Coherence of definitional equality in type theory

Rafaël Bocquet

HoTTEST, September 23, 2021

Problem

In type theory we have *typal equalities*,

$$0 + n \simeq n \qquad n + m \simeq m + n \qquad \mathsf{refl} \cdot p \simeq p$$

some of them are *definitional equalities*

$$n + 0 = n \qquad p \cdot \mathsf{refl} = p$$

Can we add new definitional equalities?

- Constructing (higher-dimensional) paths and fillers becomes easier.
  (We avoid *coherence hell*.)
- The new definitional equalities may not hold in known models.

We need conservativity/strictification/coherence theorems.

Examples: weak computation rule

We can replace the computation rules of Id-, Σ-, Π-types by weak computation rules.

$$\frac{a : A \qquad b : B(a)}{\pi_1\text{-}\beta : \pi_1(\mathsf{pair}(a, b)) \simeq_A a}$$

The path types of cubical type theory satisfy the weak computation rule of Id-types.

Are the usual computation rules conservative over the weak computation rules?

Examples: composition of paths

Can identity types satisfy the groupoid laws definitionally?

$$p \cdot \mathsf{refl} = p \qquad \mathsf{refl} \cdot p = p \qquad p \cdot (q \cdot r) = (p \cdot q) \cdot r$$

$$p^{-1} \cdot p = \mathsf{refl} \qquad (p^{-1})^{-1} = p$$

$$\dots$$

$$\mathsf{ap}(f, \mathsf{refl}) = \mathsf{refl} \qquad \mathsf{ap}(f, p \cdot q) = \mathsf{ap}(f, p) \cdot \mathsf{ap}(f, q)$$

$$\mathsf{ap}(f, p^{-1}) = \mathsf{ap}(f, p)^{-1}$$

$$\dots$$

Examples: universes of strict algebraic structures

Can we extend HoTT with a universe StrProp *of "strict" propositions and an equivalence* StrProp $\cong$ Prop*?*

$$\frac{A : \mathsf{StrProp} \qquad x, y : A}{x = y}$$

Can we also have a universe StrMonoid of strictly associative and unital monoids?

What about "strict" rings, "strict" categories, etc.?

Can we also equip StrProp with operations?

$$\frac{[a : A]\ B(a) : \mathsf{StrProp}}{\forall(A, B) : \mathsf{StrProp}}$$

Categorical semantics of type theories

A Category with Families (CwF) consists of:

- a category $\mathcal{C}$ with a terminal object;
- a presheaf of types $\mathsf{Ty}_{\mathcal{C}} : \mathrm{Psh}(\mathcal{C})$;
- a (locally representable) presheaf of terms $\mathsf{Tm}_{\mathcal{C}} : \mathsf{Ty}_{\mathcal{C}} \to \mathrm{RepPsh}(\mathcal{C})$;

A model of a type theory $\mathbb{T}$ is a CwF equipped with additional structure.

$$\frac{A \text{ type} \qquad [a : A]\ B(a) \text{ type}}{\Pi(A, B) \text{ type}} \qquad\qquad \Pi : (A : \mathsf{Ty}_{\mathcal{C}})(B : \mathsf{Tm}_{\mathcal{C}}(A) \to \mathsf{Ty}_{\mathcal{C}}) \to \mathsf{Ty}_{\mathcal{C}}$$

Locally finitely presentable 1-category $\mathbf{Mod}_{\mathbb{T}}$ of models of $\mathbb{T}$.

Syntax: initial object $\mathbf{0}_{\mathbb{T}} : \mathbf{Mod}_{\mathbb{T}}$.

Freely generated models $\mathbf{0}_{\mathbb{T}}[\cdots]$.

Hofmann's conservativity theorem

Uniqueness of Identity Proofs

$$\frac{p : \mathsf{Id}(x, x)}{\mathsf{uip}(p) : \mathsf{Id}(p, \mathsf{refl})}$$

Equality reflection

$$\frac{p : \mathsf{Id}(x, y)}{x = y}$$

Theorem (Hofmann, 1995)

Equality reflection is conservative over intensional type theory with UIP (and function extensionality).

If $(\Gamma \vdash_{\mathsf{ITT}} A \text{ type})$ *and* $(|\Gamma| \vdash_{\mathsf{ETT}} a : |A|)$*, then there exists some* $(\Gamma \vdash_{\mathsf{ITT}} a_0 : A)$ *such that* $|a_0| = a$*.*

The map $|-| : \mathbf{0}_{\mathsf{ITT}} \to \mathbf{0}_{\mathsf{ETT}}$ is surjective on types and terms.

