

Formalising the metatheory of type theory using quotient inductive types

Ambrus Kaposi
Eötvös Loránd University, Budapest

j.w.w. Thorsten Altenkirch and András Kovács

Workshop on Foundations for the
Practical Formalization of Mathematics,
Nantes,
27 April 2017

Motivation

Why type theory in type theory?

- Study the metatheory of type theory in a nice language
- Type-safe template type theory (metaprogramming)
 - ▶ generic programming
 - ▶ extensions of type theory justified by models

Plan

- 1 Extrinsic vs. intrinsic syntax for simple type theory
- 2 Extrinsic vs. intrinsic syntax for type theory
- 3 Defining functions from the intrinsic syntax
- 4 Relationship of extrinsic and intrinsic syntax
- 5 Models

Extrinsic vs. intrinsic syntax for simple type theory

Extrinsic syntax for simple type theory

4 inductive sets + 2 inductive relations.

$$x ::= \text{zero} \mid \text{suc } x$$
$$t ::= x \mid \text{lam } t \mid \text{app } t t'$$
$$A ::= \iota \mid A \Rightarrow A'$$
$$\Gamma ::= \cdot \mid \Gamma, A$$

$$\frac{}{\Gamma, A \vdash_v \text{zero} : A}$$

$$\frac{\Gamma \vdash_v x : A}{\Gamma, B \vdash_v \text{suc } x : A}$$

$$\frac{\Gamma \vdash_v x : A}{\Gamma \vdash x : A}$$

$$\frac{\Gamma, A \vdash t : B}{\Gamma \vdash \text{lam } t : A \rightarrow B}$$

$$\frac{\Gamma \vdash t : A \rightarrow B \quad \Gamma \vdash u : A}{\Gamma \vdash \text{app } t u : B}$$

Intrinsic syntax for simple type theory

4 inductively defined families of sets.

Ty	: Set
ι	: Ty
$- \Rightarrow -$: Ty \rightarrow Ty \rightarrow Ty
Con	: Set
\cdot	: Con
$-, -$: Con \rightarrow Ty \rightarrow Con
Var	: Con \rightarrow Ty \rightarrow Set
zero	: Var (Γ, A) A
suc	: Var Γ A \rightarrow Var (Γ, B) A
Tm	: Con \rightarrow Ty \rightarrow Set
var	: Var Γ A \rightarrow Tm Γ A
lam	: Tm (Γ, A) B \rightarrow Tm Γ ($A \Rightarrow B$)
app	: Tm Γ ($A \Rightarrow B$) \rightarrow Tm Γ A \rightarrow Tm Γ B

Extrinsic vs. intrinsic syntax for type theory

Extrinsic syntax

4 inductive sets + 8 inductive relations (we can't avoid talking about conversion).

$PCon, PTy, PTm, PTms : Set$

$\vdash_{Con} : PCon \rightarrow Prop$

$\vdash_{Ty} : PCon \rightarrow PTy \rightarrow Prop$

$\vdash_{Tm} : PCon \rightarrow PTy \rightarrow PTm \rightarrow Prop$

$\vdash_{Tms} : PCon \rightarrow PCon \rightarrow PTms \rightarrow Prop$

$\sim_{Con} : PCon \rightarrow PCon \rightarrow Prop$

$\sim_{Ty} : PCon \rightarrow PTy \rightarrow PTy \rightarrow Prop$

$\sim_{Tm} : PCon \rightarrow PTy \rightarrow PTm \rightarrow PTm \rightarrow Prop$

$\sim_{Tms} : PCon \rightarrow PCon \rightarrow PTms \rightarrow PTms \rightarrow Prop$

Relations are given by rules for ER, coercion, congruence, conversion.

Extrinsic syntax, PER variant (Dybjer: Undec... LCCC)

4 inductive sets + 4 inductive relations.

