How Intensional is Homotopy Type Theory?

Thomas Streicher (TU Darmstadt)

Barcelona, September 2013

Intensional vs. Extensional Type Theory

Extensional type theory (ETT) has a naive set-theoretic model and a good notion of categorical model, namely (split) *locally cartesian closed* categories.
Equality on A is interpreted as the diagonal $\delta_A = \langle id_A, id_A \rangle : A \to A \times A$.

In ETT there is just one equality: the judgement $\Gamma \vdash t = s \in A$ holds iff $\Gamma \vdash e \in \mathrm{Id}_A(t, s)$ holds for some (unique) term e.

What is the problem with ETT?

Type Checking for ETT is Undecidable!

For closed terms t and s of type A we have

$$\vdash t = s \in A \quad \text{iff} \quad \vdash r_A(t) \in \mathsf{Id}_A(t, s)$$

Suppose type checking were decidable for ETT.

Then it is decidable whether $\vdash t = s \in A$ is derivable in ETT.

Instantiating A by $N \to N$ and s by $\lambda n{:}N.0$ we observe that it is decidable whether ETT proves $\vdash \Pi n{:}N.\mathsf{Id}_N(t(n), 0)$ for closed terms t of type $N \to N$.

But for no r.e. consistent extension T of PRA the set of Π^0_1-sentence provable in T is decidable. In particular, not for ETT.

Identity Types

are the most intriguing concept of Intensional Type Theory (ITT). They are given by the following rules

$$\frac{\Gamma \vdash A}{\Gamma, x, y{:}A \vdash \mathsf{Id}_A(x,y)} \text{ (Id-F)} \qquad \frac{\Gamma \vdash A}{\Gamma, x : A \vdash r_A(x) : \mathsf{Id}_A(x,x)} \text{ (Id-I)}$$

$$\frac{\Gamma, x, y{:}A, z : \mathsf{Id}_A(x,y) \vdash C(x,y,z) \qquad \Gamma, x{:}A \vdash d : C(x,x,r_A(x))}{\Gamma, x, y{:}A, z : \mathsf{Id}_A(x,y) \vdash J((x)d)(z) : C(x,y,z)} \text{ (Id-E)}$$

together with the conversion rule $J((x)d)(r_A(t)) = d[t/x]$.

In ETT we simply have $\frac{\Gamma \vdash p : \mathsf{Id}_A(t,s)}{\Gamma \vdash t = s \in A}$ (Id-refl) instead of (Id-E).

Criteria for Intensionality

The following principles are characteristic for ITT with a universe Set

(I1) $A : \mathsf{Set}, x, y : A, z : \mathsf{Id}_A(x, y) \not\vdash x = y : A$

(I2) $A : \mathsf{Set}, B : A \to \mathbf{Set}, x, y : A, z : \mathsf{Id}_A(x, y) \not\vdash B(x) = B(y) : U$

(I3) $\vdash p : \mathsf{Id}_A(t, s)$ implies $\vdash t = s : A$

Notice that (I3) says that relative to the empty context propositional and judgemental equality coincide (for closed terms and types) whereas (I1) and (I2) say that this is not the case relative to non-empty contexts!

Uniqueness of Identity Proofs

Notice that (I1–3) are compatible with an additional eliminator K for Id-types as given by

$$\frac{\Gamma, x{:}A, z : \mathsf{Id}_A(x,x) \vdash C(x,z) \qquad \Gamma, x{:}A \vdash d : C(x, r_A(x))}{\Gamma, x{:}A, z : \mathsf{Id}_A(x,x) \vdash K((x)d)(z) : C(x,z)} \ (\text{Id-E}')$$

together with the conversion rule $K((x)d)(r_A(t)) = d[t/x]$.

The eliminator K allows one to prove *Uniqueness of Identity Proofs*

$$\text{(UIP)} \qquad A : \mathbf{Set}, x, y : A, p, q : \mathsf{Id}_A(x,y) \vdash \mathsf{Id}_{\mathsf{Id}_A(x,y)}(p,q)$$

which to refute Hofmann and S. introduced the groupoid model.

How intensional is HoTT? (1)

Voevodsky's **Univalence Axiom** (UA) for a type-theoretic universe U essentially says that isomorphic types of the universe U are equal. He has shown that UA entails the principle of **function extensionality**

$$(\mathsf{Ext}_{\mathsf{fun}}) \quad \Pi x{:}A.\mathsf{Id}_{B(x)}(f(x), g(x)) \to \mathsf{Id}_{\Pi x:A.B(x)}(f, g)$$

for $A \in U$, $B \in A \to U$ and $f, g \in \Pi x{:}A.B(x)$.

