Final Semantics for untyped λ-calculus*

Furio Honsell[1] and Marina Lenisa[1,2]

[1] Dipartimento di Matematica e Informatica
Università di Udine, Italy.
{honsell,lenisa}@dimi.uniud.it
[2] Dipartimento di Informatica
Università di Pisa, Italy.
lenisa@di.unipi.it

Abstract. *Proof principles* for reasoning about various semantics of untyped λ-calculus are discussed. The semantics are determined operationally by fixing a particular *reduction strategy* on λ-terms and a suitable set of *values*, and by taking the corresponding *observational equivalence* on terms. These principles arise naturally as co-induction principles, when the observational equivalences are shown to be induced by the *unique* mapping into a *final F-coalgebra*, for a suitable functor F. This is achieved either by *induction on computation steps* or exploiting the properties of some, computationally adequate, inverse limit denotational model. The final F-coalgebras cannot be given, in general, the structure of a "denotational" λ-model. Nevertheless the "final semantics" can count as *compositional* in that it induces a *congruence*. We utilize the intuitive categorical setting of *hypersets* and functions. The importance of the principles introduced in this paper lies in the fact that they often allow to factorize the complexity of proofs (of observational equivalence) by "straight" induction on computation steps, which are usually lengthy and error-prone.

Introduction

In this paper we present various proof-principles for reasoning about various semantics of λ-calculus which arise in the literature.

The word *semantics* is often used ambiguously. In fact, it can either refer to an *interpretation function* mapping terms into a space of denotations, or to a *procedure* for determining an equivalence relation, or to an *equivalence relation* on terms *per se* (e.g. the one induced by an interpretation function). We shall be careful in distinguishing between semantics and semantical equivalence. The proof principles that we introduce in this paper will be, therefore, rules for establishing semantical equivalence.

All the semantics that we consider originate operationally by fixing a suitable reduction strategy $\rightarrow_\sigma$ on λ-terms (denoted by Λ), or closed λ-terms (denoted

* Work supported by EEC Science contract MASK, HCM contract "Lambda Calcul Typé" and MURST 40% and 60% grants.

by Λ^0). A reduction strategy is a procedure for determining, for each term, a suitable β-redex appearing in it, to be contracted. A (possibly non-deterministic) strategy can be formalized as a relation $\rightarrow_\sigma \subseteq \Lambda \times \Lambda$ ($\Lambda^0 \times \Lambda^0$) such that, if $(M,N) \in \rightarrow_\sigma$ (also written infix as $M \rightarrow_\sigma N$), then N is a possible result of applying the reduction strategy $\rightarrow_\sigma$ to M. Reduction *paths* are defined, as usual, by repeatedly applying the reduction strategy (possibly zero times). The set of terms which do not belong to the domain of $\rightarrow_\sigma$ are partitioned into two disjoint sets: the set of *σ-values*, denoted by Val_σ and the set of *σ-deadlocks*.

Given a fixed strategy $\rightarrow_\sigma$, we can define the *evaluation relation* $\Downarrow_\sigma \subseteq \Lambda \times \Lambda$ ($\Lambda^0 \times \Lambda^0$), such that $M \Downarrow_\sigma N$ holds if and only if there exists a reduction path leading from M to N and N is a σ-value. If there exists N such that $M \Downarrow_\sigma N$, we shall say that the reduction $\rightarrow_\sigma$ *halts successfully* (terminates) on N, otherwise we shall write $M \not\Downarrow_\sigma$ and say that the reduction strategy $\rightarrow_\sigma$ *diverges* on M or reaches a *deadlock* from M. Often, the relation $\Downarrow_\sigma$ can be axiomatized pleasantly using Plotkin's S.O.S. style (see [16]).

Once we have a reduction strategy we may well say that we have an operational semantics, for we can imagine a machine which evaluates terms by implementing the given strategy. A natural semantical equivalence to define on terms is the *observational equivalence.* This arises if we consider programs as *black boxes* and take them to be equivalent if we cannot tell them apart by *observing* that for a given *program context* the machine halts successfully when one is used as a subprogram but does not halt when the other is used as a subprogram.

Definition 1 (σ-observational Equivalence). Let $\rightarrow_\sigma$ be a reduction strategy and let M, N be programs, i.e. closed λ-terms. The *observational equivalence* $\approx_\sigma$ is defined by

$$M \approx_\sigma N \ \text{ iff } \forall C[\,].(C[M], C[N] \in \Lambda^0 \Rightarrow (C[M] \Downarrow_\sigma \Leftrightarrow C[N] \Downarrow_\sigma)).$$

Notice that σ-observational equivalence is a congruence for every σ.

In this paper we consider many different reduction strategies, and corresponding observational equivalences, appearing in the literature. These strategies include those realized by the two most widespread implementations of functional programming languages, i.e. *call-by-value* and *lazy*. We shall focus in particular on the former. This was defined by Plotkin in [15] and is the one implemented by Landin's SECD machine or by the FAM. The observational equivalence induced by this strategy has been investigated in [6]. The latter strategy, which is the one used by lazy functional languages, induces an observational equivalence which has been studied in [1, 13]. Other strategies considered in the paper are:
- the "initial" leftmost strategy which stops on *head normal forms*; its observational equivalence is induced by the canonical D_∞ model of Scott (see [20]);
- the *β-normalizing* strategy, i.e. the "complete" leftmost strategy, whose observational equivalence, we conjecture coincides with the equivalence induced by the model defined in [5];
- Barendregt's *perpetual* strategy, which reduces the leftmost β-redex, which is either an I-redex, or a K-redex whose operand is in normal form, and which

can be shown to terminate exactly on *strongly normalizing* terms; we conjecture that the model defined in [10] for the λN^0-calculus is computationally adequate w.r.t. the observational equivalence induced by Barendregt's *perpetual* strategy;
- a non-deterministic strategy with a non-trivial set of deadlock terms, which induces as observational equivalence the congruence which cannot be realized by a continuously complete C.P.O. model (see [11]).

The importance of establishing observational equivalence of terms is unquestionable, e.g. it is at the core of the software development methodology of "program transformation". However, showing equivalences by induction on computation steps, is a lengthy and error-prone activity. Powerful proof-principles, which allow to factorize this difficult task, are therefore very precious.

