On the Interpretation of Type Theory in Locally Cartesian Closed Categories

Martin Hofmann*

Department of Computer Science, University of Edinburgh
JCMB, KB, Mayfield Rd., Edinburgh EH9 3JZ, Scotland

Abstract. We show how to construct a model of dependent type theory (category with attributes) from a locally cartesian closed category (lccc). This allows to define a semantic function interpreting the syntax of type theory in an lccc. We sketch an application which gives rise to an interpretation of extensional type theory in intensional type theory.

1 Introduction and Motivation

Interpreting dependent type theory in locally cartesian closed categories (lcccs) and more generally in (non split) fibrational models like the ones described in [7] is an intricate problem. The reason is that in order to interpret terms associated with substitution like pairing for Σ-types or application for Π-types one needs a semantical equivalent to syntactic substitution. To clarify the issue let us have a look at the "naive" approach described in Seely's seminal paper [14] which contains a subtle inaccuracy.

Assume some dependently typed calculus like the one defined in [10] and an lccc $\mathbf{C}$ (a category with finite limits and right adjoints to every pullback functor in order to interpret dependent product types.)

The idea is to interpret contexts as objects in $\mathbf{C}$, types in context Γ as morphisms with codomain the interpretation of Γ, and terms as sections (right inverses) of the interpretation of their types. Now the empty context gets interpreted as the terminal object and a context $\Gamma, x{:}\sigma$ gets interpreted as the domain of the interpretation of $\Gamma \vdash \sigma\ \mathit{type}$. A Σ-type $\Sigma x{:}\sigma.\tau$ in context Γ gets interpreted as the composition $s \circ t$ where s is the interpretation of σ and t is the interpretation of τ in (context $\Gamma, x{:}\sigma$). This "typechecks" because the codomain of t is the interpretation of $\Gamma, x{:}\sigma$ which is the domain of the interpretation of σ. The problem appears when we try to interpret pairing. Assume $\Gamma \vdash M : \sigma$ is a term of type σ and $\Gamma \vdash N : \tau[x := M]$ is a term of type τ with x replaced by M. We want to interpret their pairing $\Gamma \vdash (M, N) : \Sigma x{:}\sigma.\tau$. Let m and n be the interpretations of the former and the latter. The morphism m is a section of s and n is a section of the interpretation of $\tau[x := M]$ which a priori has nothing to do with t — the interpretation of τ. Seely argues that substitution

* The author is supported by a European Union HCM fellowship; contract number ERBCHBICT930420

should be interpreted as a pullback, so that the interpretation of $\tau[x := M]$ becomes the pullback of t along m. One might then interpret the pair (M, N) as the composition $m' \circ n$ where m' is the upper arrow of this pullback.

The subtle flaw of this idea is that the interpretation of $\tau[x := M]$ is already fixed by the clauses of the interpretation and there is no reason why it should equal the chosen pullback of t along m.

Curien [5] addresses the problem by making substitution a syntactic operator which may then be interpreted as (chosen) pullback. However, this changes the calculus and also results in a quite complicated interpretation function for as explained in [5] type equality must be modelled by isomorphism instead of actual semantic equality.

On the other hand interpretation of type theory is relatively straightforward if one has a model equipped with a semantic substitution operation which commutes with composition and all semantic type and term formers. In this case one can show that syntactic and semantic substitution do agree. The technique of interpreting type theory in such a model has been worked out by Streicher [15] in great detail. See also Pitts' forthcoming survey article [12].

Unfortunately, however, it seems impossible to endow an arbitrary lccc with a pullback operation which would satisfy these coherence requirements. For example the natural choice of pullbacks in the category of sets does not work. Indeed, if $f : A \to B$, $g : B \to C$, and $h : D \to C$ are set-theoretic functions then (According to the canonical choice) the pullback of h along $g \circ f$ is the set $\{(a, d) \mid a \in A \wedge d \in D \wedge g(f(a)) = h(d)\}$ whereas the iterated pullback of h first along g then along f gives the set $\{(a, (b, d)) \mid a \in A \wedge b \in B \wedge d \in D \wedge f(a) = b \wedge g(b) = h(d)\}$ which is equipollent, but not equal to the former. It seems to be open whether there exists another choice of pullbacks in the category of sets which commute with composition (and the type formers).

In this paper we propose another solution under which a type is not merely interpreted as a morphism, but as a whole family of morphisms indexed over possible substitutions. More abstractly, we describe a construction which turns an arbitrary lccc into an equivalent *category with attributes* (cwa) — a "split" notion of model introduced by Cartmell [4], see also [12], for which an interpretation function is readily available. The method we use is a very general procedure due to Bénabou (see [2] and [7, Prop. 1.3.6]) which turns an arbitrary fibration into an equivalent split fibration. Our contribution consists of the observation that the cwa obtained thus has not merely a split substitution operation, but is closed under all type formers the original lccc supported. In particular the resulting cwa has Π-types, Σ-types, and (extensional) identity types. Phoa [11, p. 14] has considered this as an open problem. Locally cartesian closed categories play the role of a running example here; the arguments immediately carry over to the more general notions of model studied by Jacobs [7, 8] and other authors.

On a more elementary level the method computes additional information along with the inductive definition of the interpretation which allows to identify the interpretation of a substituted type $\tau[x := M]$ as a pullback of the interpretation of τ albeit not the previously chosen one.

