How Intensional is Homotopy Type Theory?

Thomas Streicher

Fachbereich 4 Mathematik, TU Darmstadt, Schloßgartenstr. 7, Darmstadt D-64289 Germany
streicher@mathematik.tu-darmstadt.de

Martin-Löf's Extensional Type Theory (ETT) has a straighforward semantics in the category **Set** of sets and functions and actually in any locally cartesian closed category with a natural numbers object (nno), e.g. in any elementary topos with a nno. Dependent products are interpreted by right adjoints to pullback functors and extensional identity types are interpreted as diagonals in slice categories as explained e.g. in [**Stre1**].

Despite its intuitive flavour ETT has the defect that type checking for it is not decidable for the following reason. Since ETT identifies propositional and judgemental equality for closed terms t of type $N{\to}N$ the proposition(al type) $\Pi x{:}N.\mathrm{Id}_N(t(x),0)$ is provably inhabited in ETT if and only if ETT proves the judgemental equality $t = \lambda x{:}N.0 \in N{\to}N$. Moreover, ETT proves $\lambda x{:}N.r_N(0) \in \Pi x{:}N.\mathrm{Id}_N(t(x),0)$ iff ETT proves $t = \lambda x{:}N.0 \in N{\to}N$. Thus, if type checking for ETT were decidable one could decide which Π^0_1-sentences are derivable in ETT. But for every consistent recursively enumerable extension $\mathcal{T}$ of primitive recursive arithmetic (PRA) the set of Π^0_1 sentences provable in $\mathcal{T}$ is not decidable[1]. For this reason interactive theorem provers based on type theory (like the systems `Coq` or `ALF`) are based on Martin-Löf's Intensional Type Theory (ITT) with its characteristic separation of propositional and judgemental equality.

After having investigated the semantics of ETT in [**Stre1**], an extended version of my PhD Thesis from 1989, it became generally accepted that ITT is the appropriate kind of type theory for computer assisted interactive theorem proving. For this reason in my subsequent Habilitation Thesis [**Stre2**] I constructed models for ITT validating the following *Criteria of Intensionality*

(I1) $A : \mathrm{Set}, x, y : A, z : \mathrm{Id}_A(x,y) \not\vdash x = y : A$
(I2) $A : \mathrm{Set}, B : A \to \mathrm{Set}, x, y : A, z : \mathrm{Id}_A(x,y) \not\vdash B(x) = B(y) : \mathrm{Set}$
(I3) $\vdash p : \mathrm{Id}_A(t,s)$ implies $\vdash t = s : A$

for some universe Set. Moreover, these models refuted most of those propositions which trivially hold in ETT but cannot be derived in ITT as e.g. the *function extensionality* principle

$$\Pi x{:}A.\mathrm{Id}_B(f(x),g(x)) \to \mathrm{Id}_{A\to B}(f,g)$$

for $A, B \in \mathrm{Set}$ and $f, g \in A{\to}B$ even when A and B are base types like the type N of natural numbers or the type N_2 of booleans.

Unfortunately, the models constructed in [**Stre2**] do not refute the principle UIP

$$A : \mathrm{Set}, x, y : A, u, v : \mathrm{Id}_A(x,y) \vdash \mathrm{Id}_{\mathrm{Id}_A(x,y)}(u,v)$$

[1] Since otherwise one could recursively separate the r.e. sets $A_0 = \{n \in \mathbb{N} \mid \{n\}(n) = 0\}$ and $A_1 = \{n \in \mathbb{N} \mid \{n\}(n) = 1\}$ which can be seen as follows. For natural numbers n consider the primitive recursive predicate $P_n(k) \equiv T(n,n,k) \to U(k) = 0$. If $n \in A_0$ then $\mathcal{T} \vdash \forall k.P_n(k)$ and if $n \in A_1$ then $\mathcal{T} \vdash \neg\forall k.P_n(k)$ and thus $\mathcal{T} \not\vdash \forall k.P_n(k)$. Now let f be a total recursive function with $f(n) = 0$ iff $\mathcal{T} \vdash \forall k.P_n(k)$ which exists since the set of Π^0_1 sentences provable in $\mathcal{T}$ is assumed to be decidable. But then $f(n) = 0$ if $n \in A_0$ and $f(n) \neq 0$ if $n \in A_1$, i.e. f recursively separates the sets A_0 and A_1, which is known to be impossible

Research Perspectives CRM Barcelona, Fall 2013, vol. ??, in Trends in Mathematics
Centre de Recerca Matemàtica, Bellaterra (Barcelona)

of *Uniqueness of Identity Proofs* which can be easily derived in ETT. To overcome this shortcoming Martin Hofmann and I in 1993 introduced the *groupoid model*, see [**HS98**] for a detailed account, within which we identified a universe U of *small discrete* groupoids where $A, B \in U$ were propositionally equal iff they were isomorphic. This observation was the precursor of Voevodsky's *Univalence Axiom* (UA) lying at the heart of *Homotopy Type Theory* (HoTT) an introduction to which can be found in [**HoTT**].

In this abstract we will 1) give a simplified construction of a model for ITT satisfying the above criteria for intensionality and 2) discuss to which extent HoTT is intensional.

