

$\forall T, \sigma, k.$ if $(\exists j. \forall n \geq j. \text{Fn}_{T_n} \sigma_0 \text{ if } \forall i < k. \text{Fn}_{T_n} \sigma_{i+1})$
then $\text{HFn}(n.T_n) \sigma_0 \text{ if } \forall i < k. \text{HFn}(n.T_n) \sigma_{i+1}.$

For example, the facts about Fn expressed at the end of section 3.1, pages 39–43, carry over to the use of $\text{HFn}(n.T_n)$ when T_n is a reflective hierarchy.

Another approach to using a reflective hierarchy is to unite it into a single type system. We will not use this formulation; it is discussed only because it is similar to the theories of [HAN] and Nuprl [Constable et al. 86]. Let $\cup_n T_n$, where T_n is a sequence of equality skeletons, be the skeleton such that

$$A =_{\cup_n T_n} B \text{ iff } \exists n. A =_{T_n} B$$

and

$$a = b \in_{\cup_n T_n} A \text{ iff } \exists n. a = b \in_{T_n} A.$$

If T_n is a reflective hierarchy in n then $\cup_n T_n$ is a type system and is an extension of T_n for each n . When T_n is a reflective hierarchy, the relation $\text{Fn}_{\cup_n T_n} \sigma$ is equivalent to the relation $\text{HFn}(n.T_n) \sigma$ if every type-valued function belonging to $\cup_n T_n$ is a type-valued function of one of the T_n , that is, if

$$\forall \Delta. \text{ if } \text{Fn}_{\cup_n T_n} \Delta \text{ then } \exists n. \text{Fn}_{T_n} \Delta.$$

Note that it is not enough merely that each type inhabit a universe of the hierarchy.

As a last point, the fact that a reflective hierarchy can be united to give a type system provides us with a mechanism for producing transfinite hierarchies, but the ω -order hierarchy is sufficient to secure the use of types as objects.

4.2 An Example: Martin-Löf's Types

Here we shall define a type system hierarchy that is the non-type-theoretic analog of the types defined in Martin-Löf's paper [HAN]. A longer discussion of this definition may be found in [Allen 87]. The theory of types presented in [HAN] is open-ended in the sense that additional terms may be introduced and additional type constructors may be defined beyond those explicitly given there. In contrast, the type system to be defined here is closed, having only those terms and type constructors already given. The computation system is that embodied in the procedure given in [HAN] for evaluating terms.³

³To assure uniqueness of the results of evaluation we shall assume there is a designated variable v such that the contractum of $(T x, y, z)(\text{sup}(a, b), d)$ is always

$$d[a, b, (\lambda v)(T x, y, z)(b(v), d) / x, y, z].$$

Before proceeding with the definition proper, we shall consider the sort of definition that one might first attempt. One might define typehood and member equality by mutual recursion, where extensional type equality is simply defined by


$$T = S \text{ iff } T, S \text{ type } \& \forall t s. t = s \in T \text{ iff } t = s \in S.$$


The clauses for defining Π -types might be

$$(\Pi x \in A)B \text{ type if } A \text{ type } \& \forall a a'. B[a/x] = B[a'/x] \text{ if } a = a' \in A$$

and

$$\begin{aligned} t = t' \in (\Pi x \in A)B \text{ if } & (\Pi x \in A)B \text{ type} \\ & \& \exists u b u' b'. (\lambda u)b \leftarrow t \ \& \ (\lambda u')b' \leftarrow t' \\ & \& \forall a a'. b[a/u] = b'[a'/u'] \in B[a/x] \\ & \text{if } a = a' \in A. \end{aligned}$$

What is unusual about these clauses is that one of the definienda, member equality, occurs negatively on the right hand sides. Thus, the recursive definition does not work by presenting the usual sort of monotonic operator on, in this case, pairs of properties and relations with the aim of designating the least fixed point as the pair of definienda. 

What is intended is that we are somehow to understand that “whenever” a type is defined, its membership (equality) is completely defined as well. If we were to work out the other clauses of this definition we would find that whenever member equality is used, only its restriction to “already” defined types is needed, and the right hand sides of the clauses are (strictly) positive in typehood. Let us call the vague principle that licenses such inductive definition *half-positive* induction. Beeson gives a half-positive inductive definition in [Beeson 82] for his recursive realizability interpretation of [HAN]. He then indicates how to give a standard inductive definition of the model (by means of a device he attributes to Aczel), but this definition depends upon excluded middle. In [Beeson 85] he also mentions the stratification of typehood and member equality using classical set-theoretic ordinals; this is the classical transfinite analog of the ω -order stratification that was used to define FIN in section 2.1. 

