to appear in J. Symbolic Logic

A construction of Type:Type in Martin-Löf's partial type theory with one universe[1]

Erik Palmgren
Department of mathematics, Uppsala university
Thunbergsvägen 3, S-752 38 Uppsala, Sweden

In this note we construct Martin-Löf's inconsistent type theory, Type:Type [Martin-Löf, 1971], inside partial type theory with one universe. Thus adding a fixed point operator to type theory with one predicative universe gives impredicativity.

We may describe the theory Type:Type as follows. It contains the rules for the product construction (Π) of Martin-Löf [1984] except the η -rule and it contains the usual rules for definitional equality (=). Moreover it contains the following strongly impredicative universe

$$U^\infty \ set \qquad\qquad \frac{a \in U^\infty}{T^\infty(a)\ set} \tag{1}$$

$$u^\infty \in U^\infty \qquad\qquad T^\infty(u^\infty) = U^\infty \tag{2}$$

$$\frac{a \in U^\infty \qquad \begin{matrix}(x \in T^\infty(a))\\ b \in U^\infty\end{matrix}}{\pi'(a,(x)b) \in U^\infty} \qquad \frac{a \in U^\infty \qquad \begin{matrix}(x \in T^\infty(a))\\ b \in U^\infty\end{matrix}}{T^\infty(\pi'(a,(x)b)) = (\Pi x \in T^\infty(a))T^\infty(b).} \tag{3}$$

This theory is inconsistent (i.e. every set is inhabited), and this is seen by proving a variant of the Burali-Forti paradox – Girard's paradox – cf. Troelstra and van Dalen [1988]. Coquand [199?] has shown that by adding the wellorder type and the strong dependent sum to the universe, the fixed point operator becomes definable. It is an open problem whether it is definable without the wellorder type. The present result could be seen as a converse, namely by adding the fixed point operator to type theory with one universe, Type:Type becomes definable and, as is already known, so does the wellorder type.

It is known (cf. Amadio et al. [1986] and Coquand et al. [1989]) that we may construct a domain-theoretic model of the variant of Type:Type, where the last equality above is replaced by an explicitly given isomorphism, using finitary projections. The present construction of Type:Type together with the domain interpretation of partial type theory with universes [Palmgren, 1989], gives indirectly a domain interpretation of Type:Type where the equality is interpreted as equality. These models each show that the definitional equality is not trivial.

[1]During the preparation of this paper the author was supported by the Swedish Natural Science Research Council as a Ph.D. student in mathematical logic.

We shall do the construction in Martin-Löf's partial type theory with one universe (U, T). The qualification "partial" means that it has a general fixed point operator [Martin-Löf, 1986] given by the rules,

$$\frac{\begin{array}{c}(x \in A)\\ f \in A\end{array}}{fix((x)f) \in A} \qquad\qquad \frac{\begin{array}{c}(x \in A)\\ f \in A\end{array}}{fix((x)f) = f(fix((x)f)/x) \in A.}$$

By applying this on the universe we may solve type equations,[2] e.g.

$$T(fix((x)x \to x)) = T(fix((x)x \to x)) \longrightarrow T(fix((x)x \to x)).$$

The general fixed point operator yields a trivial proof of every proposition (i.e. $fix((x)x) \in A$), and in particular of the identity type $I(A, a, b)$. Therefore the extensional identity type of Martin-Löf [1984] is abandoned since we want a non-trivial definitional equality. We can, if desired, assume the intensional identity type, as can be seen from the domain interpretations [Palmgren and Stoltenberg-Hansen, 1990] and [Palmgren, 1989]. However, it turns out that neither identity type nor η-rule is needed in the construction.

The first (predicative) universe of total type theory could be seen as a fixed point of a monotone operator which is given by the introduction rules for the universe. In the case of an impredicative universe, such as Type:Type, this operator is no longer monotone, and we cannot guarantee the existence of a fixed point. By resorting to certain categories of Scott-domains these operators all become monotone and continuous. In the domain interpretation of partial type theory with universes [Palmgren, 1989], a large cpo of parametrizations (monotone families) of information systems was introduced and the universes were obtained as fixed points on this cpo. The idea is to internalize this construction principle into partial type theory. The universe U will correspond to the class of information systems, and $PAR := (\Sigma x \in U)[T(x) \to U]$, i.e. the set of codes for families of sets in U, correspond to parametrizations of information systems.