In the next section we define categories with attributes and sketch the standard interpretation function. Section 3 contains the main result — the construction of a cwa out of an lccc. In Section 5 we give a extension to universes which, however, does not handle the most general case. For many lcccs arising in the semantics of type theory in particular sets and ω-sets and all toposes there is already known a natural equivalent cwa. For the case of toposes see [7, Ex. 4.3.5]. In Section 6 we give an example where this is not the case and thus provide an application of the main result. Section 7 offers some concluding remarks and sketches an alternative construction of equivalent split fibration due to Power which does *not* extend to Π- and Σ-types.

Some familiarity with basic category theory and dependent type theory will be assumed. Introductory material may be found in [1] (categories) and [10] (dependent type theory). Both subjects are also well described in [12].

2 Categories with attributes

A *category with attributes* (cwa) is given by the following data

- a category $\mathbf{C}$ with terminal object 1. The unique morphism from object Γ into 1 is written $!_\star$.
- a functor $Fam : \mathbf{C}^{op} \to \mathbf{Sets}$ with morphism part written $Fam(f)(\sigma) =_{\mathrm{abbr}} \sigma[f]$. More elementarily, this means that $Fam(\Gamma)$ is a set for each $\Gamma \in Ob(\mathbf{C})$ and if $\sigma \in Fam(\Gamma)$ and $f : B \to \Gamma$ then $\sigma[f] \in Fam(\Gamma)$ and the two coherence conditions
$$\sigma[id_\star] = \sigma$$
and
$$\sigma[f \circ g] = \sigma[f][g]$$
for $g : A \to B$ are satisfied.
- an operation $\mathrm{p}(-)$ which to each $\sigma \in Fam(\Gamma)$ associates a $\mathbf{C}$-morphism $\mathrm{p}(\sigma)$ with codomain Γ — the *canonical projection* of σ. The domain of $\mathrm{p}(\sigma)$ is written $\Gamma \cdot \sigma$.
- An operation $\mathrm{q}(-,-)$ which to each $\mathbf{C}$-morphism $f : B \to \Gamma$ and $\sigma \in Fam(\Gamma)$ associates a morphism $\mathrm{q}(f,\sigma) : B \cdot \sigma[f] \to \Gamma \cdot \sigma$ such that[2]

$$\begin{tikzcd} B \cdot \sigma[f] \arrow[r, "\mathrm{q}(f,\sigma)"] \arrow[d, "\mathrm{p}(\sigma[f])"'] & \Gamma \cdot \sigma \arrow[d, "\mathrm{p}(\sigma)"] \\ B \arrow[r, "f"'] & \Gamma \end{tikzcd}$$

[2] This and the following diagrams have been typeset using Paul Taylor's diagram macros.

is a pullback and the coherence conditions

$$\mathrm{q}(id_\star, \sigma) = id_{\star \cdot \sigma}$$

and

$$\mathrm{q}(f \circ g, \sigma) = \mathrm{q}(f, \sigma) \circ \mathrm{q}(g, \sigma[f])$$

for $g : A \to B$ are satisfied.

Example 1. An important example of a cwa which also gives some intuition about the meaning of the various ingredients is the term model of some dependent type theory constructed as follows. The category $\mathbf{C}$ has as objects well formed contexts of variable declarations and equivalence classes of parallel substitutions (tuples of terms of the appropriate types as morphisms.) If Γ is a context then $Fam(\Gamma)$ is the set of types well-formed in Γ. If $f : B \to \Gamma$ is a substitution then $\sigma[f]$ is the parallel substitution of the terms of f in σ. The morphism $\mathrm{p}(\sigma)$ consists of the first $|\Gamma|$-variables of $\Gamma, x{:}\sigma$ and $\mathrm{q}(f, \sigma)$ is the substitution (f, x) where x is the last variable in $B, x{:}\sigma[f]$.

Further examples arise from families of sets or ω-sets.

Provided that suitable interpretations of base types and type constructors are given, a partial interpretation function can be defined by structural induction in such a way that every context is interpreted as a $\mathbf{C}$-object, every type is interpreted as an element of *Fam* at the interpretation of its context and finally terms are interpreted as *sections* (right inverses) of the canonical projections associated to their types. If M is a right inverse of $\mathrm{p}(\sigma)$ then by a slight abuse of language we say that M is a section of σ. The pullback requirement for $\mathrm{q}(f, \sigma)$ allows to define a semantic equivalent to substitution on terms: If M is a section of $\sigma \in Fam(\Gamma)$ and $f : B \to \Gamma$ then there is a unique section of $\sigma[f]$ written $M[f]$ which satisfies $\mathrm{q}(f, \sigma) \circ M[f] = M \circ f$.

This interpretation is sound in the sense that the interpretation of all derivable judgements is defined and that all equality judgements are validated w.r.t. the actual equality in the model. An auxiliary property of the interpretation is that syntactic substitution is interpreted as its semantic counterpart $-[f]$.

What it means that a cwa is closed under a type former can be almost directly read off from the syntactic rules. For example closure under Σ-types means that

- for every two families $\sigma \in Fam(\Gamma)$ and $\tau \in Fam(\Gamma \cdot \sigma)$ there is a family $\Sigma(\sigma, \tau) \in Fam(\Gamma)$
- for every two sections M of σ and N of $\tau[M]$ there is a section (M, N) of $\Sigma(\sigma, \tau)$ — the pairing of M and N
- for every section M of $\Sigma(\sigma, \tau)$ there is a section $M.1$ of σ and a section $M.2$ of $\tau[M.1]$ — the two projections of M

such that $(M, N).1 = M$ and $(M, N).2 = N$ and (optionally) $(M.1, M.2) = M$ and for $f : B \to \Gamma$ we have $\Sigma(\sigma, \tau)[f] = \Sigma(\sigma[f], \tau[\mathrm{q}(f, \sigma)])$ and similar coherence laws for pairing and the projections. See [15, 12] for details.