The approach we shall take here is to make precise the nature of the half-positive induction. It is simply this: the half-positive recursive definition of “ T type” and “ $t = s \in T$ ” is a less than clear definition of the relation

T is a type with equality ϕ ,

which can be defined by ordinary induction using an operation on such relations which is strictly positive, hence monotonic, in its argument. Then let $t = s \in T$ mean that for some ϕ , T is a type with equality ϕ and $t\phi s$.⁴ In the author's opinion, this form of definition is clearer than the original half-positive form as well the other suggested classical "translations."

Since type equality in our example will be extensional, let us restrict our attention to such. We may characterize an extensional type system by a relation $\tau T\phi$ between terms (T) and two-place relations on terms (ϕ) such that

$$T =_{\tau} S \text{ iff } \exists \phi. \tau T\phi \ \& \ \tau S\phi$$

and

$$t = s \in_{\tau} T \text{ iff } \exists \phi. \tau T\phi \ \& \ t\phi s.$$

We shall use σ and τ as variables ranging over such relations, which we shall call *possible type systems*. Our intention will be to characterize extensional type systems by $\tau T\phi$ that define partial functions in T , that is, τ such that

$$\forall T \phi \phi'. \phi \text{ is } \phi' \text{ if } \tau T\phi \ \& \ \tau T\phi'.$$

To define our universe hierarchy, we shall define $\mu(\sigma)$ as the least fixed-point of an operator, $\text{TyF}(\sigma; \tau)$, which is monotonic in τ . The types of $\text{TyF}(\sigma; \tau)$ are those of σ plus those gotten by applying the non-universe type constructors to the types of τ . We will pass the universes in as base types through σ .

We begin the definition by setting out the type formation methods. This is done by defining the relations $\hat{N}_?$, \hat{N} , \hat{I} , $\hat{+}$, $\hat{\Sigma}$, $\hat{\Pi}$ and \hat{W} . Each of these is an operator on possible type systems,⁵ whose value (a possible type system) has only the types that evaluate to a certain form.

The types N_n and N have no constituent types.

$$\hat{N}_? T\phi \text{ iff } \exists n. N_n \leftarrow T \ \& \ \forall a \ b. a\phi b \text{ iff } \exists m < n. m_n \leftarrow a, b.$$

Define N-equality by

Neq is the strongest ϕ such that

$$\forall a \ b. a\phi b \text{ if } 0 \leftarrow a, b \text{ or } \exists a' \ b'. \text{ suc}(a') \leftarrow a \ \& \ \text{ suc}(b') \leftarrow b \ \& \ a'\phi b'.$$

⁴The same sort of definition could be used in the construction of Frege structures in [Aczel 80] (instead of the one given there using classical ordinals) by giving an ordinary inductive definition of

$$x \text{ is a proposition which is true iff } \Phi.$$

Then, choosing some object a , let "x true" mean that x is a proposition which is true iff $a = a$.

⁵We may consider possible type systems themselves to be zero-place operators.



$\widehat{N}T\phi$ iff $N \leftarrow T$ & ϕ is Neq.

The rest of the non-universe type constructors have constituent types, and so the type formation operators need, as a parameter, a possible type system from which to get these constituent types. In the definitions of \widehat{I} and $\widehat{\dagger}$, α and β range over two-place relations on terms.

$\widehat{I}(\tau)T\phi$ iff $\exists A \alpha a b. I(A, a, b) \leftarrow T$ & $\tau A\alpha$ & $\alpha\alpha a$ & $\beta\alpha b$
& $\forall t t'. t\phi t'$ iff $\tau \leftarrow t, t'$ & $\alpha\alpha b$.

$\widehat{\dagger}(\tau)T\phi$ iff $\exists A \alpha B \beta. A + B \leftarrow T$ & $\tau A\alpha$ & $\tau B\beta$
& $\forall t t'. t\phi t'$ iff $\exists a a'. i(a) \leftarrow t$ & $i(a') \leftarrow t'$ & $\alpha\alpha a'$
or $\exists b b'. j(b) \leftarrow t$ & $j(b') \leftarrow t'$ & $\beta\beta b'$.

Now we proceed with the type constructors having families of constituent types. In the definitions below, α ranges over two-place relations between terms and γ ranges over three-place relations between terms. The application of γ to terms is indicated by $t\gamma a s$.

$\text{Fam}(\tau; A; \alpha; x; B; \gamma)$ iff $\tau A\alpha$ & $\forall a a'. \text{if } \alpha\alpha a' \text{ then } \gamma_a \text{ is } \gamma_{a'}$
& $\tau B[a/x] \gamma_a$
& $\tau B[a'/x] \gamma_{a'}$.