In recent years, much attention has been devoted to so-called "coinduction principles", see e.g. [14, 7]. These concepts arose independently in two unrelated areas: that of semantics for concurrent languages [12] and that of non-wellfounded sets [8]. Aczel made the connection and pioneered their categorical account in terms of *final coalgebras* (see [2, 3]). Recently, various authors have deeply investigated *final semantics* (see e.g. [18, 7, 19]). They have developed a general categorical methodology for deriving coinduction principles whenever the semantical equivalence is induced by a final semantics. In this paper we apply this methodology to various semantics of λ-calculus. We proceed as follows. Given a reduction strategy $\rightarrow_\sigma$ and a set Val_σ, we *endow* Λ^0 (possibly in various ways), with the structure of a F-coalgebra, for suitable functors F. These functors are then proved to be "well behaved" in the sense that they admit a final F-coalgebra and preserve "weak kernel pairs" (see Appendix A and [18] for categorical definitions). Hence the *final interpretation given* by the unique morphism into the final F-coalgebra *induces* an equivalence which can be characterized by a coinduction principle, i.e. as the union of all F-bisimulations. Finally we are left with the task of showing that this equivalence coincides with the observational equivalence *determined* by the original reduction strategy.

The final F-coalgebra cannot be given, in general, the usual structure of a "denotational" λ-model. Nevertheless the final semantics can count as *compositional* in that the equivalence induced by it is a *congruence*, w.r.t. the syntactical constructors of the language. Actually, this is precisely what implies, in all the cases discussed here, that the final semantics induces the observational equivalence. Technically, it is quite hard to show that the final semantics induces a congruence. We prove it either by *induction on computation steps* or exploiting some property of the observational equivalence itself, which, in turn, can be proved exploiting the properties of some, computationally adequate, inverse limit denotational model.

We could have considered various categorical settings, such as that of complete metric spaces or complete partial orders. We prefer to utilize that of *hypersets* and functions, because this categorical setting is natural and it allows to keep the naive set-theoretic intuition. Hypersets are *non-wellfounded sets* belonging to a Universe of $ZF_0^-(U)FCU$. This is a Zermelo Frænkel like set-theory with extensionality "up to" *Urelementen* and with the Axiom of Foundation replaced by the "unique" antifoundation axiom FCU, which generalizes the antifoundation axiom X_1 introduced by Forti and Honsell in 1982 (also called AFA by Aczel [2]).

The paper is organized as follows. In Section 1 we present the reduction strategies and the corresponding evaluation relations that we shall deal with. In Section 2, we apply the final methodology. In particular we introduce, for each reduction strategy, a suitable functor and we derive a sound coinduction principle. In Section 3 we focus on the call-by-value observational equivalence and, applying again the final methodology, we derive yet *another* coinduction principle. In Section 4 we present an "induction-coinduction" principle for the call-by-value observational equivalence inspired by [14]. Final remarks, comments and conjectures appear in Section 5. In Appendix A we recall some useful definitions.

The authors are grateful to Mariangiola Dezani, Gordon Plotkin, Simona Ronchi Della Rocca, Jan Rutten and Daniele Turi for useful discussions.

1 Operational Semantics

Throughout the paper we use standard λ-calculus concepts and notation as defined in [4].

In this Section we present six operational semantics for λ-calculus. More precisely we give six reduction strategies together with the corresponding evaluation relations. The latter are presented using Plotkin's S.O.S. style, (cfr. [16]).

Definition 2 ($\rightarrow_v$ strategy, $\Downarrow_v$ evaluation). The *lazy call-by-value* strategy $\rightarrow_v \subseteq \Lambda^0 \times \Lambda^0$ reduces the *leftmost* β-redex, not appearing within a λ-abstraction, whose argument is a λ-abstraction. $Val_v = \{\lambda x.M \mid M \in \Lambda\} \cap \Lambda^0$. The evaluation $\Downarrow_v$ is the least binary relation over $\Lambda^0 \times Val_v$ satisfying the following rules:

$$\frac{}{\lambda x.M \Downarrow_v \lambda x.M} \qquad \frac{M \Downarrow_v \lambda x.P \quad N \Downarrow_v Q \quad P[Q/x] \Downarrow_v U}{MN \Downarrow_v U}$$

Definition 3 ($\rightarrow_l$ strategy, $\Downarrow_l$ evaluation). The *lazy call-by-name* strategy $\rightarrow_l \subseteq \Lambda^0 \times \Lambda^0$ reduces the *leftmost* β-redex not appearing within a λ-abstraction. $Val_l = \{\lambda x.M \mid M \in \Lambda\} \cap \Lambda^0$. The evaluation $\Downarrow_l$ is the least binary relation over $\Lambda^0 \times Val_l$ satisfying the following rules:

$$\frac{}{\lambda x.M \Downarrow_l \lambda x.M} \qquad \frac{M \Downarrow_l \lambda x.P \quad P[N/x] \Downarrow_l Q}{MN \Downarrow_l Q}$$

Definition 4 ($\rightarrow_h$ strategy, $\Downarrow_h$ evaluation). The *eager call-by-name* strategy $\rightarrow_h \subseteq \Lambda \times \Lambda$ reduces the *leftmost* β-redex, if the term is not in head normal form. Val_h is the set of λ-terms in head normal form. The evaluation $\Downarrow_h$ is the least binary relation over $\Lambda \times Val_h$ satisfying the following rules:

$$\frac{}{xM_1 \dots M_n \Downarrow_h xM_1 \dots M_n}\ n \geq 0 \qquad \frac{M \Downarrow_h N}{\lambda x.M \Downarrow_h \lambda x.N}$$

$$\frac{M[N/x]M_1 \dots M_n \Downarrow_h P}{(\lambda x.M)NM_1 \dots M_n \Downarrow_h P}\ n \geq 0$$