This, however, contradicts (I3) for the following reason.

For closed terms t of type $N \to N$ by (I3) and $(\mathsf{Ext}_{\mathsf{fun}})$ we have $\vdash t = \lambda n{:}N.0 : N$ iff $\vdash p : \Pi x{:}N.\mathsf{Id}_N(t(x), 0)$ for some closed term p.
For type checking to be decidable judgemental equality has to be decidable, too.
Thus, the set of Π^0_1 sentences provable in ITT + $(\mathsf{Ext}_{\mathsf{fun}})$ is decidable.

How intensional is HoTT? (2)

In the groupoid model the universe U of *small groupoids* validates criteria (I1) and (I2) but not criterion (I3). However, this U does not validate UA either.

However, in the groupoid model the universe U_d of *small discrete groupoids* with $\mathrm{Id}_{U_d}(A, B) = \mathrm{Iso}(A, B)$ validates (I3) whereas (I1) and (I2) fail. Moreover, the universe U_d validates the Univalence Axiom.

Moreover, the groupoid and also the simplicial sets model validate function extensionality and (UIP) for all types which can be expressed in type theory without universes! Simply, because all such types get interpreted as small discrete groupoids.

How intensional is HoTT? (3)

In the known (non-syntactic) models of HoTT one builds a non-extensional model around the extensional *naive set-theoretic model*. They are conservative in the sense that they validate the same sentences of higher type arithmetic (HA_ω) as $\mathbf{Set}$.

This conservation result extends to *Basic Type Theory* (BTT), i.e. type theory without a universe.

However, on the level of provability, i.e. syntactic (Lindenbaum-Tarski) models, the univalence axiom allows one to prove new Π^0_1 sentences, i.e. UA is **not conservative** over BTT.

A Conservation Result (1)

It is an open problem whether **computational meaning** can be given to the Univalence Axiom. One postulates a new constant ua : UA. But there are no computation rules for ua. The constant ua cannot be syntactically defined from the rest since UA fails in the naive set-theoretic model of ETT.

The constant ua introduces new elements into *propositional* types since it allows one to prove new Π^0_1 sentences. But it should not introduce new elements into data types like simple types over N etc.

The question is **inherently vague** since there is no clear notion of "data type".

A Conservation Result (2)

As a minimal requirement Voevodsky has suggested the following criterion: whenever HoTT proves $t \in \Sigma x{:}N.P(x)$ then there is an $n \in \mathbb{N}$ such that

- HoTT proves $\mathsf{Id}_N(p_0(t), \underline{n})$ and
- type theory without UA proves $P(\underline{n})$

where $\underline{n}$ stands for $\mathsf{succ}^n(0)$, the n-th numeral.

A possible route of attack is to construct a model of HoTT inside type theory without UA.
This has failed so far for the simplicial set model of HoTT (since it is not decidable whether a simplex is degenerate or not).

A Conservation Result (3)

In ITT + Ext_{fun} with a universe one can construct the groupoid model of [HS94]. Since this model validates UA we obtain the following for *basic type theory* (BTT), i.e. type theory without a universe.

Conservation Result

If a proposition of BTT can be proved in ITT + UA then it can be proved in ITT + Ext_{fun} with a universe.

We need Ext_{fun} in the meta-theory
(1) for getting exponentials of groupoids right
(2) for defining Id-types on the universe of discrete groupoids since we need extensional equality of isomorphisms between types in the universe.

Truly Intensional Models for ITT (1)

were constructed in my Habil. Thesis (1993) [Str93].

They may be described as the $\neg\neg$-separated objects of the topos $\mathsf{Gl}(\mathcal{E}\!f\!f) = \mathbf{Set}{\downarrow}\Gamma$ obtained by *glueing* $\Gamma = \mathcal{E}\!f\!f(1, -) : \mathcal{E}\!f\!f \to \mathbf{Set}$.

For sake of simplicity one may replace Γ by the identity on $\mathbf{Set}$ giving rise to the *Sierpiński* topos $\mathcal{S} = \mathbf{Set}{\downarrow}\mathbf{Set} = \mathbf{Set}^{\mathbf{2}^{\mathrm{op}}}$. Up to isomorphism $\neg\neg$-separated objects of $\mathcal{S}$ are inclusions of subsets.