$\text{PCon}, \text{PTy}, \text{PTm}, \text{PTms} : \text{Set}$

$\sim_{\text{Con}} : \text{PCon} \rightarrow \text{PCon} \rightarrow \text{Prop}$

$\sim_{\text{Ty}} : \text{PCon} \rightarrow \text{PTy} \rightarrow \text{PTy} \rightarrow \text{Prop}$

$\sim_{\text{Tm}} : \text{PCon} \rightarrow \text{PTy} \rightarrow \text{PTm} \rightarrow \text{PTm} \rightarrow \text{Prop}$

$\sim_{\text{Tms}} : \text{PCon} \rightarrow \text{PCon} \rightarrow \text{PTms} \rightarrow \text{PTms} \rightarrow \text{Prop}$

Recovering typing relations as reflexive cases:

$\vdash_{\text{Con}} \Gamma := \Gamma \sim_{\text{Con}} \Gamma$

$\Gamma \vdash_{\text{Ty}} A := \Gamma \vdash A \sim_{\text{Ty}} A$

$\Gamma \vdash_{\text{Tm}} t : A := \Gamma \vdash t \sim_{\text{Tm}} t : A$

$\Gamma \vdash_{\text{Tms}} \sigma : \Delta := \Gamma \vdash \sigma \sim_{\text{Tms}} \sigma : \Delta$

Congruence rules and typing rules are identified. E.g.

$$\frac{\vdash_{\text{Con}} \Gamma \quad \Gamma \vdash_{\text{Ty}} A}{\vdash_{\text{Con}} \Gamma, A} \text{ is expressed by } \frac{\Gamma \sim_{\text{Con}} \Gamma' \quad \Gamma \vdash A \sim_{\text{Ty}} A'}{\Gamma, A \sim_{\text{Con}} \Gamma', A'}$$

Intrinsic syntax (James Chapman: TT should eat itself)

An inductive inductive definition of 4 families of sets + 4 families of relations.

$\text{Con} : \text{Set}$

$\text{Ty} : \text{Con} \rightarrow \text{Set}$

$\text{Tm} : (\Gamma : \text{Con}) \rightarrow \text{Ty } \Gamma \rightarrow \text{Set}$

$\text{Tms} : \text{Con} \rightarrow \text{Con} \rightarrow \text{Set}$

$\sim_{\text{Con}} : \text{Con} \rightarrow \text{Con} \rightarrow \text{Prop}$

$\sim_{\text{Ty}} : (\Gamma : \text{Con}) \rightarrow \text{Ty } \Gamma \rightarrow \text{Ty } \Gamma \rightarrow \text{Prop}$

$\sim_{\text{Tm}} : (\Gamma : \text{Con})(A : \text{Ty } \Gamma) \rightarrow \text{Tm } \Gamma A \rightarrow \text{Tm } \Gamma A \rightarrow \text{Prop}$

$\sim_{\text{Tms}} : (\Gamma \Delta : \text{Con}) \rightarrow \text{Tms } \Gamma \Delta \rightarrow \text{Tms } \Gamma \Delta \rightarrow \text{Prop}$

No more separation of pre-things and things. One can only talk about well-typed terms.

Quotient intrinsic syntax

A quotient inductive inductive definition of 4 families of sets.

$\text{Con} : \text{Set}$

$\text{Ty} : \text{Con} \rightarrow \text{Set}$

$\text{Tm} : (\Gamma : \text{Con}) \rightarrow \text{Ty } \Gamma \rightarrow \text{Set}$

$\text{Tms} : \text{Con} \rightarrow \text{Con} \rightarrow \text{Set}$

- Conversion relation is the identity type for each set. Conversion rules (e.g. β, η) are given as equality constructors.
- Rules for ER, coercion and congruence are properties of the identity type.
- No more separation of convertible things. One can only do constructions on the syntax up to equality.

