

# Metatheoretic proofs internally to presheaf categories\*

Rafaël Bocquet<sup>1</sup>, Ambrus Kaposi<sup>1</sup>, and Christian Sattler<sup>2</sup>

Eötvös Loránd University, Budapest, Hungary  
bocquet@inf.elte.hu and akaposi@inf.elte.hu  
University of Nottingham, United Kingdom  
sattler.christian@gmail.com

**Introduction** Proofs of the metatheoretic properties of dependent type theories and other typed logics and languages, such as proofs of canonicity, normalization, gluing, parametricity or various translations between theories, involve complicated inductions over the syntax of the theory. We present a way to structure such proofs by working in the internal type-theoretic languages of suitable presheaf categories.

**Internal models of type theory** We use categories with families (CwFs) [4, 2] equipped with additional structure as the models of our type theories. They consist of a category  $\mathcal{C}$ , equipped with presheaves of types and terms, objects representing the empty context and context extensions, along with natural transformations for each type-theoretic operation and equations between them. Most of the additional structure on  $\mathcal{C}$  can concisely be described in the type-theoretic internal language of the presheaf category  $\widehat{\mathcal{C}}$ . This observation is used in some of the existing general definitions of type theories [1, 9]. For example, the presheaves of types and terms, and the  $\mathbb{N}$  and  $\Pi$  type formers can be specified in the internal language of  $\widehat{\mathcal{C}}$  as follows:

$$\begin{aligned} \text{Ty} & : \text{Set} \\ \text{Tm} & : \text{Ty} \rightarrow \text{Set} \\ \mathbb{N} & : \text{Ty} \\ \Pi & : (A : \text{Ty}) \rightarrow (B : \text{Tm } A \rightarrow \text{Ty}) \rightarrow \text{Ty} \end{aligned}$$

The type  $(\text{Tm } A \rightarrow \text{Ty})$  of the argument  $B$  of  $\Pi$  is a presheaf function type: using the internal language of presheaf categories is a way to interpret higher-order abstract syntax (HOAS).  $\Pi$ -types are usually given externally by a map  $\Pi_\Gamma : (A : \text{Ty}_\Gamma) \rightarrow \text{Ty}_{\Gamma \triangleright A} \rightarrow \text{Ty}_\Gamma$ , natural in  $\Gamma$ , but the properties of the context extension operation  $(- \triangleright -)$  imply that the internal and external definitions are equivalent.

Only the empty context and the context extension operations can not directly be described internally, unless we use the interpretation of *crisp type theory* [8, 6] in  $\widehat{\mathcal{C}}$  and its comonadic modality  $\flat$ .

**Internal dependent models** The generalized algebraic presentation of CwFs automatically provides an initial model  $\mathcal{S}$  satisfying an induction principle: there is a dependent section from  $\mathcal{S}$  to any dependent model over  $\mathcal{S}$ . The definition of dependent model can be derived mechanically from the QIIT-signature presenting the type theory [5]. By applying a similar transformation

---

\*The first author was supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.3-VEKOP-16-2017-00002). The second author was supported by the ÚNKP-19-4 New National Excellence Program of the Ministry for Innovation and Technology and by the Bolyai Fellowship of the Hungarian Academy of Sciences. The third author was supported by USAF grant FA9550-16-1-0029.

to the internal definition of models, we define a notion of dependent model internal to  $\widehat{\mathcal{S}}$ :

$$\begin{aligned} \mathbf{Ty}^\bullet & : \mathbf{Ty} \rightarrow \mathbf{Set} \\ \mathbf{Tm}^\bullet & : \{A\}(A^\bullet : \mathbf{Ty}^\bullet A)(a : \mathbf{Tm} A) \rightarrow \mathbf{Set} \\ \mathbf{N}^\bullet & : \mathbf{Ty}^\bullet \mathbf{N} \\ \mathbf{\Pi}^\bullet & : \{A\}(A^\bullet : \mathbf{Ty}^\bullet A)\{B\}(B^\bullet : \{a\}(a^\bullet : \mathbf{Tm}^\bullet A^\bullet a) \rightarrow \mathbf{Ty}^\bullet (B a)) \rightarrow \mathbf{Ty}^\bullet (\mathbf{\Pi} A B) \end{aligned}$$

Internal and external dependent models do not exactly correspond to each other, but we can still reconstruct an external dependent model from any internal one, and then obtain, externally, a dependent section of the reconstructed external model.