The construction. We refer to Martin-Löf [1984] for explications of the constants and rules of type theory. In the following we use $\langle \cdot, \cdot \rangle$ for construction of pairs; p and q are the first and second projection respectively. We define the following operations on $PAR := (\Sigma x \in U)[T(x) \to U]$: If $c, d \in PAR$ let

$$c \hat{+} d := \langle p(c) + p(d), (\lambda z)D(z, (x)Ap(q(c), x), (y)Ap(q(d), y))\rangle,$$

the sum of two families, and if $c \in PAR$ let

$$\begin{aligned} s_\pi(c) := \langle &\sigma(p(c),(x)Ap(q(c), x) \to p(c)),\\ &\qquad (\lambda w)\pi(Ap(q(c), p(w)), (x)Ap(q(c), Ap(q(w), x)))\rangle, \end{aligned}$$

the operator corresponding to the closure of the universe under Π-formation (3). We leave it to the reader to check that $c \hat{+} d \in PAR$.

[2]Note that the *indexed* fixed point operator of non-standard type theory [Martin-Löf, 1989] will not solve such equations, and hence can not be used to carry out the construction below.

To see that $s_\pi(c) \in PAR$ we perform the following informal deduction. Assume that $c \in PAR$. Then $p(c) \in U$ and $q(c) \in T(p(c)) \longrightarrow U$ by Σ-elimination. From the assumption that $x \in T(p(c))$ we obtain $Ap(q(c), x) \in U$, and hence $Ap(q(c), x) \to p(c) \in U$ by U-introduction. Thus $\sigma(p(c), (x)Ap(q(c), x) \to p(c)) \in U$, again by U-introduction.

Now assume that $w \in T(\sigma(p(c), (x)Ap(q(c), x) \to p(c)))$. By the commutation of T with Σ and Π and by Σ-elimination we get $p(w) \in T(p(c))$ and $q(w) \in T(Ap(q(c), p(w))) \longrightarrow T(p(c))$. We have $Ap(q(c), p(w)) \in U$ and may assume $x \in T(Ap(q(c), p(w)))$. Then $Ap(q(w), x) \in T(p(c))$, and thus $Ap(q(c), Ap(q(w), x)) \in U$. By U-introduction we get

$$\pi(Ap(q(c), p(w)), (x)Ap(q(c), Ap(q(w), x))) \in U.$$

Abstracting on w and pairing with $\sigma(p(c), (x)Ap(q(c), x) \to p(c))$ in the first coordinate yields

$$\begin{aligned}&\langle \sigma(p(c), (x)Ap(q(c), x) \to p(c)),\\ &\qquad (\lambda w)\pi(Ap(q(c), p(w)), (x)Ap(q(c), Ap(q(w), x)))\rangle \in PAR,\end{aligned}$$

i.e. $s_\pi(c) \in PAR$.

We define the operator that builds the universe (U^∞, T^∞) by putting

$$f(c) := s_\pi(c) \hat{+} \langle n_1, (\lambda x)R_1(x, p(c))\rangle,$$

for $c \in PAR$, and let $e := fix((c)f(c))$. Hence $e \in PAR$ is a fixed point of f, $e = f(e)$. The right summand of f corresponds to the rules (2).

We now interpret Type:Type. The universe (U^∞, T^∞) is defined by letting

$$U^\infty := T(p(e))$$

and

$$T^\infty(a) := T(Ap(q(e), a)),$$

for $a \in U^\infty$. Thus the rules (1) are verified.

Using the equality $e = f(e)$ and the commutation of T with Σ, Π and $+$ we get

$$\begin{aligned}(4) \qquad U^\infty = T(p(e)) &= T(p(f(e)))\\ &= T(\sigma(p(e), (x)Ap(q(e), x) \to p(e))) + T(n_1)\\ &= (\Sigma x \in T(p(e)))[T(Ap(q(e), x)) \longrightarrow T(p(e))] + N_1\\ &= (\Sigma x \in U^\infty)[T^\infty(x) \longrightarrow U^\infty] + N_1\end{aligned}$$

and hence $j(0_1) \in U^\infty$. Furthermore we have $T^\infty(j(0_1)) = T(Ap(q(e), j(0_1))) = T(p(e)) = U^\infty$, and we let $u^\infty := j(0_1)$. Hence we have $u^\infty \in U^\infty$ and $T^\infty(u^\infty) = U^\infty$, i.e. (2) is verified.