The syntax for a type theory with Π and an empty universe

\cdot	: Con	$[\text{id}]$: $A[\text{id}] \equiv A$
$-, -$: $(\Gamma : \text{Con}) \rightarrow \text{Ty } \Gamma \rightarrow \text{Con}$	$[\]$: $A[\sigma][\nu] \equiv A[\sigma \circ \nu]$
$-[-]$: $\text{Ty } \Delta \rightarrow \text{Tms } \Gamma \Delta \rightarrow \text{Ty } \Gamma$	$\text{id} \circ$: $\text{id} \circ \sigma \equiv \sigma$
id	: $\text{Tms } \Gamma \Gamma$	oid	: $\sigma \circ \text{id} \equiv \sigma$
$- \circ -$: $\text{Tms } \Theta \Delta \rightarrow \text{Tms } \Gamma \Theta \rightarrow \text{Tms } \Gamma \Delta$	$\circ \circ$: $(\sigma \circ \nu) \circ \delta \equiv \sigma \circ (\nu \circ \delta)$
ϵ	: $\text{Tms } \Gamma \cdot$	$\epsilon \eta$: $\{\sigma : \text{Tms } \Gamma \cdot\} \rightarrow \sigma \equiv \epsilon$
$-, -$: $(\sigma : \text{Tms } \Gamma \Delta) \rightarrow \text{Tm } \Gamma A[\sigma] \rightarrow \text{Tms } \Gamma (\Delta, A)$	$\pi_1 \beta$: $\pi_1(\sigma, t) \equiv \sigma$
π_1	: $\text{Tms } \Gamma (\Delta, A) \rightarrow \text{Tms } \Gamma \Delta$	$\pi \eta$: $(\pi_1 \sigma, \pi_2 \sigma) \equiv \sigma$
$-[-]$: $\text{Tm } \Delta A \rightarrow (\sigma : \text{Tms } \Gamma \Delta) \rightarrow \text{Tm } \Gamma A[\sigma]$	$, \circ$: $(\sigma, t) \circ \nu \equiv (\sigma \circ \nu), (\text{[]}_* t[\nu])$
π_2	: $(\sigma : \text{Tms } \Gamma (\Delta, A)) \rightarrow \text{Tm } \Gamma A[\pi_1 \sigma]$	$\pi_2 \beta$: $\pi_2(\sigma, t) \equiv \pi_1^\beta t$
Π	: $(A : \text{Ty } \Gamma) \rightarrow \text{Ty } (\Gamma, A) \rightarrow \text{Ty } \Gamma$	$\Pi \]$: $(\Pi A B)[\sigma] \equiv \Pi A[\sigma] B[\sigma \uparrow]$
lam	: $\text{Tm } (\Gamma, A) B \rightarrow \text{Tm } \Gamma (\Pi A B)$	$\Pi \beta$: $\text{app } (\text{lam } t) \equiv t$
app	: $\text{Tm } \Gamma (\Pi A B) \rightarrow \text{Tm } (\Gamma, A) B$	$\Pi \eta$: $\text{lam } (\text{app } t) \equiv t$
U	: $\text{Ty } \Gamma$	$\text{lam } \]$: $(\text{lam } t)[\sigma] \equiv \text{[]}^\Pi \text{lam } (t[\sigma \uparrow])$
El	: $\text{Tm } \Gamma \text{U} \rightarrow \text{Ty } \Gamma$	$\text{U } \]$: $\text{U}[\sigma] \equiv \text{U}$
		$\text{El } \]$: $(\text{El } \hat{A})[\sigma] \equiv \text{El } (\text{U} \]_* \hat{A}[\sigma])$

Defining functions from the intrinsic syntax

Nondependent eliminator (recursor)

The inductive type of natural numbers:

$$\mathbb{N} : \text{Set}$$
$$\text{zero} : \mathbb{N}$$
$$\text{suc} : \mathbb{N} \rightarrow \mathbb{N}$$

Arguments of the recursor (a natural number algebra):

$$\mathbb{N}_1 : \text{Set}$$
$$\text{zero}_1 : \mathbb{N}_1$$
$$\text{suc}_1 : \mathbb{N}_1 \rightarrow \mathbb{N}_1$$

The recursor is a function which respects the operations (an algebra morphism).

$$\text{Rec}\mathbb{N} : \mathbb{N} \rightarrow \mathbb{N}_1$$
$$\text{Rec}\mathbb{N} \text{ zero} = \text{zero}_1$$
$$\text{Rec}\mathbb{N} (\text{suc } n) = \text{suc}_1 (\text{Rec}\mathbb{N} n)$$

Recursor for a higher inductive type

Example:

Constructors:

I : Set

left : I

right : I

segment : left \equiv right

Arguments of the recursor:

I_1 : Set

left₁ : I_1

right₁ : I_1

segment₁ : left₁ \equiv right₁

The recursor:

RecI : $I \rightarrow I_1$

RecI left = left₁

RecI right = right₁

ap RecI segment = segment₁

Recursor for intrinsic type theory

- An algebra for the quotient intrinsic syntax is a categories with families (CwF, a notion of model of type theory).
 - ▶ An algebra for the intrinsic syntax is a more relaxed CwF, where conversion can be interpreted by relations other than equality.
- The recursor is a strict morphism of models from the initial model (the syntax) to the model given by the arguments of the recursor.
- The recursor for an inductive inductive type is recursive recursive (Forsberg).