3 From lcccs to categories with attributes

Our aim in this section is to construct a category with attributes supporting Π- and Σ-types, and extensional identity types from a given locally cartesian closed category.

Preliminaries. Let $\mathbf{C}$ be a category with finite limits (terminal object and pullbacks) and $\Gamma \in Ob(\mathbf{C})$. The *slice category* $\mathbf{C}/\Gamma$ has as objects $\mathbf{C}$-morphisms with codomain Γ and a $\mathbf{C}/\Gamma$-morphism from $s : dom(s) \to \Gamma$ to $t : dom(t) \to \Gamma$ is a $\mathbf{C}$-morphism $\alpha : dom(s) \to dom(t)$ with $t \circ \alpha = s$. Notice the important triviality that any $\mathbf{C}$-morphism α with codomain $dom(t)$ is a $\mathbf{C}/\Gamma$-morphism with codomain t (and domain $t \circ \alpha$.) For each $\mathbf{C}$-morphism $f : B \to \Gamma$ there is a functor $f^* : \mathbf{C}/\Gamma \to \mathbf{C}/B$ sending $s : dom(s) \to \Gamma$ to the left vertical arrow of the pullback of s along f. The action of f^* on morphisms is defined by the universal property of the pullback. The functor f^* has a right adjoint Σ_f which sends $s : dom(s) \to B$ to the composition $f \circ s$. The arrow category $\mathbf{C}^{\to}$ has as objects all morphisms of $\mathbf{C}$ and commuting squares as morphisms. Equivalently, a $\mathbf{C}^{\to}$-morphism from $s : dom(s) \to B$ to $t : dom(s) \to \Gamma$ is a $\mathbf{C}$-morphism $f : B \to \Gamma$ and a $\mathbf{C}/B$-morphism $\alpha : s \to f^*t$. Taking the domain of a morphism extends to a functor $dom : \mathbf{C}^{\to} \to \mathbf{C}$.

Categories with finite limits loosely correspond to dependent type theories if one views morphisms as families of types the morphisms denoting the projection from the disjoint union of all fibres to the indexing type. For example in the lccc of sets the type of m, n-matrices indexed over the set $\mathbf{N} \times \mathbf{N}$ would be modelled as the function $format : Mat \to \mathbf{N} \times \mathbf{N}$ which maps an arbitray matrix to its "format" a pair of natural numbers indicating the numbers of rows and columns.

Substitution then corresponds (up to isomorphism) to pullback and composition to disjoint union. For example we obtain the set of square matrices indexed over $\mathbf{N}$ as the pullback of $format$ along the diagonal function $\mathbf{N} \to \mathbf{N} \times \mathbf{N}$ and similarly the set of matrices with variable number of columns indexed over the number of rows as the composition of $format$ with the first projection $\mathbf{N} \times \mathbf{N} \to \mathbf{N}$.

Using equalisers one can also model extensional identity types. In order to have dependent product types one also needs right adjoints to pullback functors which leads to the following definition.

Definition 1. A locally cartesian closed category (lccc) is a category with finite limits and right adjoints Π_f to every pullback functor $f^* : \mathbf{C}/\Gamma \to \mathbf{C}/B$ for $f : B \to \Gamma$.

Examples of lcccs are the categories of sets and ω-sets, all toposes, and the term model of extensional Martin-Löf type theory as constructed in [14]. For the rest of this section assume a fixed lccc $\mathbf{C}$. In order to derive an interpretation of dependent type theory in $\mathbf{C}$ we construct a cwa with base category $\mathbf{C}$ as follows. For $\Gamma \in Ob(\mathbf{C})$ the set $Fam(\Gamma)$ is defined as the set of those functors σ from the slice category $\mathbf{C}/\Gamma$ to the arrow category $\mathbf{C}^{\to}$ which map every morphism to a pullback square and for which $cod \circ \sigma = dom$. More precisely $\sigma \in Fam(\Gamma)$

associates to every morphism $s : B \to \Gamma$ a $\mathbf{C}$-morphism $\sigma(s)$ with codomain B and to $\alpha : B' \to B$ a morphism $\sigma(s, \alpha)$ such that

$$
\begin{tikzcd}
dom(\sigma(s \circ \alpha)) \arrow[r, "\sigma(s,\alpha)"] \arrow[d, "\sigma(s \circ \alpha)"'] & dom(\sigma(s)) \arrow[d, "\sigma(s)"] \\
B' \arrow[r, "\alpha"'] & B
\end{tikzcd}
$$

is a pullback. Moreover, the assignment of the morphism $\sigma(s, \alpha)$ is functorial in the sense that $\sigma(s, id_B) = id_{dom(\sigma(s))}$ and $\sigma(s, \alpha \circ \beta) = \sigma(s, \alpha) \circ \sigma(s \circ \alpha, \beta)$ for $\beta : B'' \to B'$. An element of the thus defined set $Fam(\Gamma)$ is called a *functorial family* over Γ.