Note that $\text{Fam}(\tau; A; \alpha; x; B; \gamma)$ is strictly positive in τ .

$\widehat{\Sigma}(\tau)T\phi$ iff $\exists A \alpha x B \gamma. (\Sigma x \in A) B \leftarrow T$ & $\text{Fam}(\tau; A; \alpha; x; B; \gamma)$
& $\forall t t'. t\phi t'$ iff $\exists a b a' b'. (a, b) \leftarrow t$ & $(a', b') \leftarrow t'$
& $\alpha\alpha a'$ & $b\gamma_a b'$.

$\widehat{\Pi}(\tau)T\phi$
iff $\exists A \alpha x B \gamma. (\Pi x \in A) B \leftarrow T$ & $\text{Fam}(\tau; A; \alpha; x; B; \gamma)$
& $\forall t t'. t\phi t'$ iff $\exists u b u' b'. (\lambda u) b \leftarrow t$ & $(\lambda u') b' \leftarrow t'$
& $\forall a a'. b[a/u] \gamma_a b'[a'/u']$ if $\alpha\alpha a'$.

Let us define the equality for W types.

$\text{Weq}(\alpha; \gamma)$ is the strongest ϕ such that
 $\forall t t'. t\phi t'$ if $\exists a f u s a' f' u' s'. \text{sup}(a, f) \leftarrow t$ & $(\lambda u) s \leftarrow f$
& $\text{sup}(a', f') \leftarrow t'$ & $(\lambda u') s' \leftarrow f'$
& $\alpha\alpha a'$ & $\forall b b'. s[b/u] \phi s'[b'/u']$ if $b\gamma_a b'$.

$\widehat{W}(\tau)T\phi$
iff $\exists A \alpha x B \gamma. (W x \in A) B \leftarrow T$ & $\text{Fam}(\tau; A; \alpha; x; B; \gamma)$ & ϕ is $\text{Weq}(\alpha; \gamma)$.

In each of these definitions of type formation, the definition of the member equality of a type depends only on the member equalities of its constituent types.

We may now define type formation under these constructors plus any base types.

$$\text{TyF}(\sigma; \tau)T\phi \text{ iff } \sigma T\phi \text{ or } \hat{N}_?T\phi \text{ or } \hat{N}T\phi \text{ or } \hat{I}(\tau)T\phi \text{ or } \hat{\dagger}(\tau)T\phi \\ \text{or } \hat{\Sigma}(\tau)T\phi \text{ or } \hat{\Pi}(\tau)T\phi \text{ or } \hat{W}(\tau)T\phi.$$

The relation $\text{TyF}(\sigma; \tau)T\phi$ is strictly positive, hence monotonic, in τ . Let us introduce a convenient notation for closure under $\text{TyF}(\sigma; \cdot)$.

$$\text{CTyF}(\sigma; \tau) \text{ iff } \forall T\phi. \tau T\phi \text{ if } \text{TyF}(\sigma; \tau)T\phi.$$

Now we define μ .

$$\mu(\sigma) \text{ is the strongest } \tau \text{ such that } \text{CTyF}(\sigma; \tau).$$

Before discussing the validity of this definition, let us finish up the definition of the hierarchy. We define our hierarchy using universes U_i and generator μ .

$$\text{HAN}_n \text{ is } \mu^*(i.U_i)_n.$$

Defining spine_n by

$$\text{spine}_n T\phi \text{ iff } \exists m < n. U_m \leftarrow T \ \& \ \phi \text{ is } =_{\text{HAN}_m},$$

we may say that HAN_n is $\mu(\text{spine}_n)$.

Returning to the definition of μ , it is clearly valid set-theoretically (in a theory with the power set axiom such as ZF or IZF); the terms can be represented as members of a set T , and for any subset σ of $T \times \text{Pow}(T \times T)$, $\text{TyF}(\sigma; \tau)$ is monotonic in τ on the subsets of $T \times \text{Pow}(T \times T)$, hence

$$\mu(\sigma) \text{ is } \bigcap \{ \tau \subseteq T \times \text{Pow}(T \times T) \mid \text{TyF}(\sigma; \tau) \subseteq \tau \}.$$

Standard intuitionistic theory of inductive definition directly licenses inductive definitions of

$$\text{the strongest } P \text{ such that } \forall \bar{x}. P\bar{x} \text{ if } \theta(\bar{x}; P),$$

where $\theta(\bar{x}; P)$ is a relation between individuals and properties of individuals that is (strictly) positive in P . The definition of $\mu(\sigma)$ does not quite conform to this standard since it is not a relation between individuals, but rather, for each σ , $\mu(\sigma)$ is a relation between individuals and two-place relations between individuals. Still, the intuitionist might be convinced of the validity of our definition since $\text{TyF}(\sigma; \tau)T\phi$ is strictly positive in τ ; it might also

help to note that the equality of each type depends only on the equalities of constituent types.