$\text{Con} : \text{Set}$

$\text{Ty} : \text{Con} \rightarrow \text{Set}$

$\text{Con}_1 : \text{Set}$

$\text{Ty}_1 : \text{Con}_1 \rightarrow \text{Set}$

$\text{RecCon} : \text{Con} \rightarrow \text{Con}_1$

$\text{RecTy} : \text{Ty } \Gamma \rightarrow \text{Ty}_1 (\text{RecCon } \Gamma)$

Recursor for quotient inductive types in Agda

```
{-# OPTIONS --rewriting #-}
```

```
postulate
```

```
  Con : Set
```

```
  Ty  : Con → Set
```

```
  _,_ : (Γ : Con) → Ty Γ → Con
```

```
  RecCon : Con → Con1
```

```
  RecTy  : Ty Γ → Ty1 (RecCon Γ)
```

```
  β, : RecCon (Γ , A) ≡ RecCon Γ ,1 RecTy A
```

```
{-# REWRITE β, #-}
```

```
...
```

Relationship of extrinsic and intrinsic syntax

From intrinsic to extrinsic syntax (work in progress)

- Inductive inductive types can be represented by normal inductive types and “typing relations” (noticed by Altenkirch and Capriotti). Similar to representing indexed W -types by plain ones.
 - ▶ If we start with intrinsic syntax, we get back an extrinsic syntax which is
 - ★ fully annotated (e.g. $- \circ -$ has 5 arguments)
 - ★ paranoid (e.g. typing for lam needs well-formedness of Γ)
 - ▶ The usual syntax comes after some ad-hoc constructions (removing assumptions that are admissible).
- Going from quotient intrinsic syntax to intrinsic is doing an internal setoid-interpretation. For example, the inductively defined setoid equality relation for \mathbb{N} :

$$\sim_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{N} \rightarrow \text{Prop}$$

$$\sim_{\text{zero}} : \text{zero} \sim_{\mathbb{N}} \text{zero}$$

$$\sim_{\text{suc}} : (n_0 \sim_{\mathbb{N}} n_1) \rightarrow \text{suc } n_0 \sim_{\mathbb{N}} \text{suc } n_1$$

From extrinsic to intrinsic syntax (Streicher: Semantics of TT)

A model is given by a category with families \mathcal{C} .

A preterm carries enough information to reconstruct its precontext (ctx) and pretype (ty). Similarly for the other pre-things.

Partial functions by recursion on the presyntax:

$$\llbracket - \rrbracket_{\text{Con}} : \text{PCon} \rightarrow |\mathcal{C}|$$

$$\llbracket - \rrbracket_{\text{Ty}} : (A : \text{PTy}) \rightarrow \text{Ty}_{\mathcal{C}}(\llbracket \text{ctx } A \rrbracket_{\text{Con}})$$

$$\llbracket - \rrbracket_{\text{Tm}} : (t : \text{PTm}) \rightarrow \text{Tm}_{\mathcal{C}}(\llbracket \text{ctx } t \rrbracket_{\text{Con}}, \llbracket \text{ty } t \rrbracket_{\text{Ty}})$$

$$\llbracket - \rrbracket_{\text{Tms}} : (\sigma : \text{PTms}) \rightarrow \mathcal{C}(\llbracket \text{dom } \sigma \rrbracket_{\text{Con}}, \llbracket \text{cod } \sigma \rrbracket_{\text{Con}})$$

By induction on the typing and conversion relations we have:

$$\Gamma \vdash_{\text{Tm}} t : A \rightarrow \llbracket t \rrbracket_{\text{Tm}} \text{ is defined}$$

$$\Gamma \vdash t \sim_{\text{Tm}} t' : A \rightarrow \llbracket t \rrbracket_{\text{Tm}} \text{ and } \llbracket t' \rrbracket_{\text{Tm}} \text{ are defined and are equal}$$

Similarly for contexts, types and substitutions.

