Revisiting the categorical interpretation of dependent type theory

Pierre-Louis Curien
πr^2 team, PPS Laboratory, CNRS, Université Paris Diderot, and INRIA

Richard Garner
Macquarie University, Sydney

Martin Hofmann
Ludwig-Maximilian's-Universität, München

February 25, 2013

We relate Hofmann's and Curien's interpretations of Martin-Löf's type theory, which were both designed to cure a mismatch between syntax and semantics in Seely's original interpretation in locally cartesian closed categories.

As an outcome, we obtain a new proof of the coherence theorem that Curien had proved, which implied that Seely's interpretation in the end was sound.

Seely 84 explained how to interpret extensional Martin-Löf's type theory in locally cartesian closed categories, using the substitution-as-pullback paradigm of categorical logic. But there was a coherence issue arising in this interpretation from the pseudo-functoriality of pullbacks.

Curien 90 proved the soundness of Seely's interpretation by first designing a syntax with explicit coercions (thus mirroring the pseudo-functoriality at the level of the language being modelled), and then by showing the coherence as a syntactic result.

The proof of coherence was made by using rewriting techniques. The observation made by Huet in his (unpublished) lecture notes on category theory that Mac Lane's proof of coherence for monoidal categories was a "categorification" of Knuth-Bendix lemma was instrumental for this proof.

Hofmann 94 circumvented the coherence issue by showing how to obtain a split model (that is, a model in which composition of substitutions in types and terms is associative "on the nose") of Martin-Löf's type theory from a locally cartesian closed category. Then the original type theory can be interpreted straightforwardly in this "strictified" model.

The strictification consisted in taking a construction going back to Giraud and Bénabou, of a right adjoint to the forgetful functor from split fibrations and strict morphisms on a fixed base category to the category of fibrations and fibration morphisms (which are required to preserve the structure, but not on the nose), and in showing that this construction carries over to deal with the required additional structure for the interpretation of Martin-Löf's type theory.

Goal of the present work. In retrospect, these were "dual" routes to cure the mismatch between the (strict) syntax and the (non-strict) models: either "unstrictify" the syntax, or strictify the model. We wanted to understand the conceptual architecture in which these two approaches can be linked. Three large categories were involved:

- $\rightsquigarrow$ of non-strict structures and functors preserving the structure up to iso: this is where locally cartesian closed categories and Seely's original interpretation live;
- $\dashrightarrow$ of strict structures and strict morphisms (i.e., preserving the structure exactly): this is where Hofmann's interpretation lives;
- $\longrightarrow$ of non-strict structures and strict morphisms: this is where Curien's modified syntax lives as a free structure.

This three-fold superstructure comes in various contexts, starting with monoidal categories.

In our case, the structures under consideration are Jacobs' comprehension categories with products and sums, or $\mathbb{ML}$-categories for short. Let us denote the corresponding three large categories by $\mathbb{ML}$, $\mathbf{S}\mathbb{ML}_{\mathbf{s}}$ and $\mathbb{ML}_{\mathbf{s}}$, respectively.

Let $\mathtt{Synt}^e$ be the classifying $\mathbb{ML}$-category (i.e., term model) built up from the syntax with explicit coercions, and let $\mathtt{Synt}$ be the classifying strict $\mathbb{ML}$-category built up from the original syntax of Martin-Löf's type theory. Our story then goes as follows.

I. Let p_1 and p_2 be $\mathbb{ML}$-categories. Let $[\![\_]\!]_1$ and $[\![\_]\!]_2$ be the interpretation functions of the explicit syntax in p_1, p_2, respectively. Thus we have

$$[\![\_]\!]_1 \in \mathbb{ML}_{\mathbf{s}}[\mathtt{Synt}^e, p_1] \qquad\qquad [\![\_]\!]_2 \in \mathbb{ML}_{\mathbf{s}}[\mathtt{Synt}^e, p_2]$$

(note that by design interpretation functions are strict). Consider further a morphism

$$F \in \mathbb{ML}[p_1, p_2]$$

Since F is not required to be strict, we do not have $F \circ [\![\_]\!]_1 = [\![\_]\!]_2$, but the two functors are still related through a natural isomorphism γ (a fact that can be neatly encapsulated through the construction of a third $\mathbb{ML}$-category $[F]$ – the iso-glueing of p_1 and p_2 along F).