We write $\mathcal{LP}$ for the ensuing category of *logical predicates*. Its objects are pairs $X = (|X|, P_X)$ where $|X|$ is a set and $P_X \subseteq |X|$. Morphisms from X to Y are functions $f : |X| \to |Y|$ such that

$$\begin{array}{ccc} P_X & \cdots\cdots\to & P_Y \\ \cap\downarrow & & \cap\downarrow \\ |X| & \xrightarrow{\quad f \quad} & |Y| \end{array}$$

Truly Intensional Models for ITT (2)

Like $\mathbf{Set}$ the category $\mathcal{LP}$ provides a model for ETT. For getting a truly intensional model of ITT we have to choose an appropriate universe U.

Let $\mathcal{U}$ be a Grothendieck universe. Then U consists of all objects $X \in \mathcal{LP}$ with $|X| \in \mathcal{U}$ and $0 \in |X|$.

Idea : $|X|$ are the **potential** objects and P_X are the **actual** objects. Elements of $|X| \setminus P_X$ will "simulate free variables" as we will see.

Let 2 be the set $\{0, 1\}$. For $X \in U$ we define its identity type as

$$\mathrm{Id}_X(x, y) = (2, \{1\}) \qquad \text{if } x = y$$

$$\mathrm{Id}_X(x, y) = (2, \emptyset) \qquad \text{if } x \neq y$$

One interprets r_X as 1.

Truly Intensional Models for ITT (3)

For $C \in \Pi x, y{:}X.\,\mathsf{Id}_X(x, y) \to U$ and $d \in \Pi x{:}X.C(x, x, r_X(x))$ we put $J((x)d)(x, x, 1) = d(x)$ and $J((x)d)(x, y, 0) = 0 \in C(x, y, 0)$.
Similarly one interprets the eliminator K.

It is straightforward to check that (I1)-(I3) hold for U in $\mathcal{LP}$.

For (I1) and (I2) the reason is that $0 \in \mathsf{Id}_X(x, y)$ even if $x \neq y$ and (I3) holds since the interpretation of $\vdash t \in \mathsf{Id}_X(x, y)$ is necessarily $1 \in \mathsf{Id}_X(x, y)$ (since $(\{0\}, \{0\})$ is terminal in $\mathcal{LP}$) and thus $x = y$.

For $f, g : X \to Y$ in U we have
$x{:}X \vdash f(x) = g(x) : Y$ iff $f = g$ and
$x{:}X \vdash \mathsf{Id}_Y(f(x), g(x))$ iff $f_{|P_X} = g_{|P_X}$
for which reason $\mathsf{Ext}_{\mathsf{fun}}$ fails in $\mathcal{LP}$ for U.

Truly Intensional Models for ITT (3a)

NB When interpreting $X \to Y$ for $X, Y \in U$ one has to replace $\lambda x.0$ by 0. Otherwise, it is the full function space. Elements f of $X \to Y$ are actual iff they preserve actual elements.

This bureaucracy can be avoided when working in the category of $\neg\neg$-separated objects of the glueing of $\Gamma : \mathcal{E}\!f\!f \to \mathbf{Set}$.
One takes for U those X where $|X|$ is a modest set containing an element 0_X realized by 0.

By appropriate choice of Gödel numbering one has $\{0\}(n) = 0$ for all $n \in \mathbb{N}$. Thus $0_{X \to Y}$ sends all elements of $|X|$ to 0_Y.

Truly Intensional Models for ITT (4)

The type N of naturals numbers is interpreted as $(\mathbb{N}, \mathbb{N} \setminus \{0\})$. We put $0_N = 1$ and $\mathsf{succ}_N(0) = 0$ and $\mathsf{succ}_N(n{+}1) = n + 2$.

Similarly, one interprets the finite type N_k.

Then $\mathsf{Ext}_{\mathsf{fun}}$ fails already for $X = Y = N_1$ because $f = id_2$ and $g = \lambda x.1$ are different although $x : N_1 \vdash \mathsf{Id}_N(f(x), g(x))$ is witnessed by the identity on $2 = \{0, 1\}$.

Some Questions to the Experts (1)

There is the idea that HoTT is the internal language for ∞-toposes.

But for interpreting ITT one needs something like a category with **display maps**, e.g. a (lcc) model category where families of types are interpreted as **fibrations**.

Most known models are of this kind, i.e. the interpretation of Id-types just **hides** the real equality.

But what should one do when starting from an ∞-topos?

Some Questions to the Experts (2)

Starting from a ∞-*category* $\mathcal{K}$ one may replace it by a *simplicial model category* $\mathcal{C}$. This is a kind of **strictification**.

But what then is the relevance of the **simplicial enrichment** of $\mathcal{C}$?

Can it be defined syntactically in HoTT?

Is it the **simplicial localization** of the non-enriched model?