How to prove that STLC is normalizing

Carlo Angiuli

May 18, 2015

Consider the simply typed lambda calculus with a base type **b** containing any number of uninterpreted constants c_i.

$$A := \mathbf{b} \mid A \to A$$
$$M := c_i \mid x \mid \lambda x{:}A.M \mid M\ M$$

$$\frac{}{\Gamma \vdash c_i : \mathbf{b}} \qquad \frac{}{\Gamma, x{:}A \vdash x : A} \qquad \frac{\Gamma, x{:}A \vdash M : B}{\Gamma \vdash \lambda x{:}A.M : A \to B} \qquad \frac{\Gamma \vdash M : A \to B \quad \Gamma \vdash N : A}{\Gamma \vdash M\ N : B}$$

$$\frac{}{(\lambda x{:}A.M)\ N \to_\beta [N/x]M} \qquad \frac{M \to_\beta M'}{\lambda x{:}A.M \to_\beta \lambda x{:}A.M'} \qquad \frac{M \to_\beta M'}{M\ N \to_\beta M'\ N} \qquad \frac{N \to_\beta N'}{M\ N \to_\beta M\ N'}$$

We say a term M is *normalizing*, $\mathsf{Norm}(M)$, if it has some terminating sequence of $\to_\beta$ reductions. In this note, we reinvent the proof that all well-typed terms are normalizing.

Acknowledgements

Thanks to Bob Harper, who explained this proof, and its reinvention, to me.

Conjecture 1.1. *If* $\Gamma \vdash M : C$*, then* $\mathsf{Norm}(M)$*.*

Proof. By induction on the derivation of $\Gamma \vdash M : C$.

- Case Const. Want to show $\mathsf{Norm}(c_i)$, which is true.
- Case Var. Want to show $\mathsf{Norm}(x)$, which is true.
- Case Lam. Want to show $\mathsf{Norm}(\lambda x{:}A.M)$, given $\mathsf{Norm}(M)$. We simply perform the same sequence of reductions under the binder.
- Case App. Want to show $\mathsf{Norm}(M\ N)$, given $\mathsf{Norm}(M)$ and $\mathsf{Norm}(N)$. ???

We certainly need to use the fact that M is well-typed in order to prove it is normalizing, because some terms in the *untyped* lambda calculus terms are not normalizing; for example, $(\lambda x.x\ x)\ (\lambda x.x\ x)$.

So the simplest thing which could possibly work is to induct on the typing derivation, and show that each typing rule yields a normalizing term. But this fails in the App case, because (a priori) both terms could be normalizing without the application being normalizing—this is the case for the untyped term above.

In the App case, we need the additional information that terms at function type take normalizing arguments to normalizing results. Therefore, we patch the theorem statement so that it demands normalization at base type, but "normalization-preservation" (or "hereditary normalization") at higher type. The resulting inductive hypothesis is called a *logical* predicate.

Definition 2.1.
$\mathsf{HN}_C(-)$ is a predicate on terms $\Gamma \vdash M : C$, defined by induction on C:

- $\mathsf{HN}_{\mathbf{b}}(M)$ iff $\mathsf{Norm}(M)$.
- $\mathsf{HN}_{A\to B}(M)$ iff for any N such that $\mathsf{HN}_A(N)$, $\mathsf{HN}_B(M\,N)$.

Conjecture 2.2. *If* $\Gamma \vdash M : C$, *then* $\mathsf{HN}_C(M)$.

Proof. By induction on the derivation of $\Gamma \vdash M : C$.

- Case Const. Want to show $\mathsf{Norm}(c_i)$, which is true.
- Case Var. Want to show $\mathsf{HN}_A(x)$. ???
- Case Lam. Want to show $\mathsf{HN}_{A\to B}(\lambda x{:}A.M)$, given $\mathsf{HN}_B(M)$. It suffices to show that, for any $\mathsf{HN}_A(N)$, $\mathsf{HN}_B((\lambda x{:}A.M)\,N)$. ???
- Case App. Want to show $\mathsf{HN}_B(M\,N)$, given $\mathsf{HN}_{A\to B}(M)$ and $\mathsf{HN}_A(N)$. This follows from the definition of $\mathsf{HN}_{A\to B}(M)$, for the given N.

This time, we attempt to prove that all well-typed terms are hereditarily normalizing. We defined HN in order to make the App case go through, but unfortunately we broke the Var and Lam cases.

For Var, we know that $\mathsf{HN}_{\mathbf{b}}(x)$, but we don't know anything about variables at function type. For Lam, we know $(\lambda x{:}A.M)\,N$ reduces to $[N/x]M$, but we don't know anything about substitution instances of M. (In fact, we also don't know that $\mathsf{HN}_B([N/x]M)$ implies $\mathsf{HN}_B((\lambda x{:}A.M)\,N)$.)

The root of both of these problems is that we don't know anything about the free variables in a term, either $\Gamma, x{:}A \vdash x : A$ or $\Gamma, x{:}A \vdash M : B$. For Lam, we need to know that substituting a HN term for x yields a HN term. The same is trivially true for Var: substituting a HN term for x would yield that same HN term.

Because our contexts are unordered, it doesn't make sense to ask this property just for the "last" variable in the context; we need to demand that any substitution of HN terms for Γ yields an HN term. So we change the inductive hypothesis yet again, and expect that the theorem will now go through.

Definition 3.1.
$\mathsf{HN}_C(-)$ is a predicate on terms $\cdot \vdash M : C$, defined by induction on C:

- $\mathsf{HN}_{\mathbf{b}}(M)$ iff $\mathsf{Norm}(M)$.
- $\mathsf{HN}_{A \to B}(M)$ iff for any N such that $\mathsf{HN}_A(N)$, $\mathsf{HN}_B(M\ N)$.

$\mathsf{HN}_\Gamma(-)$ is a predicate on substitutions $\gamma : \cdot \to \Gamma$, which holds when for each $x{:}A \in \Gamma$, $\mathsf{HN}_A(\widehat{\gamma}(x))$.