In picture:

$$\begin{tikzcd}
\mathtt{Synt}^e \arrow[dr, "\llbracket _ \rrbracket_1"] \arrow[ddr, "\llbracket _ \rrbracket_2"'] & \\
& p_1 \arrow[d, squiggly, "F"] \\
& p_2
\end{tikzcd}$$

where the triangle commutes up to the isos γ, which collectively form a 2-cell cell between $F \circ \llbracket \_ \rrbracket_1$ and $\llbracket \_ \rrbracket_2$.

II. Let us denote with $\{\!\{\_\}\!\}$ the interpretation function of the original syntax in a strict structure p, and let us write more suggestively $|\_|$ for $[\![\_]\!]_{\mathtt{Synt}} \in \mathbb{ML}_{\mathbf{s}}[\mathtt{Synt}^e, \mathtt{Synt}]$ (indeed, $|\_|$ removes the explicit coercions from the syntax). Then we have the following factorisation:

$$[\![\_]\!] = \{\!\{|\_|\}\!\}$$

$$\begin{tikzcd}
\mathtt{Synt} \arrow[drr, dashed, "\{\!\{_\}\!\}"'] & \mathtt{Synt}^e \arrow[l, "|_|"'] \arrow[dr, "{[\![_]\!]}"] & \\
 & & p
\end{tikzcd}$$

III. Let $\mathbb{C}$ be a locally cartesian closed category, which one can view as a fibration $p_1 = p = \mathtt{cod} : \mathbb{C}^{\rightarrow} \rightarrow \mathbb{C}$, endowed with the trivial identity comprehension structure. Let $p_2 = Rp$ be the Giraud-Bénabou-Hofmann strictification of p. Then there is a faithful functor $F : p \rightarrow Rp$ over $\mathbb{C}$ (i.e., such that $Rp \circ F = p$).

Then we can instantiate the situation in I as indicated in III: via II, we obtain natural isos relating Curien's interpretation $[\![\_]\!]$ and Hofmann's interpretation $\{\!\{\}\!\}$, and as a bonus we get a *new proof* of the coherence theorem for the interpretation of Martin-Löf's type theory in locally cartesian closed categories.

$$
\begin{tikzcd}
\mathtt{Synt} \arrow[ddrr, dashed, "\{\!\{\text{-}\}\!\}"'] & \mathtt{Synt}^e \arrow[l, "|\text{-}|"'] \arrow[dr, "[\![\text{-}]\!]"] \arrow[ddr] & \\
& & p \arrow[d, squiggly, "F"] \\
& & Rp
\end{tikzcd}
$$

Indeed, if we have

$$\Gamma \vdash M_1 : \sigma \qquad \Gamma \vdash M_2 : \sigma \qquad |M_1| = |M_2|$$

then

$$
\begin{aligned}
[\![M_1]\!] \; = \; F_b([\![M_1]\!]) \; &= \; \gamma_{\Gamma,\sigma} \circ \{\!\{|M_1|\}\!\} \circ \gamma_\Gamma^{-1} \\
&= \; \gamma_{\Gamma,\sigma} \circ \{\!\{|M_2|\}\!\} \circ \gamma_\Gamma^{-1} \; = \; F_b([\![M_2]\!]) = [\![M_2]\!]
\end{aligned}
$$

Type theory with variable-free syntax.

$$\sigma ::= \text{base types, possibly dependent} \parallel \sigma[s]$$
$$\Gamma ::= \emptyset \parallel (\Gamma, \sigma)$$
$$M ::= 1 \parallel M[s] \parallel \text{cases given by the signature}$$
$$s ::= id \parallel \uparrow \parallel M \cdot s \parallel s \circ s$$

The typing judgements are:

$$\Gamma\mathtt{context} \qquad \Gamma \vdash \sigma : \mathtt{type} \qquad \Gamma \vdash M : \sigma \qquad \Gamma' \vdash s : \Gamma$$

The typing rules are on next slide.

