A SYNTACTIC CHARACTERIZATION OF THE EQUALITY IN SOME MODELS FOR THE LAMBDA CALCULUS

MARTIN HYLAND

0. *Introduction*

An equality relation on the terms of the λ-calculus is an equivalence relation closed under the (syntactical) operations of application and λ-abstraction. We may distinguish between syntactic and semantic ways of introducing equality relations. β-equality is introduced syntactically; it is the least equality relation satisfying the equations for α- and β-conversion. For a more subtle way of introducing equality relations syntactically, consider the relations $=_f$ and $=_h$ of §5 of this paper. To give a semantic characterization of an equality relation, we simply take the relation ' has the same value in D', where D is some model for the λ-calculus. Of course, no equality relation is of interest to the intended interpretation of the λ-calculus, unless it extends β-equality.

An equality relation is inconsistent if and only if it sets all terms equal; otherwise it is consistent. It is maximal consistent if and only if it is consistent and has no consistent proper extensions.

In this paper we consider a class of continuous lattice models for the λ-calculus, and a particular model, the Graph model. The same equality is induced by all the continuous lattice models; we shall refer to them as the Scott models (see [3], where they were first constructed). For the history of the Graph model see [4]. We shall give, in this paper, syntactic characterizations of the equality induced by the Scott models, and by the Graph model; and we shall show that the equality induced by the Scott models is the unique maximal consistent equality relation, extending the relation $=_H$, which was proved consistent in [1].

We use x, y, z, w ... for variables, and M, N, P ... for terms of the λ-calculus (with subscripts as necessary). D will refer to whatever model or models are under consideration.

The content of our Theorem 5.4 (a) has been discovered independently by C. P. Wadsworth.

1. *Preliminaries*

First, we collect together those facts that the reader needs to know about the Scott models. D is a complete lattice; we write $\leqslant$ for the partial order, and $\bigcup$ for the least upper bound. For each natural number n, there is $D_n \subset D$, such that

(1) $D_n \subset D_{n+1}$,

(2) D is the completion of the union of the D_n's,

(3) for each $n \in \omega$, $d \in D$, there is a maximal $c \leqslant d$, with $c \in D_n$; we write this c, as $(d)_n$,

Received 1 October, 1973; revised 1 March, 1974.

(4) (by the above) $d = \bigcup\{(d)_n | n \in \omega\}$,

(5) each D_{n+1} can be regarded as the set of continuous maps from D_n to D_n; so there is defined a continuous application: $D_{n+1} \times D_n \to D_n$,

(6) D can be identified with the set of continuous maps from D to D, where $d \in D$ acts as a map by taking $c \in D$ to $d(c) = \bigcup\{d_{n+1}(c_n) | n \in \omega\}$,

(7) the following equations of Scott are satisfied:

$$d_{n+1}(c) = d_{n+1}(c_n) = (d(c_n))_n,$$
$$d_0(c) = d_0(\bot) = (d(\bot))_0.$$

Here $\bot$ denotes the bottom element of the lattice D.

Details of the construction are in [3]. There a lattice isomorphic to its function space is determined by (a) the choice of D_0, and (b) the maps $\phi : D_0 \to D_1$ and $\psi : D_1 \to D_0$. For our purpose D_0 may be chosen arbitrarily, but to ensure (7) above we take $\phi(d_0) = \lambda x.\ d_0$, $\psi(d_1) = d_1(\bot)$. This determines our class of continuous lattice models, the Scott models of this paper.

The natural interpretation of a term M in a Scott model D, (called its value, and written $[\![M]\!]$), is implicitly a value $[\![M]\!]_\rho$, depending on a valuation ρ of the free variables. It is defined by the induction.

$$[\![x]\!]_\rho = \rho(x),$$
$$[\![MN]\!]_\rho = [\![M]\!]_\rho([\![N]\!]_\rho),$$
$$[\![\lambda x . P]\!]_\rho = \lambda d . [\![P]\!]_{\rho[d/x]},$$

where $\rho[d/x]$ denotes the valuation obtained from ρ by valuing x as d; the expression λd ... (is not a λ-term, but) represents a continuous map from D to D and so by (6) above can be taken as an element of D.