Proof of Hofmann's conservativity theorem

Equivalence relations $(\sim)$ on types and terms of ITT:

$$\begin{aligned} (A \sim B) &\iff \exists p : \mathsf{Tm}_{\mathsf{ITT}}(\mathsf{Id}(\mathcal{U}, A, B)) \\ ((a : A) \sim (b : B)) &\iff \exists p : \mathsf{Tm}_{\mathsf{ITT}}(\mathsf{Id}((X : \mathcal{U}) \times X, (A, a), (B, b))) \end{aligned}$$

By UIP, if $(a : A) \sim (b : A)$, then there exists $p : \mathsf{Tm}_{\mathsf{ITT}}(\mathsf{Id}(A, a, b))$.

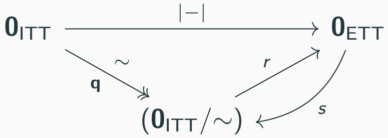
Furthermore, $(\mathsf{Tm}_{\mathsf{ITT}}, \sim) \twoheadrightarrow (\mathsf{Ty}_{\mathsf{ITT}}, \sim)$ is a setoid fibration:
If $(A \sim B)$, then for $a : \mathsf{Tm}_{\mathsf{ITT}}(A)$, there exists $b : \mathsf{Tm}_{ITT}(B)$ such that $(a \sim b)$.

All type- and term- formers respect $(\sim)$. For $\lambda(-)$ (and other binders) this requires function extensionality.

Proof of Hofmann's conservativity theorem

Quotients $(\mathsf{Ty}_{\mathsf{ITT}}/\sim)$ and $(\mathsf{Tm}_{\mathsf{ITT}}/\sim)$.

We can construct a quotient model $(\mathbf{0}_{\mathsf{ITT}}/\sim)$.



Since $|-|$ is a retract of $\mathbf{q}$, $|-|$ is surjective on types and terms.

(Alternative: Use the relative induction principle for $\mathcal{R}en(\mathbf{0}_{\mathsf{ITT}}) \to \mathbf{0}_{\mathsf{ETT}}$)

Mac Lane's coherence theorem for monoidal categories

In **MonCat**:

$$\begin{aligned} \alpha_{x,y,z} &: (x \otimes y) \otimes z \simeq x \otimes (y \otimes z) \\ \lambda_x &: (I \otimes x) \simeq x \\ \rho_x &: (x \otimes I) \simeq x \end{aligned}$$

In **StrMonCat**:

$$\begin{aligned} \alpha_{x,y,z} &= \mathrm{id} \\ \lambda_x &= \mathrm{id} \\ \rho_x &= \mathrm{id} \end{aligned}$$

$$\mathbf{MonCat} \underset{L}{\overset{R}{\rightleftarrows}} \mathbf{StrMonCat} \quad (L \dashv R)$$

(strictification) For every monoidal category $\mathcal{C}$, the unit $\eta : \mathcal{C} \to R(L(\mathcal{C}))$ is an equivalence.

(coherence) Every formal composition of associators and unitors commutes.

Formal compositions of associators and unitors form a groupoid.

Main theorem

Let $\mathbb{T}_s$ be an extension of $\mathbb{T}_w$ in which a collection E of type equivalences and typal equalities are replaced by definitional equalities.

Theorem

Assume that the following two conditions hold:

1. *The type theory $\mathbb{T}_w$ satisfies external univalence;*
2. *Any formal composition of equalities in E is trivial.*

Then $\mathbb{T}_s$ is conservative over $\mathbb{T}_w$.

Equivalences between models of type theory

Kapulkin and Lumsdaine, *The homotopy theory of type theories* (2016).

Isaev, *Model Structures on Categories of Models of Type Theories* (2016).

Definition

A morphism $F : \mathcal{C} \to \mathcal{D}$ in $\mathbf{CwF}_{\mathsf{Id}}$ is a weak equivalence if it is essentially surjective on types and terms:

(weak type lifting) for every $A : \mathsf{Ty}_{\mathcal{D}}(F(\Gamma))$, there exists $A_0 : \mathsf{Ty}_{\mathcal{C}}(\Gamma)$ and a type equivalence $\alpha : F(A_0) \cong A$;

(weak term lifting) for every $a : \mathsf{Tm}_{\mathcal{D}}(F(\Gamma), F(A))$, there exists $a_0 : \mathsf{Tm}_{\mathcal{C}}(\Gamma, A)$ and a typal equality $p : F(a_0) \simeq a$.