Conjecture 3.2. *If $\Gamma \vdash M : C$, then for any γ such that $\mathsf{HN}_\Gamma(\gamma)$, $\mathsf{HN}_C(\widehat{\gamma}(M))$.*

Proof. By induction on the derivation of $\Gamma \vdash M : C$.

- Case Const. Want to show $\mathsf{Norm}(\widehat{\gamma}(c_i))$. But $\widehat{\gamma}(c_i)$ is c_i, and $\mathsf{Norm}(c_i)$.
- Case Var. Want to show $\mathsf{HN}_A(\widehat{\gamma}(x))$. This is implied by the hypothesis $\mathsf{HN}_{\Gamma,x:A}(\gamma)$.
- Case Lam. Want to show $\mathsf{HN}_{A\to B}(\widehat{\gamma}(\lambda x{:}A.M))$. It suffices to show that, for any $\mathsf{HN}_A(N)$, $\mathsf{HN}_B(\widehat{\gamma}(\lambda x{:}A.M)\ N)$. By the inductive hypothesis, for any γ' such that $\mathsf{HN}_{\Gamma,x:A}(\gamma')$, $\mathsf{HN}_B(\widehat{\gamma'}(M))$. We choose γ' to extend γ by sending x to N. Clearly $\mathsf{HN}_{\Gamma,x:A}(\gamma')$, and $\widehat{\gamma'}(M) = [N/x]\widehat{\gamma}(M)$. (The result follows if $\mathsf{HN}_B([N/x]\widehat{\gamma}(M))$ implies $\mathsf{HN}_B((\lambda x{:}A.\widehat{\gamma}(M))\ N)$.)
- Case App. Want to show $\mathsf{HN}_B(\widehat{\gamma}(M)\ \widehat{\gamma}(N))$. The inductive hypotheses, instantiated at this same γ, give us $\mathsf{HN}_{A\to B}(\widehat{\gamma}(M))$ and $\mathsf{HN}_A(\widehat{\gamma}(N))$. The result follows from the definition of $\mathsf{HN}_{A\to B}(\widehat{\gamma}(M))$, for the given $\widehat{\gamma}(N)$.

$\gamma : \cdot \to \Gamma$ is a total substitution of closed terms for each variable in Γ (that is, a *closing substitution* for Γ). We say $\mathsf{HN}_\Gamma(\gamma)$ holds if γ is pointwise HN, and write $\widehat{\gamma}(M)$ for the application of the substitution γ to the term M.

We choose closing substitutions, rather than arbitrary ones, because they are the simplest choice which allows the Var and Lam cases to go through, and furthermore, they simplify the definition of HN by allowing us to consider only closed terms. Our theorem statement is now that all closed substitution instances, by a HN substitution, of a well-typed term are HN.

Now the proof actually works, again modulo a lemma in the Lam case that being HN is preserved when moving backwards through a β reduction. The Const and App cases are essentially the same as they were, but Var and Lam work because we strengthened the inductive hypothesis to require that all open terms map HN substitutions to HN closed terms. (Previously, we didn't actually make any demands of the free variables of an open term.) Because Lam is the only rule in which the context changes above the line, it is the only case in which we extend the substitution γ.

Unfortunately, this theorem doesn't imply that all well-typed terms are normalizing, even at base type! Consider $\Gamma \vdash x : \mathbf{b}$. We only know that if $\mathsf{HN}_\Gamma(\gamma)$ then $\mathsf{Norm}(\widehat{\gamma}(x))$, where $\widehat{\gamma}(x)$ is any *closed*, normalizing term of type $\mathbf{b}$. (At higher type, the theorem statement doesn't say anything whatsoever about normalization, so we'll eventually also need to know that $\mathsf{HN}_{A\to B}(M)$ implies $\mathsf{Norm}(M)$. But we will get to this later.)

In summary, we need to know something about substitution instances of open terms for the Lam case, but we also need the theorem statement to imply open terms are HN in order to prove normalization of open terms at $\mathbf{b}$. The solution is to consider not only closing substitutions, but *any* substitutions in another context Δ.

Definition 4.1.
$\mathsf{HN}^{\Delta}_{C}(-)$ is a predicate on terms $\Delta \vdash M : C$, defined by induction on C:

- $\mathsf{HN}^{\Delta}_{\mathbf{b}}(M)$ iff $\mathsf{Norm}(M)$.
- $\mathsf{HN}^{\Delta}_{A\to B}(M)$ iff for any N such that $\mathsf{HN}^{\Delta}_{A}(N)$, $\mathsf{HN}^{\Delta}_{B}(M\ N)$.

$\mathsf{HN}^{\Delta}_{\Gamma}(-)$ is a predicate on substitutions $\gamma : \Delta \to \Gamma$, which holds when for each $x{:}A \in \Gamma$, $\mathsf{HN}^{\Delta}_{A}(\widehat{\gamma}(x))$.

Conjecture 4.2. *If $\Gamma \vdash M : C$, then for any $\gamma : \Delta \to \Gamma$ such that $\mathsf{HN}^{\Delta}_{\Gamma}(\gamma)$, $\mathsf{HN}^{\Delta}_{C}(\widehat{\gamma}(M))$.*

Proof. By induction on the derivation of $\Gamma \vdash M : C$.

