that in the classifying category of the type theory, if we have $\Gamma \vdash A$, then we defined:

$$\Gamma.x, y : A \vdash w_P(x, y) : \mathsf{Id}_A(x, y) \to \mathsf{Equiv}(P(x), P(y)).$$

Then P is univalent if we have $\Gamma.x, y : A \vdash \mathsf{isEquiv}(w_P(x, y))$. This means that $\mathsf{IsUnivalent}(P)$ is exactly the space of sections of the display map:

$$\Gamma.(x, y : A).\, \mathsf{isEquiv}(w_P(x, y)) \to \Gamma.(x, y : A).$$

This point of view is particularly useful when deciding whether a particular model satisfies univalence.

Univalence holds in the simplicial model

Given a model of type theory we get an interpretation of the univalence axiom. In general the axiom might or might not hold, and a very important application of the study of models of type theory is the proof that the univalence axiom is consistent with Martin-Löf type theory. In [KL12, Theorem 3.4.3] a univalent model in simplicial sets is constructed, we will now describe some of the basic concepts involved in the construction.

The first step is to construct a model using Definition 26. For this one needs a universe of Kan complexes, constructed in [KL12, Section 2]. This construction yields an internal universe with the necessary logical structure ([KL12, Theorem 2.3.4]). At this point it already makes sense to ask whether the model is univalent or not.

To prove that the universe is univalent one introduces the notion of simplicial univalence, which we do next. In the rest of the section all the spaces will be simplicial sets.

Given two fibrations E_1, E_2 over a base B denote the internal hom from E_1 to E_2 in sSet/B by $\mathrm{Hom}_B(E_1, E_2) \to B$. Explicitly an n-simplex in $\mathrm{Hom}_B(E_1, E_2)$ is an n-simplex in B, say $b : \Delta^n \to B$, together with a map $u : b^*E_1 \to b^*E_2$. Consider the subobject $\mathsf{Eq}_B(E_1, E_2)$ of $\mathrm{Hom}_B(E_1, E_2)$ given by pairs (b, u) where again $b : \Delta^n \to B$ and $u : b^*E_1 \to b^*E_2$ but u is a weak equivalence.

Lemma 40 ([KL12, Corollary 3.2.8]). *The induced map* $\mathsf{Eq}_B(E_1, E_2) \to B$ *is a fibration.* ■

Given a fibration $E \to B$ we can consider its pullback along the projections $\pi_i : B \times B \to B$. Together with the above definition we define an object that represents equivalences between fibers of the fibration $E \to B$:

$$\mathsf{Eq}(E) := \mathsf{Eq}_{B\times B}(\pi_1^* E, \pi_2^* E).$$

Notice that we have a fiberwise map:

$$b \longmapsto (b, b, 1_{b^*E})$$

$$\begin{array}{ccc} B & \xrightarrow{\ \delta_E\ } & \mathsf{Eq}(E) \\ & {\scriptstyle\Delta}\searrow \quad \swarrow & \\ & B \times B & \end{array}$$

Similarly to the type theoretic case, the fibration is $E \to B$ is simplicially univalent if $\mathsf{Eq}(E)$ classifies the equivalences of B. The following definition makes this precise.

Definition 41. A fibration $E \to B$ is (simplicially) *univalent* if δ_E is a weak equivalence.

The next main theorems is the following.

Theorem 42 ([KL12, Theorem 3.3.7]). *If B is a Kan complex and $E \to B$ a fibration, E is simplicially univalent if and only if* $\mathsf{IsUnivalent}(E)$ *is inhabited in the simplicial model.*

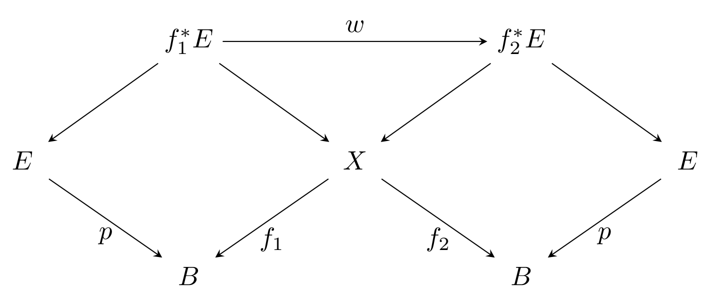
The final step is [KL12, Theorem 3.4.1] where the universe is shown to be simplicially equivalent. In [Shu15b, Section 2] some aspects of the proof are abstracted, which is especially relevant to the next section.