Example 2. The intuition behind these families is that instead of making substitution (viz. pullback) an arbitrarily chosen structure, every family comes equipped with its own behaviour under substitution. Thus in $\sigma(s)$ one should view s as a requested substitution and $\sigma(s)$ itself as the result of performing this substitution. Indeed, given a (not necessarily split) choice of pullbacks in $\mathbf{C}$ we can see that every $\mathbf{C}$-morphism σ with codomain Γ induces a family $\hat{\sigma}$ over Γ. For $s : B \to \Gamma$ we put $\hat{\sigma}(s) := s^*\sigma$ where s^* is the pullback functor defined above. If in addition $\alpha : B' \to B$ we define $\hat{\sigma}(s, \alpha)$ as the unique mediating morphism in

$$
\begin{tikzcd}
\cdot \arrow[rr] \arrow[ddd, "(s\alpha)^*\sigma"'] & & \cdot \arrow[ddd, "\sigma"] \\
& & \\
& \cdot \arrow[uur, dotted] \arrow[d, "s^*\sigma"] & \\
B' \arrow[r, "\alpha"'] & B \arrow[r, "s"'] & \Gamma
\end{tikzcd}
$$

where the lower right trapezium and the outer square are pullbacks. It follows from a simple diagram chase that the resulting lower left trapezium is also a pullback as required. Since $\hat{\sigma}(s, \alpha)$ is defined by a universal property it must be functorial.

We continue with the definition of the cwa of functorial families. If $\sigma \in Fam(\Gamma)$ then the canonical projection $\mathrm{p}(\sigma)$ is defined as $\sigma(id_\star)$. Thus $\Gamma \cdot \sigma = dom(\sigma(id_\star))$. If in addition $f : B \to \Gamma$ we define the substitution $\sigma[f]$ by $\sigma[f](s) := \sigma(f \circ s)$ for $s : A \to B$ and by $\sigma[f](s, \alpha) := \sigma(f \circ s, \alpha)$ for $\alpha : A' \to A$. Since this substitution is defined by composition the functor laws for Fam are immediate. Finally, the morphism $\mathrm{q}(f, \sigma)$ is given by $\sigma(id_\star, f)$ which indeed yields the required pullback square. The coherence law for $\mathrm{q}(-,-)$ follows from the functoriality of σ.

Notice that by the definition of canonical projection a section of some family σ is merely a right inverse to $\sigma(id)$. Thus terms do not carry any intensional information with respect to substitution. See also Section 5.

We have now constructed a cwa over $\mathbf{C}$ which can be shown to be equivalent to $\mathbf{C}$ in some suitable 2-categorical sense. We shall content ourselves by noticing that the hat-construction and canonical projection (p) establish an equivalence between the category $Fam(\Gamma)$ where a morphism from σ to τ is a map f with $\mathrm{p}(\tau) \circ f = \mathrm{p}(\sigma)$, and the slice category $\mathbf{C}/\Gamma$ for every $\Gamma \in Ob(\mathbf{C})$.

Theorem 2. *The category with attributes constructed above admits Σ-types, Π-types, and extensional identity types.*

Proof. We give the full proof for Σ-types which conveys the idea and sketch the interpretation of Π-types and identity types. Let $\sigma \in Fam(\Gamma)$ and $\tau \in Fam(\Gamma \cdot \sigma)$. The family $\Sigma(\sigma, \tau)$ is defined by $\Sigma(\sigma, \tau)(s) := \sigma(s) \circ \tau(\mathrm{q}(s, \sigma))$ and $\Sigma(\sigma, \tau)(s, \alpha) := \tau(\mathrm{q}(s, \sigma), \sigma(s, \alpha))$. Thus to obtain the value of $\Sigma(\sigma, \tau)$ at some substitution $s : B \to \Gamma$ we first perform the substitution inside σ yielding $\sigma(s)$ and τ yielding $\tau(\mathrm{q}(s, \sigma))$ and then calculate the sum of the resulting morphisms in $\mathbf{C}$ as usual by composition.

$$
\begin{tikzcd}
dom(\Sigma(\sigma,\tau)(s\alpha)) \arrow[r, "\tau(\mathrm{q}(s,\sigma),\sigma(s,\alpha))"] \arrow[d] & dom(\Sigma(\sigma,\tau)(s)) \arrow[d, "\tau(\mathrm{q}(s,\sigma))"'] & \\
\cdot \arrow[r, "\sigma(s,\alpha)"] \arrow[d, "\sigma(s)"'] & \cdot \arrow[r, "\mathrm{q}(s,\sigma)"] \arrow[d, "\sigma(s)"'] & \Gamma \cdot \sigma \arrow[d, "\mathrm{p}(\sigma)"] \\
B' \arrow[r, "\alpha"'] & B \arrow[r, "s"'] & \Gamma
\end{tikzcd}
$$

The fact that $\Sigma(\sigma, \tau)(s, \alpha) = \tau(\mathrm{q}(s, \sigma), \sigma(s, \alpha))$ forms a pullback with α and the vertical arrows follows because the vertical composition of two pullback squares is a pullback. Functoriality follows from functoriality of σ and τ and the coherence laws for $\mathrm{q}(-,-)$.

Next, we check that the thus defined Σ-type is indeed stable under substitution. If $f : B \to \Gamma$ and $s : A \to B$ then $\Sigma(\sigma, \tau)[f](s) = \Sigma(\sigma, \tau)(fs) = \sigma(fs) \circ \tau(\mathrm{q}(fs, \sigma)) = \sigma[f](s) \circ \tau(\mathrm{q}(f, \sigma) \circ \mathrm{q}(s, \sigma[f])) = \sigma[f](s) \circ \tau[\mathrm{q}(f, \sigma)](\mathrm{q}(s, \sigma[f])) = \Sigma(\sigma[f], \tau[\mathrm{q}(f, \sigma)])(s)$ as required. For the morphism part we calculate similarly.