- Case Const. Want to show $\mathsf{Norm}(\widehat{\gamma}(c_i))$. But $\widehat{\gamma}(c_i)$ is c_i, and $\mathsf{Norm}(c_i)$.
- Case Var. Want to show $\mathsf{HN}^{\Delta}_{A}(\widehat{\gamma}(x))$. This is implied by the hypothesis $\mathsf{HN}^{\Delta}_{\Gamma,x:A}(\gamma)$.
- Case Lam. Want to show $\mathsf{HN}^{\Delta}_{A\to B}(\widehat{\gamma}(\lambda x{:}A.M))$. It suffices to show that, for any $\mathsf{HN}^{\Delta}_{A}(N)$, $\mathsf{HN}^{\Delta}_{B}(\widehat{\gamma}(\lambda x{:}A.M)\ N)$. By the inductive hypothesis, for any γ' such that $\mathsf{HN}^{\Delta}_{\Gamma,x:A}(\gamma')$, $\mathsf{HN}^{\Delta}_{B}(\widehat{\gamma'}(M))$. We choose γ' to extend γ by sending x to N. Clearly $\mathsf{HN}^{\Delta}_{\Gamma,x:A}(\gamma')$, and $\widehat{\gamma}'(M) = [N/x]\widehat{\gamma}(M)$. (The result follows if $\mathsf{HN}^{\Delta}_{B}([N/x]\widehat{\gamma}(M))$ implies $\mathsf{HN}^{\Delta}_{B}((\lambda x{:}A.\widehat{\gamma}(M))\ N)$.)
- Case App. Want to show $\mathsf{HN}^{\Delta}_{B}(\widehat{\gamma}(M)\ \widehat{\gamma}(N))$. The inductive hypotheses, instantiated at this same γ, give us $\mathsf{HN}^{\Delta}_{A\to B}(\widehat{\gamma}(M))$ and $\mathsf{HN}^{\Delta}_{A}(\widehat{\gamma}(N))$. The result follows from the definition of $\mathsf{HN}^{\Delta}_{A\to B}(\widehat{\gamma}(M))$, for the given $\widehat{\gamma}(N)$.

Now the theorem states that all substitution instances, by a HN substitution, of a well-typed term are HN. The proof goes through as before, so now we can turn our attention to the lemma that if the result of a β reduction is HN, then so is the redex.

Let's think about proving that lemma. It is obvious at base type, where HN is normalization. At $A \to B$, we will need to prove if $\mathsf{HN}^{\Delta}_{A\to B}([N/x]M)$ then $\mathsf{HN}^{\Delta}_{A\to B}((\lambda x{:}A.M)\ N)$. This assumption tells us that, for any N' with $\mathsf{HN}^{\Delta}_{A}(N')$, $\mathsf{HN}^{\Delta}_{B}(([N/x]M)\ N')$. We need to show $\mathsf{HN}^{\Delta}_{B}(((\lambda x{:}A.M)\ N)\ N')$.

Ideally we would apply the inductive hypothesis here, but the β reduction isn't occurring at the root. We can solve this by strengthening the theorem to say that HN is preserved by reverse *leftmost* β reductions—that is, either at the root or on the left side of an application (recursively). This notion is called *weak head reduction*, and is defined as follows:

$$\frac{}{(\lambda x{:}A.M)\ N \to_{\mathrm{wh}} [N/x]M} \qquad \frac{M \to_{\mathrm{wh}} M'}{M\ N \to_{\mathrm{wh}} M'\ N}$$

Definition 5.1.
$\mathsf{HN}^{\Delta}_{C}(-)$ is a predicate on terms $\Delta \vdash M : C$, defined by induction on C:

- $\mathsf{HN}^{\Delta}_{\mathbf{b}}(M)$ iff $\mathsf{Norm}(M)$.
- $\mathsf{HN}^{\Delta}_{A\to B}(M)$ iff for any N such that $\mathsf{HN}^{\Delta}_{A}(N)$, $\mathsf{HN}^{\Delta}_{B}(M\ N)$.

$\mathsf{HN}^{\Delta}_{\Gamma}(-)$ is a predicate on substitutions $\gamma : \Delta \to \Gamma$, which holds when for each $x{:}A \in \Gamma$, $\mathsf{HN}^{\Delta}_{A}(\widehat{\gamma}(x))$.

Lemma 5.2. *If $\Delta \vdash M : C$, $\Delta \vdash M' : C$, $M \to_{\mathrm{wh}} M'$, and $\mathsf{HN}^{\Delta}_{C}(M')$, then $\mathsf{HN}^{\Delta}_{C}(M)$.*

Proof. By induction on C.

- Case $\mathbf{b}$. Want to show $\mathsf{Norm}(M)$, given $\mathsf{Norm}(M')$ and $M \to_{\mathrm{wh}} M'$. It is easy to see that if $M \to_{\mathrm{wh}} M'$ then $M \to_{\beta} M'$, so we obtain a terminating sequence of reductions for M by prepending the one for $\mathsf{Norm}(M')$ by $M \to_{\beta} M'$.
- Case $A \to B$. Want to show $\mathsf{HN}^{\Delta}_{A\to B}(M)$. Show for any N such that $\mathsf{HN}^{\Delta}_{A}(N)$, $\mathsf{HN}^{\Delta}_{B}(M\ N)$. But $M\ N \to_{\mathrm{wh}} M'\ N$, so this follows by the induction hypothesis at B. ☐

Theorem 5.3. *If $\Gamma \vdash M : C$, then for any $\gamma : \Delta \to \Gamma$ such that $\mathsf{HN}^{\Delta}_{\Gamma}(\gamma)$, $\mathsf{HN}^{\Delta}_{C}(\widehat{\gamma}(M))$.*

The head expansion lemma goes through without any difficulty, using the notion of weak head reduction defined on the previous page. (We explicitly require that M and M' have the same type only to avoid proving a preservation lemma.) Then the main theorem is proven exactly as before, with an appeal to this lemma in the necessary place.

Since our goal is actually to prove that all terms are normalizing, we need to show that HN terms are normalizing. Can we prove this? At base type, this is the definition of HN. At $A \to B$, we only know that for any N, if $\mathsf{HN}^{\Delta}_{A}(N)$ then $\mathsf{Norm}(M\ N)$. Can we choose N in such a way that $\mathsf{Norm}(M\ N)$ implies $\mathsf{Norm}(M)$? Since we need such an N at every type A, our only real option is a variable x. Luckily, as we will check shortly, $\mathsf{Norm}(M\ x)$ does imply $\mathsf{Norm}(M)$. (One might wonder why we didn't avoid this entire difficulty by defining $\mathsf{HN}^{\Delta}_{A\to B}(M)$ to mean that M normalizes to a lambda term. Since M is an open term, this isn't actually true; it could be a variable, for example.)