Now we describe the Graph model. D consists of all subsets of ω. Let us take a pairing function, defined by $(k, m) = \frac{1}{2}(k+m)(k+m+1)+m$; and if $k = 2^{r}+\ldots+2^{r}$, with $r_0 < \ldots < r_n$, let e_k be the finite set $\{r_0, \ldots, r_k\}$. On D, we define application and λ-abstraction as follows;

$$c(d) = \{m | (\exists\, k)\,(e_k \subset d \,\&\, (k, m) \in c)\}$$
$$\lambda x . \tau(x) = \{(k, m) | m \in \tau[e_k/x]\}.$$

For our purposes it is sufficient to take $\tau(x)$ to be a λ-term (though the definition is good whenever $\tau(x)$ is continuous in x in the sense of [4]), and then just repeating the definition we gave for the Scott models, we can define a value $[\![M]\!]$, of M in the Graph model.

For $d \in D$, we define $(d)_n = \{m | m \in d \,\&\, m \leqslant n\}$. Then the facts (1)–(4), above, hold for the Graph model (now $\bigcup$ is real union etc.). Observe that

(a) if $k \in e_n$, then $k < n$,

(b) $m \leqslant (n, m)$ with equality only if $m = n = 0$,

(c) $n \leqslant (n, m)$ with equality only if $m = 0$ and $n = 0$ or 1.

These enable us to derive by straightforward computation the following basic equations to correspond to (7) above;

$$d_{n+2}(c) = d_{n+2}(c_{n+1}) \subset (d(c_{n+1}))_{n+1},$$
$$d_1(c) = d_1(c_0) = (d(c_0))_0,$$
$$d_0(c) = d_0(\varnothing) = (d(\varnothing))_0.$$

This is all we need to know about the Graph model.

As in **[5]** we adjoin to the λ-calculus a constant term Ω, whose value shall be $\perp$ in the Scott models, and $\varnothing$ in the Graph model. This is a technical and heuristic device, but it adds nothing to the λ-definable elements in either model, as $\lambda x.(xx)\,(\lambda x.(xx))$ has the same value as Ω (this is an immediate consequence of the approximation theorems of §2). Henceforth, " term " will mean " term of the extended calculus ". It is clear how to define the value of an (extended) term.

2. *The Approximation Theorems*

We say that a term M is in normal form, if it has no β-subredexes, and no Ω-subredexes; an Ω-redex is a term of the form ΩN or $\lambda x.\Omega$. The appropriate notion of Ω-reduction is that any Ω-redex reduces to Ω. We define a set $\omega(M)$, the set of Ω-*approximants* to the term M, by $\omega(M) = \{L \mid L$ is in normal form and is obtained from some N with $N =_\beta M$ by replacing subterms of N by $\Omega\}$.
The purpose of this section is to show that the value $[\![M]\!]$ of M in our models is the limit (i.e. least upper bound or union respectively) of the values of the Ω-approximants of M.

We introduce the notion of an indexed term. An *indexed term* (M, I), is a term M, together with a map I from the subterms of M to the natural numbers. The value, $[\![M]\!]^I_\rho$, of an indexed term (for a given valuation ρ, of the free variables) is given by

$$[\![x]\!]^I_\rho = (\rho(x))_{I(x)},$$

$[\![\Omega]\!]^I_\rho = \perp$ or $\varnothing$, in the Scott models or Graph model, respectively,

$$[\![MN]\!]^I_\rho = ([\![M]\!]^I_\rho\,([\![N]\!]^I_\rho))_{I(MN)},$$

$$[\![\lambda x.P]\!]^I_\rho = (\lambda d.[\![P]\!]^I_{\rho[d/x]})_{I(\lambda x.P)}.$$

Here, the remarks following the similar definition in §1 apply; subscripts refer to the taking of the (finite) approximations $(d)_n$ as defined in §1. Finally, there should be no confusion caused by our using I both as an index map for MN and as its restriction to M (say).