$$\frac{}{\emptyset\ \texttt{context}} \qquad \frac{\Gamma \vdash \sigma\ \texttt{type}}{\Gamma, \sigma\ \texttt{context}} \qquad \frac{\Gamma' \vdash s : \Gamma \quad \Gamma \vdash \sigma :\ \texttt{type}}{\Gamma' \vdash \sigma[s]\ \texttt{type}}$$

$$\frac{}{\Gamma, \sigma \vdash 1 : \sigma[\uparrow]} \qquad \frac{\Gamma' \vdash s : \Gamma \qquad \Gamma \vdash M : \sigma}{\Gamma \vdash M[s] : \sigma[s]}$$

$$\frac{}{\Gamma \vdash id : \Gamma} \qquad \frac{}{\Gamma, \sigma \vdash \uparrow : \Gamma} \qquad \frac{\Gamma_1 \vdash s_1 : \Gamma_2 \qquad \Gamma_2 \vdash s_2 : \Gamma_3}{\Gamma_1 \vdash s_2 \circ s_1 : \Gamma_3}$$

$$\frac{\Gamma' \vdash s : \Gamma \qquad \Gamma \vdash \sigma\ \texttt{type} \quad \Gamma' \vdash M' : \sigma[s]}{\Gamma' \vdash M' \cdot s : \Gamma, \sigma}$$

$$\frac{\vdash \Gamma = \Gamma' \quad \Gamma \vdash M : \sigma}{\Gamma' \vdash M : \sigma} \qquad \frac{\Gamma \vdash M : \sigma \quad \Gamma \vdash \sigma = \sigma'}{\Gamma \vdash M : \sigma'}$$

The axioms for equality are:

$$
\begin{aligned}
&\sigma[s][t] = \sigma[s \circ t] \\
&\sigma[id] = \sigma \\
\\
&1[M \cdot s] = M \\
&M[s][t] = M[s \circ t] \\
\\
&\uparrow \circ (M \cdot s) = s \\
&(s_1 \circ s_2) \circ s_3 = s_1 \circ (s_2 \circ s_3) \\
&id \circ s = s \circ id = s \\
&(M \cdot s) \circ t = M[t] \cdot (s \circ t) \\
&1[s] \cdot (\uparrow \circ s) = s
\end{aligned}
$$

Reviewing Seely's interpretation in LCCCs. A locally cartesian closed category (LCCC) is a category $\mathbb{C}$ with a terminal object and such that all slice categories $\mathbb{C}/C$ are cartesian closed (in particular, $\mathbb{C}$ is cartesian closed).

- Contexts are mapped to objects, substitutions to morphisms.
- Types are interpreted as objects in a slice category, and terms as sections. More precisely the interpretations of $\Gamma \vdash \sigma\ \texttt{type}$ and of $\Gamma, \sigma \vdash \uparrow : \Gamma$ coincide, and the interpretation of $\Gamma \vdash M : \sigma$ is a morphism in $\mathbb{C}[\Gamma, (\Gamma, \sigma)]$ such that $\uparrow \circ M = id$.

We represent the interpretation of the respective judgements

$$\Gamma' \vdash s : \Gamma \qquad\qquad \Gamma \vdash \sigma\ \texttt{type} \qquad\qquad \Gamma \vdash M : \sigma$$

as follows:

$$\begin{tikzcd} \Gamma & \Gamma' \arrow[l, "s"'] \end{tikzcd} \qquad \begin{tikzcd} \Gamma \arrow[r, bend left, "M"] & (\Gamma, \sigma) \arrow[l, bend left, "\uparrow = \sigma"] \end{tikzcd}$$

$$\begin{tikzcd} \Gamma, \sigma \arrow[d, bend right, "\uparrow"'] & \Gamma', \sigma[s] \arrow[l] \arrow[d, bend right, "\uparrow"'] \\ \Gamma \arrow[u, bend right, "M"'] & \Gamma' \arrow[l, "s"'] \arrow[u, bend right, "M[s] = \langle id, M \circ s \rangle"'] \end{tikzcd} \qquad \begin{tikzcd} \Gamma, \sigma \arrow[d, bend right, "\uparrow"'] & \Gamma, \sigma, \sigma[\uparrow] \arrow[l] \arrow[d, bend right, "\uparrow"'] \\ \Gamma & \Gamma, \sigma \arrow[l, "\uparrow"'] \arrow[u, bend right, "1 = \langle id, id \rangle"'] \end{tikzcd}$$