There may not already be a variable $x{:}A \in \Delta$. We can weaken Δ and use $\mathsf{HN}^{\Delta,x:A}_{A}(x)$, but then the context doesn't match $\mathsf{HN}^{\Delta}_{A\to B}(M)$, so we can't conclude that $M\ x$ is HN (at which context?).

We can solve this by changing the definition of HN one final time: $\mathsf{HN}^{\Delta}_{A\to B}(M)$ iff for any weakening Δ' of Δ, and any N such that $\mathsf{HN}^{\Delta'}_{A}(N)$, $\mathsf{HN}^{\Delta'}_{B}(M\ N)$. For the App case to work, we also need to change the theorem to state that for $\gamma : \Delta \to \Gamma$, $\widehat{\gamma}(M)$ is HN in any weakening of Δ. We will write $\Delta' \geq \Delta$ when Δ' is obtained by zero or more weakenings of Δ. $\geq$, or *context extension*, is a preorder.

To summarize, our plan is to:

1. Change the definition of HN and redo the proofs;
2. Prove that $\mathsf{Norm}(M\ x)$ implies $\mathsf{Norm}(M)$; and
3. Prove that $\mathsf{HN}^{\Delta}_{C}(M)$ implies $\mathsf{Norm}(M)$.

Definition 6.1.
$\mathsf{HN}_C^\Delta(-)$ is a predicate on terms $\Delta \vdash M : C$, defined by induction on C:

- $\mathsf{HN}_{\mathbf{b}}^\Delta(M)$ iff $\mathsf{Norm}(M)$.
- $\mathsf{HN}_{A\to B}^\Delta(M)$ iff for any $\Delta' \geq \Delta$ and any N such that $\mathsf{HN}_A^{\Delta'}(N)$, $\mathsf{HN}_B^{\Delta'}(M\ N)$.

$\mathsf{HN}_\Gamma^\Delta(-)$ is a predicate on substitutions $\gamma : \Delta \to \Gamma$, which holds when for each $x{:}A \in \Gamma$, $\mathsf{HN}_A^\Delta(\widehat{\gamma}(x))$.

Lemma 6.2. *If $\Delta \vdash M : C$, $\Delta \vdash M' : C$, $M \to_{\mathrm{wh}} M'$, and $\mathsf{HN}_C^\Delta(M')$, then $\mathsf{HN}_C^\Delta(M)$.*

Proof. By induction on C.

- Case $\mathbf{b}$. Want to show $\mathsf{Norm}(M)$, given $\mathsf{Norm}(M')$ and $M \to_{\mathrm{wh}} M'$. It is easy to see that if $M \to_{\mathrm{wh}} M'$ then $M \to_\beta M'$, so we obtain a terminating sequence of reductions for M by prepending the one for $\mathsf{Norm}(M')$ by $M \to_\beta M'$.
- Case $A \to B$. Want to show $\mathsf{HN}_{A\to B}^\Delta(M)$. Show for any $\Delta' \geq \Delta$ and N such that $\mathsf{HN}_A^{\Delta'}(N)$, $\mathsf{HN}_B^{\Delta'}(M\ N)$. But $M\ N \to_{\mathrm{wh}} M'\ N$, so this follows by the induction hypothesis at B. □

Theorem 6.3. *If $\Gamma \vdash M : C$, then for any $\gamma : \Delta \to \Gamma$ and $\Delta' \geq \Delta$ such that $\mathsf{HN}_\Gamma^{\Delta'}(\gamma)$, $\mathsf{HN}_C^{\Delta'}(\widehat{\gamma}(M))$.*

Proof. By induction on the derivation of $\Gamma \vdash M : C$.

- Case Const. Want to show $\mathsf{HN}_{\mathbf{b}}^{\Delta'}(\widehat{\gamma}(c_i))$. But $\widehat{\gamma}(c_i)$ is c_i, and $\mathsf{Norm}(c_i)$.
- Case Var. Want to show $\mathsf{HN}_A^{\Delta'}(\widehat{\gamma}(x))$. This is implied by the hypothesis $\mathsf{HN}_{\Gamma,x:A}^{\Delta'}(\gamma)$, since $\Delta' \geq \Delta'$.
- Case Lam. Want to show $\mathsf{HN}_{A\to B}^{\Delta'}(\widehat{\gamma}(\lambda x{:}A.M))$. Show for any $\Delta'' \geq \Delta'$ and N such that $\mathsf{HN}_A^{\Delta''}(N)$, $\mathsf{HN}_B^{\Delta''}(\widehat{\gamma}(\lambda x{:}A.M)\ N)$. By head expansion, it suffices to show $\mathsf{HN}_B^{\Delta''}([N/x]\widehat{\gamma}(M))$. Let us denote γ extended by $x \mapsto N$ by γ'.

  Notice that if $\mathsf{HN}_\Gamma^{\Delta'}(\gamma)$ then $\mathsf{HN}_\Gamma^{\Delta''}(\gamma)$, because at all types C, $\mathsf{HN}_C^{\Delta'}(-)$ implies $\mathsf{HN}_C^{\Delta''}(-)$. Therefore, we can instantiate the inductive hypothesis at Δ'' (by transitivity of context extension) and γ', implying what we wanted to show: $\mathsf{HN}_B^{\Delta''}(\widehat{\gamma'}(M))$.
- Case App. Want to show $\mathsf{HN}_B^{\Delta'}(\widehat{\gamma}(M\ N))$. The inductive hypothesis for N (at $\Delta' \geq \Delta$) says that $\mathsf{HN}_A^{\Delta'}(\widehat{\gamma}(N))$. The inductive hypothesis for M says that $\mathsf{HN}_{A\to B}^{\Delta'}(\widehat{\gamma}(M))$, which implies $\mathsf{HN}_B^{\Delta'}(\widehat{\gamma}(M)\,N')$ for any $\mathsf{HN}_A^{\Delta'}(N')$. Choosing $N' = \widehat{\gamma}(N)$ finishes this case. □

Now HN demands that terms of type $A \to B$ send HN arguments from any larger context, to HN results in that context. As discussed on the previous page, this allows us to always apply such terms to a variable in order to prove normalization.