We also have (Cofibrations, Trivial fibrations) and (Trivial cofibration, Fibrations) weak factorization systems.

Hofmann's conservativity theorem: $\mathbf{0}_{\mathsf{ITT}} \to \mathbf{0}_{\mathsf{ETT}}$ is a trivial fibration.

Morita equivalences

Isaev, *Morita equivalences between algebraic dependent type theories* (2018).

$$\mathbf{Mod}_{\mathbb{T}_w} \underset{L}{\overset{R}{\leftrightarrows}} \mathbf{Mod}_{\mathbb{T}_s} \qquad \top$$

Definition

The extension $\mathbb{T}_w \to \mathbb{T}_s$ is a **Morita equivalence** if for every cofibrant $\mathcal{C} : \mathbf{Mod}^{\mathsf{cxl}}_{\mathbb{T}_w}$, the unit $\eta : \mathcal{C} \to R(L(\mathcal{C}))$ is a weak equivalence.

In particular $\mathbf{0}_{\mathbb{T}_w} \to \mathbf{0}_{\mathbb{T}_s}$ is a weak equivalence.

Type-theoretic 1-categories

We have biequivalences:

$$\mathbf{CwF}^{\mathsf{dem}}_{\Sigma,\mathsf{Eq}} \cong \{\text{finitely complete 1-categories}\} \cong \{\text{essentially algebraic theories}\}$$

$$\mathbf{CwF}^{\mathsf{dem}}_{\Sigma,\Pi,\mathsf{Eq}} \cong \{\text{locally cartesian closed 1-categories}\}$$

$$\mathbf{CwF}^{\mathsf{dem}}_{\Sigma} \cong \{\text{display map 1-categories}\} \cong \{\text{generalized algebraic theories}\}$$

(Type-theoretic) representable map 1-categories

Taichi Uemura, *A General Framework for the Semantics of Type Theory* (2019) introduces representable map categories.

$$\mathbf{CwF}^{\mathrm{dem}}_{\Sigma,\overline{\Pi},\mathrm{Eq}} \overset{?}{\cong} \{\text{representable map 1-categories}\} \cong \{(\text{essentially algebraic}) \text{ type theories}\}$$

Where $\overline{\Pi}$-types are Π-types with arities in a subfamily of *representable types*.

$$\frac{A \text{ rep type}}{A \text{ type}} \qquad \frac{A \text{ rep type} \qquad [a : A]\ B(a) \text{ type}}{\overline{\Pi}(A, B) \text{ type}}$$

Internal models

Take $\mathcal{C} : \mathbf{CwF}_{\Sigma,\overline{\Pi}}$. It is a CwF $(\mathcal{C}, \mathrm{Sort}, \mathrm{Elem})$ with $\mathbf{1}$-, Σ- and $\overline{\Pi}$- type structures.

Elements of Sort are called **sorts** (or outer types).

Elements of RepSort are called **representable sorts** (or outer representable types).

Definition

An internal model of $\mathbb{T}$ in $\mathcal{C}$ consists of:

- a sort $\mathbf{ty} : \mathrm{Sort}$ of (inner) types;
  $\mathbf{Ty} \triangleq \mathrm{Elem}(\mathbf{ty})$;
- a representable sort family $\mathbf{tm} : \mathbf{Ty} \to \mathrm{RepSort}$ of (inner) terms;
  $\mathbf{Tm}(A) \triangleq \mathrm{Elem}(\mathbf{tm}(A))$;
- the structure of a model of $\mathbb{T}$ over the CwF $(\mathcal{C}, \mathbf{Ty}, \mathbf{Tm})$.

$$\mathbf{Id} : (A : \mathbf{Ty})(x, y : \mathbf{Tm}(A)) \to \mathbf{Ty} \qquad \mathbf{\Pi} : (A : \mathbf{Tm})(B : \mathrm{Elem}(\overline{\Pi}(\mathbf{tm}(A), \mathbf{ty}))) \to \mathbf{Ty}$$

$$\ldots$$

The walking model

Definition

The walking model $\mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}]$ is the initial type-theoretic representable map category equipped with an internal model of $\mathbb{T}$.

Some contexts of $\mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}]$:

$$() \qquad (A : \mathbf{ty}) \qquad (A : \mathbf{ty}, x : \mathbf{tm}(A))$$

$$(A : \mathbf{ty}, B : \mathbf{tm}(A) \to \mathbf{ty}, b : (a : \mathbf{tm}(A)) \to \mathbf{tm}(B(a)))$$

$$\partial \mathsf{Id} = (A : \mathbf{ty}, x : \mathbf{tm}(A), y : \mathbf{tm}(A)) \qquad \partial \mathbf{\Pi} = (A : \mathbf{ty}, B : \mathbf{tm}(A) \to \mathbf{ty})$$

Proposition

The category $(\mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}])^{\mathrm{op}}$ is equivalent to the category of finitely generated models of $\mathbb{T}$.