A proof of canonicity based on logical predicates can be given as an internal dependent model in the internal language of  $\widehat{\mathcal{S}}$ :

$$\begin{aligned} \mathbf{Ty}^\bullet A & \quad \equiv \quad \mathbf{Tm} A \rightarrow \mathbf{Set} \\ \mathbf{Tm}^\bullet A^\bullet & \quad \equiv \quad \lambda(a : \mathbf{Tm} A) \mapsto A^\bullet a \\ \mathbf{N}^\bullet & \quad \equiv \quad \lambda(n : \mathbf{Tm} \mathbf{N}) \mapsto (m : \mathbf{N}) \times (n = \mathbf{suc}^m \mathbf{zero}) \\ \mathbf{\Pi}^\bullet A^\bullet B^\bullet & \quad \equiv \quad \lambda(f : \mathbf{Tm} (\mathbf{\Pi} A B)) \mapsto (a : \mathbf{Tm} A)(a^\bullet : A^\bullet a) \rightarrow B^\bullet (\mathbf{app} f a) \end{aligned}$$

**Internal induction principles** The presheaf category  $\widehat{\mathcal{S}}$  is not a nice setting for more complicated proofs, such as normalization proofs: it forces all of our constructions to be stable under all substitutions, but normal forms are only stable under renamings. To fix this, we change the base category. For normalization, we work in the presheaf category  $\widehat{\mathcal{G}}$  over the comma category  $\mathcal{G} = (\mathcal{S} \downarrow F)$ , where  $F : \mathcal{R} \rightarrow \mathcal{S}$  is the CwF morphism from the CwF of renamings  $\mathcal{R}$  to  $\mathcal{S}$ .

The presheaf category  $\widehat{\mathcal{G}}$  has many good properties: the CwF structures of  $\mathcal{R}$  and  $\mathcal{S}$  can faithfully be transported over  $\mathcal{G}$ , and the CwF morphism  $F : \mathcal{R} \rightarrow \mathcal{S}$  can also faithfully be encoded. Furthermore, we can distinguish in  $\widehat{\mathcal{G}}$  the presheaves that come from  $\widehat{\mathcal{R}}$  or  $\widehat{\mathcal{S}}$ ; and call them  $\mathcal{R}$ -discrete or  $\mathcal{S}$ -discrete presheaves. Type-theoretically, they are accessible reflective subuniverses of the universe of all presheaves. Moreover, the  $\mathcal{R}$ -/ $\mathcal{S}$ -discrete presheaves can be identified with the discrete types arising from interpretations of spatial type theory [8] in  $\widehat{\mathcal{G}}$ . In particular, we have an adjoint pair of modalities  $(\Box \dashv \Diamond)$ , where the comonadic modality  $\Box$  classifies the  $\mathcal{R}$ -discrete presheaves. This means that we can reuse the theory of modalities developed in [7, 8] in this setting.

We can then use the  $\Box$  modality to define *F-relative internal dependent models*, which encode inductions over the syntax whose results are only stable under renamings. We can define and prove an induction principle for *F*-relative dependent models.

A recent normalization proof by Coquand [3] can be translated to this framework by defining a suitable *F*-relative dependent model. We can prove normalization and the decidability of equality for types and terms fully internally, without ever working explicitly with contexts, substitutions or renamings in the proof.

**Agda formalization<sup>1</sup>** We have formalized in Agda internal proofs of canonicity and normalization for a reasonably large dependent type theory (including  $\Pi$ -types with the  $\eta$  rule, booleans, natural numbers, identity types and a universe closed under the other type formers).

## References

- [1] Paolo Capriotti. *Models of type theory with strict equality*. PhD thesis, University of Nottingham, UK, 2017.

<sup>1</sup>[https://gitlab.com/RafaelBocquet/internal\\_metatheory/tree/master/Agda](https://gitlab.com/RafaelBocquet/internal_metatheory/tree/master/Agda)

- [2] Simon Castellan, Pierre Clairambault, and Peter Dybjer. Categories with families: Untyped, simply typed, and dependently typed. *CoRR*, abs/1904.00827, 2019.
- [3] Thierry Coquand. Canonicity and normalization for dependent type theory. *Theor. Comput. Sci.*, 777:184–191, 2019.
- [4] Peter Dybjer. Internal type theory. In Stefano Berardi and Mario Coppo, editors, *Types for Proofs and Programs, International Workshop TYPES'95, Torino, Italy, June 5-8, 1995, Selected Papers*, volume 1158 of *Lecture Notes in Computer Science*, pages 120–134. Springer, 1995.
- [5] Ambrus Kaposi, András Kovács, and Thorsten Altenkirch. Constructing quotient inductive-inductive types. *PACMPL*, 3(POPL):2:1–2:24, 2019.
- [6] Daniel R. Licata, Ian Orton, Andrew M. Pitts, and Bas Spitters. Internal universes in models of homotopy type theory. In Hélène Kirchner, editor, *3rd International Conference on Formal Structures for Computation and Deduction, FSCD 2018, July 9-12, 2018, Oxford, UK*, volume 108 of *LIPICs*, pages 22:1–22:17. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018.
- [7] Egbert Rijke, Michael Shulman, and Bas Spitters. Modalities in homotopy type theory. *CoRR*, abs/1706.07526, 2017.
- [8] Michael Shulman. Brouwer’s fixed-point theorem in real-cohesive homotopy type theory. *Mathematical Structures in Computer Science*, 28(6):856–941, 2018.
- [9] Taichi Uemura. A general framework for the semantics of type theory. *CoRR*, abs/1904.04097, 2019.