The head expansion lemma is essentially unchanged from last time. The proof of the main theorem has some additional subtleties, but the core ideas are the same. Notice that we use both the reflexivity and transitivity of $\geq$, in the Var and Lam cases respectively.

Now let's prove $\mathsf{Norm}(M\ x)$ implies $\mathsf{Norm}(M)$, and $\mathsf{HN}_C^\Delta(M)$ implies $\mathsf{Norm}(M)$.

Definition 7.1.
$\mathsf{HN}^{\Delta}_{C}(-)$ is a predicate on terms $\Delta \vdash M : C$, defined by induction on C:

- $\mathsf{HN}^{\Delta}_{\mathbf{b}}(M)$ iff $\mathsf{Norm}(M)$.
- $\mathsf{HN}^{\Delta}_{A\to B}(M)$ iff for any $\Delta' \geq \Delta$ and any N such that $\mathsf{HN}^{\Delta'}_{A}(N)$, $\mathsf{HN}^{\Delta'}_{B}(M\ N)$.

$\mathsf{HN}^{\Delta}_{\Gamma}(-)$ is a predicate on substitutions $\gamma : \Delta \to \Gamma$, which holds when for each $x{:}A \in \Gamma$, $\mathsf{HN}^{\Delta}_{A}(\widehat{\gamma}(x))$.

Lemma 7.2. *If $\Delta \vdash M : C$, $\Delta \vdash M' : C$, $M \to_{\mathrm{wh}} M'$, and $\mathsf{HN}^{\Delta}_{C}(M')$, then $\mathsf{HN}^{\Delta}_{C}(M)$.*

Theorem 7.3. *If $\Gamma \vdash M : C$, then for any $\gamma : \Delta \to \Gamma$ and $\Delta' \geq \Delta$ such that $\mathsf{HN}^{\Delta'}_{\Gamma}(\gamma)$, $\mathsf{HN}^{\Delta'}_{C}(\widehat{\gamma}(M))$.*

Lemma 7.4. *If $\mathsf{Norm}(M\ x)$ then $\mathsf{Norm}(M)$.*

Proof. By induction on the derivation of $\mathsf{Norm}(M\ x)$.

- Case $(\lambda x{:}A.N)\ x \to_\beta N$. Want to show $\mathsf{Norm}(\lambda x{:}A.N)$, given $\mathsf{Norm}(N)$. Perform the same sequence of reductions under the binder.
- Case $\lambda x{:}A.N \to_\beta \lambda x{:}A.N'$. Cannot apply.
- Case $M\ x \to_\beta M'\ x$. Want to show $\mathsf{Norm}(M)$, given $M \to_\beta M'$ and $\mathsf{Norm}(M'\ x)$. By the induction hypothesis, $\mathsf{Norm}(M')$. But then $\mathsf{Norm}(M)$ by prepending this sequence with $M \to_\beta M'$.
- Case $M\ x \to_\beta M\ N'$. Cannot apply, since x does not reduce.
- Case Irreducible. Then M must also be irreducible, so $\mathsf{Norm}(M)$. ☐

Conjecture 7.5. *If $\mathsf{HN}^{\Delta}_{C}(M)$ then $\mathsf{Norm}(M)$.*

Proof. By induction on C.

- Case $\mathbf{b}$. Want to show $\mathsf{Norm}(M)$, which is the definition of $\mathsf{HN}^{\Delta}_{\mathbf{b}}(M)$.
- Case $A \to B$. Want to show $\mathsf{Norm}(M)$, given $\mathsf{HN}^{\Delta}_{A\to B}(M)$. For any N and $\Delta' \geq \Delta$ such that $\mathsf{HN}^{\Delta'}_{A}(N)$, $\mathsf{HN}^{\Delta'}_{B}(M\ N)$. Choose $N = x$ and $\Delta' = \Delta, x{:}A$. (But we don't know $\mathsf{HN}^{\Delta,x:A}_{A}(x)$! If we did, then by the inductive hypothesis at B, $\mathsf{Norm}(M\ x)$, which by the previous lemma implies $\mathsf{Norm}(M)$.)

In order to prove that $\mathsf{Norm}(M\ x)$ implies $\mathsf{Norm}(M)$, we have to carefully define $\mathsf{Norm}(-)$. We say that $\mathsf{Norm}(M)$ if either $M \to_\beta M'$ and $\mathsf{Norm}(M')$, or M is irreducible. The lemma then follows from a straightforward induction on the derivation of $\mathsf{Norm}(M\ x)$.

In the proof of $\mathsf{HN}^{\Delta}_{C}(M)$ implies $\mathsf{Norm}(M)$, we hit a subtle, unforeseen roadblock: we don't know that $\mathsf{HN}^{\Delta,x:A}_{A}(x)$! The main theorem only tells us that any HN substitution instance of x is HN. We can try applying the identity substitution $\mathsf{id}_{\Delta,x:A}$, but then we need to show that $\mathsf{HN}^{\Delta,x:A}_{\Delta,x:A}(\mathsf{id}_{\Delta,x:A})$; in particular, we need $\mathsf{HN}^{\Delta,x:A}_{A}(x)$, which is what we were trying to show.

Instead, we will directly prove that variables are HN. (One consequence of this is that identity substitutions are HN, which means for any $\Delta \vdash M : C$, $\mathsf{HN}^{\Delta}_{C}(M)$.) As usual, this proof would immediately fail at higher type, since we will only know that $\mathsf{HN}^{\Delta}_{B}(x\ N)$ for $\mathsf{HN}^{\Delta}_{A}(N)$. Instead, we prove that any k-fold application of x to normalizing terms is HN. (Why to *normalizing* terms instead of HN terms? Otherwise the proof does not even go through at base type!)