LEMMA 2.1. *In both models,* $[\![M]\!] = \bigcup\{[\![M]\!]^I \mid (M, I)$ *is an indexed term*$\}$.

The proof is by a routine structural induction.

Now, adapting a useful idea (unpublished) of Wadsworth's which serves to simplify our original proof of (2.5), we introduce the notion of the *indexed reduction* of indexed terms. We use superscripts on terms to indicate the index associated with them (i.e. their value under I).

Indexed β-reduction

$(\lambda x.P^n)^{m+1}.Q^p$ reduces to $(P[Q^a/x])^b$, where b is the minimum of m and n, while a varies amongst substitution instances of Q and is the minimum of m, p, and the index of the x for which Q has been substituted;

$$(\lambda x.P^n)^0.Q^p \text{ reduces to } (P[\Omega^0/x])^0.$$

Indexed Ω-reduction

We define what it is for an *occurrence* of Ω to be *active* in a subterm of a term by

(i) Ω is active in itself;

(ii) if Ω is active in P, then it is active in $\lambda x.P$ and PN.

We remark that if Ω is active in L, then a series of Ω-reductions will reduce L to Ω. An indexed Ω-reduction consists in reducing (in one step) a maximal subterm in which a given occurrence of Ω is active to Ω. (Doing all possible Ω-reductions for a given occurrence at once simplifies the proof of (2.2)).

LEMMA 2.2. *Some indexed reduction of (M, I) terminates.*

Proof. Define $u(0) = \text{card}(\{A|\ (M,I)\ A$ is a subterm of m which may be (indexed) Ω-reduced$\})$,

$$u(n+1) = \text{card}(\{A|\ (M,I)\ A \text{ is a subterm of } m, \text{ of form } (\lambda x.P)^n.Q\}).$$

u is an eventually zero function from ω to ω. For such functions, u and v, let $u < v$ if and only if the greatest $i \in \omega$ such that $u(i) \neq v(i)$ is such that $u(i) < v(i)$. Plainly, $<$ is a well-ordering on these functions. Also, by reducing from the " inside-out ", (M, I) index reduces to (M', I') with corresponding u and u' such that $u' < u$. Hence some reduction sequence terminates.

LEMMA 2.3. (a) *In the Scott models, indexed reduction preserves the value of an indexed term.*

(b) *In the Graph model, if (M, I) index reduces to (M', I'), then we have $[\![M]\!]^I \subset [\![M']\!]^{I'}$.*

Proof. This is simply a reproof of the fact that β- (and Ω-) reduction do not alter the values of terms. One uses equations (7) of §1 (or their counterparts for the Graph model) to keep track of the indices.

LEMMA 2.4. *For any (M, I), there is (L, J), with $L \in \omega(M)$, and $[\![M]\!]^I \subset [\![L]\!]^J$.*

Proof. The proof is by (2.2) and (2.3), since indexed reductions terminate in Ω-approximants. (Of course, $\subset$ can be replaced by $=$ for the Scott models in (2.4).)

THEOREM 2.5. *In both models, $[\![M]\!] = \bigcup\{[\![L]\!]|\ L \in \omega(M)\}$.*

Proof. If $L \in \omega(M)$, L is obtained from a term with the same value as M, by replacing subterms by Ω, and so $[\![L]\!] \subset [\![M]\!]$. But by (2.1) and (2.4), we have $[\![M]\!] \subset \bigcup\{[\![L]\!]|\ L \in \omega(M)\}$. Hence the result.

COROLLARY. *If M has no* hnf (*see* §3), *then* $[\![M]\!]$ *is* $\bot$ *or* $\varnothing$.