Definition 5 ($\rightarrow_n$ **strategy,** $\Downarrow_n$ **evaluation).** The *normalizing* strategy $\rightarrow_n \subseteq \Lambda \times \Lambda$ reduces the leftmost β-redex. Val_n is the set of λ-terms in normal form. The evaluation $\Downarrow_n$ is the least binary relation over $\Lambda \times Val_n$ satisfying the following rules:

$$\frac{M_1 \Downarrow_n M_1' \ \ldots \ M_n \Downarrow_n M_n'}{xM_1 \ldots M_n \Downarrow_n xM_1' \ldots M_n'} \ n \geq 0 \qquad \frac{M \Downarrow_n N}{\lambda x.M \Downarrow_n \lambda x.N}$$

$$\frac{M[N/x]M_1 \ldots M_n \Downarrow_n P}{(\lambda x.M)NM_1 \ldots M_n \Downarrow_n P} \ n \geq 0$$

Definition 6 ($\rightarrow_i$ **strategy,** $\Downarrow_i$ **evaluation).** Barendregt's *perpetual* strategy $\rightarrow_i \subseteq \Lambda \times \Lambda$ reduces the leftmost β-redex not in the operator of a redex, which is either an $I\beta$-redex, or a $K\beta$-redex whose argument is a normal form. Val_i is the set of λ-terms in normal form. The evaluation $\Downarrow_i$ is the least binary relation over $\Lambda \times Val_i$ satisfying the following rules:

$$\frac{M_1 \Downarrow_i M_1' \ \ldots \ M_n \Downarrow_i M_n'}{xM_1 \ldots M_n \Downarrow_i xM_1' \ldots M_n'} \ n \geq 0 \qquad \frac{M \Downarrow_i N}{\lambda x.M \Downarrow_i \lambda x.N}$$

$$\frac{(\lambda x.M)N \ \textit{I}\beta\textit{-redex} \quad M[N/x]M_1 \ldots M_n \Downarrow_i V}{(\lambda x.M)NM_1 \ldots M_n \Downarrow_i V} \ n \geq 0$$

$$\frac{(\lambda x.M)N \ \textit{K}\beta\textit{-redex} \quad N \Downarrow_i P \quad MM_1 \ldots M_n \Downarrow_i V}{(\lambda x.M)NM_1 \ldots M_n \Downarrow_i V} \ n \geq 0$$

Definition 7 ($\rightarrow_e$ **strategy,** $\Downarrow_e$ **evaluation).** Let Ω be a new constant. The non-deterministic strategy $\rightarrow_e \subseteq \Lambda^0(\{\Omega\}) \times \Lambda^0(\{\Omega\})$ rewrites λ-terms which contain occurrences of the constant Ω by reducing any β-redex. $Val_e = \Lambda^0$. Normal forms which are not in Val_e are the $\rightarrow_e$-deadlock terms. The evaluation relation $\Downarrow_e$ is the least binary relation over $\Lambda^0(\{\Omega\}) \times Val_e$ satisfying the following rules:

$$\frac{M \in Val_e}{M \Downarrow_e M} \qquad \frac{C[(\lambda x.M)N] \notin Val_e \quad C[M[N/x]] \Downarrow_e P}{C[(\lambda x.M)N] \Downarrow_e P}$$

2 Final Descriptions of observational Equivalences

In this section we give a first series of "final" accounts of the observational equivalences induced by the evaluation relations defined in the previous section. Each of these accounts gives rise to a particular coinductive characterization of the observational equivalence under consideration.

We work in the category *Class* whose objects are the classes of non-well-founded sets belonging to a Universe of $ZF_0^-(U)FCU$, and the arrows are the functional classes. The theory $ZF_0^-(U)FCU$ is a Zermelo-Frænkel-like set theory with extensionality, "up to" the proper class U of Urelementen (atoms), and with the axiom of Foundation replaced by the anti-foundation axiom FCU of [9] (see Definition 23 in Appendix A). This axiom is the version of X_1 (AFA) "up to" Urelementen, see [8, 2]. The Axiom FCU implies that the Universe is *strongly extensional*, i.e. that sets are unique "up to" bisimulations which preserve atoms (see Definition 23 in Appendix A, and [8, 2, 9] for more details). Alternatively we could have used the category $Class^*$ of [18], the category of *C.P.O.*'s and *strict functions*, or the category of *complete metric spaces* and *non-distance increasing functions.*

Given an evaluation relation $\Downarrow_\sigma$, we will proceed uniformly as follows (see [18]):

1. we endow the set Λ^0 with a structure of F-coalgebra, for a suitable functor $F : Class^* \to Class^*$;
2. we prove that the functor has a final F-coalgebra;
3. we define the interpretation function $\mathcal{M}_\sigma$ as the unique F-coalgebra morphism from the F-coalgebra on Λ^0 into the final F-coalgebra;
4. we prove that the the equivalence induced by $\mathcal{M}_\sigma$ is given by the union of all F-bisimulations on the F-coalgebra on Λ^0;
5. we prove that the equivalence induced by $\mathcal{M}_\sigma$ is $\approx_\sigma$.

The steps 1 and 5 above are motivated and simplified if we introduce and discuss the notion of *applicative equivalence.*

Definition 8. Let $\approx_\sigma^{app} \subseteq \Lambda^0 \times \Lambda^0$, for $\sigma \in \{v, l, h, n, i, e\}$, be the *applicative equivalence* defined by

$$M \approx_\sigma^{app} N \iff \forall P_1, \ldots, P_n \in \Lambda^0.\ (MP_1 \ldots P_n \Downarrow_\sigma \Leftrightarrow NP_1 \ldots P_n \Downarrow_\sigma)\ .$$

In general $\approx_\sigma^{app}$ is not a congruence, but for all $\sigma \in \{v, l, h, n, e\}$, we can prove that $\approx_\sigma^{app}$ coincides with $\approx_\sigma$ and hence it is a congruence.[3] The proofs of these facts will be outlined in the sequel.

The relation $\approx_\sigma^{app}$, for $\sigma \in \{v, l, h, n, i, e\}$ can be characterized coinductively as the greatest fixed point of a suitable monotone operator. Namely:

Definition 9. Let X be a set and $\Psi : \mathcal{P}(X \times X) \to \mathcal{P}(X \times X)$ be an operator. A Ψ*-bisimulation* is a relation $R \subseteq X \times X$ s.t. $R \subseteq \Psi(R)$.