A context (or closed sort) $\Gamma : \mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}]$ generates a model $\mathbf{0}_{\mathbb{T}}[\Gamma] : \mathbf{Mod}_{\mathbb{T}}$.

Syntactic characterization of Morita equivalences

Recall that $\mathbb{T}_w \to \mathbb{T}_s$ is a Morita equivalence if for every cofibrant $\mathcal{C} : \mathbf{Mod}^{\mathsf{cxl}}_{\mathbb{T}_w}$, the unit $\eta : \mathcal{C} \to R(L(\mathcal{C}))$ is a weak equivalence.

Proposition

An extension $\mathbb{T}_w \to \mathbb{T}_s$ is a Morita equivalence if and only if

$$\mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}_w] \to \mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}_s]$$

is a weak equivalence (in $\mathbf{Mod}_{\mathbb{T}_w}$).

Other walking models

We also have $\mathbf{0}_{\Sigma,\Pi}[\mathbb{T}]$, $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Eq}}[\mathbb{T}]$, $\mathbf{0}_{\Sigma,\Pi,\mathsf{Eq}}[\mathbb{T}]$.

Some contexts of $\mathbf{0}_{\Sigma,\Pi}[\mathbb{T}]$:

$$(P : \mathbf{ty} \to \mathbf{ty}, A : \mathbf{ty}, a : P(P(A))) \qquad (P : \mathbf{ty} \to \mathbf{ty}, A : \mathbf{ty}, B : \mathbf{ty}, \alpha : A \cong B)$$

Proposition

The category $(\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Eq}}[\mathbb{T}])^{\mathsf{op}}$ is equivalent to the category of finitely presented models of $\mathbb{T}$.

Type-theoretic ∞-categories

$$\mathbf{CwF}^{\mathsf{cxl}}_{\Sigma,\mathsf{Id}} \cong \{\text{finitely complete } \infty\text{-categories}\}$$

$$\mathbf{CwF}^{\mathsf{cxl}}_{\Sigma,\Pi,\mathsf{Id}} \overset{?}{\cong} \{\text{locally cartesian closed } \infty\text{-categories}\}$$

$$\mathbf{CwF}^{\mathsf{cxl}}_{\Sigma,\overline{\Pi},\mathsf{Id}} \overset{?}{\cong} \{\text{representable map } \infty\text{-categories}\}$$

We have $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}]$ and $\mathbf{0}_{\Sigma,\Pi,\mathsf{Id}}[\mathbb{T}]$.

We will construct $\mathcal{D} : \mathbf{CwF}_{\Sigma,\overline{\Pi},\mathsf{Id}}$ equipped with an internal model of $\mathbb{T}_w$.

$$\begin{array}{ccc} \mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}_w] & \xrightarrow{\quad\eta\quad} & \mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}_s] \\ & {\scriptstyle F}\searrow \quad \nearrow{\scriptstyle G} & \\ & \mathcal{D} & \end{array}$$

Elements of $\mathsf{Elem}_{\mathcal{D}}(x \simeq y)$ will be the formal compositions of equalities in E.

Univalent internal models

Take $\mathcal{C} : \mathbf{CwF}_{\Sigma,\overline{\Pi},\mathsf{Id}}$ with an internal model of $\mathbb{T}$.

We have comparison maps:

$$\begin{aligned} \mathsf{coe}_{\mathbf{ty}} &: (A \simeq_{\mathbf{ty}} B) \to (A \cong B) \\ \mathsf{coe}_{\mathbf{tm}} &: (x \simeq_{\mathbf{tm}(A)} y) \to \mathbf{Tm}(x \simeq_A y) \end{aligned}$$

Definition

The internal model of $\mathbb{T}$ is **univalent** if $\mathsf{coe}_{\mathbf{ty}}$ and $\mathsf{coe}_{\mathbf{tm}}$ have homotopy sections (equivalently if they are homotopy equivalences).

We also say that $\mathcal{C}$ is **saturated**, or that the outer identity types of $\mathcal{C}$ satisfy **saturation**.

We have $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}, \mathsf{univ}]$, etc.