$$\begin{tikzcd} \Gamma, \sigma \arrow[d, bend right, "\uparrow"'] & \Gamma', \sigma[s] \arrow[l, "s'"] \arrow[d, bend right, "\uparrow"'] \\ \Gamma & \Gamma' \arrow[l, "s"] \arrow[ul, "M' \cdot s = s' \circ M'"'] \arrow[u, bend right, "M'"'] \end{tikzcd}$$

The coherence issue. The conversion rule $\sigma[s][t] = \sigma[s \circ t]$ cannot be modelled by an equality, since in an LCCC the composition of chosen pullbacks is only isomorphic to the chosen pullback of the composite, in general.

It follows that the syntax of dependent type theory cannot be interpreted soundly in an LCCC in this direct way. One way of rectifying this is to interpret the (proofs of) type equalities as *isomorphisms*.

Let us now examine the "equality" $1[M' \cdot s] = M'$. The left-hand side has type $\sigma[\uparrow][M' \cdot s]$ while M' has type $\sigma[s]$: these types are only isomorphic.

This suggests to introduce explicit coercions in the syntax, so as to be able to track down, already at the syntactic level, all the isos involved in the interpretation of a dependent judgement.

Type theory with explicit coercions.

$$M ::= \ldots \parallel \mathsf{c}(M, \sigma, \sigma')$$

with the following new typing rule:

$$\frac{\Gamma \vdash M : \sigma \qquad \Gamma \vdash \sigma \cong \sigma'}{\Gamma \vdash \mathsf{c}(M, \sigma, \sigma') : \sigma'}$$

where $\Gamma \vdash \sigma \cong \sigma'$ is a new judgement.

Besides equality conversions that concern all syntactic categories (contexts, substitutions, types, and terms), denoted as before with the symbol $=$, there are now so-called iso conversions for types, denoted with the symbol $\cong$.

Axioms for equality and for isomorphisms are given on the next slide.

$$\sigma[s][t] \cong \sigma[s \circ t]$$
$$\sigma[id] \cong \sigma$$

$$\mathsf{c}(1[M \cdot s], \sigma[\uparrow][M \cdot s], \sigma[s]) = M$$
$$\mathsf{c}(M[s][t], \sigma[s][t], \sigma[s \circ t]) = M[s \circ t]$$

$$\uparrow \circ (M \cdot s) = s$$
$$(s_1 \circ s_2) \circ s_3 = s_1 \circ (s_2 \circ s_3)$$
$$id \circ s = s \circ id = s$$
$$(M \cdot s) \circ t = \mathsf{c}(M[t], \sigma[s][t], \sigma[s \circ t]) \cdot (s \circ t)$$
$$\mathsf{c}(1[s], \sigma[\uparrow][s], \sigma[\uparrow \circ s]) \cdot (\uparrow \circ s) = s$$

$$\mathsf{c}(\mathsf{c}(M, \sigma_1, \sigma_2), \sigma_2, \sigma_3) = \mathsf{c}(M, \sigma_1, \sigma_3)$$
$$\mathsf{c}(M, \sigma, \sigma) = M$$
$$\mathsf{c}(M, \sigma_1, \sigma_2)[t] = \mathsf{c}(M[t], \sigma_1[t], \sigma_2[t])$$

Fibrations. Let $p : \mathbb{E} \to \mathbb{B}$ be a Grothendieck fibration:

$$\begin{tikzcd} & A \arrow[dl, dashed, "\forall"'] \arrow[d, dashed, "\exists!"] \\ B \arrow[d, dotted, no head] & B[u] \arrow[l, dashed, "\overline{u}_B"] \arrow[d, dotted, no head] \\ Y & X \arrow[l, "u"] \end{tikzcd}$$