Univalence in homotopy theory

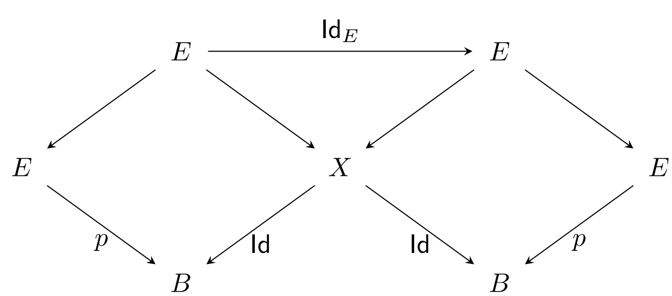
The definition of univalent (simplicial) fibration can be generalized to any arrow in a quasicategory with finite limits. We now sketch the constructions involved in the definition, the details can be found in [GK12].

So let $\mathcal{C}$ be a quasicategory with finite limits and let $\mathcal{S}$ denote the quasicategory of spaces (defined as the simplicial nerve of the simplicially enriched category of Kan complexes). Given an arrow $E \xrightarrow{p} B$ in $\mathcal{C}$ we want to define what it means for this arrow to be univalent.

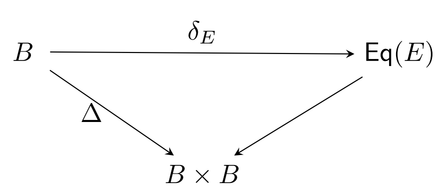
Instead of defining an object $\mathsf{Eq}(E)$ like in the simplicial case we start by defining a functor $\mathbb{E}q(E) : (\mathcal{C}/B \times B)^{op} \to \mathcal{S}$, the functor represented by the hypothetical object $\mathsf{Eq}(E)$. Given a fibration $X \xrightarrow{\langle f_1, f_2 \rangle} B \times B$, consider the space of fiberwise equivalences $w : f_1^* E \to f_2^* E$ making the following diagram commutative:



Suppose $\mathbb{E}q(E)$ is represented by an object $\mathsf{Eq}(E)$. Then the identity $\mathsf{Id}_E : E \to E$ in the diagram:



corresponds to a fiberwise map:



In that case we define $E \to B$ to be univalent if δ_E is an equivalence. If we don't want to assume that $\mathbb{E}q(E)$ is representable for every $E \to B$ we can define:

Definition 43. A map $E \to B$ in a quasicategory with finite limits is *univalent* if $\mathbb{E}q(E)$ is representable, represented by $B \xrightarrow{\Delta} B \times B$.

Tribes and Type theoretic fibration categories

In this section we introduce two closely related categorical models *tribes* (Joyal) and *type theoretic fibration categories*. Although the two concepts are essentially equivalent, tribes are a little bit more modular.

Definition 44. Given a category $\mathcal{C}$, a *fibration structure* on $\mathcal{C}$ consists of:

- A terminal object $1 \in \mathcal{C}$;
- a wide subcategory $\mathcal{F} \subseteq \mathcal{C}$ whose morphisms are called *fibrations* and whose left orthogonal morphisms are called *anodyne*.

Such that:

- isomorphisms are fibrations;
- pullbacks of fibrations exist and are fibrations;
- for every object $X \in \mathcal{C}$ the unique map $X \to 1$ is a fibration.

Definition 45. A *tribe* structure on a category $\mathcal{C}$ is a fibration structure on $\mathcal{C}$ such that:

- every morphism can be factored as an anodyne morphism followed by a fibration;
- anodyne morphisms are stable under pullback along fibrations.

As usual when modelling Π-types we will need a right adjoint for pullbacks along fibrations, defined at least on fibrations.

Definition 46. A *π-tribe* structure on a category $\mathcal{C}$ is a tribe structure on $\mathcal{C}$ such that for every pair of composable fibrations $X \xrightarrow{g} A \xrightarrow{f} B$ the dependent product $\Pi_f g$ exists and is again a fibration.

The second definition is a combination of the above properties.

Definition 47. A *type-theoretic fibration category* is a category $\mathcal{C}$ together with the following structure:

- A distinguished terminal object $1 \in \mathcal{C}$;
- a wide subcategory $\mathcal{F} \subseteq \mathcal{C}$ whose morphisms are called *fibrations* and whose left orthogonal morphisms are called *acyclic cofibrations*;

such that:

- isomorphisms are fibrations;
- morphisms $X \to 1$ are fibrations;
- pullbacks of fibrations exist and are fibrations;
- for every pair of composable fibrations $X \xrightarrow{g} A \xrightarrow{f} B$ the dependent product $\Pi_f g$ exists and is again a fibration.
- every morphism can be factored as an anodyne morphism followed by a fibration;

Note that the existence of dependent products for pairs of composable fibrations implies that acyclic cofibrations are stable under pullback so that a type-theoretic fibration category is the same as a π-tribe.