The pairing and projection combinators are defined as usual in an lccc: If M is a section of σ, *i.e.* a right inverse of $\sigma(id_\star)$ and N is a section of $\tau[M]$, *i.e.* a right inverse to $\tau[M](id_\star) = \tau(M)$ then we define the pairing (M, N) as $\mathrm{q}(M, \tau) \circ N$ which is a section of $\Sigma(\sigma, \tau)$ by simple equality reasoning. On the other hand, if M is a section of $\Sigma(\sigma,\tau)$ then $M.1 := \mathrm{p}(\tau)\circ M$ is a section of M and $M.2$ is the unique section of $\tau[M.1]$ with $\mathrm{q}(M.1,\tau)\circ M.2 = M$ determined by the universal property of the pullback. Now we have $(M.1, M.2) = M$ by definition, $(M,N).1 = M$ by equational reasoning, and $(M,N).2 = N$ by uniqueness of the second projection.

It remains to show that these operations are stable under substitution. We do the calculation for pairing, the two other cases may be verified similarly or can be deduced from the case of pairing and the $\beta$and η-equations. Let M and N be as above in the definition of pairing and $f : B \to \Gamma$. Our aim is to show that

$$(M,N)[f] = (M[f], N[f])$$

The participating sections are defined uniquely by the equations

$$\begin{array}{l} \mathrm{q}(f,\Sigma(\sigma,\tau))\circ(M,N)[f] = (M,N)\circ f \\ \mathrm{q}(f,\sigma)\circ M[f] = M\circ f \\ \mathrm{q}(f,\tau[M])\circ N[f] = N\circ f \end{array}$$

Now in view of the unique characterisation of $(M,N)[f]$ stability follows if we can show

$$\mathrm{q}(f,\Sigma(\sigma,\tau))\circ(M[f],N[f]) = (M,N)\circ f$$

Here the left hand side equals

$$\mathrm{q}(\mathrm{q}(f,\sigma),\tau)\circ \mathrm{q}(M[f],\tau[\mathrm{q}(f,\sigma)])\circ N[f]$$

by expanding the definitions of Σ and pairing. This in turn equals

$$\mathrm{q}(\mathrm{q}(f,\sigma)\circ M[f],\tau)\circ N[f]$$

using the coherence law for $\mathrm{q}(-,-)$. Now using the defining equation for $M[f]$ and applying the coherence law in the other direction we arrive at

$$\mathrm{q}(M,\tau)\circ \mathrm{q}(f,\tau[M])\circ N[f]$$

Using the defining equation for $N[f]$ and the definition of (M,N) we arrive at the right hand side.

The type constructors Π and identity type are defined in a similar fashion. For families σ, τ as above the value of the family $\Pi(\sigma,\tau)$ at substitution $s : B \to \Gamma$ is $\Pi_{\sigma(s)}(\tau(\mathrm{q}(s,\sigma)))$. We leave the messy, but essentially forced definitions of the morphism part and the associated combinators to the reader.

For $\sigma \in \mathit{Fam}(\Gamma)$ and M,N sections of $\mathrm{p}(\sigma)$ we define the identity type $\mathit{Eq}_\sigma(M,N)$ at s as the (chosen) equaliser of $M[s]$ and $N[s]$ where $M[s]$ is the unique section of $\sigma(s)$ for which $\mathrm{q}(s,\sigma)\circ M[s] = M\circ s$. Compatibility of substitution of the associated combinators requires again some lengthy calculation which in the case of Π basically amount to reproving the Beck-Chevalley condition for lcccs [14].

It is worth pointing out that a certain choice of pullbacks and equalisers albeit not a split one is required to interpret identity types which are the basic source of type dependency.

In a similar way we can show that the cwa of families supports lists or natural numbers if the category **C** supports them in a coherent way. Instead of carrying out these (rather laborious) examples we attempt to clarify the ideas a bit further by elaborating the conditions on **C** which are necessary in order that in the associated category with attributes we can interpret an (admittedly contrived) type former governed by the rules

$$\frac{\Gamma \vdash \sigma\ \mathit{type}}{\Gamma \vdash T(\sigma)\ \mathit{type}}\quad \text{T-FORM}$$

$$\frac{\Gamma \vdash M : \sigma}{\Gamma \vdash T\text{-}\mathit{intro}(M) : T(\sigma)}\quad \text{T-INTRO}$$

and the associated congruence rules. This can in general be interpreted if there is an operation T which to every morphism σ with codomain Γ associates a morphism $T(\sigma)$ also with codomain Γ and to every pullback square another pullback square

$$\begin{tikzcd} \cdot \arrow[r, "f'"] \arrow[d, "\sigma'"'] & \cdot \arrow[d, "\sigma"] \\ B \arrow[r, "f"'] & \Gamma \end{tikzcd} \quad \mapsto \quad \begin{tikzcd} \cdot \arrow[r, "T\langle f'\rangle"] \arrow[d, "T(\sigma')"'] & \cdot \arrow[d, "T(\sigma)"] \\ B \arrow[r, "f"'] & \Gamma \end{tikzcd}$$

functorial in the sense that $T\langle id\rangle = id$ and $T\langle f' \circ g'\rangle = T\langle f'\rangle \circ T\langle g'\rangle$. This action on pullback squares is another way of stating that T is compatible with the chosen pullback up to isomorphism *and* admits a functorial action on isomorphisms, but of course not necessarily on arbitrary morphisms.

Moreover, for each section M of σ we need a section $T\text{-}\mathit{intro}(M)$ of $T(\sigma)$ in such a way that in the above pullback situation we have

$$T\langle f'\rangle \circ T\text{-}\mathit{intro}(M') = T\text{-}\mathit{intro}(M) \circ f$$

where M' is the unique section of σ' with $f' \circ M' = M \circ f$. This is the coherence condition one would reasonably expect.

