Sheaf model of type theory

The set theoretic model of type theory interprets universes à la Russell. The (pre)sheaf models do not validate these universes. However we can validate a simpler version than universes à la Tarski, and this is what we present here. We present it in the case of presheaf models, but essentially the same holds for sheaf models.

1 Syntax

We list the rules of type theory, using a name-free syntax.

$$\frac{\Gamma \vdash}{1 : \Gamma \to \Gamma} \qquad \frac{\sigma : \Delta \to \Gamma \quad \delta : \Theta \to \Delta}{\sigma\delta : \Theta \to \Gamma}$$

$$\frac{\Gamma \vdash A \ \mathsf{Type}_n \qquad \sigma : \Delta \to \Gamma}{\Delta \vdash A\sigma \ \mathsf{Type}_n} \qquad \frac{\Gamma \vdash t : A \qquad \sigma : \Delta \to \Gamma}{\Delta \vdash t\sigma : A\sigma}$$

$$\frac{}{() \vdash} \qquad \frac{\Gamma \vdash \qquad \Gamma \vdash A \ \mathsf{Type}_n}{\Gamma.A \vdash} \qquad \frac{\Gamma \vdash A \ \mathsf{Type}_n}{\mathsf{p} : \Gamma.A \to \Gamma} \qquad \frac{\Gamma \vdash A \ \mathsf{Type}_n}{\Gamma.A \vdash \mathsf{q} : A\mathsf{p}}$$

$$\frac{\sigma : \Delta \to \Gamma \qquad \Gamma \vdash A \qquad \Delta \vdash u : A\sigma}{(\sigma, u) : \Delta \to \Gamma.A}$$

$$\frac{\Gamma \vdash A \ \mathsf{Type}_n \quad \Gamma.A \vdash B \ \mathsf{Type}_n}{\Gamma \vdash \ \Pi \ A \ B \ \mathsf{Type}_n} \qquad \frac{\Gamma.A \vdash b : B}{\Gamma \vdash \lambda b : \Pi \ A \ B} \qquad \frac{\Gamma \vdash w : \Pi \ A \ B \qquad \Gamma \vdash u : A}{\Gamma \vdash \mathsf{app}(w, u) : B[u]}$$

$$\frac{\Gamma \vdash A \ \mathsf{Type}_n \quad \Gamma.A \vdash B \ \mathsf{Type}_n}{\Gamma \vdash \ \Sigma \ A \ B \ \mathsf{Type}_n} \qquad \frac{\Gamma \vdash u : A \qquad \Gamma \vdash v : B[u]}{\Gamma \vdash (u, v) : \Sigma \ A \ B} \qquad \frac{\Gamma \vdash w : \Sigma \ A \ B}{\Gamma \vdash \mathsf{p}w : A} \qquad \frac{\Gamma \vdash w : \Sigma \ A \ B}{\Gamma \vdash \mathsf{q} \ w : B[\mathsf{p}w]}$$

where $[u] = (1, u) : \Gamma \to \Gamma.A$ if $\Gamma \vdash u : A$.

$$1\sigma = \sigma = \sigma 1 \qquad (\sigma\delta)\nu = \sigma(\delta\nu)$$

$$(\sigma, u)\delta = (\sigma\delta, u\delta) \qquad \mathsf{p}(\sigma, u) = \sigma \qquad \mathsf{q}(\sigma, u) = u$$

$$(A\sigma)\delta = A(\sigma\delta) \qquad A1 = A \qquad (a\sigma)\delta = a(\sigma\delta) \qquad a1 = a$$

$$\mathsf{app}(w, u)\delta = \mathsf{app}(w\delta, u\delta) \qquad \mathsf{app}(\lambda b, u) = b[u] \qquad (\lambda b)\sigma = \lambda(b(\sigma\mathsf{p}, \mathsf{q}))$$

$$u, v)\delta = (u\delta, v\delta) \qquad \mathsf{p}(u, v) = u \qquad \mathsf{q}(u, v) = v \qquad (\mathsf{p}u)\sigma = \mathsf{p}(u\sigma) \qquad (\mathsf{q}u)\sigma = \mathsf{q}(u\sigma)$$

$$1 = (\mathsf{p}, \mathsf{q}) \qquad v = \lambda\mathsf{app}(v\mathsf{p}, \mathsf{q})$$