The fibration is called *split* when composites of chosen morphisms are the chosen morphisms, i.e., when

$$\begin{array}{ll} C[v][u] = C[v \circ u] & v_C \circ u_{C[v]} = (v \circ u)_C \\ A[id] = A & (id_X)_A = id_A \end{array}$$

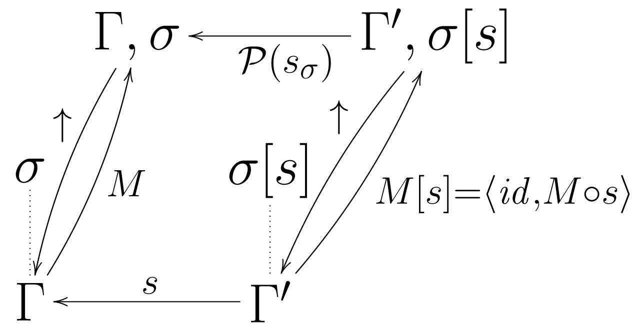
Comprehension categories. A comprehension structure on a fibration p is given by a morphism from p to $\mathtt{cod}$ over $\mathbb{B}$, i.e., a functor $\mathcal{P} : \mathbb{E} \to \mathbb{B}^{\to}$ such that $p = \mathtt{cod} \circ \mathcal{P}$, that moreover preserves cartesian morphisms (i.e., that sends the chosen cartesian morphisms to pullback squares).

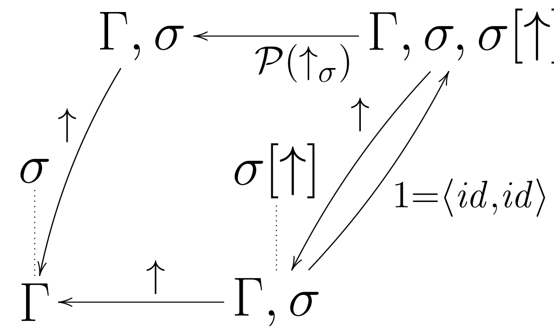
Obviously, the fibration $\mathtt{cod}$ has a trivial comprehension structure, given by the identity functor.

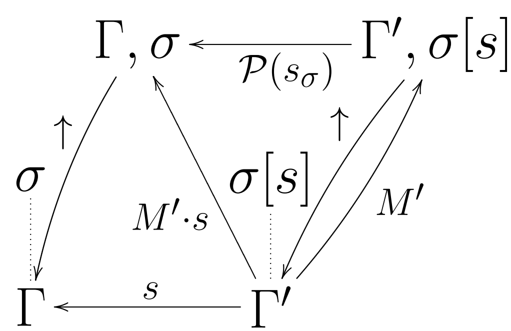
interpretation in comprehension categories. Types are no longer base morphims but *local objects*, while the other syntactic categories are interpreted as before. Also, $\uparrow$ is no longer assimilated to σ but is rather defined as $\mathcal{P}(\sigma)$.

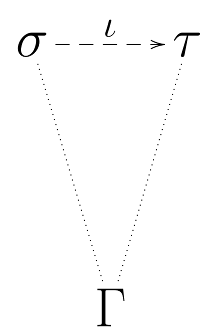
$$\begin{tikzcd} & & \sigma \arrow[d, dash, dotted] & \\ \Gamma & \Gamma' \arrow[l, "s"'] & \Gamma \arrow[r, bend left, "M"] & (\Gamma, \sigma) \arrow[l, bend left, "\uparrow = \mathcal{P}(\sigma)"] \end{tikzcd}$$

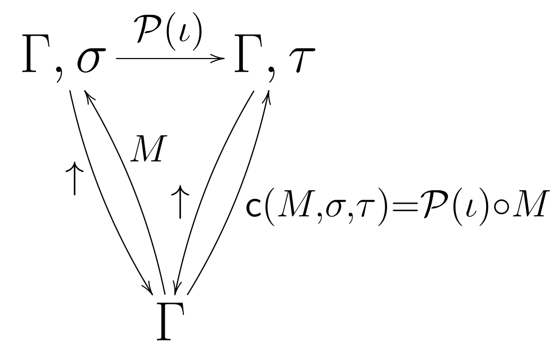
Given $\Gamma' \vdash s : \Gamma$ and $\Gamma \vdash \sigma$ `type`, $\sigma[s]$ is interpreted as the domain of s_σ, whence our notation chosen for cartesian morphisms.