Notice that the notion of type system, as opposed to possible type system, does not enter into the definition of HAN_n . Although it may already be clear that HAN_n is a reflective hierarchy of type systems, we are in a position to prove it explicitly by using induction on type formation. We shall not carry out that proof here, but the three key lemmas are that $\mu(\sigma)$ is $\text{TyF}(\sigma; \mu(\sigma))$, that $\mu(\sigma)$ is monotonic in σ , and that if every type of σ evaluates to some U_i , then $\mu(\sigma)T\phi$ defines a partial function in T if $\sigma T\phi$ does.

In $[\text{HAN}]$ such a hierarchy is not the end result of the type definitions. All the universes are taken as types of a single system whose non-type-theoretic analog would be $\mu(\cup_n \text{spine}_n)$, which we may call HAN_ω . For future reference, let us use spine_ω to mean $\cup_n \text{spine}_n$. It turns out that HAN_ω is $\cup_n \text{HAN}_n$. This would not be so if, say, there were a term constructor $\text{univ}(e)$, with principle argument e , such that $U_0 \leftarrow \text{univ}(0)$ and $U_{i+1} \leftarrow \text{univ}(\text{suc}(t))$ if $U_i \leftarrow \text{univ}(t)$; in that case, $(\prod x \in N)\text{univ}(x)$ would be a type of HAN_ω but would not be a type of $\cup_n \text{HAN}_n$. The argument to be presented here turns on the fact that the U_i are, in a sense, computationally inert. In the course of evaluating a term that has a value, occurrences of U_i are just dragged around or abandoned without any notice taken of which term is being used, and their lineage can always be traced back to occurrences of U_i in the original term. We shall see that if all indices of universes occurring in a type T of HAN_ω are less than n , then T is a type of HAN_n .

We shall exploit the fact that there are inert canonical terms that are not types of HAN_ω and in which universes do not occur. One such term is $I(0,0,0)$. Let $t-n$ be the term gotten from t by replacing each occurrence of U_{n+i} (for every i) by $I(0,0,0)$. Let us say that terms t and s are *variants up from n* , or $t \text{ var}_n s$, when $t-n$ is $s-n$, that is, when t and s differ only in occurrences of $I(0,0,0)$ and universes at or above U_n . Clearly,

$$t[\bar{s}/\bar{x}] \text{ var}_n t'[\bar{s}'/\bar{x}] \text{ if } t \text{ var}_n t' \text{ \& } \bar{s} \text{ var}_n \bar{s}',$$

where $\bar{s} \text{ var}_n \bar{s}'$ is the obvious analog of $t \text{ var}_n t'$. By induction on evaluation,

$$\text{if } t \leftarrow s \text{ var}_n s' \text{ then } \exists t'. t' \leftarrow s' \text{ \& } t \text{ var}_n t'.$$

We shall now characterize a certain kind of immunity to variation up from n which types of HAN_n may have. Define $\text{VAR}_n T\phi$ by

$$\begin{aligned} \text{VAR}_n T\phi \text{ iff } \forall T'. \text{HAN}_n T'\phi \text{ if } T' \text{ var}_n T \\ \text{\& } \forall t s. \text{ if } t\phi s \text{ then } \forall s'. t\phi s' \text{ if } s' \text{ var}_n s. \end{aligned}$$

In fact, every type of HAN_n has this immunity (and so, since every term is a variant of itself, HAN_n is VAR_n):

if $\text{HAN}_n T\phi$ then $\text{VAR}_n T\phi$.

It is enough to show (but we will not) that $\forall n. \text{CTyF}(\text{spine}_n; \text{VAR}_n)$, which may be proved by induction on n .⁶ The inductive hypothesis (for numbers less than n) is applied only in a certain case of the induction over type formation, namely, when showing that $\text{VAR}_n T\phi$ if $\text{spine}_n T\phi$. This is also the only point in the proof at which is applied the fact that universes below U_n are left intact under variation up from n . If $\text{spine}_n T\phi$ then $\exists m < n. U_m - T$. Any variation T' of T (up from n) must evaluate to a variation of U_m . But U_m is the only variant of itself, thus, $\text{spine}_n T'\phi$. Since equality in U_m is $=_{\text{HAN}_m}$, the elimination of the inductive hypothesis on m establishes the second conjunct of $\text{VAR}_n T\phi$.