If we also had product types in this language, then $\mathsf{HN}^{\Delta}_{A\times B}(M)$ would mean that both projections of M are HN, and we would instead need to prove that any any sequence of function applications *and projections* of x is HN. Such a sequence is called an *evaluation context.*

Definition 8.1.
$\mathsf{HN}_C^{\Delta}(-)$ is a predicate on terms $\Delta \vdash M : C$, defined by induction on C:

- $\mathsf{HN}_{\mathbf{b}}^{\Delta}(M)$ iff $\mathsf{Norm}(M)$.
- $\mathsf{HN}_{A\to B}^{\Delta}(M)$ iff for any $\Delta' \geq \Delta$ and any N such that $\mathsf{HN}_A^{\Delta'}(N)$, $\mathsf{HN}_B^{\Delta'}(M\ N)$.

$\mathsf{HN}_\Gamma^{\Delta}(-)$ is a predicate on substitutions $\gamma : \Delta \to \Gamma$, which holds when for each $x{:}A \in \Gamma$, $\mathsf{HN}_A^{\Delta}(\widehat{\gamma}(x))$.

Lemma 8.2. *If* $\Delta \vdash M : C$, $\Delta \vdash M' : C$, $M \to_{\text{wh}} M'$, *and* $\mathsf{HN}_C^{\Delta}(M')$, *then* $\mathsf{HN}_C^{\Delta}(M)$.

Theorem 8.3. *If* $\Gamma \vdash M : C$, *then for any* $\gamma : \Delta \to \Gamma$ *and* $\Delta' \geq \Delta$ *such that* $\mathsf{HN}_\Gamma^{\Delta'}(\gamma)$, $\mathsf{HN}_C^{\Delta'}(\widehat{\gamma}(M))$.

Lemma 8.4. *If* $\mathsf{Norm}(M\ x)$ *then* $\mathsf{Norm}(M)$.

Conjecture 8.5. *If for* $0 \leq i < k$, $\Delta \vdash N_i : A_i$ *and* $\mathsf{Norm}(N_i)$, *then for any* $x{:}A_0 \to \cdots \to A_{k-1} \to C \in \Delta$, $\mathsf{HN}_C^{\Delta}(x\ N_0\ \ldots\ N_{k-1})$.

Proof. By induction on C.

- Case $\mathbf{b}$. Want to show $\mathsf{Norm}(x\ N_0\ \ldots\ N_{k-1})$, which is true.
- Case $A \to B$. Want to show $\mathsf{HN}_{A\to B}^{\Delta}(x\ N_0\ \ldots\ N_{k-1})$. Show for any $\Delta' \geq \Delta$ and $\mathsf{HN}_A^{\Delta'}(N)$ that $\mathsf{HN}_B^{\Delta'}(x\ N_0\ \ldots\ N_{k-1}\ N)$. (But $\mathsf{HN}_A^{\Delta'}(N)$ isn't helpful! If we instead knew $\mathsf{Norm}(N)$, then we could apply the inductive hypothesis to $N_0, \ldots, N_{k-1}, N$.)

Conjecture 8.6. *If* $\mathsf{HN}_C^{\Delta}(M)$ *then* $\mathsf{Norm}(M)$.

Proof. By induction on C.

- Case $\mathbf{b}$. Want to show $\mathsf{Norm}(M)$, which is the definition of $\mathsf{HN}_{\mathbf{b}}^{\Delta}(M)$.
- Case $A \to B$. Want to show $\mathsf{Norm}(M)$, given $\mathsf{HN}_{A\to B}^{\Delta}(M)$. For any N and $\Delta' \geq \Delta$ such that $\mathsf{HN}_A^{\Delta'}(N)$, $\mathsf{HN}_B^{\Delta'}(M\ N)$. Choose $N = x$ and $\Delta' = \Delta, x{:}A$. (But we don't know $\mathsf{HN}_A^{\Delta,x:A}(x)$! If we did, then by the inductive hypothesis at B, $\mathsf{Norm}(M\ x)$, which by the previous lemma implies $\mathsf{Norm}(M)$.)

The theorem statement looks complicated, but it just says that any k-fold application of a variable to normalizing terms is HN. At base type, we can normalize the application by normalizing each argument in turn. At higher type, we only know that applying the k-fold application to a HN argument yields a HN result. If we knew that argument was normalizing, then the inductive hypothesis would say that the resulting $(k+1)$-fold application of x is HN.

We need to know that HN terms are normalizing in order to prove variables are HN, and we need to know that variables are HN in order to prove that HN terms are normalizing. The solution is to prove both of these facts simultaneously, so we have *both* facts at smaller types as our inductive hypotheses.

Definition 9.1.
$\mathsf{HN}_C^\Delta(-)$ is a predicate on terms $\Delta \vdash M : C$, defined by induction on C:

- $\mathsf{HN}_{\mathbf{b}}^\Delta(M)$ iff $\mathsf{Norm}(M)$.
- $\mathsf{HN}_{A\to B}^\Delta(M)$ iff for any $\Delta' \geq \Delta$ and any N such that $\mathsf{HN}_A^{\Delta'}(N)$, $\mathsf{HN}_B^{\Delta'}(M\ N)$.

$\mathsf{HN}_\Gamma^\Delta(-)$ is a predicate on substitutions $\gamma : \Delta \to \Gamma$, which holds when for each $x{:}A \in \Gamma$, $\mathsf{HN}_A^\Delta(\widehat{\gamma}(x))$.

Lemma 9.2. *If $\Delta \vdash M : C$, $\Delta \vdash M' : C$, $M \to_{\text{wh}} M'$, and $\mathsf{HN}_C^\Delta(M')$, then $\mathsf{HN}_C^\Delta(M)$.*

Theorem 9.3. *If $\Gamma \vdash M : C$, then for any $\gamma : \Delta \to \Gamma$ and $\Delta' \geq \Delta$ such that $\mathsf{HN}_\Gamma^{\Delta'}(\gamma)$, $\mathsf{HN}_C^{\Delta'}(\widehat{\gamma}(M))$.*

Lemma 9.4. *If $\mathsf{Norm}(M\ x)$ then $\mathsf{Norm}(M)$.*

Theorem 9.5.