3. *Head normal forms*

We define which terms, M, are *head normal forms* (we write, "M is hnf"), by:

(a) all variables are hnf;

(b) if $X_1, \ldots, X_k$ are terms and x is a variable, then $xX_1 \ldots X_k$ is hnf;

(c) if P is hnf, then $\lambda x.P$ is hnf.

M has hnf if and only if there is $N =_\beta M$ and N is hnf. Otherwise, *M has no hnf.* An hnf has the form, $\lambda x_1 \ldots x_k . zX_1 \ldots X_l$, and z is the *head variable.* A non-hnf has the form, $\lambda x_1 \ldots x_k . (\lambda y . P) X_1 \ldots X_l$; the *head redex* is $(\lambda y . P)X_1$, and the (possibly infinite) reduction of a term, obtained by always reducing the head redex if any, is the *head reduction.* By the Standardization Theorem, M has hnf if and only if its head reduction terminates; and hence the set of terms with no hnf has strong closure properties (see [**5**]). A term has hnf if and only if its closure is soluble in the sense of Barendregt [**1**].

Let $\lambda x_1 \ldots x_m . zX_1 \ldots X_k$ and $\lambda y_1 \ldots y_n . wY_1 \ldots Y_l$, be two hnfs. We may assume that (say) $n \leqslant m$, and then by α-conversion, that each y_i is x_i. The hnfs are *similar* if and only if z is w and $(m-k) = (n-l)$. The point of this definition is that when two hnfs are not similar, then we can find a context which reduces them to distinct variables (see §4 for elucidation). Thus two dissimilar hnfs can never be set equal consistently with β-equality.

Now we define, for $k \geqslant 1$, (a) M and N have the same *k-normal form* (written $M =_k N$), and (b) the set of *k-pairs* of (M, N), as follows. If both M and N have no hnf, then $M =_1 N$, and there are no 1-pairs. Otherwise, $M =_1 N$ if and only if they have similar hnfs; then let M' and N' be the hnfs obtained by head reduction; M' is $\lambda x_1 \ldots x_m . zX_1 \ldots X_k$ and N' is $\lambda y_1 \ldots y_n . wY_1 \ldots Y_l$; making the same assumptions as above, we write $M' x_1 \ldots x_m =_\beta zX_1 \ldots X_k$ and $N' x_1 \ldots x_m =_\beta wY_1 \ldots Y_l x_{n+1} \ldots x_m$ which is $Y_1' \ldots Y_k'$, say; then the 1-pairs of (M, M) are the (X_i, Y_i') for $i \leqslant k$. Now for the induction step of the definition, we set $M =_{k+1} N$ if and only if $M =_1 N$ and for any 1-pairs of (M, N), (X, Y) say, we have $X =_k Y$; the $k+1$-pairs of (M, N) are the k-pairs of its 1-pairs.

Remarks:

(1) In the induction we could interchange the role of 1 and k.

(2) If $k > 1$, then (M, N) can have k-pairs, without $M =_k N$.

(3) $M =_k N$ means roughly that M and N are similar down to depth k (with the agreement that non-hnfs are similar).

We now introduce relations $<_k^s$, $k \geqslant 1$, and $<_k^g$, $k \geqslant 1$. The superscripts s and g indicate that the relations are useful for the Scott and Graph models, respectively.

We set $M <_1^s N$ if and only if either $M =_1 N$ or M has no hnf (i.e. $M =_1 \Omega$); and by induction, $M <_{k+1}^s N$ if and only if $M <_1^s N$ and for any 1-pairs (X, Y) we have $X <_k^s Y$.

Remarks:

(1) Plainly $M =_k N$ if and only if $M <_k^s N$ and $N <_k^s M$.

(2) Roughly, $M <_k^s N$ means that M and N are similar down to depth k, except "at subterms of M which have no hnf".