This can be used to map extrinsic syntax to quotient intrinsic syntax.

Conjecture

$$\text{Con} \cong \left((\Gamma : \text{PCon}) \times \vdash \Gamma \right) / \sim_{\text{Con}}$$

$$\begin{aligned} & (\Gamma : \text{Con}) \times \text{Ty } \Gamma \\ \cong & \left((\Gamma : \text{PCon}) \times \vdash \Gamma \times (A : \text{PTy}) \times \Gamma \vdash A \right) / \sim_{\text{Con}} / \sim_{\text{Ty}} \end{aligned}$$

...

Typechecking (not yet formalised)

$\text{inferTm} : (\Gamma : \text{Con})(t : \text{PTm}) \rightarrow (A' : \text{Ty } \Gamma) \times \text{Tm } \Gamma A'$

$\text{checkTm} : (\Gamma : \text{Con})(t : \text{PTm}) \rightarrow (A' : \text{Ty } \Gamma) \rightarrow \text{Tm } \Gamma A'$

$\text{checkTy} : (\Gamma : \text{Con}) \rightarrow \text{PTy} \rightarrow \text{Ty } \Gamma$

$\text{inferTm } \Gamma x := \text{lookup } \Gamma x$

$\text{inferTm } \Gamma (t u) := \text{case inferTm } \Gamma t \text{ of}$

$((x : A') \rightarrow B', t') \mapsto \text{case checkTm } \Gamma u A' \text{ of}$
 $u' \mapsto (B'[x \mapsto u'], t' u')$

$\text{inferTm } \Gamma (t : A) := \text{case checkTy } \Gamma A \text{ of}$

$A' \mapsto \text{case checkTm } \Gamma t A' \text{ of}$
 $t' \mapsto (A', t')$

$\text{checkTm } \Gamma (\lambda x. t) ((x : A') \rightarrow B') := \text{lam } (\text{checkTm } (\Gamma, x : A') t B')$

$\text{checkTm } \Gamma t A := \text{case inferTm } \Gamma t \text{ of}$

$(A', t') \mapsto \text{if } A \stackrel{?}{=} A' \text{ then } t'$

Models

Models formalised in Agda (with K and funext)

For a theory with Π , a base type and a family over the base type.

- Non-dependent eliminator:
 - ▶ standard model
 - ▶ presheaf model
 - ▶ setoid model
- Dependent eliminator:
 - ▶ logical predicate translation of Bernardy
 - ▶ presheaf logical predicate interpretation

Standard model

- A sanity check
- Every syntactic construct is interpreted as the corresponding metatheoretic construction.

$$\begin{aligned} \text{Con}_1 &:= \text{Set} \\ \text{T}\gamma_1 \llbracket \Gamma \rrbracket &:= \llbracket \Gamma \rrbracket \rightarrow \text{Set} \\ \llbracket \Gamma \rrbracket, 1 \llbracket A \rrbracket &:= (\gamma : \llbracket \Gamma \rrbracket) \times \llbracket A \rrbracket \gamma \\ &\dots \\ \Pi_1 \llbracket A \rrbracket \llbracket B \rrbracket \gamma &:= (x : \llbracket A \rrbracket \gamma) \rightarrow \llbracket B \rrbracket (\gamma, x) \\ \text{lam}_1 \llbracket t \rrbracket \gamma &:= \lambda x \rightarrow \llbracket t \rrbracket (\gamma, x) \\ &\dots \\ \Pi\beta_1 &:= \text{refl} \end{aligned}$$

- We defined this for a syntax with Σ , \perp , \top , Bool , \mathbb{N} , Id as well.