Now we can define a T-operator on families by putting

$$T(\sigma)(s) := T(\sigma(s))$$

and

$$T(\sigma)(s, \alpha) := T\langle \sigma(s, \alpha)\rangle$$

for $\sigma \in \mathit{Fam}(\Gamma)$ and $s : B \to \Gamma$ and $\alpha : B' \to B$. Functoriality follows from functoriality of σ and $T\langle - \rangle$. The operation $T\text{-}\mathit{intro}$ is defined as in **C**. Stability of same under substitution follows directly from the above coherence condition by instantiating with the pullback square

$$\begin{array}{ccc} B \cdot f[\sigma] = dom(\sigma(f)) & \xrightarrow{\mathrm{q}(f,\sigma) = \sigma(id_\star, f)} & \Gamma \cdot \sigma = dom(\sigma(id_\star)) \\ {\scriptstyle \mathrm{p}(\sigma[f]) = \sigma(f)} \downarrow & & \downarrow {\scriptstyle \mathrm{p}(\sigma) = \sigma(id_\star)} \\ B & \xrightarrow[f]{} & \Gamma \end{array}$$

This example shows that the described method carries over to other type constructors like *e.g.* lists or natural numbers provided they are present in $\mathbf{C}$ in a coherent way. We also see that a type former need not necessarily be given by a universal construction as is the case for Π- and Σ-types. The lesson to be learned is that whenever a type former admits a functorial action on pullback squares which is compatible with the associated structure then it may be lifted to the cwa of families.

4 The interpretation in the family model

The general interpretation function for categories with attributes now gives rise to a semantic function mapping contexts to objects in $\mathbf{C}$, types to families over their context, *etc.* Now if $\Gamma \vdash \sigma\ type$ then $[\![\Gamma \vdash \sigma]\!](id_{[\![\star]\!]})$ is an object in the slice category $\mathbf{C}/[\![\Gamma]\!]$ which we may view as the intended interpretation of σ in $\mathbf{C}$. This intended semantics is not "compositional" since for example in the interpretation of pairing we use substitutions other than the identity. A reader familiar with theory of functional programming may notice here some similarity with the continuation-passing-style translation where semantics is inductively defined with respect to an arbitrary continuation, but in the end one is only interested in the instance of the identity continuation.

5 Universes

In the construction described above types get interpreted as functions associating substitutions to morphisms. Terms, however, get interpreted simply as sections and do not carry any intensional information about their behaviour under substitution, it being forced upon by the universal property of the pullbacks associated with families. This implies that our construction does not carry over to universes (which mix terms and types) unless the universe was "split" in the first place. What this means is exemplified by the following definition specialising the notion of model for the Calculus of Constructions given in [6].

Definition 3. A *split dictos* is an lccc $\mathbf{C}$ with a morphism $gen : \top \to \Omega$ and an operation which to every two morphisms $s : S \to \Gamma$ and $p : S \to \Omega$ associates a morphism $\forall_s(p) : \Gamma \to \Omega$ such that $\forall_s(p)^* gen$ and $\Pi_s(p^* gen)$ are isomorphic in $\mathbf{C}/\Gamma$ and for every pullback square

$$\begin{tikzcd} S \arrow[r, "f'"] \arrow[d, "s"'] & T \arrow[d, "t"] \\ B \arrow[r, "f"'] & \Gamma \end{tikzcd}$$

and morphism $p : T \to \Omega$ we have $\forall_t(p) \circ f = \forall_s(p \circ f')$.

In [6] the last requirement is weakened to isomorphism in $\mathbf{C}/\Gamma$ of the morphisms associated by $-^* gen$. The stricter condition imposed here means that the $\forall$ operator is stable under substitution up to equality. The two most prominent examples of dictoses, namely the category of sets with $\Omega = \{0, 1\}$ and the category of ω-sets with Ω equal the set of partial equivalence relations on ω are split dictoses. In split dictoses we can interpret the Calculus of Constructions

Theorem 4. *To every split dictos there exists an equivalent cwa with enough structure to interpret the Calculus of Constructions.*

Proof. Let us first define what it means for a cwa with Π-types to be a model of the Calculus of Constructions. Following [15] we need a family $Prop$ over 1 and a family Prf over $1 \cdot Prop$ in such a way that two morphisms $s, s' : \Gamma \to 1 \cdot Prop$ are equal if $Prf[s] = Prf[s']$. Moreover, if σ is a family over Γ and $p : \Gamma \cdot \sigma \to 1 \cdot Prop$ then there is a morphism $\forall_\sigma(p) : \Gamma \to 1 \cdot Prop$ such that $Prf[\forall_\sigma(s)] = \Pi(\sigma, Prf[s])$. One could stay even closer to the syntax but only at the expense of clarity.

Now let a split dictos $\mathbf{C}$ be given. We construct a cwa with base $\mathbf{C}$ as follows. The set of families over Γ is defined as the disjoint union of the set of functorial families as defined in Section 3 and the homset $\mathbf{C}(\Gamma, \Omega)$. We call the elements of $\mathbf{C}(\Gamma, \Omega)$ *propositional families* (over Γ). The operations of substitution and canonical projection are extended to propositional families by defining for $\sigma : \Gamma \to \Omega$ and $f : B \to \Gamma$:

$$\begin{array}{ll} f[\sigma] & = \sigma \circ f \\ \mathrm{p}(\sigma) & = \sigma^* gen \\ \mathrm{q}(f, \sigma) & \text{defined by universal property like in Ex. 2} \end{array}$$

It follows by straightforward calculation that this is a cwa. Every propositional family $\sigma : \Gamma \to \Omega$ induces a functorial family $\tilde{\sigma}$ defined by applying the hat-construction from Ex. 2 to $\sigma^* gen$. We may then extend Σ-types and other possible type formers except Π-types to propositional families by precomposition with $\tilde{\ }$. We have a functorial family $Prop$ over 1 defined by $Prop = \widehat{!_\Omega}$. W.l.o.g. we may identify $1 \cdot Prop$ with Ω. A propositional family Prf over $1 \cdot Prop$ is then defined as the identity on Ω. Notice that if $s : \Gamma \to 1 \cdot Prop$ then $Prf[s]$ equals s. Therefore $Prf[-]$ is injective as required.