We set $M <_1^g N$ if and only if either M has no hnf or $M =_1 N$ and if we obtain by head reduction $M =_\beta \lambda x_1 \ldots x_m . z X_1 \ldots X_k$ and $N =_\beta \lambda x_1 \ldots x_n . z Y_1 \ldots Y_l$, then $m \leqslant n$ (i.e. M is essentially less functorial than N); again by induction, $M <_{k+1}^g N$ if and only if $M <_1^g N$ and for any 1-pairs (X, Y), we have $X <_k^g Y$.

These relations seem the right ones with which to attempt to approximate $\subset$ in the models, because (a) they take into account that terms with no hnf have value $\perp$, or $\varnothing$ and (b) they look deeper and deeper into the terms as k increases; the definition of $<_k^g$ takes into account the fact that η-expansion increases the value of a term in the Graph model. The ensuing results make this clearer. (For the rest of this section, $<_k$ refers to the relation appropriate to the model considered).

LEMMA 3.1. *If L is a normal form, and for all k, $L <_k N$, then $[\![L]\!] \subset [\![N]\!]$.*

Proof. The proof is by induction on the structure of normal forms. We shall not need to distinguish between the models until the very end.

(i) L is Ω, and the result is trivial.

(ii) L is x; we are assuming $x <_k N$ for all k, so in particular $N =_\beta \lambda y_1 \ldots y_r . x Y_1 \ldots Y_r$; by (2.5), it is sufficient to show that for any index I, for x, $[\![x]\!]^I \subset [\![N]\!]$; we do this by induction:

Case $I(x) = 0$; then $[\![x]\!]^I = ([\![x]\!])_0 = ([\![\lambda y_1 \ldots y_r . (x)_0 \, \Omega \ldots \Omega]\!])_0$, by the usual basic equations; thus $[\![x]\!]^I \subset [\![N]\!]$.

Case $I(x) = n+1$ (the induction step); using superscripts to indicate indices again, we have $[\![(x)^{n+1}]\!] \subset [\![\lambda y_1 \ldots y_r . (x)^{n+1} \, (y_1)^n \ldots (y_r)^{n-r+1}]\!]$, by the usual equations; (in the Scott model case we would have equality here); now for all k, $y_i <_k Y_i$, and so by induction hypothesis, $[\![(y_1)^n]\!] \subset [\![Y_1]\!]$, ... $[\![(y_r)^{n-r+1}]\!] \subset [\![Y_k]\!]$; hence $[\![x]\!]^I \subset [\![N]\!]$.

(iii) Let L be $\lambda x_1 \ldots x_m . x X_1 \ldots X_r$, where the X_i are normal forms for which we already have our result. We take our models separately.

The Scott model case: we can take

$$N =_{\beta\eta} \lambda x_1 \ldots x_{m+s} . x X_1 \ldots X_{r+s}, \quad \text{and} \quad L =_{\beta\eta} \lambda x_1 \ldots x_{m+s} . x X_1 \ldots X_{r+s}$$

(where if $s > 0$, X_{r+s} is x_{m+s} and so on); for all k, $X_i <_k Y_i$ and so $[\![X_i]\!] \subset [\![Y_i]\!]$; hence $[\![L]\!] \subset [\![N]\!]$. This completes the proof in the Scott model case.

The Graph model case: we have $N =_\beta \lambda x_1 \ldots x_{m+s} . x Y_1 \ldots Y_{r+s}$; by the basic equations, we have $[\![L]\!] \subset [\![\lambda x_1 \ldots x_{m+s} . x X_1 \ldots X_{r+s}]\!]$; now we can argue as in the Scott model case. This completes the proof of Lemma (3.1).

THEOREM 3.2. *If, for all k, $M <_k N$, then $[\![M]\!] \subset [\![N]\!]$.*

Proof. If not $[\![M]\!] \subset [\![N]\!]$, then by (2.5) there is $L \in \omega(M)$ with not $[\![L]\!] \subset [\![N]\!]$. But by the Church-Rosser theorem, we find that if $L \in \omega(M)$, then for all k, $L <_k M$. Then by obvious transitivity, $L <_k N$ for all k, which will contradict (3.1).

