Coherence of definitional equality in type theory

Rafaël Bocquet
HoTTEST, September 23, 2021
Problem

In type theory we have typal equalities,

\[ 0 + n \simeq n \quad n + m \simeq m + n \quad \text{refl} \cdot p \simeq p \]

some of them are definitional equalities

\[ n + 0 = n \quad p \cdot \text{refl} = p \]

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.
We can replace the computation rules of $\text{Id}$-, $\Sigma$-, $\Pi$-types by weak computation rules.

$$
\begin{array}{c}
a : A \\
b : B(a)
\end{array} \quad \Rightarrow \\
\pi_1-\beta : \pi_1(\text{pair}(a, b)) \simeq_A a
$$

The path types of cubical type theory satisfy the weak computation rule of $\text{Id}$-types.

Are the usual computation rules conservative over the weak computation rules?
Examples: composition of paths

Can identity types satisfy the groupoid laws definitionally?

\[ p \cdot \text{refl} = p \quad \text{refl} \cdot p = p \quad p \cdot (q \cdot r) = (p \cdot q) \cdot r \]

\[ p^{-1} \cdot p = \text{refl} \quad (p^{-1})^{-1} = p \]

\[ \ldots \]

\[ \text{ap}(f, \text{refl}) = \text{refl} \quad \text{ap}(f, p \cdot q) = \text{ap}(f, p) \cdot \text{ap}(f, q) \]

\[ \text{ap}(f, p^{-1}) = \text{ap}(f, p)^{-1} \quad \ldots \]
Can we extend HoTT with a universe StrProp of “strict” propositions and an equivalence $\text{StrProp} \simeq \text{Prop}$?

$A : \text{StrProp} \quad x, y : A$

$x = y$

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?

$[a : A] \quad B(a) : \text{StrProp}$

$\forall (A, B) : \text{StrProp}$
A Category with Families (CwF) consists of:

- a category \( \mathcal{C} \) with a terminal object;
- a presheaf of types \( \mathcal{T}_\mathcal{C} : \text{Psh}(\mathcal{C}) \);
- a (locally representable) presheaf of terms \( \mathcal{T}_\mathcal{m}_\mathcal{C} : \mathcal{T}_\mathcal{C} \to \text{RepPsh}(\mathcal{C}) \);

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

\[
\text{A type } [a : A] B(a) \text{ type } \quad \Pi : (A : \mathcal{T}_\mathcal{C})(B : \mathcal{T}_\mathcal{m}_\mathcal{C}(A) \to \mathcal{T}_\mathcal{C}) \to \mathcal{T}_\mathcal{C}
\]

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

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

Freely generated models \( 0_\mathbb{T}[\cdots] \).
Hofmann’s conservativity theorem

**Uniqueness of Identity Proofs**

\[ p : \text{Id}(x, x) \]
\[ \text{uip}(p) : \text{Id}(p, \text{refl}) \]

**Equality reflection**

\[ p : \text{Id}(x, y) \]
\[ x = y \]

**Theorem (Hofmann, 1995)**

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

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

The map \( |\cdot| : \text{ITT} \to \text{ETT} \) is surjective on types and terms.
Proof of Hofmann’s conservativity theorem

Equivalence relations (∼) on types and terms of ITT:

\[(A \sim B) \iff \exists p : \text{Tm}_{\text{ITT}}(\text{Id}(U, A, B))\]
\[((a : A) \sim (b : B)) \iff \exists p : \text{Tm}_{\text{ITT}}(\text{Id}((X : U) \times X, (A, a), (B, b)))\]

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

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

All type- and term- formers respect (∼). For λ(−) (and other binders) this requires function extensionality.
Proof of Hofmann’s conservativity theorem

Quotients \((Ty_{\text{ITT}}/\sim)\) and \((Tm_{\text{ITT}}/\sim)\).

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

Since \(|-|\) is a retract of \(q\), \(|-|\) is surjective on types and terms.

(Alternative: Use the relative induction principle for \(\mathcal{R}en(0_{\text{ITT}}) \rightarrow 0_{\text{ETT}}\))
Mac Lane’s coherence theorem for monoidal categories

\[ \alpha_{x,y,z} : (x \otimes y) \otimes z \simeq x \otimes (y \otimes z) \]
\[ \lambda_x : (I \otimes x) \simeq x \]
\[ \rho_x : (x \otimes I) \simeq x \]

\[ \alpha_{x,y,z} = \text{id} \]
\[ \lambda_x = \text{id} \]
\[ \rho_x = \text{id} \]

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

(\textbf{coherence}) Every formal composition of associators and unitors commutes.