For the definition of the Π-type $\Pi(\sigma, \tau)$ we first replace σ by $\tilde{\sigma}$ if σ is propositional. So let's assume that σ is functorial. Then we proceed by case distinction on whether τ is functorial or propositional. In the former case we use the Π-type for functorial families as defined in Section 3. If τ is propositional, *i.e.* $\tau : \Gamma \cdot \sigma \to 1 \cdot Prop$ then we define $\Pi(\sigma, \tau)$ as the propositional family $\forall_{\mathrm{p}(\sigma)}(\tau)$. Abstraction and application are defined by suitably interspersing the isomorphism between $\forall_s(p)^* gen$ and $\Pi_s(p^* gen)$ assumed in the definition of a split dictos.

By lengthy but straightforward calculation it follows that this satisfies all the properties of dependent products. In particular to see that Π is stable under substitution we instantiate the coherence property for $\forall$ with the pullback square formed out of $\mathrm{p}(\sigma[f])$, $\mathrm{p}(\sigma)$, $\mathrm{q}(f, \sigma)$, and f for some $f : B \to \Gamma$.

The $\forall$-operator is defined in exactly the same way using the fact that propositional families and morphisms into $1 \cdot Prop$ coincide.

As in Ex. 2 the hat-construction and canonical projection define an equivalence between $\mathbf{C}$ and the constructed cwa.

It deserves attention that the coherence requirement imposed on the $\forall$-operation was crucial for the definition of Π-types by case distinction and that the methods described in this paper do not seem to generalize to arbitrary dictoses or more generally lcccs which support universes in a non split way.

6 Application: A category of setoids

As mentioned in the Introduction for many lcccs an equivalent cwa is known already. However, there is an interesting example motivated by a construction in [3] for which the construction described in this paper seems to be the only viable way. Consider the syntax of intensional Martin-Löf type theory with natural numbers as described *e.g.* in [10]. We write $\Gamma \vdash \sigma\ true$ to mean that there exists a term $\Gamma \vdash M : \sigma$. We write $\times$ and $\to$ for the nondependent special cases of Σ and Π, resp. A category $\mathbf{C}$ of "setoids" (Types with equivalence relations) is formed as follows.

- An object of $\mathbf{C}$ is a quintuple $X = (X_{set}, X_{rel}, r, s, t)$ such that the following hold.
  i) X_{set} is a closed type.
  ii) $x, x' {:} X_{set} \vdash X_{rel}(x, x')\ type$.
  iii) $x {:} X_{set} \vdash r(x) : X_{rel}(x, x)$.
  iv) $x, x' {:} X_{set}\ ,\ p {:} X_{rel}(x, x') \vdash s(p) : X_{rel}(x', x)$.
  v) $x, x', x'' {:} X_{set}\ ,\ p {:} X_{rel}(x, x')\ ,\ q {:} X_{rel}(x', x'') \vdash t(p, q) : X_{rel}(x, x'')$.

  So r, s, t are "proofs" that X_{rel} is an equivalence relation on X_{set}. If no confusion can arise the subscripts $_{set}$ and $_{rel}$ may be omitted.

– A morphism from X to Y is a term $x\colon X_{set} \vdash f(x) : Y_{set}$ such that

$$x, x'\colon X_{set}\,,\ -\colon X_{rel}(x,x') \vdash Y_{rel}(f(x), f(x'))\ true$$

Moreover, two morphisms f and f' are identified if

$$x\colon X_{set} \vdash Y_{rel}(f(x), f'(x))\ true$$

It is easy to check that equality on morphisms is an equivalence relation and that morphisms are closed under composition and contain the identity so that indeed a category has been defined. Essentially, this construction is the same as the one described in [3] although there the category is defined categorically rather than syntactically and one starts out with an lccc in the first place. By mimicking the argument given there we obtain the following proposition.

Proposition 5. *The category* $\mathbf{C}$ *of setoids is locally cartesian closed and contains a natural numbers object.*

Proof. We only give the required objects leaving the verifications to the reader. Let $f : Y \to X$ and $g : Z \to X$. The pullback of f and g is defined as the object W given by $W_{set} = \Sigma y\colon Y.\Sigma z\colon Z.X(f(y), g(z))$ and $W_{rel}((y,z,-)\,,\,(y',z',-)) = Y(y,y') \times Z(z,z')$. The two pullback projections send $(y,z,-)$ to y and z respectively.