4. *Two technical lemmas*

A context $C[\ \ ]$ gives for every term M a term $C[M]$, which is the result of substituting M for the blank in $C[\ \ ]$; in this process free variables of M may get bound (i.e. no attempt is made to prevent this). In other words $C[M]$ and $C[N]$ are obtained from M and N by the same series of applications and λ-abstractions. Equality relations as defined in §0 are closed under context substitution. The following lemmas exploit this fact to give conditions which guarantee that terms M and N *cannot* be set equal. They are the essential content of [**2**].

LEMMA 4.1. *Given M and N, with (X, Y) k-pairs of (M, N), there is a context $C[\ \ ]$, such that $C[M] =_\beta X(R/x, \ldots)$, some substitution instance of X, while $C[N] =_\beta Y(R/x, \ldots)$, the same substitution instance of Y; the terms substituted are all of the form $\lambda x_1 \ldots x_h . x_h x_1 \ldots x_{h-1}$, for h sufficiently large; R_h will denote this term.*

Proof. We give the induction step on k, which also establishes the case $k = 1$. Suppose (X, Y) are 1-pairs of (U, V) which are $(k-1)$-pairs of (M, N), and we have the lemma for $(k-1)$. Thus we have a context $C_1[\ \ ]$, such that

$$C_1[M] =_\beta U(\ldots) =_\beta (\lambda x_1 \ldots x_m . z A_1 \ldots A_r)(\ldots),$$

and

$$C_1[N] =_\beta V(\ldots) = (\lambda x_1 \ldots x_n . z B_1 \ldots B_s)(\ldots),$$

where $m \geqslant n$ say, and where $(\ldots)$ will continue to denote some substitutions about which we have no need to be explicit. Now,

$C_1[M]x_1 \ldots x_m =_\beta z A_1 \ldots A_r(\ldots)$, and $C_1[N]x_1 \ldots x_m =_\beta z B_1' \ldots B_r'(\ldots)$, and (X, Y) is (A_i, B_i'), for some i.

We separate into cases:

(a) Suppose that z is not free in X or Y, and that $(\ldots)$ does not substitute anything for z; set $C[M]$ to be

$$(\lambda z . C_1[M]x_1 \ldots x_m)(\lambda y_1 \ldots y_r . y_i) =_\beta X(\ldots),$$

and then $C[N] =_\beta Y(\ldots)$.

(b) Suppose that z is free in X or in Y, but $(\ldots)$ does not substitute anything for z; then set $C[M]$ to be

$$\Big(\big((\lambda z . C_1[M]x_1 \ldots x_m)R_h\big)\, w_{r+1} \ldots w_{h-1}\Big)\, (\lambda x_1 \ldots x_{h-1} . x_i) =_\beta X(\ldots)\,(R_h/z),$$

and then $C[N] =_\beta Y(\ldots)\,(R_h/z)$. The primary intention of this manoeuvre is to prevent our substituting $(\lambda x_1 \ldots x_r . x_i)$ into X or Y; so we must take $h > r$.

(c) Suppose that $(\ldots)$ does substitute for z; we may assume that R_h was substituted for z, where $h > r$; then set $C[M]$ to be $(C_1[M]w_{r+1} \ldots w_{h-1})(\lambda x_1 \ldots x_{h-1} . x_i) =_\beta X(\ldots)$, and then $C[N] =_\beta Y(\ldots)$. This completes the proof.

The point of (4.1) is that if M and N are to be set equal (together with β-equality), then (substitution instances of) k-pairs of (M, N) must be set equal. We now prove a lemma which shows what terms can never be set equal (consistently with β-equality).