Formal compositions of associators and unitors form a groupoid.
Main theorem

Let $T_s$ be an extension of $T_w$ 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 $T_w$ satisfies external univalence;
2. Any formal composition of equalities in $E$ is trivial.

Then $T_s$ is conservative over $T_w$. 
Equivalences between models of type theory


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

**Definition**

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

**(weak type lifting)** for every $A : \text{Ty}_D(F(\Gamma))$, there exists $A_0 : \text{Ty}_C(\Gamma)$ and a type equivalence $\alpha : F(A_0) \simeq A$;

**(weak term lifting)** for every $a : \text{Tm}_D(F(\Gamma), F(A))$, there exists $a_0 : \text{Tm}_C(\Gamma, A)$ 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: $0_{\text{ITT}} \to 0_{\text{ETT}}$ is a trivial fibration.
Morita equivalences


![Diagram](image)

**Definition**

The extension $T_w \to T_s$ is a **Morita equivalence** if for every cofibrant $C : Mod_{T_w}^{cxl}$, the unit $\eta : C \to R(L(C))$ is a weak equivalence.

In particular $0_{T_w} \to 0_{T_s}$ is a weak equivalence.
Type-theoretic 1-categories

We have biequivalences:

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

\[ \mathbf{CwF}^{\text{dem}}_{\Sigma, \text{Eq}} \cong \{ \text{locally cartesian closed 1-categories} \} \]

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

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

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

\[
\begin{array}{c}
\text{A rep type} \\
\overline{\Pi}(A, B) \text{ type}
\end{array}
\]

\[
\begin{array}{c}
A \text{ type} \\
[\text{A rep type}] [a : A] B(a) \text{ type}
\end{array}
\]
Internal models

Take $C : \mathbf{CwF}_{\Sigma, \Pi}$. It is a CwF $(C, \text{Sort}, \text{Elem})$ with $1$, $\Sigma$- and $\Pi$- 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 $\mathcal{T}$ in $C$ consists of:

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

\[ \text{Id} : (A : \mathbf{Ty})(x, y : \mathbf{Tm}(A)) \to \mathbf{Ty} \quad \Pi : (A : \mathbf{Tm})(B : \text{Elem}(\Pi(\mathbf{tm}(A), \mathbf{ty}))) \to \mathbf{Ty} \]

\[ \ldots \]
The walking model

**Definition**

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

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

- $()$
- $(A : ty)$
- $(A : ty, x : tm(A))$
- $(A : ty, B : tm(A) \rightarrow ty, b : (a : tm(A)) \rightarrow tm(B(a)))$

\[ \partial \text{Id} = (A : ty, x : tm(A), y : tm(A)) \]

\[ \partial \Pi = (A : ty, B : tm(A) \rightarrow ty) \]

**Proposition**

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

A context (or closed sort) $\Gamma : 0_{\Sigma, \Pi}[\mathbb{T}]$ generates a model $0_{\mathbb{T}}[\Gamma] : \text{Mod}_{\mathbb{T}}$. 
Recall that $T_w \to T_s$ is a Morita equivalence if for every cofibrant $C : \text{Mod}_{T_w}^{cxl}$, the unit $\eta : C \to R(L(C))$ is a weak equivalence.

**Proposition**

An extension $T_w \to T_s$ is a Morita equivalence if and only if

$$0_{\Sigma, \bar{\Pi}[T_w]} \to 0_{\Sigma, \bar{\Pi}[T_s]}$$

is a weak equivalence (in $\text{Mod}_{T_w}$).
We also have $0_{\Sigma,\Pi[\mathbb{T}]}$, $0_{\Sigma,\Pi,\mathrm{Eq}[\mathbb{T}]}$, $0_{\Sigma,\Pi,\mathrm{Eq}[\mathbb{T}]}$.

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

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

**Proposition**

The category $(0_{\Sigma,\Pi,\mathrm{Eq}[\mathbb{T}]})^\text{op}$ is equivalent to the category of finitely presented models of $\mathbb{T}$. 
Type-theoretic $\infty$-categories

\[
\text{CwF}^{\text{cxl}}_{\Sigma,\Pi,\text{Id}} \cong \{\text{finitely complete } \infty\text{-categories}\}
\]

\[
\text{CwF}^{\text{cxl}}_{\Sigma,\Pi,\text{Id}} \cong \{\text{locally cartesian closed } \infty\text{-categories}\}
\]

\[
\text{CwF}^{\text{cxl}}_{\Sigma,\Pi,\text{Id}} \cong \{\text{representable map } \infty\text{-categories}\}
\]