If Ψ is monotone, then the greatest fixed point of Ψ is the greatest Ψ-bisimulation.

Lemma 1. *The applicative equivalence $\approx_\sigma^{app}$, for $\sigma \in \{v, l, h, n, i, e\}$, can be viewed as the greatest fixed point of the monotone operator $\Psi_\sigma : \mathcal{P}(\Lambda^0 \times \Lambda^0) \to \mathcal{P}(\Lambda^0 \times \Lambda^0)$ defined by*

$$\begin{aligned}\Psi_\sigma(R) = \{(M, N) \mid\ & (M \not\Downarrow_\sigma\ \wedge N \not\Downarrow_\sigma\ \wedge \forall P \in \Lambda^0.\ ((MP, NP) \in R)) \vee \\ & (M \Downarrow_\sigma\ \wedge N \Downarrow_\sigma\ \wedge \forall P \in \Lambda^0.\ ((MP, NP) \in R))\}.\end{aligned}$$

Proof. Let $\nu R.\Psi_\sigma(R)$ be the greatest fixed point of Ψ_σ. It is immediate to show that $\approx_\sigma^{app}$ is a Ψ_σ-bisimulation, hence $\approx_\sigma^{app} \subseteq \nu R.\Psi_\sigma(R)$. In order to show the converse, i.e. $\approx_\sigma^{app} \supseteq \nu R.\Psi_\sigma(R)$, we prove first, by induction on the length of $\overrightarrow{P}$ ($\overrightarrow{P}$ abbreviates $P_1 \ldots P_n$ for $n \geq 0$), that:
$(M, N) \in \nu R.\Psi_\sigma(R) \implies (M\overrightarrow{P}, N\overrightarrow{P}) \in \nu R.\Psi_\sigma(R)$.
Hence, reasoning by contradiction, we get $\approx_\sigma^{app} \supseteq \nu R.\Psi_\sigma(R)$.

[3] We conjecture that the same holds also for $\sigma = i$.

□

Notation. Throughout the paper we will denote by $X \to Y$ the class of all functions defined on X taking values in Y; and we will denote by $X + Y$ the "disjoint sum" of X and Y, e.g. $\{v\} \times X \cup \{u\} \times Y$, where v and u are two distinct "fresh" atoms.

2.1 A final Description of $\approx_v$ in the Hyperset Setting

The set Λ^0 can be endowed with a coalgebra structure appropriate for dealing with $\approx_v$ as follows:

Definition 10. i) Let $F_v : Class \to Class$ be the endofunctor defined by

$$F_v(X) = (\Lambda^0 \to X) + \{*\} \ ,$$

where $*$ is a generic atom; the definition of F_v on morphisms is canonical.
ii) Let $(\Lambda^0, \ \alpha_v)$ be the F_v-coalgebra defined by

$$\alpha_v(M) = \begin{cases} (u, *) & \text{if } M \not\Downarrow_v \\ (v, \{(N, MN) \mid N \in \Lambda^0\}) & \text{if } M \Downarrow_v \ . \end{cases}$$

Lemma 2. *The functor F_v has a greatest fixed point $\overline{X_v}$ such that $(\overline{X_v}, \ id)$ is a final F_v-coalgebra.*

Proof. One can easily extend the "Special Final Coalgebra Theorem" (see [2] and Corollary 4.23 of [18]) to $ZF_0^-(U)FCU$. The functor F_v can be easily seen to satisfy the appropriate generalizations, to $ZF_0^-(U)FCU$, of the hypotheses of the above theorem: i.e. it is set-continuous, inclusion preserving and uniform on maps.

□

Definition 11. Let $\mathcal{M}_v : \Lambda^0 \to \overline{X_v}$ be the unique F_v-morphism from the F_v-coalgebra $(\Lambda^0, \ \alpha_v)$ to the F_v-coalgebra $(\overline{X_v}, \ id)$, i.e.:

$$\mathcal{M}_v(M) = \begin{cases} (u, *) & \text{if } M \not\Downarrow_v \\ (v, \{(N, \ \mathcal{M}_v(MN)) \mid N \in \Lambda^0\}) & \text{if } M \Downarrow_v \ . \end{cases}$$

The following lemma can be proved straightforwardly.

Lemma 3. *R is a Ψ_v-bisimulation if and only if R is a F_v-bisimulation*

Lemma 4. *Let $M, N \in \Lambda^0$. Then:*

$$\mathcal{M}_v(M) = \mathcal{M}_v(N) \iff M \approx_v^{app} N \ .$$

Proof. One can easily see that F_v weakly preserves kernel pairs, and hence, by Corollary 3.9 of [18], the equivalence induced by the final morphism is the greatest F_v-bisimulation. Now, the thesis follows immediately from Lemma 3.

□

Lemma 5 (Theorem 33 of [6]). $\approx_v = \approx_v^{app}$

Using the above lemmata we can establish the validity of the following proof principle:

Theorem 6. *Let $M, N \in \Lambda^0$, then the following coinduction principle holds:*

$$\frac{(M,N) \in R \quad R \textit{ is a } F_v\textit{-bisimulation}}{M \approx_v N}$$

Remark. A more general coinduction principle can be given introducing the notion of Ψ_v-bisimulation up to $\approx_v$ following [12]. Since for any R, S, if $R \subseteq \Psi_v(R_1)$ and $S \subseteq \Psi_v(S_1)$ then $R \circ S \subseteq \Psi_v(R_1 \circ S_1)$, the following principle holds

$$\frac{(M,N) \in R \quad R \subseteq \Psi_v(\approx_v \circ R \circ \approx_v)}{M \approx_v N}$$

2.2 A final Description of $\approx_l$ in the Hyperset Setting

The set Λ^0 can be endowed with a coalgebra structure appropriate for dealing with $\approx_l$ as follows:

Definition 12. i) Let $F_l : Class \to Class$ be the endofunctor defined by

$$F_l(X) = (\Lambda^0 \to X) + \{*\} ,$$

where $*$ is a generic atom; the definition of F_l on morphisms is canonical.
ii) Let $(\Lambda^0,\ \alpha_l)$ be the F_l-coalgebra defined by

$$\alpha_l(M) = \begin{cases} (u, *) & \text{if } M \not\Downarrow_l \\ (v, \{(N, MN) \mid N \in \Lambda^0\}) & \text{if } M \Downarrow_l . \end{cases}$$

Following the same lines of reasoning as in Section 2.1, provided we prove $\approx_l = \approx_l^{app}$, we can eventually show that:

Theorem 7. *Let $M, N \in \Lambda^0$.*
i) Then

$$\mathcal{M}_l(M) = \mathcal{M}_l(N) \iff M \approx_l N .$$

ii) The following coinduction principle holds:

$$\frac{(M,N) \in R \quad R \textit{ is a } F_l\textit{-bisimulation}}{M \approx_l N}$$

The coincidence of the observational equivalence with the applicative equivalence can be proved in various ways (see e.g. [1]). Here we give a syntactical proof, similar to one of those in [1].

Lemma 8. $\approx_l = \approx_l^{app}$

Proof. Clearly $\approx_l \subseteq \approx_l^{app}$. In order to show the converse, we proceed by induction on computation steps. Suppose by contradiction that there is a context $C[\ ]$ such that $C[P] \Downarrow_l$ and $C[P'] \not\Downarrow_l$. Choose a context $C_{min}[\ ]$ satisfying the property above such that the length of a path starting from $C_{min}[P]$, converging to a value, is minimal. Since $C_{min}[P] \Downarrow_l$ and $C_{min}[P'] \not\Downarrow_l$, an occurrence of P must necessarily appear as the head of a term in the $\rightarrow_l$-reduction path leading from $C_{min}[P]$ to a value. Consider the first time in which an occurrence of P appears in the head in the reduction starting from $C_{min}[P]$. Then we have $C_{min}[P] \rightarrow_l^* PC'[P]$ and $C_{min}[P'] \rightarrow_l^* P'C'[P']$, for some context $C'[\ ]$. By definition of applicative equivalence, we have $P'C'[P'] \approx_l^{app} PC'[P']$. Now $P \Downarrow_l$, otherwise we have immediately a contradiction. Therefore, suppose $P \Downarrow_l \lambda x.M$, we have
$PC'[P] \rightarrow_l^* (\lambda x.M)C'[P] \rightarrow_l M[C'[P]/x] = C''[P] \Downarrow_l$
$PC'[P'] \rightarrow_l^* (\lambda x.M)C'[P'] \rightarrow_l M[C'[P']/x] = C''[P'] \not\Downarrow_l$,
for some context $C''[\ ]$ such that $C''[P]$ converges to a value with a path whose length is strictly less than the length of the converging path of $C_{min}[P]$.
□

2.3 A final Description of $\approx_h$ in the Hyperset Setting

The set Λ^0 can be endowed with a coalgebra structure appropriate for dealing with $\approx_h$ as follows:

Definition 13. i) Let $F_h : Class \rightarrow Class$ be the endofunctor defined by

$$F_h(X) = (\Lambda^0 \rightarrow X) + \{*\} \ ,$$

where $*$ is a generic atom; the definition of F_h on morphisms is canonical.
ii) Let $(\Lambda^0,\ \alpha_h)$ be the F_h-coalgebra defined by

$$\alpha_h(M) = \begin{cases} (u, *) & \text{if } M \not\Downarrow_h \\ (v, \{(N, MN) \mid N \in \Lambda^0\}) & \text{if } M \Downarrow_h \ . \end{cases}$$

Following the same lines of reasoning as in Section 2.1, we can show:

Theorem 9. *Let $M, N \in \Lambda^0$.*
i) Then

$$\mathcal{M}_h(M) = \mathcal{M}_h(N) \iff M \approx_h N \ .$$

ii) The following coinduction principle holds:

$$\frac{(M,N) \in R \quad R \textit{ is a } F_h\textit{-bisimulation}}{M \approx_h N}$$

The proofs of the lemmata necessary for showing Theorem 9 are similar to those in Section 2.1, but for:

Lemma 10. $\approx_h = \approx_h^{app}$

A proof of this lemma can be achieved along the lines of the corresponding lemma in Section 2.2, extending the notion of $\approx_h^{app}$ to open terms. It can be also obtained using Wadsworth's extension of Böhm's "separability" Theorem (see [20]).

2.4 A final Description of $\approx_n$ in the Hyperset Setting

The $\approx_n$ equivalence does not equate all $\rightarrow_n$-divergent terms, in fact it is not true that a divergent term applied to any argument always diverges. Hence, in order to define a non well-founded final semantics which induces the $\approx_n$ equivalence, we consider the following functor and coalgebra on λ-terms, which allow us to distinguish correctly between divergent λ-terms whose applicative behavior is different:

Definition 14. i) The endofunctor $F_n : Class \rightarrow Class$ is defined by

$$F_n(X) = (\Lambda^0 \rightarrow X) + (\Lambda^0 \rightarrow X) \ ,$$

the definition of F_n on morphisms is canonical.
ii) Let $(\Lambda^0,\ \alpha_n)$ be the F_n-coalgebra defined by

$$\alpha_n(M) = \begin{cases} (u, \{(N, MN) \mid N \in \Lambda^0\}) & \text{if } M \not\Downarrow_n \\ (v, \{(N, MN) \mid N \in \Lambda^0\}) & \text{if } M \Downarrow_n \ . \end{cases}$$

Following the same lines of reasoning as in Section 2.1, we can show:

Theorem 11. *Let $M, N \in \Lambda^0$.*
i) Then

$$\mathcal{M}_n(M) = \mathcal{M}_n(N) \iff M \approx_n N \ .$$

ii) The following coinduction principle holds:

$$\frac{(M, N) \in R \quad R \textit{ is a } F_n\textit{-bisimulation}}{M \approx_n N}$$

The proofs of the lemmata necessary for showing Theorem 11 are similar to those used in Section 2.1, but for:

Lemma 12. $\approx_n = \approx_n^{app}$

A proof of this lemma can be achieved along the lines of the corresponding lemma in Section 2.2, extending the notion of $\approx_n^{app}$ to open terms.