LEMMA 4.2. *If not $M <_1^s N$, and M' and N' are substitution instances of M and N of the form mentioned in* (4.1), *then either* (a) *both M and N have* hnf (*but these are not* *similar), and then there is a context* $C[\ \ ]$*, such that* $C[M'] =_\beta x$ *and* $C[N'] =_\beta y$ *where* x *and* y *are distinct variables, or* (b) M *has* hnf *but* $N =_1 \Omega$*, and there is a context* $C[\ \ ]$*, such that* $C[M'] =_\beta x$ *and* $C[N'] =_1 \Omega$.

Proofs. We only do (a), as (b) is similar but easier because of the strong closure properties of non-hnfs. Cases where substitutions have been made for the head variable of the hnfs of M and N can be reduced to those we consider, by the methods of (4.1); so this leaves us with the following two cases:

Case (i): Head variables distinct; so $M =_\beta \lambda x_1 \dots x_m . z X_1 \dots X_k$, $N =_\beta \lambda x_1 \dots x_n . w Y_1 \dots Y_l$, z is not w and $m \geqslant n$, say. Since to get M' and N', nothing has been substituted for z or w, set $C[M']$ to be $(\lambda zw . M' x_1 \dots x_m xy)\ (\lambda a_1 \dots a_{k+2} . a_{k+1})\ (\lambda b_1 \dots b_r . b_r)$, where $r = l+m-n+2$. Then $C[M'] =_\beta x$, and $C[N'] =_\beta y$.

Case (ii): Head variables the same;

$$M =_\beta \lambda x_1 \dots x_m . z X_1 \dots X_k, \quad N =_\beta \lambda x_1 \dots x_n . z Y_1 \dots Y_l, \quad \text{and} \quad (m-k) \neq (n-l);$$

$m \geqslant n$, say; then

$$M' x_1 \dots x_m =_\beta z X_1' \dots X_k', \text{ and } N' x_1 \dots x_m =_\beta z Y_1' \dots Y_r',$$

where $r = m+l-n$; take $s > |r-k|$ and set $C'[M']$ to be $(\lambda z . M' x_1 \dots x_m y_1 \dots y_s) R_h$, where $h = \min(r, k)+s$; then $C'[M']$ and $C'[N']$ are suitable for applying the technique of case (i). So the proof is completed for this case also.

5. *The main results*

First we define (syntactically) some equality relations:

(i) $M <_h N$ if and only if for all contexts $C[\ \ ]$, $C[M] <_1{}^s C[N]$; $M =_h N$ if and only if $M <_h N$ and $N <_h M$.

(ii) $M <_f N$ if and only if for all contexts $C[\ \ ]$, $C[M] <_1{}^g C[N]$; $M =_f N$ if and only if $M <_f N$ and $N <_f M$.

(iii) let $=_H$ be the equality generated by the axioms for β-equality, together with $M = N$ whenever both M and N have no hnf.

Remarks:

(1) By (4.2), $M <_h N$ if and only if whenever $C[M]$ has hnf, then so has $C[N]$.

(2) The main result of [**2**] is that if M and N are distinct $\beta\eta$-normal forms (of the pure calculus), then there is a context $C[\ \ ]$ such that $C[M] =_\beta x$, while $C[N] =_\beta y$ (x and y distinct variables). Thus the consistency of $=_h$ and $=_f$ is assured. We prove much more below.

(3) The consistency of $=_H$ was proved in [**1**].

Lemma 5.1. (a) *If* $[\![X]\!] \subset [\![Y]\!]$ *in the Scott models, then* $X <_1{}^s Y$.

(b) *If* $[\![X]\!] \subset [\![Y]\!]$ *in the Graph model, then* $X <_1{}^g Y$.