We now give the essential homotopy theoretic definitions needed to state the main theorem that lets us construct a model to prove homotopy canonicity.

We start with the definition of path object, familiar to anyone who read about model categories.

Definition 48. A *path object* of an object B in a type-theoretic fibration category is an object PB together with an acyclic cofibration-fibration factorization $B \to PB \to B \times B$ of the diagonal $B \to B \times B$.

Using path objects we get a notion of homotopy between two maps.

Definition 49. A *right homotopy* between two maps $f, g : A \to B$ in a type-theoretic fibration category is a lifting of $f \times g : A \to B \times B$ to a path object PB of B.

Notice that since $B \to PB$ is an acyclic cofibration, every path object factors through every other. This implies that being homotopical is a well defined equivalence relation on the hom sets of a type-theoretic fibration category.

Notation 50. Given two maps $f, g : A \to B$ in a type-theoretic fibration the expression $f \sim g$ denotes the fact that f and g are (right) homotopical.

Next we need a notion of "well behaved" functor between type-theoretic fibration categories.

Definition 51. A functor between type-theoretic fibration categories is a *strong fibration functor* if it preserves terminal objects, fibrations, pullbacks of fibrations, and homotopy equivalences.

Given such a functor we can form a new category as follows.

Definition 52. Given a strong fibration functor $F : \mathcal{C} \to \mathcal{D}$ between type-theoretic fibration categories we define the *gluing* $(\mathcal{D} \downarrow F)_f$ as the subcategory of the comma category $\mathcal{D} \downarrow F$ that has:

- As objects: morphisms $A_1 \to F(A_0)$ that are fibrations in $\mathcal{D}$.
- As morphisms: morphisms $A \to B$ such that $A_0 \to B_0$ is a fibration in $\mathcal{C}$ and the induced map $A_1 \to F(A_0) \times_{F(B_0)} B_1$ is a fibration in $\mathcal{D}$.

The main theorem is that this category again carries the structure of a type-theoretic fibration category.

Theorem 53 ([Shu15a, Theorem 2.7]). *If $F : \mathcal{C} \to \mathcal{D}$ is a strong fibration functor, the gluing $(\mathcal{D} \downarrow F)_f$ is a type-theoretic fibration category with as many univalent universes $\mathcal{C}$ and $\mathcal{D}$ have. Moreover, the forgetful functor $(\mathcal{D} \downarrow F)_f \to \mathcal{C}$ preserves all the structure strictly.*

Homotopy canonicity

A type theory with a natural number type satisfies *canonicity* if every term of natural number type is judgmentally equal to a *numeral*, i.e. a term of the form $\mathsf{succ}(\ldots(\mathsf{succ}\,0)\ldots)$. Once we include axioms like function extensionality and univalence this property is usually lost—although we can have type theories in which univalence and fuction extensionality are *provable*, that satisfy canonicity (see for example [Hub16]). Since we have identity types we can consider a weaker property.

Definition 54. A type theory with a natural number type and identity types satisfies *homotopy canonicity* if every term of natural number type is *propositionally* equal to a numeral.

Voevodsky's conjecture for type theory is the statement that Martin-Löf's intensional type theory with univalence satisfies homotopy canonicity. In this section we sketch Shulman's partial answer to this question ([Shu15c]). Let $\mathcal{C}$ be a type theoretic fibration category.

As stated in Theorem 53 given a strong fibration functor $F : \mathcal{C} \to \mathcal{D}$ between type-theoretic fibration categories the category $(\mathcal{D} \downarrow F)_f$ is again a type theoretic fibration category.

The idea is that if $\mathcal{C}$ is split and has a natural number object, the same holds for $(\mathcal{D} \downarrow F)_f$, and the forgetful functor $U : (\mathcal{D} \downarrow F)_f \to \mathcal{C}$ is strict and preserves the natural number object.