Γ, σ
Γ′, σ[s]
P(s_σ)
↑
σ
M
σ[s]
↑
M[s]=⟨id,M∘s⟩
Γ
s
Γ′


Γ, σ
Γ, σ, σ[↑]
P(↑_σ)
↑
σ
σ[↑]
↑
1=⟨id,id⟩
Γ
↑
Γ, σ


Γ, σ
Γ′, σ[s]
P(s_σ)
↑
σ
σ[s]
↑
M′·s
M′
Γ
s
Γ′


σ
ι
τ
Γ


Γ, σ
P(ι)
Γ, τ
M
↑
↑
c(M,σ,τ)=P(ι)∘M
Γ

The Giraud-Bénabou-Hofmann construction. Let $p : \mathbb{E} \to \mathbb{B}$ be a (non-split) fibration. We define a new fibration (on the same basis)

$$R(p : \mathbb{E} \to \mathbb{B}) = p' : \mathbb{E}' \to \mathbb{B}$$

as follows:

- $Obj(\mathbb{E}')$ consists of the pairs (X, ϕ), where $\phi : \mathtt{dom}_X \to p$, where $\mathtt{dom}_X$ is the domain fibration $\mathbb{B}/X \to \mathbb{B}$. We require ϕ to be a morphism of fibrations. We set $p'(X, \phi) = X$.
- $\mathbb{E}'[(X, \phi), (Y, \psi)]$ consists of the pairs (t, μ) where $t \in \mathbb{B}[X, Y]$ and μ is a natural transformation from ϕ to $\psi \circ \mathtt{dom}_t$ over $\mathbb{B}$.

We now define the comprehension structure as follows:

$$\mathcal{P}'(\phi) = \mathcal{P}(\phi(id)) \qquad \mathcal{P}'(t, \mu) = \mathcal{P}(\phi(t : t \to id) \circ \mu_{id}$$

The faithful embedding. Let A a local object over X. We set

$$F(A) = (X, \phi), \quad \text{where } \phi(f : Y \to X) = A[f]$$

For $f : A_1 \to A_2$, we set $F(f) = (u, \mu)$ where $u = p(f)$ and where, for every $v : X \to X_1 = p(A_1)$, μ_v is the unique morphism from $A_1[v]$ to $A_2[u \circ v]$ such that $(u \circ v)_{A_2} \circ \mu_v = (f \circ v_{A_1})$.

This functor is faithful since one can recover f from μ_{id}.

Future work. The analogy with monoidal categories, which are algebras for a certain 2-monad, further suggests that some of the constructions described here (glueing, strictification) for $\mathbb{ML}$-categories could be carried out in the even larger context of two-dimensional monad theory.

While we have focussed our attention on extensional type theory and its coherent interpretation in locally cartesian closed categories, we believe that our framework should have wider applicability. In particular, the presence of two levels of equality in the explicit syntax has a higher-dimensional category-theoretic flavour. We would like to investigate how to adapt our approach to intensional type theory [**?**] and more widely to homotopy type theory [**?**], by further varying the notions of equality in the modelled type theory and by examining the corresponding coherence issues.

Related work. Clairambault and Dybjer establish a biequivalence between the 2-category of LCCC's and a 2-category of categories with families, which amounts in our setting to a biequivalence between $\mathbb{ML}$ and $\mathbf{S}\mathbb{ML}$, where the latter is the full sub-2-category of $\mathbb{ML}$ spanned by the strict $\mathbb{ML}$-categories. This latter category does not play a role in our analysis here, but fits clearly as the fourth item in the matrix strict/non-strict for objects/morphisms.

Lumsdaine and Warren exploit the other strictification arising as a left, rather than right, adjoint to the inclusion functor from $\mathbf{S}\mathbb{ML}_\mathbf{s}$ to $\mathbb{ML}$ (over a fixed basis), which they nicely explain in terms of local universes. Again, this is relevant in the context of the present homotopic developments in type theory.