1 Truly Intensional Models of ITT

i.e. models of ITT validating the criteria (I1), (I2) and (I3) above were constructed in [**Stre2**]. In more modern terminology they may be described as living within the $\neg\neg$-separated objects of the topos $\mathsf{Gl}(\mathcal{E}\!f\!f) = \mathbf{Set}{\downarrow}\Gamma$ obtained by *glueing* the global elements functor $\Gamma = \mathcal{E}\!f\!f(1,-) : \mathcal{E}\!f\!f \to \mathbf{Set}$. The book [**vO08**] is an excellent reference for all things related to realizability models and, in particular, the effective topos $\mathcal{E}\!f\!f$.

For sake of simplicity we replace Γ by the identity functor on **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 $f(x) \in P_Y$ whenever $x \in P_X$. It is easy to see that like **Set** the category $\mathcal{LP}$ gives rise to a model for ETT. However, for obtaining a truly intensional model of ITT we have to choose an appropriate universe U within $\mathcal{LP}$ which serves the purpose of interpreting the constant Set in (I1), (I2) and (I3). Let $\mathcal{U}$ be a Grothendieck universe. Then U consists of all objects $X \in \mathcal{LP}$ with $|X| \in \mathcal{U}$ and $0 = \emptyset \in |X|$.

The intuition behind this definition of U is that for $X \in U$ the set $|X|$ is the set of **potential** objects of X and P_X is the subset of **actual** objects of X. Elements of $|X| \setminus P_X$ will serve the purpose of *simulating the syntactic notion of free variables on the level of semantics*.

For showing that U in $\mathcal{LP}$ gives rise to a truly intensional model of ITT we next describe the interpretation of identity types within U. As usual let 2 be the set $\{0, 1\}$. For $X \in U$ we define its identity type as

$$\begin{aligned} &\mathrm{Id}_X(x,y) = (2,\{1\}) && \text{if } x = y \\ &\mathrm{Id}_X(x,y) = (2,\emptyset) && \text{if } x \neq y \end{aligned}$$

for $x, y \in |X|$. We interpret $r_X(x)$ as 1 for all $x \in |X|$. For $C \in \Pi x,y{:}X.\,\mathrm{Id}_X(x,y) \to U$ and $d \in \Pi x{:}X.C(x,x,r_X(x))$ we put

$$J((x)d)(x,x,1) = d(x) \qquad \text{and} \qquad J((x)d)(x,y,0) = 0 \in C(x,y,0)$$

for $x, y \in |X|$. Similarly one may interpret the eliminator K of [**Stre2**] allowing one to prove UIP.

Theorem 1 *For the above interpretation of identity types the universe U in $\mathcal{LP}$ validates the criteria of intensionality* (I1)*,* (I2) *and* (I3) *and refutes the principle of function extensionality.*

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

Notice that for $X, Y \in U$ we have

(1) $x{:}X \vdash f(x) = g(x) : Y$ iff $f = g$ and
(2) $x{:}X \vdash \mathrm{Id}_Y(f(x), g(x))$ iff $f_{|P_X} = g_{|P_X}$

for $f, g \in X{\to}Y$. There are types X and Y and different elements f and g in $P_{X\to Y}$ which, however, coincide on P_X. For this reason the principle of function extensionality fails for U in $\mathcal{LP}$. □

Notice that when interpreting $X{\to}Y$ for $X, Y \in U$ one has to replace $\lambda x.0$ by 0 and redefine the application function appropriately. But otherwise $X{\to}Y$ is interpreted as the full function space in the sense of **Set**. Moreover, elements f of $|X{\to}Y|$ are actual, i.e. $f \in P_{X\to Y}$, iff they preserve actual elements, i.e. $f(x) \in P_Y$ whenever $x \in P_X$. In [**Stre2**] this kind of bureaucracy could be avoided since one was working in the category of $\neg\neg$-separated objects of the glueing of $\Gamma : \mathcal{E}\!f\!f \to \mathbf{Set}$. There for U one took those X where $|X|$ is a *modest set* (see [**vO08**]) containing an element 0_X realized by 0 and P_X still was an arbitrary subset of (the underlying set of) $|X|$. By appropriate choice of Gödel numbering one has $\{0\}(n) = 0$ for all $n \in \mathbb{N}$ and, accordingly, the function $0_{X\to Y}$ sends all elements of $|X|$ to 0_Y.

Finally, we discuss our interpretation of the types N and N_k in the universe U in $\mathcal{LP}$. The type N of natural numbers is interpreted as $(\mathbb{N}, \mathbb{N} \setminus \{0\})$. We put $0_N = 1$ and the successor operation *succ* is given by $succ_N(0) = 0$ and $succ_N(n{+}1) = n{+}2$. Similarly, one interprets the finite types N_k. Thus, the principle of function extensionality fails already for $X = Y = N_1$ because if f is the identity on 2 and g is the constant map with value $1 \in 2$ then f and g are different elements of $P_{N_1\to N_1}$ although $x : N_1 \vdash \mathrm{Id}_{N_1}(f(x), g(x))$ is witnessed (essentially) by the identity on $2 = \{0, 1\}$.