Now let $f : Y \to X$ and $g : Z \to Y$. We define $\Pi_f(g) : W \to X$ by

$W_{set} := \Sigma x\colon X.\Sigma u\colon \Pi y\colon Y.X(f(y),x) \to \Sigma z\colon Z.Y(g(z),y).Resp(u)$
$W_{rel}((x,u,-)\,,\,(x',u',-)) :=$
$\qquad X(x,x') \times \Pi y\colon Y.\Pi p\colon X(f(y),x).\Pi p'\colon X(f(y),x').Z(u\ \ y\ \ p\,,\ u'\ \ y\ \ p')$
where
$Resp(u) :=$
$\qquad \Pi y,y'\colon Y.\Pi p\colon X(f(y),x).\Pi p'\colon X(f(y'),x).Y(y,y') \to Z(u\ \ y\ \ p, u\ \ y'\ \ p')$

The morphism $\Pi_f(g)$ itself is then the first projection from W to X. Finally the natural numbers object has as underlying type the type of natural numbers and the intensional identity type Id_N as relation.

Unfortunately, this proof only shows the existence of pullbacks and Π and no canonical choice arises from it because the constructions given in the proof are not independent of the particular representatives chosen for the involved morphisms. Therefore, in order to endow the category of setoids with chosen structure the use of the axiom of choice seems unavoidable. However, given such a choice we can use the method described in this paper to obtain an interpretation of extensional type theory in the category of setoids and thus in a certain sense in intensional type theory. The category of setoids is a worthwhile object for further study. In particular it appears to have coequalisers of equivalence relations and thus provides a model for the extensional quotient types studied by Mendler in [9]. Moreover, we believe that the full subcategory of the category of setoids consisting of those objects taken on by the interpretation function is actually equivalent to the lccc of types and terms in extensional type theory defined in [14] and presented there as the initial one. Incidentally, the precise proof of initiality (up to natural isomorphism) of this syntactic lccc is another field of application for our methods.

7 Summary and Concluding Remarks

We have described a method for obtaining an equivalent category with attributes from a locally cartesian closed category. This solves the problem of interpreting (at least first order) dependent type theory in lcccs. The method consists of applying Bénabou's construction of a split fibration from an arbitrary one to the particular case of the codomain fibration associated to an lccc. The observation that the thus obtained cwa is closed under various type operators is to our knowledge original.

Incidentally, for another somewhat dual construction of split fibrations due to Power [13] this is not the case. Using it $Fam(\Gamma)$ would be the set of pairs (s, σ) where s and σ are morphisms with common codomain and $dom(s) = \Gamma$. The associated canonical projection to such a family is the morphism $s^*\sigma$ with codomain Γ. If $f : B \to \Gamma$ then we define $(s, \sigma)[f]$ as $(s \circ f, \sigma)$. This gives rise to a cwa which, however, is not even closed under Σ-types in a natural way. Intuitively, the reason is that (s, σ) can be viewed as a type σ together with a delayed substitution which is meant to be carried out upon taking the canonical projection. But if two types have different associated substitutions we cannot compute their sum or product without performing the substitutions which destroys the split property.

Power's very elegant result applies to much more general coherence problems than the one considered here; in fact it requires some effort to extract the above concrete description from the general construction. The aim of the previous paragraph is by no means to criticise his beautiful work, but to pinpoint the particular properties of Bénabou's construction which make our result go through.

In view of the lack of generality with respect to universes pointed out in Section 5 one might want to endow the meaning of terms with behaviour under substitution, too. Then, however, the framework of cwas is no longer sufficient and more generally no model in which substitution on terms is modelled by a universal property could work. We do not know of any notion of model where this is not the case, so maybe some further research into the abstract semantics of dependent types is called for.

In conclusion we may remark that a certain gap in the literature has been filled, but that the practical usefulness of the result remains unclear until more examples like the one from Section 6 are found and investigated.

Acknowledgement

My thanks go to Thomas Streicher for explaining Bénabou's construction to me.

References

1. Michael Barr and Charles Wells. *Category Theory for Computing Science*. International Series in Computer Science. Prentice Hall, 1990.
2. J. Bénabou. Fibred categories and the foundations of naive category theory. *Journal of Symbolic Logic*, 50:10–37, 1985.
3. A. Carboni. Some free constructions in realizability and proof theory. *Journal of Pure and Applied Algebra*, to appear.
4. J. Cartmell. *Generalized algebraic theories and contextual categories*. PhD thesis, Univ. Oxford, 1978.
5. Pierre-Louis Curien. Substitution up to isomorphism. *Fundamenta Informaticae*, 19:51–86, 1993.
6. Thomas Ehrhard. Dictoses. In *Proc. Conf. Category Theory and Computer Science, Manchester, UK*, pages 213–223. Springer LNCS vol. 389, 1989.
7. Bart Jacobs. *Categorical Type Theory*. PhD thesis, University of Nijmegen, 1991.
8. Bart Jacobs. Comprehension categories and the semantics of type theory. *Theoretical Computer Science*, 107:169–207, 1993.
9. Nax P. Mendler. Quotient types via coequalisers in Martin-Löf's type theory. in the informal proceedings of the workshop on Logical Frameworks, Antibes, May 1990.
10. B. Nordström, K. Petersson, and J. M. Smith. *Programming in Martin-Löf's Type Theory, An Introduction*. Clarendon Press, Oxford, 1990.
11. Wesley Phoa. An introduction to fibrations, topos theory, the effective topos, and modest sets. Technical Report ECS-LFCS-92-208, LFCS Edinburgh, 1992.
12. Andrew Pitts. Categorical logic. In *Handbook of Logic in Computer Science (Vol. VI)*. Oxford University Press, 199? to appear.
13. A. J. Power. A general coherence result. *Journal of Pure and Applied Algebra*, 57:165–173, 1989.
14. Robert A. G. Seely. Locally cartesian closed categories and type theory. *Mathematical Proceedings of the Cambridge Philosophical Society*, 95:33–48, 1984.
15. Thomas Streicher. *Semantics of Type Theory*. Birkhäuser, 1991.