If we choose $\mathcal{C}$ to be the classifying category of a type theory with a natural number type we get interpretations:

$$\begin{array}{ccc} & & \mathcal{D} \\ & \overset{!}{\nearrow} & \uparrow \\ \mathcal{C} & & \\ & \underset{!}{\searrow} & \\ & & (\mathcal{D} \downarrow F)_f \end{array}$$

by the initiality of $\mathcal{C}$. Now, the objects of $\mathcal{D} \downarrow F$ are triples $(x \in \mathcal{C}, d \in \mathcal{D}, \alpha : d \to F(x))$, so we have a natural transformation $\Theta :! \Rightarrow F$. Given a term $t : \mathbb{N}$ in our type theory, we have a corresponding morphism $t : () \to (n : \mathbb{N})$ in $\mathcal{C}$ and a corresponding interpretation $[\![t]\!] : 1_{\mathcal{D}} \to \mathbb{N}_{\mathcal{D}}$ in $\mathcal{D}$. The natural transformation Θ gives us a commuting square:

$$\begin{array}{ccc} 1_{\mathcal{D}} & \xrightarrow{\simeq} & F() \\ {\scriptstyle [\![t]\!]}\downarrow & & \downarrow{\scriptstyle F(t)} \\ \mathbb{N}_{\mathcal{D}} & \longrightarrow & F(\mathbb{N}) \end{array}$$

Suppose F is the following functor.

Definition 55. Define the functor $\Gamma_0 : \mathcal{C} \to \mathsf{Set}$ by sending A to $\mathcal{C}(1, A)/\sim$.

[Shu15c, Lemma 13.3] says that Γ_0 is a strong fibration functor if our type theory is 0-truncated (every type is a set). Then the commuting square above tells us that every term of natural number type is homotopy equivalent to a numeral.

More generally, let F be the following functor.

Definition 56. Define the functor $\Gamma_1 : \mathcal{C} \to \mathsf{Gpd}$ by sending A to the groupoid that has as objects maps $1 \to A$ in $\mathcal{C}$ and, as morphisms, homotopy classes of homotopies between these maps.

[Shu15c, Lemma 13.8] says that Γ_1 is a strong fibration functor if our type theory is 1-truncated. So the same argument as above gives us homotopy canonicity for such a type theory.

The idea is that this construction should still work for a suitable functor $F : \mathcal{C} \to \infty\text{-}\mathsf{Gpd}$, without the need of a truncation restriction on the type theory.

References

[AHR99] J. Adámek, M. Hébert, and J. Rosický. On essentially algebraic theories and their generalizations. *Algebra Universalis*, 41(3):213–227, 1999.

[Car78] John Cartmell. *Generalised Algebraic Theories and Contextual Categories*. PhD thesis, 01 1978.

[CGH14] Pierre-Louis Curien, Richard Garner, and Martin Hofmann. Revisiting the categorical interpretation of dependent type theory. *Theoret. Comput. Sci.*, 546:99–119, 2014.

[GK12] D. Gepner and J. Kock. Univalence in locally cartesian closed infinity-categories. *ArXiv e-prints*, August 2012.

[Hof95] Martin Hofmann. On the interpretation of type theory in locally Cartesian closed categories. In *Computer science logic (Kazimierz, 1994)*, volume 933 of *Lecture Notes in Comput. Sci.*, pages 427–441. Springer, Berlin, 1995.

[Hof97] Martin Hofmann. Syntax and semantics of dependent types. In *Semantics and logics of computation (Cambridge, 1995)*, volume 14 of *Publ. Newton Inst.*, pages 79–130. Cambridge Univ. Press, Cambridge, 1997.

[Hub16] S. Huber. Canonicity for Cubical Type Theory. *ArXiv e-prints*, July 2016.

[Jac93] Bart Jacobs. Comprehension categories and the semantics of type dependency. *Theoret. Comput. Sci.*, 107(2):169–207, 1993.

[KL12] Chris Kapulkin and Peter LeFanu Lumsdaine. The simplicial model of univalent foundations (after Voevodsky). *arXiv preprint arXiv:1211.2851*, 2012.

[LW15] Peter LeFanu Lumsdaine and Michael A. Warren. The local universes model: an overlooked coherence construction for dependent type theories. *ACM Trans. Comput. Log.*, 16(3):Art. 23, 31, 2015.

[Pro13] The Univalent Foundations Program. *Homotopy type theory—univalent foundations of mathematics*. The Univalent Foundations Program, Princeton, NJ; Institute for Advanced Study (IAS), Princeton, NJ, 2013.

[Shu15a] M. Shulman. Univalence for inverse EI diagrams. *ArXiv e-prints*, August 2015.

[Shu15b] Michael Shulman. The univalence axiom for elegant Reedy presheaves. *Homology Homotopy Appl.*, 17(2):81–106, 2015.

[Shu15c] Michael Shulman. Univalence for inverse diagrams and homotopy canonicity. *Math. Structures Comput. Sci.*, 25(5):1203–1277, 2015.