2.5 A final Description of $\approx_i$ in the Hyperset Setting

A final semantics inducing the $\approx_i^{app}$ equivalence can be given similarly to the previous cases where all divergent terms are equated, say $\rightarrow_h$. We conjecture the coincidence between $\approx_i$ and $\approx_i^{app}$.

2.6 A final Description of $\approx_e$ in the Hyperset Setting

The equivalence $\approx_e$ does not equates all $\rightarrow_e$-divergent terms. A final semantics for handling this observational equivalence can be given along the lines of that for $\approx_n$ in Section 2.4. The coincidence between $\approx_e$ and $\approx_e^{app}$ can be proved using a model theoretic argument similar to that utilized for $\approx_v$ in [6]. More details will be given in a forthcoming paper.

3 Yet another final Description of $\approx_v$ in the Hyperset setting

In this section we present another final semantics, inducing the $\approx_v$ equivalence, which makes use of a functor different from the one considered in the previous section. From this semantics we derive yet another coinduction principle for establishing $\approx_v$. In particular we prove that $\approx_v^{app}$ can be viewed as the greatest fixed point of the following monotone operator:

Definition 15. Let $\Phi_v : \mathcal{P}(\Lambda^0 \times \Lambda^0) \to \mathcal{P}(\Lambda^0 \times \Lambda^0)$ be the operator defined by $\Phi_v(R) = \{(M, N) \mid (M \not\Downarrow_v \ \wedge N \not\Downarrow_v) \ \vee \ (M \Downarrow_v \ \wedge N \Downarrow_v \ \wedge \forall P \exists Q.\ ((P, Q) \in R \ \wedge \ (MP, NQ) \in R) \ \wedge \ \forall P \exists Q.\ ((P, Q) \in R \ \wedge \ (NP, MQ) \in R))\}$.

Definition 16. i) Let $G_v : Class \to Class$ be the endofunctor defined by

$$G_v(X) = \mathcal{P}(X \times X) + \{*\} \ ,$$

where $*$ is a generic atom; the definition on morphisms is canonical.
ii) Let $(\Lambda^0, \ \beta_v)$ be the G_v-coalgebra defined by

$$\beta_v(M) = \begin{cases} (u, *) & \text{if } M \not\Downarrow_v \\ (v, \{(N, MN) \mid N \in \Lambda^0\}) & \text{if } M \Downarrow_v \ . \end{cases}$$

Lemma 13. *R is a Φ_v-bisimulation if and only if R is a G_v-bisimulation.*

Proof. ($\Rightarrow$) The assertion follows from the definition of G_v-bisimulation.
($\Leftarrow$) By contradiction.

$\square$

The following lemma is proved as Lemma 2.

Lemma 14. *The functor G_v has a greatest fixed point $\overline{Y_v}$ such that $(\overline{Y_v}, \ id)$ is a final G_v-coalgebra.*

Definition 17. Let $\mathcal{N}_v : \Lambda^0 \to \overline{Y_v}$ be the unique G_v-morphism from the G_v-coalgebra $(\Lambda^0, \ \beta_v)$ to the G_v-coalgebra $(\overline{Y_v}, \ id)$, i.e.:

$$\mathcal{N}_v(M) = \begin{cases} (u, *) & \text{if } M \not\Downarrow_v \\ (v, \{(\mathcal{N}_v(N), \ \mathcal{N}_v(MN)) \mid N \in \Lambda^0\}) & \text{if } M \Downarrow_v \ . \end{cases}$$

Now we prove that the greatest Φ_v-bisimulation, $\sim_v$, coincides with the applicative equivalence $\approx_v^{app}$, and hence that the equivalence induced by $\mathcal{N}_v$ coincides with $\approx_v$. To this end we introduce the syntactical counterparts of the relations $\equiv_n$ introduced in Section 3.1 of [6].

Definition 18. Let P be the initial solution in the category $C.P.O.^E$ of the domain equation $D = [D \to_\perp D]_\perp$, i.e. $P = lim_{\leftarrow} P_n$. Let Λ_n^0 be the set of all closed λ-terms whose interpretation, up to isomorphism, belongs to P_n.
i) For all n, let $\approx_v^n \subseteq \Lambda_n^0 \times \Lambda_n^0$ be the relation inductively defined as follows
$\approx_v^0 = \Lambda_0^0 \times \Lambda_0^0$

$\approx_v^{n+1}=\{(M,N)\in\Lambda^0_{n+1}\times\Lambda^0_{n+1}\mid\forall P,Q\in\Lambda^0_n.(P\approx_v^n Q\Rightarrow MP\approx_v^n NQ)\}$.
ii) For all n, let $\subseteq_v^n\subseteq\Lambda^0_n\times\Lambda^0_n$ be the relation inductively defined as follows
$\subseteq_v^0=\Lambda^0_0\times\Lambda^0_0$
$\subseteq_v^{n+1}=\{(M,N)\in\Lambda^0_{n+1}\times\Lambda^0_{n+1}\mid M\in\Lambda^0_0\ \vee\ (M\notin\Lambda^0_0\ \wedge\ N\notin\Lambda^0_0\ \wedge\ \forall P,Q\in\Lambda^0_n.(P\subseteq_v^n Q\Rightarrow MP\subseteq_v^n NQ))\}$.
iii) Since the projections $\pi_n: P\to P_n$ are all λ-definable (see [6]), say by Π_n, we define, for all $M\in\Lambda^0$, M_n to be the term $(\Pi_n M)\in\Lambda^0_n$.

In the sequel we use freely that for all $M,N\in\Lambda^0$, $(M_{n+1}N)=_P(M_{n+1}N_n)=_P(M_{n+1}N_n)_n=_P(MN_n)_n$, where $=_P$ denotes equality in the model P; and also that application is monotone w.r.t. $\subseteq_P$.