Logical predicate interpretation

Parametricity expressed as a syntactic translation.

$$\frac{\Gamma \vdash}{\Gamma^P \vdash} \quad \frac{\Gamma \vdash A : U}{\Gamma^P \vdash A^P : A \rightarrow U} \quad \frac{\Gamma \vdash t : A}{\Gamma^P \vdash t^P : A^P t}$$

All of the following equations need to be well-typed (and preserve conversion).

$$(\Gamma, x : A)^P \quad := \Gamma^P, x : A, x^M : A^P x$$

$$x^P \quad := x^M$$

$$U^P A \quad := A \rightarrow U$$

$$((x : A) \rightarrow B)^P f \quad := (x : A)(x^M : A^P x) \rightarrow B^P (f x)$$

$$(\lambda x. t)^P \quad := \lambda x x^M. t^P$$

$$(f a)^P \quad := f^P a a^P$$

NBE for dependent types

Presheaf logical predicate

$$P_{\Delta} : \forall \Psi. \text{Tms } \Psi \Delta \rightarrow \text{Set}$$

$$P_A : \forall \Psi. (\rho : \text{Tms } \Psi \Gamma) \rightarrow P_{\Gamma} \rho \rightarrow \text{Tm } \Psi A[\rho] \rightarrow \text{Set}$$

$$P_{\sigma} : \forall \Psi. (\rho : \text{Tms } \Psi \Gamma) \rightarrow P_{\Gamma} \rho \rightarrow P_{\Delta} (\sigma \circ \rho)$$

$$P_t : \forall \Psi. (\rho : \text{Tms } \Psi \Gamma) (\rho : P_{\Gamma} \rho) \rightarrow P_A \rho \rho (t[\rho])$$

At the base type:

$$P_{\iota} \rho t = \text{isNf } \Psi \iota t$$

Quote and unquote:

$$q_A : (\rho : P_{\Gamma} \rho) (t : \text{Tm } \Psi A[\rho]) \rightarrow P_A \rho \rho t \rightarrow \text{isNf } \Psi A[\rho] t$$

$$u_A : (\rho : P_{\Gamma} \rho) (t : \text{Tm } \Psi A[\rho]) \rightarrow \text{isNe } \Psi A[\rho] t \rightarrow P_A \rho \rho t$$

NBE for a universe and Bool with large elimination (not yet formalised)

$$P_{\Delta} : \forall \Psi. (\rho : \text{Tms } \Psi \Delta) \rightarrow (r : \text{Set}) \times (u : \text{isNes } \Psi \Delta \rho \rightarrow r)$$

$$P_A : \forall \Psi. (\rho : \text{Tms } \Psi \Delta) \rightarrow P_{\Gamma} \rho. r \rightarrow (t : \text{Tm } \Psi A[\rho]) \\ \rightarrow (r : \text{Set}) \times (q : r \rightarrow \text{isNf } \Psi A[\rho] t) \\ \times (u : \text{isNe } \Psi A[\rho] t \rightarrow r)$$

$$P_t : \forall \Psi. (\rho : \text{Tms } \Psi \Delta) \rightarrow P_{\Gamma} \rho. r \rightarrow P_A \rho q (t[\rho]). r$$

$$P_U \Psi (\rho : \text{Tms } \Psi \Gamma)(\rho : P_{\Gamma} \rho)(\hat{A} : \text{Tm } \Psi U). r$$

$$:= \text{isNf } \Psi U \hat{A} \times \forall \Omega. (\beta : \text{REN}(\Omega, \Psi))(t : \text{Tm } \Omega (\text{El } \hat{A}[\beta]))$$

$$\rightarrow (r : \text{Set}) \times (q : r \rightarrow \text{isNf } \Omega (\text{El } \hat{A}[\beta]) t) \times (u : \text{isNe } \Omega (\text{El } \hat{A}[\beta]) t \rightarrow r)$$

Summary

Quotient intrinsic syntax has the following properties:

- more abstract: close to categorical models; analogy with HITs and setoids (c.f. Peter Dybjer's talk)
- get back old-style syntax using general methods (WIP)
- typechecking and normalisation fit well
- definition of operations on the syntax in a type-safe way

Future work:

- finish unfinished things
- extend the syntax with QIITs to do full internalisation
- formalisation is hard
- we need cubical type theory or similar to compute with quotient types

Formalisation: <http://bitbucket.org/akaposi/tt-in-tt>