Proof. For (a): by (4.2), if not $X <_1^s Y$, there is context $C[\ \ ]$ with $C[X] =_\beta x$, while $C[Y] =_\beta y$ or Ω. Thus $[\![C[X]]\!]$ is not less than $[\![[CY]]\!]$, so not $[\![X]\!] \subset [\![Y]\!]$. For (b): if not $X <_1^g Y$, either not $X <_1^s Y$ and we apply the argument for (a), or we have $X =_\beta \lambda x_1 \dots x_{n+r}.xX_1 \dots X_{k+r}$, and $Y =_\beta \lambda x_1 \dots x_n.xY_1 \dots Y_k$, with $r > 0$. In this latter case, set X' to be $Xx_1 \dots x_n =_\beta \lambda x_{n+1} \dots x_{n+r}.xX_1 \dots X_{k+r}$, and Y' to be $Yx_1 \dots x_n =_\beta xY_1 \dots Y_k$. Now set $\rho(x) = \{0\}$, and then $[\![Y']\!] = \{0\}$, while $[\![X']\!]$ is an infinite set. Again not $[\![X]\!] \subset [\![Y]\!]$.

LEMMA 5.2. (a) *If* $[\![M]\!] \subset [\![N]\!]$ *in the Scott models, then* $M <_h N$.

(b) *If* $[\![M]\!] \subset [\![N]\!]$ *in the Graph model, then* $M <_f N$.

Proof. In either case, if the right hand side does not hold, we have a context $C[\ \ ]$, and we apply (5.1) to $C[M]$ and $C[N]$.

LEMMA 5.3. (a) *If* $M <_h N$, *then, for all* k, $M <_k^s N$.

(b) *If* $M <_f N$, *then, for all* k, $M <_k^g N$.

Proof. The cases are identical, so we write $<_k$ without superscript. If there is k with not $M <_k N$, then there are k-pairs (X, Y) of (M, N) with not $X <_1 Y$. By (4.1), we have a context $C[\ \ ]$, and $C[M] =_\beta X(\dots)$, $C[N] =_\beta Y(\dots)$, where $(\dots)$ denotes various substitutions. But it is easy to see (*cf.* the proof of (4.2)) that $X(\dots) <_1 Y(\dots)$ cannot hold. Hence not $M <_h N$.

THEOREM 5.4. (a) *For the Scott models* $[\![M]\!] \subset [\![N]\!]$ *if and only if* $M <_h N$ *if and only if, for all* k, $M <_k^s N$.

(b) *For the Graph model* $[\![M]\!] \subset [\![N]\!]$ *if and only if* $M <_f N$ *if and only if, for all* k, $M <_k^g N$.

Proof. The proof is by (3.2), (5.2) and (5.3).

COROLLARY. (a) *For the Scott models* $[\![M]\!] = [\![N]\!]$ *if and only if* $M =_h N$.

(b) *For the Graph model* $[\![M]\!] = [\![N]\!]$ *if and only if* $M =_f N$.

This corollary presents our syntactic characterization of the equalities in our models. We prove in conclusion a remarkable further characterization of the equality in the Scott models.

THEOREM 5.5. $=_h$ *(and hence the equality in the Scott models) is the unique maximal consistent equality extending* $=_H$.

Proof. Consistency is clear. It remains to show that if not $M =_h N$, then adding $M = N$ to the axioms for $=_H$ produces an inconsistency. But if, say, not $M <_h N$, then there is a context $C[\ \ ]$ such that $C[M]$ has hnf while $C[N]$ has no hnf. Then (in the standard combinators) $C[N] =_H Y(K)$. But a lemma of **[1]** shows how to derive inconsistency from the equality of any hnf with $Y(K)$. This completes the proof.

References

1. H. P. Barendregt, *Some extensional term models for combinatory logics and λ-calculi* (Utrecht, 1971).
2. C. Bohm, *Alcune proprieta della forme βη-normali nel λK-calcolo* (Consiglio Nazionale della Ricerche, No. 696, Rome).
3. D. Scott, *Continuous lattices* (Springer Lecture Notes in Mathematics, No. 274, pp. 97–136).
4. ———, *Lambda calculus and recursion theory* (Oxford University Computing Laboratory, PRG Technical Report, to appear).
5. C. P. Wadsworth, *Semantics and pragmatics of the λ-calculus* (Oxford D.Phil. Thesis).

Mathematical Institute,
Oxford.