We add the following rules for universes.

$$\frac{\Gamma \vdash A \ \mathsf{Type}_n}{\Gamma \vdash |A| : U_n} \qquad \frac{\Gamma \vdash T : U_n}{\Gamma \vdash El \ T \ \mathsf{Type}_n}$$

$$\frac{\Gamma \vdash A \ \mathsf{Type}_n}{\Gamma \vdash A \ \mathsf{Type}_{n+1}} \qquad \frac{\Gamma \vdash T : U_n}{\Gamma \vdash T : U_{n+1}}$$

$$\frac{}{\Gamma \vdash U_n \ \mathsf{Type}_{n+1}}$$

$$El \ |A| = \ A \qquad\qquad |El \ T| = T$$

With this presentation, we can define $\pi \ T \ V = |\Pi \ (El \ T) \ (El \ V)|$ if $\Gamma \vdash T : U_n$ and $\Gamma.El \ T \vdash V : U_n$. This satisfies $El \ (\pi \ T \ V) = \Pi \ (El \ T) \ (El \ V)$.

2 Presheaf model

If $\mathcal{C}$ is any small category, the presheaf model of type theory over $\mathcal{C}$ can be described as follows.

To simplify the presentation, we don't consider the question of size.

We write $X, Y, Z, \ldots$ the objects of $\mathcal{C}$ and $f, g, h, \ldots$ the maps of $\mathcal{C}$. If $f : X \to Y$ and $g : Y \to Z$ we write gf the composition of f and g. We write $1_X : X \to X$ or simply $1 : X \to X$ the identity map of X. Thus we have $(fg)h = f(gh)$ and $1f = f1 = f$.

A context is interpreted by a presheaf Γ: for any object X of $\mathcal{C}$ we have a set $\Gamma(X)$ and if $f : Y \to X$ we have a map $\rho \longmapsto \rho f,\ \Gamma(X) \to \Gamma(Y)$. This should satisfy $\rho 1 = \rho$ and $(\rho f)g = \rho(fg)$ for $f : Y \to X$ and $g : Z \to Y$.

A type $\Gamma \vdash A$ over Γ is given by a set $A\rho$ for each $\rho : \Gamma(X)$. Furthermore if $f : Y \to X$ we have $\rho f : \Gamma(Y)$ and we can consider the set $A\rho f$. We should have a map $u \longmapsto uf,\ A\rho \to A\rho f$ which should satisfy $u1 = u$ and $(uf)g = u(fg)$.

An element $\Gamma \vdash a : A$ is interpreted by a family $a\rho : A\rho$ such that $(a\rho)f = a(\rho f)$ for any $\rho : \Gamma(X)$ and $f : Y \to X$.

This can be seen as a concrete description of what is respectively a fibration and a section of this fibration.

If $\Gamma \vdash A$ we can define a new presheaf $\Gamma.A$ by taking $(\rho, u) : (\Gamma.A)(X)$ to mean $\rho : \Gamma(X)$ and $u : A\rho$. We define $(\rho, u)f = \rho f, uf$.

If we have a map $\sigma : \Delta \to \Gamma$ and $\Gamma \vdash A$ we define $\Delta \vdash A\sigma$ by $(A\sigma)\rho = A(\sigma\rho)$.

If $\Gamma \vdash A$ and $\rho : \Gamma(X)$ we define $|A|\rho$ to be the family $(A\rho f,\ f : Y \to X)$ with restriction map $A\rho f \to A\rho fg,\ u \longmapsto ug$ for $g : Z \to Y$.