The following lemma is instrumental:

Lemma 15. *o) For all $n\in\omega$, $\approx_v^n$ is P-saturated, i.e. if $M\approx_v^n N$, $M=_P M'$ and $N=_P N'$, then $M'\approx_v^n N'$. Moreover for all $n\in\omega$, $\subseteq_v^n$ is P-saturated, i.e. if $M\subseteq_v^n N$, $M\subseteq_P M'$ and $N\subseteq_P N'$, then $M'\subseteq_v^n N'$.*
i) For all $n\in\omega$, $\approx_v^n=\subseteq_v^n\cap\supseteq_v^n$.
ii) For all $n\in\omega$, $\approx_v^n$ is an equivalence relation.
iii) For all $M,N\in\Lambda^0$, $\quad M\approx_v^{app}N\iff\forall n\in\omega.\ M_n\approx_v^n N_n$.
iv) $\approx_v^{app}\subseteq\sim_v$.
v) For all n: $\sim_v^n=\approx_v^n$, where $\sim_v^n\subseteq\Lambda^0_n\times\Lambda^0_n$ is inductively defined as follows:
$\sim_v^0=\Lambda^0_0\times\Lambda^0_0$
$\sim_v^{n+1}=$
$\{(M,N)\in\Lambda^0_{n+1}\times\Lambda^0_{n+1}\mid\forall P\in\Lambda^0_n\exists Q\in\Lambda^0_n((P,Q)\in\sim_v^n\ \wedge(MP,NQ)\in\sim_v^n)\ \wedge$
$\forall Q\in\Lambda^0_n\exists P\in\Lambda^0_n((Q,P)\in\sim_v^n\ \wedge(NQ,MP)\in\sim_v^n)\}$.
vi) For all $M,N\in\Lambda^0$, $\quad M\sim_v N\iff\forall n\in\omega.\ M_n\sim_v^n N_n$.

Proof. o) Both assertions are easily proved by induction on n. We show only the first one. The case $n=0$ is trivial. Suppose $M\approx_v^{n+1}N$, $M=_P M'$ and $N=_P N'$, then if $P\approx_v^n Q$, we have $MP\approx_v^n NQ$, hence by induction hypothesis $M'P\approx_v^n N'Q$.
i) The assertion is proved by induction on n, using the fact that, for all n and m such that $n+m\geq 1$,
$M\approx_v^{n+m}N\ \wedge\ P_1\subseteq_v^{n+m-1}Q_1\ \wedge\ldots\wedge\ P_m\subseteq_v^n Q_m\ \Rightarrow$
$MP_1\ldots P_m\subseteq_v^n NQ_1\ldots Q_m$,
which is easily provable by induction on n, regrouping $n+1+m$.
ii) Symmetry is proved straightforwardly by induction on n. Reflexivity follows from the fact that, for all n and m such that $n+m\geq 1$,
$(M\in\Lambda^0_{n+m}\ \wedge\ P_1\approx_v^{n+m-1}Q_1\ \wedge\ldots\wedge\ P_m\approx_v^n Q_m)\ \Rightarrow$
$MP_1\ldots P_m\approx_v^n MQ_1\ldots MQ_n$,
which is easily provable by induction on n, regrouping $n+1+m$.
Finally, transitivity is proved straightforwardly by induction on n, using reflexivity.
iii) Immediate from Theorem 33 of [6].
iv) The assertion is easily proved by coinduction.
v) The assertion is proved by induction on n. If $n=0$ the thesis is trivially true.

Let $\sim_v^n = \approx_v^n$. Then using the induction hypothesis and reflexivity of $\approx_v^n$, one gets immediately $\approx_v^{n+1} \subseteq \sim_v^{n+1}$. Now suppose by contradiction that $\sim_v^{n+1} \not\subseteq \approx_v^{n+1}$. Then there exist M, N such that $M \sim_v^{n+1} N$ and $M \not\approx_v^{n+1} N$, i.e. there are $P, T \in \Lambda_n^0$ such that $P \approx_v^n T$ and $MP \not\approx_v^n NT$. But there is $Q \in \Lambda_n^0$ such that $P \sim_v^n Q$ and $MP \sim_v^n NQ$. By induction hypothesis, $P \approx_v^n Q$ and $MP \approx_v^n NQ$, hence, by (ii), $NQ \approx_v^n NT$ and $MP \approx_v^n NT$, which is a contradiction.
vi) The implication ($\Leftarrow$) follows immediately from (iii), (iv) and (v). The other implication is proved by induction on n. If $n = 0$ the thesis is trivially true. Let $M_n \sim_v^n N_n$. We will show that $\forall P \in \Lambda_n^0\ \exists Q \in \Lambda_n^0$ such that $P \sim_v^n Q$ and $M_{n+1}P \sim_v^n N_{n+1}Q$. Let $P \in \Lambda_n^0$, then there is Q such that $P \sim_v Q$ and $MP \sim_v NQ$. By induction hypothesis, (v) and (o) $P \approx_v^n P_n \approx_v^n Q_n$, $(MP)_n \approx_v^n (NQ)_n$ and $M_{n+1}P \approx_v^n (NQ)_n$. Now $(NQ)_n \supseteq_v^n N_{n+1}Q_n \approx_v^n N_{n+1}P$, and in particular $M_{n+1}P \supseteq_v^n N_{n+1}Q_n$. Now we show the converse, i.e. $M_{n+1}P \subseteq_v^n N_{n+1}Q_n$. By definition of $\sim_v$, there exists T such that $Q_n \sim_v T$ and $MT \sim_v NQ_n$. By induction hypothesis and (v), $Q_n \approx_v^n T_n$, hence $T_n \approx_v^n P$, and $(MT)_n \approx_v^n (NQ_n)_n \approx_v^n N_{n+1}Q_n$. Hence, $N_{n+1}Q_n \approx_v^n (MT)_n \supseteq_v^n M_{n+1}P$. Summing up we have $N_{n+1}Q_n \approx_v^n M_{n+1}P$; and hence $P \sim_v^n Q_n$ and $M_{n+1}P \sim_v^n N_{n+1}Q_n$.

$\square$

Now we can give:

Theorem 16. *The greatest Φ_v-bisimulation, $\sim_v$, coincides with the applicative equivalence $\approx_v^{app}$.*

Proof. The thesis follows immediately from points (iii), (v) and (vi) of lemma 15.