Finally we shall define the code $\pi'(a,(x)b)$ and verify the rules (3). Suppose that $a \in U^\infty$ and that $b \in U^\infty$ under the assumption that $x \in T^\infty(a)$. Hence $(\lambda x)b \in T^\infty(a) \longrightarrow U^\infty$. Σ-introduction yields

$$\langle a,(\lambda x)b\rangle \in (\Sigma x \in U^\infty)[T^\infty(x) \longrightarrow U^\infty],$$

and $+$-introduction gives

$$i\langle a,(\lambda x)b\rangle \in T(p(f(e))) = T(p(e)) = U^\infty,$$

by the equalities (4) above. Now define $\pi'(a,(x)b) := i\langle a,(\lambda x)b\rangle$. Then $\pi'(a,(x)b) \in U^\infty$, as desired, and finally

$$\begin{aligned}
T^\infty(\pi'(a,(x)b)) &= T(Ap(q(f(e)), i\langle a,(\lambda x)b\rangle))\\
&= T(\pi(Ap(q(e),a),(x)Ap(q(e),Ap((\lambda x)b,x))))\\
&= (\Pi x \in T(Ap(q(e),a)))T(Ap(q(e),Ap((\lambda x)b,x)))\\
&= (\Pi x \in T(Ap(q(e),a)))T(Ap(q(e),b))\\
&= (\Pi x \in T^\infty(a))T^\infty(b).
\end{aligned}$$

We have proved the following

THEOREM. *Type:Type is definable in partial type theory with one universe.* □

Note that the strong sum (Σ) is easily added to the universe (U^∞, T^∞) by inserting

$$\begin{aligned}
s_\sigma(c) := \langle &\sigma(p(c),(x)Ap(q(c),x) \rightarrow p(c)),\\
&(\lambda w)\sigma(Ap(q(c),p(w)),(x)Ap(q(c),Ap(q(w),x)))\rangle,
\end{aligned}$$

as a summand of f, by defining a code for $\sigma'(a,(x)b)$ and modifying the codes $\pi'(a,(x)b)$ and u^∞ accordingly.

Using the method described we may construct universes for other impredicative type theories such as the second order λ-calculus. Partial type theory with one universe seems indeed to be "universal" for type theories.

Acknowledgement. I wish to thank the referee for his comments.

References

1. Amadio, R., K.B. Bruce, G. Longo, The finitary projection model for second order lambda calculus and solutions to higher order domain equations, In: *Logic in Computer Science* (ed. A.R. Meyer), IEEE Computer Society Press, 1986, pp. 122-130.
2. Coquand, T., C. Gunter, G. Winskel, *Domain theoretic models of polymorphism,* Information and Computation vol. 81, 1989, pp. 123-167.
3. Coquand, T., *The paradox of trees in Type Theory,* BIT, to appear, 199?.
4. Martin-Löf, P., A theory of types, unpublished manuscript, 1971.
5. Martin-Löf, P., *Intuitionistic type theory,* Bibliopolis, 1984.
6. Martin-Löf, P., Unifying Scott's theory of domains for denotational semantics and intuitionistic type theory (abstract), *Atti del congresso logica e filsofia della scienza, oggi. San Gimignano 7-11 dicembre 1983, vol. I – Logica.* CLUEB, Bologna 1986.

7. Martin-Löf, P., Mathematics of infinity, In: *Proceedings of the COLOG 88 conference, Tallinn* (eds. P. Martin-Löf, G.E. Mints), Springer Lecture Notes in Computer Science, Springer, 1989.
8. Palmgren, E., V. Stoltenberg-Hansen, *Domain interpretations of Martin-Löf's partial type theory,* Annals of Pure and Applied Logic, 48 (1990) pp. 135-196.
9. Palmgren, E., A domain interpretation of Martin-Löf's partial type theory with universes, U.U.D.M. Report 1989:11, Uppsala 1989.
10. Troelstra, A.S., D. van Dalen, *Constructivism in mathematics, vol. II,* North-Holland, 1988.