We define $U(X)$ as the set of families of sets $Pf,\ f : Y \to X$ together with restriction maps $Pf \to Pfg,\ u \longmapsto ug$ satisfying $u1 = u$ and $(ug)h = u(gh)$. We define then $\Gamma \vdash U$ by taking $U\rho = U(X)$ if $\rho : \Gamma(X)$.

If we have $\Gamma \vdash T : U$ we define $\Gamma \vdash El\ T$ by the equation $(El\ T)\rho = T\rho 1_X$ for $\rho : \Gamma(X)$.

We validate then $|El\ T| = T$ and $El\ |A| = A$.

If $\Gamma \vdash A$ we have $(El\ |A|)\rho = |A|\rho 1_X$ and $|A|\rho$ is the family $A\rho f,\ f : \to X$, so that $|A|\rho 1_X = A\rho 1_X = A\rho$. The restriction map $u \longmapsto uf,\ (El\ |A|)\rho \to (El\ |A|)\rho f$ is the restriction map defined by $A\rho \to A\rho f$.

If $\Gamma \vdash T : U$ the family $(El\ T)\rho f,\ f : Y \to X$ is defined by $T(\rho f)1_Y = T\rho f$, and so $|El\ T| = T$.

We can interpret dependent products $\Gamma \vdash \Pi\ A\ B$ and sums $\Gamma \vdash \Sigma\ A\ B$ if we have $\Gamma \vdash A$ and $\Gamma.A \vdash B$. For $\rho : \Gamma(X)$ we define $(u, v) : (\Sigma\ A\ B)\rho$ to mean $u : A\rho$ and $v : B(\rho, u)$. We define $(u, v)f = uf, vf$ for $f : Y \to X$. On the other hand an element of $(\Pi\ A\ B)\rho$ is a family w indexed by $h : Y \to X$ with

$$wh : \prod_{u : A\rho h} B(\rho h, u)$$

and such that $\mathsf{app}(wh, u)g = \mathsf{app}(whg, ug)$ if $h : Y \to X$ and $g : Z \to Y$. We define then $(wh)f = w(hf)$. We write $w = w1$.

We can interpret $\Gamma \vdash \lambda t : \Pi\ A\ B$ whenever $\Gamma.A \vdash t : B$ and $\Gamma \vdash \mathsf{app}(v, u) : B[u]$ if $\Gamma \vdash u : A$ and $\Gamma \vdash v : \Pi\ A\ B$. Here we write $[u]$ the map $\Gamma \to \Gamma.A$ defined by $[u]\rho = \rho, u\rho$. If $\rho : \Gamma(X)$ and $f : Y \to X$ we define $\mathsf{app}((\lambda t)\rho f, a) = t(\rho f, a) : B(\rho f, a)$ for $a : A\rho f$. We take $\mathsf{app}(v, u)\rho = \mathsf{app}(v\rho, u\rho) : B(\rho, u\rho)$. We can then check that we have

$$\mathsf{app}(\lambda t, u)\rho = t(\rho, u\rho) = t[u]\rho : B(\rho, u\rho)$$

if $\Gamma.A \vdash t : B$ and $\Gamma \vdash u : A$ and $\rho : \Gamma(X)$, which shows that the model validates the conversion rule $\Gamma \vdash \mathsf{app}(\lambda t, u) = t[u] : B[u]$.

3 Sheaf model

The previous definitions extend in the case of sheaf models over a site. We shall consider only the case of disjoint covering $f_0 : X_0 \to X,\ f_1 : X_1 \to X$ and in this case the sheaf condition is that the map $\Gamma(X) \to \Gamma(X_0) \times \Gamma(X_1),\ x \longmapsto (xf_0, xf_1)$ is an isomorphism.

We shall consider two examples.