Univalent internal models

In $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T},\mathsf{univ}]$ we can transport structures over type equivalences:

If $P : \mathbf{Ty} \to \mathbf{Ty}$ and $\alpha : A \cong B$, then

$$
\begin{aligned}
\mathsf{coe}_{\mathsf{ty}}^{-1}(\alpha) &: A \simeq_{\mathsf{ty}} B, \\
\mathsf{ap}(P, \mathsf{coe}_{\mathsf{ty}}^{-1}(\alpha)) &: P(A) \simeq_{\mathsf{ty}} P(B), \\
\mathsf{coe}_{\mathsf{ty}}(\mathsf{ap}(P, \mathsf{coe}_{\mathsf{ty}}^{-1}(\alpha))) &: P(A) \cong P(B).
\end{aligned}
$$

External univalence

Theorem

The following conditions are equivalent:

1. *The map* $\mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}] \to \mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T},\mathsf{univ}]$ *is essentially surjective on elements (outer terms).*
2. *The category* $\mathbf{Mod}_{\mathbb{T}}^{\mathsf{cxl}}$ *satisfies the axioms of a left semi-model category.*

If they hold, we say that $\mathbb{T}$ *satisfies* **external univalence**.

Partial saturation

Take $\mathcal{C} : \mathbf{CwF}_{\Sigma,\overline{\Pi},\mathsf{Id}}$ with an internal model of $\mathbb{T}$.

A lift $(\widehat{p}, \widetilde{p}) : \mathsf{lift}(p)$ of $p : \mathbf{Tm}(x \simeq_A y)$ is a witness that p lies in the essential image of $\mathsf{coe}_{\mathbf{tm}}$:

$$\widehat{p} : (x \simeq_{\mathbf{tm}(A)} y)$$
$$\widetilde{p} : (\mathsf{coe}_{\mathbf{tm}}(\widehat{p}) \simeq p)$$

Say that $\mathcal{C}$ is partially saturated with respect to E if we have lift of every type equivalence / typal equality in E.

Partial saturation

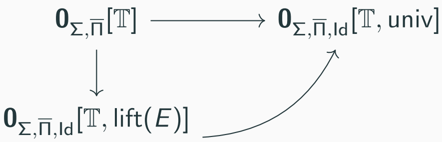
We have $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}, \mathsf{lift}(E)]$.
An element of $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}, \mathsf{lift}(E)]$ is a formal composition of equalities from E.

Theorem

If $\mathbb{T}$ satisfies external univalence, then

$$\mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}] \to \mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}, \mathsf{lift}(E)]$$

is essentially surjective on elements (outer terms).



Acyclicity

Factorization:

$$\begin{array}{ccc} \mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}_w] & \xrightarrow{\quad\eta\quad} & \mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}_s] \\ & {\scriptstyle F}\searrow \quad \nearrow{\scriptstyle G} & \\ & \mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}_w, \mathsf{lift}(E)] & \end{array}$$

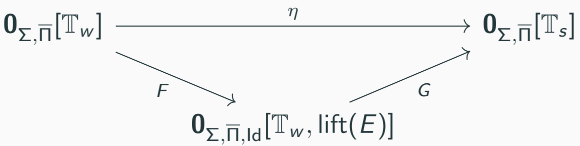
Definition

We say that $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}_w, \mathsf{lift}(E)]$ is acyclic in the image of F if for every $p : \mathsf{Tm}(F(\Gamma), x \simeq_A x)$, there exists some $p' : \mathsf{Tm}(F(\Gamma), p \simeq \mathsf{refl})$.

Lemma

If $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}_w, \mathsf{lift}(E)]$ is acyclic in the image of F, then G is surjective on types and terms, when restricted to the image of F.

Main theorem



Theorem

Assume that the following two conditions hold:

1. *The type theory* $\mathbb{T}_w$ *satisfies external univalence;*
2. *The model* $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}_w, \mathsf{lift}(E)]$ *is acyclic in the image of* F*.*

Then $\mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}_w] \to \mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}_s]$ *is a weak equivalence.*

Concluding remarks

- The two conditions of the theorem do not always hold.
- The fact that $\mathbb{T}_w$ satisfies external univalence can usually be proven using homotopical diagram models.
- It remains to prove acyclicity.
  I expect that acyclicity follows from a normalization argument: for every normal form of $\mathbf{0}_{\Sigma,\overline{\Pi}}[\mathbb{T}_s]$ there should be a contractible space of terms of $\mathbf{0}_{\Sigma,\overline{\Pi},\mathsf{Id}}[\mathbb{T}_w, \mathsf{lift}(E)]$ corresponding to that normal form.