1. *If for $0 \leq i < k$, $\Delta \vdash N_i : A_i$ and $\mathsf{Norm}(N_i)$, then for any $x{:}A_0 \to \cdots \to A_{k-1} \to C \in \Delta$, $\mathsf{HN}_C^\Delta(x\ N_0\ \ldots\ N_{k-1})$.*
2. *If $\mathsf{HN}_C^\Delta(M)$ then $\mathsf{Norm}(M)$.*

Proof. By induction on C.

- Case $\mathbf{b}$.
  1. Want to show $\mathsf{Norm}(x\, N_0\ \ldots\ N_{k-1})$. Reduce each N_i in turn; afterwards, no other reduction rule applies, so this reduction sequence terminates.
  2. Want to show $\mathsf{Norm}(M)$, which is the definition of $\mathsf{HN}_{\mathbf{b}}^\Delta(M)$.
- Case $A \to B$.
  1. Want to show $\mathsf{HN}_{A\to B}^\Delta(x\ N_0\ \ldots\ N_{k-1})$. Show for any $\Delta' \geq \Delta$ and $\mathsf{HN}_A^{\Delta'}(N)$ that $\mathsf{HN}_B^{\Delta'}(x\ N_0\ \ldots\ N_{k-1}\ N)$. By the second inductive hypothesis on N, $\mathsf{Norm}(N)$. Then apply the first inductive hypothesis to $N_0, \ldots, N_{k-1}, N$.
  2. Want to show $\mathsf{Norm}(M)$, given $\mathsf{HN}_{A\to B}^\Delta(M)$. By the first inductive hypothesis with $k = 0$, $\mathsf{HN}_A^{\Delta,x:A}(x)$. Then, because $\Delta, x{:}A \geq \Delta$, $\mathsf{HN}_B^{\Delta,x:A}(M\ x)$. By the second inductive hypothesis, $\mathsf{Norm}(M\ x)$, which by the previous lemma implies $\mathsf{Norm}(M)$. □

Corollary 9.6. *If $\Gamma \vdash M : A$, then $\mathsf{Norm}(M)$.*

Proof. The fundamental theorem, specialized to id_Γ (which is HN because $\mathsf{HN}_A^\Gamma(x)$ for each $x{:}A \in \Gamma$) and $\Gamma \geq \Gamma$, says that $\mathsf{HN}_A^\Gamma(M)$. It follows from the previous theorem that $\mathsf{Norm}(M)$. □

At last, the proof is complete.

Full proof

In the proof that follows, we will need several auxiliary definitions. *Weak head reduction*, $\to_{\text{wh}}$, is a restriction of $\to_\beta$ to only leftmost β redexes.

$$\frac{}{(\lambda x{:}A.M)\ N \to_{\text{wh}} [N/x]M} \qquad \frac{M \to_{\text{wh}} M'}{M\ N \to_{\text{wh}} M'\ N}$$

We say that a context Δ' is an *extension* of a context Δ, written $\Delta' \geq \Delta$, if it can be obtained from Δ by a series of weakenings; that is, if Δ is a subset of Δ' when regarded as sets.

Definition 10.1 (Hereditary normalization)**.**
$\mathsf{HN}_A^\Delta(-)$ is a predicate on terms $\Delta \vdash M : A$, defined by induction on A:

- $\mathsf{HN}_{\mathbf{b}}^\Delta(M)$ iff $\mathsf{Norm}(M)$.
- $\mathsf{HN}_{A\to B}^\Delta(M)$ iff for any $\Delta' \geq \Delta$ and any N such that $\mathsf{HN}_A^{\Delta'}(N)$, $\mathsf{HN}_B^{\Delta'}(M\ N)$.

$\mathsf{HN}_\Gamma^\Delta(-)$ is a predicate on substitutions $\gamma : \Delta \to \Gamma$, which holds when for each $x{:}A \in \Gamma$, $\mathsf{HN}_A^\Delta(\widehat{\gamma}(x))$.

Lemma 10.2 (Head expansion)**.** *If $\Delta \vdash M : C$, $\Delta \vdash M' : C$, $M \to_{\text{wh}} M'$, and $\mathsf{HN}_C^\Delta(M')$, then $\mathsf{HN}_C^\Delta(M)$.*

Proof. By induction on C.

- Case $\mathbf{b}$. Want to show $\mathsf{Norm}(M)$, given $\mathsf{Norm}(M')$ and $M \to_{\text{wh}} M'$. It is easy to see that if $M \to_{\text{wh}} M'$ then $M \to_\beta M'$, so we obtain a terminating sequence of reductions for M by prepending the one for $\mathsf{Norm}(M')$ by $M \to_\beta M'$.
- Case $A \to B$. Want to show $\mathsf{HN}_{A\to B}^\Delta(M)$. Show for any $\Delta' \geq \Delta$ and N such that $\mathsf{HN}_A^{\Delta'}(N)$, $\mathsf{HN}_B^{\Delta'}(M\ N)$. But $M\ N \to_{\text{wh}} M'\ N$, so this follows by the induction hypothesis at B. □

Theorem 10.3 (Fundamental theorem)**.** *If $\Gamma \vdash M : C$, then for any substitution $\gamma : \Delta \to \Gamma$ and $\Delta' \geq \Delta$ such that $\mathsf{HN}_\Gamma^{\Delta'}(\gamma)$, $\mathsf{HN}_C^{\Delta'}(\widehat{\gamma}(M))$.*

Proof. By induction on the derivation of $\Gamma \vdash M : C$.