3.1 Sheaf model over Cantor spaces

The category $\mathcal{C}$ is a poset. The objects $X, Y, Z, \ldots$ are basic open of Cantor space. We represent them as finite amount of informations

$$\omega(k_0) = b_0, \ldots, \omega(k_{p-1}) = b_{p-1}$$

about a generic function $\omega : N \to N_2$. If k is not in the domain of X then a basic covering of X is given by $X, \omega(k) = 0$ and $X, \omega(k) = 1$. We can then define inductively what is a partition $X \lhd P$ of X

1. $X \lhd X$
2. $X \lhd P_0, P_1$ whenever $X, \omega(k) = 0 \lhd P_0$ and $X, \omega(k) = 1 \lhd P_1$

An element of $N(X)$ is a formal sum $u_0X_0 + \ldots + u_{n-1}X_{n-1}$ if $u_0, \ldots, u_{n-1}$ are numeral and $X \lhd X_0, \ldots, X_{n-1}$. Similarly an element of $N_2(X)$ is a formal sum $u_0X_0 + \ldots + u_{n-1}X_{n-1}$ if $u_0, \ldots, u_{n-1}$ are Booelan and $X \lhd X_0, \ldots, X_{n-1}$. We require $uX = uX_0 + \ldots + uX_{n-1}$. This defines the interpretation of N and N_2 as sheaves over Cantor space.

We define a natural transformation $\omega : N \to N_2$ by taking

$$\omega(u_0X_0 + \ldots + u_{n-1}X_{n-1}) = \omega(u_0X_0) + \ldots + \omega(u_{n-1}X_{n-1})$$

and $\omega(uX) = bX$ if u is in the domain of X and $\omega(u) = b$ is in X and

$$\omega(uX) \quad = \quad 0X(\omega(u) = 0) + 1X(\omega(u) = 1)$$

otherwise.

One suggestive way to describe this model is that we have added to type theory one generic infinite binary sequence ω, and each stage X describes the value about some finite amount of information about this generic sequence. At any stage of knowledge, we can require the information about a new value $\omega(k)$ and we should get this information in a finite amount of time.

We define in type theory with universes $T : N_2 \to U_0$ by $T\ 0 = |N_0|$ and $T\ 1 = |N_1|$ where N_0 is the empty type and N_1 the unit type, and N_2 the type with two elements $0, 1$. We define $\neg A = A \to N_0$.

Theorem 3.1 *The following statement expressing Markov's principle*

$$\Pi p : N \to N_2.(\neg\neg\Sigma n : N.El\ (T\ (p\ n))) \to \Sigma n : N.El\ (T\ (p\ n))$$

is not provable in dependent type theory with universes.

Proof. We show that this statement is not valid in the sheaf model over Cantor space. We are going to analyse the type

$$S = \Sigma n : N.El\ (T\ (\omega\ n))$$

We have that $S(X)$ is inhabited as soon as we have k in the domain of X and $\omega(k) = 1$ in X for then $(kX, 0)$ is an element of $S(X)$. On the other hand $S()$ is empty since if $X_0, \ldots, X_{n-1}$ is a partition of $()$ then there is always one X_i of the form $\omega(k_0) = 0, \ldots, \omega(k_{p-1}) = 0$. For any X we can choose k not in the domain of X and we have

$$X, \omega(k) = 1 \quad \leqslant \quad X$$

It follows from this that $\neg S = S \to N_0$ is empty at all stages X and so that $\neg\neg S$ is a singleton at all stages. Since $S()$ is empty, it follows that $(\neg\neg S \to S)()$ is empty and so

$$\Pi p : N \to N_2.(\neg\neg\Sigma n : N.El\ (T\ (p\ n))) \to \Sigma n : N.El\ (T\ (p\ n))$$

has no global element. $\square$

3.2 Interpretation of the fan functional

The category $\mathcal{C}$ has for objects basic open subset of finite power of Cantor space. The maps are uniformely continuous maps. The covering are finite partition in disjoint clopen subsets. In this model, we define $N(X)$ to be the uniformely continuous map from X to N, and similarly $N_2(X)$ is the set of uniformely continuous maps from X to N_2.