We have $0_{\Sigma,\Pi,\text{Id}[T]}$ and $0_{\Sigma,\Pi,\text{Id}[T]}$. 
We will construct $\mathcal{D} : \text{CwF}_{\Sigma, \Pi, \text{Id}}$ equipped with an internal model of $\mathbb{T}_w$.

$$
0_{\Sigma, \Pi}[\mathbb{T}_w] \xrightarrow{\eta} 0_{\Sigma, \Pi}[\mathbb{T}_s]
$$

$\mathcal{D}$

$F$ $G$

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

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

We have comparison maps:

$$\text{coe}_{\text{ty}} : (A \simeq_{\text{ty}} B) \to (A \approx B)$$
$$\text{coe}_{\text{tm}} : (x \simeq_{\text{tm}(A)} y) \to Tm(x \simeq_{A} y)$$

**Definition**

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

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

We have $0_{\Sigma, \Pi, \text{Id}}[\mathbb{T}, \text{univ}]$, etc.
In $0_{\Sigma, \Pi, \text{Id}}[T, \text{univ}]$ we can transport structures over type equivalences:

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

\[
\begin{align*}
\text{coe}_{ty}^{-1}(\alpha) &: A \cong_{ty} B, \\
ap(P, \text{coe}_{ty}^{-1}(\alpha)) &: P(A) \cong_{ty} P(B), \\
\text{coe}_{ty}(ap(P, \text{coe}_{ty}^{-1}(\alpha))) &: P(A) \cong P(B).
\end{align*}
\]
External univalence

Theorem

The following conditions are equivalent:

1. The map $0_{\Sigma, \Pi}[\mathbb{T}] \to 0_{\Sigma, \Pi, \text{Id}}[\mathbb{T}, \text{univ}]$ is essentially surjective on elements (outer terms).

2. The category $\text{Mod}_{cxl}^{\mathbb{T}}$ satisfies the axioms of a left semi-model category.

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

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

A lift $(\hat{p}, \tilde{p}) : \text{lift}(p)$ of $p : \text{Tm}(x \equiv_A y)$ is a witness that $p$ lies in the essential image of $\text{coe}_{\text{tm}}$:

$$\hat{p} : (x \equiv_{\text{tm}(A)} y)$$
$$\tilde{p} : (\text{coe}_{\text{tm}}(\hat{p}) \equiv p)$$

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 $0_{\Sigma,\Pi,\text{Id}[T,\text{lift}(E)]}$.

An element of $0_{\Sigma,\Pi,\text{Id}[T,\text{lift}(E)]}$ is a formal composition of equalities from $E$.

**Theorem**

If $T$ satisfies external univalence, then

$$0_{\Sigma,\Pi}[T] \rightarrow 0_{\Sigma,\Pi,\text{Id}[T,\text{lift}(E)]]$$

is essentially surjective on elements (outer terms).

\[0_{\Sigma,\Pi}[T] \rightarrow 0_{\Sigma,\Pi,\text{Id}[T,\text{univ}]} \]

\[0_{\Sigma,\Pi,\text{Id}[T,\text{lift}(E)]]} \rightarrow \]
Acyclicity

Factorization:

\[
0_{\Sigma, \Pi}[T_w] \xrightarrow{\eta} 0_{\Sigma, \Pi}[T_s] \xrightarrow{F} 0_{\Sigma, \Pi, \text{Id}}[T_w, \text{lift}(E)] \xrightarrow{G}
\]

**Definition**

We say that \(0_{\Sigma, \Pi, \text{Id}}[T_w, \text{lift}(E)]\) is acyclic in the image of \(F\) if for every \(p : \text{Tm}(F(\Gamma), x \simeq_A x)\), there exists some \(p' : \text{Tm}(F(\Gamma), p \simeq \text{refl})\).

**Lemma**

*If \(0_{\Sigma, \Pi, \text{Id}}[T_w, \text{lift}(E)]\) 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 \( 0_{\Sigma, \bar{\Pi}, \text{Id}}[\mathbb{T}_w, \text{lift}(E)] \) is acyclic in the image of \( F \).

Then \( 0_{\Sigma, \bar{\Pi}}[\mathbb{T}_w] \to 0_{\Sigma, \bar{\Pi}}[\mathbb{T}_s] \) 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 $0_{\Sigma,\Pi}[\mathbb{T}_s]$ there should be a contractible space of terms of $0_{\Sigma,\Pi,\text{Id}}[\mathbb{T}_w, \text{lift}(E)]$ corresponding to that normal form.