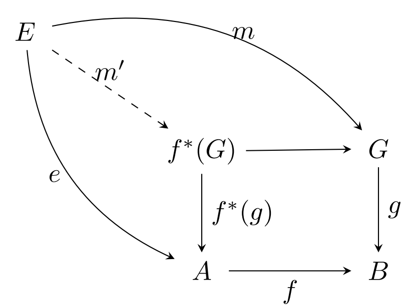
[Str91] Thomas Streicher. *Semantics of Type Theory: Correctness, Completeness, and Independence Results*. Birkhauser Boston Inc., Cambridge, MA, USA, 1991.

[Voe15] Vladimir Voevodsky. A C-system defined by a universe category. *Theory Appl. Categ.*, 30:No. 37, 1181–1215, 2015.

Appendix A – Locally cartesian closed categories

In any category $\mathcal{C}$ with pullbacks, given a morphism $f : A \to B$ we have an induced pullback functor $f^* : \mathcal{C}/B \to \mathcal{C}/A$. Moreover, this functor always has a left adjoint $f_! : \mathcal{C}/A \to \mathcal{C}/B$ given by composition with f. Concretelly the functor sends $E \xrightarrow{e} A$ to $E \xrightarrow{f \circ e} B$.

To see that this is indeed a left adjoint consider $m : f_!(e) \to g$ with $e : E \to A$ and $g : G \to B$. We can form the diagram:



where $m' \in e \to f^*(g)$ is given by the existence condition in the universal property of the pullback. Conversely, given m' we get m by composition. The uniqueness condition in the universal property of pullbacks guarantees that these operations are inverses to each other. Naturality (in g and e) is also given by the universal property of pullbacks.

Observe that in the case $\mathcal{C} = \mathsf{Set}$, given $x : X \to A$, we can write $X = \{X_a \mid a \in A\}$, where $X_a := x^{-1}(a)$. Then, given $f : A \to B$ we have:

$$f_!(X) = \left\{ \coprod_{a \in f^{-1}(b)} X_a \;\middle|\; b \in B \right\}.$$

This motivates the following definition.

Definition 57. Given morphisms $f : A \to B$ and $x : X \to A$ in a category with pullbacks, the object $f_!(X)$ is called the *dependent sum* of X along f and is also denoted by $\Sigma_f x$ or $\Sigma[f, x]$.

In some cases the functor f^* has also a right adjoint, which is usually denoted by $f_* : \mathcal{C}/A \to \mathcal{C}/B$. Let's construct this functor for $\mathcal{C} = \mathsf{Set}$. In this case we can represent $f^*(g) : f^*(G) \to A$ by:

$$\pi_1 : \{(a, \gamma) \mid a \in A, \gamma \in G, f(a) = g(\gamma)\} \to A.$$

We claim that $f_*(e) : f_*(E) \to B$ is given by:

$$\pi_1 : \left\{ (b, \varepsilon) \;\middle|\; b \in B,\ \varepsilon \in \prod_{a \in f^{-1}(b)} E_a \right\} \to B.$$

To see this, given $m : f^*(g) \to e$ consider $m' : g \to f_*(e)$ defined by $m'(\gamma) = (g(\gamma), (m(a, \gamma))_{a \in A_{g(\gamma)}})$, which is well defined since if $a \in A_{g(\gamma)}$, then $f(a) = g(\gamma)$. For the converse, given m' define $m(a, \gamma) = \pi_2(m'(\gamma))(a)$. A straightforward check shows that these are inverses to each other. Naturality is also straightforward.

Notice that we have

$$f_*(G) = \left\{ \prod_{a \in f^{-1}(b)} X_a \;\middle|\; b \in B \right\},$$

which motivates the following definition.

Definition 58. Given morphisms $g : G \to A$ and $f : A \to B$ in a category with pullbacks $\mathcal{C}$ the *dependent product* of g along f and is an object $\Pi[f, g]$ together with natural isomorphisms:

$$\mathcal{C}_{/B}\left(c, \Pi[f, g]\right) \simeq \mathcal{C}_{/A}\left(c \times_B f, g\right)$$

for every $c : C \to A$. The dependent product of g along f is also denoted by $\Pi_f g$.

Observe that for a given morphism f in a category with pullbacks, the dependent product $\Pi[f, g]$ exists for all g exactly when the change of basis functor f^* has a right adjoint.

Definition 59. A locally cartesian closed category is a category with pullbacks such that for every morphism f, the change of basis functor f^* has a right adjoint Π_f.