One intuitive description of this model is the following. Each stage of knowledge X represents the values of initial segment of finitely many generic sequences $\omega^0, \dots, \omega^{m-1}$. At any stage of knowledge, we can choose some of the sequence, and require to know what is its value at a choosen element, information we should obtain in a finite amount of time. It may also happen that we discover that these sequences can actually be obtained as continuous functions of other sequences, for which we know some finite initial segment, which is represent by a new stage of knowledge Y and a continuous function $f : Y \to X$. But, contrary to the change of information that is specified by the covering, this change of information is not bound to happen in a finite amount of time (cf. Kripke's discussion of the difference between his notion of model and Beth's notion of model). This explains in what sense this model can be seen as a refinement of the previous model of sheaves over Cantor space.

In this model, we have a direct description of a sheaf for representing the space $N \to N_2$ by defining $C(X)$ to be the set of all uniformly continuous functions $X \to C$ where C is the Cantor space.

Lemma 3.2 *The sheaf $C(X)$ represents $(N \to N_2)(X)$.*

Proof. We build a natural isomorphism between the functors $C(X)$ and $(N \to N_2)(X)$. In one direction, if we have an element $\varphi : X \to C$ then we can define a family $\varphi_f : N(Y) \to N_2(Y)$ for $f : Y \to X$ by taking $\varphi_f \alpha = \lambda y.\varphi(f(y))(\alpha(y))$. In the other direction, if we have a family $\varphi_f : N(Y) \to N_2(Y)$ satisfying $(\varphi_f \alpha)g = \varphi_{fg}(\alpha g)$ for $g : Z \to Y$ then we can define $\varphi : X \to C$ by taking $\varphi(x, n) = \varphi_1(\lambda y.n)(x)$. These two functions define a natural isomorphism between the functors $C(X)$ and $(N \to N_2)(X)$. $\square$

We use $\alpha, \beta, \gamma, \dots$ to range over elements of type C and $n, m, \dots$ to range over elements of N.

We define a function $(\leqslant) : N \to N \to U$ by

$$0 \leqslant m \; = \; |N_1| \qquad S(n) \leqslant 0 \; = \; |N_0| \qquad S(n) \leqslant S(m) \; = \; n \leqslant m$$

We define an element

$$FT : (\Pi\alpha : C.\Sigma n : N.\varphi(\alpha, n)) \to \Sigma M : N.\Pi\alpha : C.\Sigma n : N.\varphi(\alpha, n) \times El\ (n \leqslant M)$$

For this we assume to have

$$h : \Pi\alpha : C.\Sigma n : N.\varphi(\alpha, n)$$

at some stage of knowledge X. This means that if $g : Z \to X$ and α is in $C(Z)$ then $h(g, \alpha)$ is a pair n, u where n is in $N(Z)$ and u is an element of $\varphi(\alpha, n)$. We can then consider $p : X \times C \to X$ and $q : C(X \times C)$ and $h(p, q) = (n, u) : \Sigma n.\varphi(q, n)$ with n in $N(X \times C)$. This gives a finite partition $U_0, \dots, U_{l-1}$ of $X \times C$ with a finite number of associated values $n_0, \dots, n_{l-1}$. We define then M to be the maximum value of $n_0, \dots, n_{l-1}$. If we have $f : Y \to X$ and $\alpha : C(Y)$ we can then consider $(f, \alpha) : Y \to X \times C$ and $n(f, \alpha)$ is an element of $N(Y)$ and $u(f, \alpha)$ is a proof of $\varphi(\alpha, n(f, \alpha))$. Furthermore the function $n(f, \alpha)$ is always less then the constant function M on Y. So the definition of FT is $FT(h) = (M, h')$ where M is the maximum values of $n_0, \dots, n_{l-1}$ obtained by computing $h(p, q)$ and $h'(f, \alpha) = (n, 0, u)$ where $(n, u) = h(f, \alpha)$.