- Case Const. Want to show $\mathsf{HN}_{\mathbf{b}}^{\Delta'}(\widehat{\gamma}(c_i))$. But $\widehat{\gamma}(c_i)$ is c_i, and $\mathsf{Norm}(c_i)$.
- Case Var. Want to show $\mathsf{HN}_A^{\Delta'}(\widehat{\gamma}(x))$. We know $\mathsf{HN}_{\Gamma,x:A}^{\Delta'}(\gamma)$, which by definition implies $\mathsf{HN}_A^{\Delta'}(\widehat{\gamma}(x))$ (since $\Delta' \geq \Delta'$).
- Case Lam. Want to show $\mathsf{HN}_{A\to B}^{\Delta'}(\widehat{\gamma}(\lambda x{:}A.M))$. Show for any $\Delta'' \geq \Delta'$ and N such that $\mathsf{HN}_A^{\Delta''}(N)$, $\mathsf{HN}_B^{\Delta''}(\widehat{\gamma}(\lambda x{:}A.M)\ N)$. By head expansion, it suffices to show $\mathsf{HN}_B^{\Delta''}([N/x]\widehat{\gamma}(M))$. Let us denote γ extended by $x \mapsto N$ by γ'.

  Notice that if $\mathsf{HN}_\Gamma^{\Delta'}(\gamma)$ then $\mathsf{HN}_\Gamma^{\Delta''}(\gamma)$, because at all types C, $\mathsf{HN}_C^{\Delta'}(-)$ implies $\mathsf{HN}_C^{\Delta''}(-)$. Therefore, we can instantiate the inductive hypothesis at Δ'' (by transitivity of context extension) and γ', implying what we wanted to show: $\mathsf{HN}_B^{\Delta''}(\widehat{\gamma'}(M))$.
- Case App. Want to show $\mathsf{HN}_B^{\Delta'}(\widehat{\gamma}(M\ N))$. The inductive hypothesis for N (at $\Delta' \geq \Delta$) says that $\mathsf{HN}_A^{\Delta'}(\widehat{\gamma}(N))$. The inductive hypothesis for M says that $\mathsf{HN}_{A\to B}^{\Delta'}(\widehat{\gamma}(M))$, which implies $\mathsf{HN}_B^{\Delta'}(\widehat{\gamma}(M)\,N')$ for any $\mathsf{HN}_A^{\Delta'}(N')$. Choosing $N' = \widehat{\gamma}(N)$ finishes this case. □

Lemma 10.4. *If $\mathsf{Norm}(M\ x)$ then $\mathsf{Norm}(M)$.*

Proof. By induction on the derivation of $\mathsf{Norm}(M\ x)$.

- Case $(\lambda x{:}A.N)\; x \to_\beta N$. Want to show $\mathsf{Norm}(\lambda x{:}A.N)$, given $\mathsf{Norm}(N)$. Perform the same sequence of reductions under the binder.
- Case $\lambda x{:}A.N \to_\beta \lambda x{:}A.N'$. Cannot apply.
- Case $M\; x \to_\beta M'\; x$. Want to show $\mathsf{Norm}(M)$, given $M \to_\beta M'$ and $\mathsf{Norm}(M'\; x)$. By the induction hypothesis, $\mathsf{Norm}(M')$. But then $\mathsf{Norm}(M)$ by prepending this sequence with $M \to_\beta M'$.
- Case $M\; x \to_\beta M\; N'$. Cannot apply, since x does not reduce.
- Case Irreducible. Then M must also be irreducible, so $\mathsf{Norm}(M)$. □

Theorem 10.5.

1. *If for* $0 \leq i < k$, $\Delta \vdash N_i : A_i$ *and* $\mathsf{Norm}(N_i)$*, then for any* $x{:}A_0 \to \cdots \to A_{k-1} \to C \in \Delta$, $\mathsf{HN}^{\Delta}_{C}(x\, N_0\; \ldots\; N_{k-1})$.
2. *If* $\mathsf{HN}^{\Delta}_{C}(M)$ *then* $\mathsf{Norm}(M)$.

Proof. By induction on C.

- Case $\mathbf{b}$.
  1. Want to show $\mathsf{Norm}(x\, N_0\; \ldots\; N_{k-1})$. Reduce each N_i in turn; afterwards, no other reduction rule applies, so this reduction sequence terminates.
  2. Want to show $\mathsf{Norm}(M)$, which is the definition of $\mathsf{HN}^{\Delta}_{\mathbf{b}}(M)$.
- Case $A \to B$.
  1. Want to show $\mathsf{HN}^{\Delta}_{A\to B}(xN_0 \ldots N_{k-1})$. Show for any $\Delta' \geq \Delta$ and $\mathsf{HN}^{\Delta'}_{A}(N)$ that $\mathsf{HN}^{\Delta'}_{B}(xN_0 \ldots N_{k-1}N)$. By the second inductive hypothesis on N, $\mathsf{Norm}(N)$. Then apply the first inductive hypothesis to $N_0, \ldots, N_{k-1}, N$.
  2. Want to show $\mathsf{Norm}(M)$, given $\mathsf{HN}^{\Delta}_{A\to B}(M)$. By the first inductive hypothesis with $k = 0$, $\mathsf{HN}^{\Delta,x:A}_{A}(x)$. Then, because $\Delta, x{:}A \geq \Delta$, $\mathsf{HN}^{\Delta,x:A}_{B}(M\; x)$. By the second inductive hypothesis, $\mathsf{Norm}(M\; x)$, which by the previous lemma implies $\mathsf{Norm}(M)$. □

One consequence is that for any $x{:}A \in \Delta$, $\mathsf{HN}^{\Delta}_{A}(x)$. In particular, $\mathsf{HN}^{\Gamma}_{\Gamma}(\mathsf{id}_\Gamma)$.

Corollary 10.6 (Normalization)**.** *If* $\Gamma \vdash M : A$*, then* $\mathsf{Norm}(M)$.

Proof. The fundamental theorem, specialized to id_Γ and $\Gamma \geq \Gamma$, says that $\mathsf{HN}^{\Gamma}_{A}(M)$. It follows from the previous lemma that $\mathsf{Norm}(M)$. □