2 How intensional is HoTT ?

Since the models of the previous section 1 and the groupoid model were both constructed for the purpose of showing that certain propositions cannot be derived in ITT one might dream of combining both ideas in order to construct a model of ITT which is truly intensional and at the same time refutes UIP. The most immediate idea is to construct a groupoid model inside one of the models described in section 1. However, as became clear to me in discussion with S. Awodey for constructing a universe U of small discrete groupoids in (a model of) ITT where $\mathrm{Id}_U(A, B)$ is the set, i.e. discrete groupoid, of isomorphisms from A and B one needs the principle of function extensionality in order to organize U into a groupoid (bijections are equal iff they are pointwise equal).

Thus, since the groupoid model validates UA one might ask whether UA is compatible with our criteria for intensionality. The answer to this question, however, is negative since as shown in [**HoTT**] the univalence axiom UA allows one to derive from it the principle of function extensionality. The latter, however, is in contradiction with condition (I3) since together with function extensionality it has the consequence that for closed terms t of type $N{\to}N$ the proposition $\Pi x{:}N.\mathrm{Id}_N(t(x), 0)$ is derivable if and only if $t = \lambda x{:}N.0 \in N{\to}N$ is derivable which is impossible since the set of Π^0_1 sentences provable in ITT + UA is not decidable (since ITT + UA extends PRA).

Thus, adding UA to ITT is an extension which is not conservative w.r.t. Basic Type Theory (BTT), i.e. ITT without universes, since this extension is not even conservative w.r.t. to Π^0_1 sentences.[2] However, we have the following conservation result for ITT extended by function extensionality.

Theorem 2 *If a proposition of* BTT *can be proved in* ITT + UA *then it can be proved in* ITT *with a universe, the principle* $\mathrm{Ext}_{\mathrm{fun}}$ *of function extensionality and UIP in form of the eliminator* K*.*

Proof. In ITT + $\mathrm{Ext}_{\mathrm{fun}}$ + K with a universe one can construct the groupoid model of [HS94]. Notice that we need $\mathrm{Ext}_{\mathrm{fun}}$ in the meta-theory for

(1) getting exponentials of groupoids right
(2) defining Id-types on the universe of discrete groupoids since we need extensional equality of isomorphisms between types in the universe.

The eliminator K is needed for avoiding problems with intensional identity types. In ITT + $\mathrm{Ext}_{\mathrm{fun}}$ + K with a universe one can prove that all types of BTT are interpreted by their corresponding discrete groupoid. Accordingly, in this theory one can prove for every type A of BTT that if the interpretation of A is inhabited in the groupoid model by some element a then the type A is inhabited actually by "stripping" the element a from additional information. □

Thus, as far as BTT is concerned the univalence axiom does not contribute more than the principle $\mathrm{Ext}_{\mathrm{fun}}$ of function extensionality and the eliminator K. In [**Hof97**] using "setoid" models M. Hofmann has investigated to which extent extensional concepts can be interpreted within intensional type theory. Actually, he managed to interpret ITT + $\mathrm{Ext}_{\mathrm{fun}}$ + K without universes in ITT (with a universe). If one could also interpret universes this way Theorem 2 would give a positive answer to Voevodsky's "no junk" conjecture which claims that for every closed term t of type N (natural numbers) there exists an $n \in \mathbb{N}$ such that HoTT proves $t = \underline{n} \in N$ where $\underline{n}$ stands for the numeral $\mathsf{succ}^n(0)$. Thus, in light of Theorem 2 the problem rather is to prove this "no junk" conjecture for ITT + $\mathrm{Ext}_{\mathrm{fun}}$ + K.

Summarizing we observe that the answer to our question is twofold. HoTT is inconsistent with *equality reflection* and thus with ETT but on the other hand it is conservative over ITT + $\mathrm{Ext}_{\mathrm{fun}}$ + K w.r.t. basic type theory.

References

[Stre1] T. Streicher *Semantics of Type Theory* Birkhäuser Publ. Comp. 1991.

[Stre2] T. Streicher *Investigations into Intensional Type Theory* Habilitation Thesis Ludwig-Maximilans-Universität Munich, 1993. electronically available from `www.mathematik.tu-darmstadt.de/~streicher/HabilStreicher.pdf`

[HS98] M. Hofmann and T. Streicher *The Groupoid Interpretation of Type Theory* in *Twenty-five years of Martin-Löf type theory* Oxford Logic Guides vol. 36, pp. 83-111, OUP (1998).

[Hof97] M. Hofmann *Extensional Concepts in Intensional Type Theory* Springer (1997).

[vO08] J. van Oosten *Realizability. An Introduction to its Categorical Side.* Elsevier (2008).

[HoTT] *Homotopy Type Theory. Univalent Foundations of Mathematics.* many authored book, spring 2013, electronically available from `homotopytypetheory.org/book/`

[2] This has to be seen in sharp contrast with the fact that most known non-syntactic models for ITT + UA (as e.g. the groupoid and the simplicial sets model) validate the same propositions of BTT as the model in **Set** does.