It follows from the monotonicity of $\mu(\sigma)$ in σ that $\cup_n \text{HAN}_n$ is as strong as HAN_ω . Thus, since T is $T - n$ for some n , to show that HAN_ω is $\cup_n \text{HAN}_n$ it is enough to show that

if $\text{HAN}_\omega T\phi$ & T is $T - n$ then $\text{HAN}_n T\phi$,

for which in turn it enough to show that $\text{CTyF}(\text{spine}_\omega; \text{BD}_n)$, where $\text{BD}_n T\phi$ means that $\text{HAN}_n T\phi$ if T is $T - n$. The only interesting aspect of the proof is a lemma,

if $\text{Fam}(\text{BD}_n; A; \alpha; x; B; \gamma)$ & A is $A - n$ & B is $B - n$
then $\text{Fam}(\text{HAN}_n; A; \alpha; x; B; \gamma)$.

Proof:

arb $n, A, \alpha, x, B, \gamma$ s.t. the antecedent holds.

$\text{HAN}_n A\alpha$ since $\text{BD}_n A\alpha$ & A is $A - n$.

arb a, a' s.t. $a\alpha a'$.

γ_a is $\gamma_{a'}$.

enough to show $\text{HAN}_n B[a/x] \gamma_a$ & $\text{HAN}_n B[a'/x] \gamma_{a'}$.

$a\alpha(a' - n)$ & $(a - n)\alpha a'$ since HAN_n is VAR_n .

γ_a is γ_{a-n} & $\gamma_{a'}$ is $\gamma_{a'-n}$ since $a\alpha(a - n)$ & $a'\alpha(a' - n)$.

enough to show $\text{HAN}_n B[a/x] \gamma_{a-n}$ & $\text{HAN}_n B[a'/x] \gamma_{a'-n}$.

$\text{HAN}_n B[a-n/x] \gamma_{a-n}$ & $\text{HAN}_n B[a'-n/x] \gamma_{a'-n}$ by the first assumption, since $(a - n)\alpha(a' - n)$.

QED since HAN_n is VAR_n .

⁶A helpful observation is that

if $\text{Fam}(\text{VAR}_n; A; \alpha; x; B; \gamma)$ & A' var_n A & B' var_n B then $\text{Fam}(\text{HAN}_n; A'; \alpha; x; B'; \gamma)$.

Finally, we shall see that $F_{n\text{HAN}_\omega}$ is $\text{HF}_{n\text{HAN}_n}$. As was indicated in the previous section, it is enough that, for all Δ , if $F_{n\text{HAN}_\omega} \Delta$ then $\exists n. F_{n\text{HAN}_n} \Delta$. And this reduces to showing that

$$F_{n\text{HAN}_n} \Delta \text{ if } F_{n\text{HAN}_\omega} \Delta - n,$$

(where $\Delta - n$ is the obvious analog of $t - n$) since for every Δ there is an n such that Δ is $\Delta - n$.

Proof by induction on the length of Δ :

We shall abbreviate HAN_n and HAN_ω by n and ω in subscripts to F_n , $=$, and \in .

arb n :

The base case is trivial.

arb Δ s.t. the inductive hypothesis holds for $|\Delta|$.

arb A, B, x s.t. $F_\omega (A, B : x \Delta) - n$.

$$(A - n) =_\omega (B - n).$$

$(A - n) =_n (B - n)$ since HAN_ω is a subrelation of BD_n .

$A =_n B$ since HAN_n is VAR_n .

arb a, b s.t. $a = b \in_n A$.

show $F_n \Delta[a; b/x]$.

enough by the inductive hypothesis

to show $F_\omega (\Delta - n)[a - n; b - n/x]$.

$a - n = b - n \in_n A - n$ since HAN_n is VAR_n .

$a - n = b - n \in_\omega A - n$.

QED .

4.3 Universe Polymorphism

Normally, we may expect to design reflective hierarchies whose constituent type systems are very similar. The principal similarity is that there will be certain type constructors under which every level of the hierarchy is closed. Let us call such type constructors *uniform*. Beyond this, there is the regularity of universe construction. Often our assertions relating several universes depend not on the particular universes mentioned, but rather, only on certain simple arithmetical relations between the indices of those universes. For example, we know that for any i and j ,

$$\text{HF}_{n\text{HAN}_n} U_i : X \ U_j : Y \ X + Y \in U_{\max(i,j)}.$$