$\square$

Proceeding as in the previous section we can now prove:

Theorem 17. *Let $M, N \in \Lambda^0$.*
i) Then

$$\mathcal{N}_n(M) = \mathcal{N}_n(N) \iff M \approx_v N\ .$$

ii) The following coinduction principle holds:

$$\frac{(M,N) \in R \quad R \textit{ is a } G_v\textit{-bisimulation}}{M \approx_v N}$$

4 A syntactical induction-coinduction Principle for $\approx_v$

In this section we prove the soundness of an induction-coinduction principle for establishing $\approx_v$-equivalence. This principle should deserve more investigation; however, it can be viewed as a syntactical version of the semantical induction-coinduction principle appearing in [14]. We use the notation introduced in the previous section.

Definition 19. Let $T : \mathcal{P}(\Lambda^0 \times \Lambda^0) \times \mathcal{P}(\Lambda^0 \times \Lambda^0) \to \mathcal{P}(\Lambda^0 \times \Lambda^0) \times \mathcal{P}(\Lambda^0 \times \Lambda^0)$ be the operator defined by

$T(R^-, R^+) =$
$(\{(M,N) \mid (M \not\Downarrow_v \ \wedge N \not\Downarrow_v) \ \vee \ (M \Downarrow_v \ \wedge N \Downarrow_v \ \wedge \forall (P,Q) \in R^+. \ (MP, NQ) \in R^-)\}$,
$\{(M,N) \mid (M \not\Downarrow_v \ \wedge N \not\Downarrow_v) \ \vee \ (M \Downarrow_v \ \wedge N \Downarrow_v \ \wedge \forall (P,Q) \in R^-. \ (MP, NQ) \in R^+)\})$.

Definition 20. A relation $R \subseteq \Lambda^0 \times \Lambda^0$ is λ-*inclusive* if for all $M, N \in \Lambda^0$, if for all $n \in \omega$ there exist M', N' such that $(M', N') \in R$, $M' \approx_v M_n$ and $N' \approx_v N_n$, then $(M, N) \in R$.

Theorem 18. *Let R^-, R^+ be two relations on $\Lambda^0 \times \Lambda^0$ such that R^+ is λ-inclusive. Then the following principle holds*

$$\frac{R^- \subseteq \pi_1(T(R^-, R^+)) \quad \pi_1(T(R^+, R^-)) \subseteq R^+}{R^- \subseteq \approx_v \subseteq R^+}$$

Proof. First of all we prove by induction on n that, for all n,

$$\widetilde{R}^-_n \subseteq \approx^n_v \subseteq R^+ \ ,$$

where $\widetilde{R}^-_n = \{(M,N) \in \Lambda^0_n \times \Lambda^0_n \mid \exists (P,Q) \in R^-.(P_n \approx_v M \ \wedge \ Q_n \approx_v N)\}$.

The base case ($n = 0$) follows trivially from the hypotheses of the principle. Suppose that $\widetilde{R}^-_n \subseteq \approx^n_v \subseteq R^+$. Then, since $(\approx^{n+1}_v, \approx^{n+1}_v) = T(\approx^n_v, \approx^n_v) \cap (\Lambda^0_{n+1} \times \Lambda^0_{n+1})$ and T is monotone in the first component and antimonotone in the second component, we have:

1. $\approx^{n+1}_v \subseteq \pi_1(T(\approx^n_v, \widetilde{R}^-_n)) \cap (\Lambda^0_{n+1} \times \Lambda^0_{n+1}) = \pi_1(T(\approx^n_v, R^-)) \subseteq \pi_1(T(R^+, R^-)) \cap (\Lambda^0_{n+1} \times \Lambda^0_{n+1}) \subseteq R^+$,
   where the equality is established using $(MN_n)_n \approx_v (M_{n+1}N_n)$.
2. $\widetilde{R}^-_{n+1} \subseteq \pi_1(T(\widetilde{R}^-_n, \approx^n_v)) \cap (\Lambda^0_{n+1} \times \Lambda^0_{n+1}) \subseteq \pi_1(T(\approx^n_v, \approx^n_v)) \cap (\Lambda^0_{n+1} \times \Lambda^0_{n+1}) = \approx^{n+1}_v$,
   where the first inclusion follows, using $(MN_n)_n \approx_v (M_{n+1}N_n)$, from $R^- \subseteq \pi_1(T(R^-, \approx^n_v))$, which in turn is a consequence of the left hypothesis of the principle and the induction hypothesis.

Now the inclusion $\approx_v \subseteq R^+$ follows immediately from the fact that R^+ is λ-inclusive, while the inclusion $R^- \subseteq \approx_v$ can be directly obtained by contradiction. □

5 Final Remarks

The constructions carried out in this paper raise many open questions and all should deserve more investigations. For lack of space, we can give here only a list of conjectures, claims and concise remarks. We shall elaborate on them in a forthcoming paper.

1. Coinductive characterizations of $\approx_\sigma^{app}$ are useful in factoring out the complexity of establishing observational equivalences between λ-terms. It would be interesting to compare the strength of coinduction principles to that of other tools, e.g. "approximation theorems" such as arise from "computationally adequate" mathematical models (see [20, 6, 11]). Here are some equivalences on which to test the power of coinduction principles:
   - divergent terms are $\approx_\sigma$ for $\sigma \in \{v, l, h, i\}$;
   - *black holes* (i.e. closed λ-terms M s.t. $\forall P.\ MP \to_\sigma^* M$) are $\approx_\sigma$ for $\sigma \in \{v, l, h, i, n, e\}$;
   - appropriate classes of fixed point operators are $\approx_\sigma$ for each σ;
   - many identities involving fixed points (e.g. the double iteration identity, i.e. $Fix(\lambda x.(Fix(\lambda y.fxy))) \approx_\sigma Fix(\lambda x.fxx)$), hold, for each σ, for appropriate classes of fixed points.
2. We have considered only coinductive characterizations of equivalence relations. We could have discussed, more generally, coinductive principles for establishing partial orders such as *observational approximation* and *applicative approximation*. These are obtained by replacing the bi-implication in the "equitermination" predicate by a simple implication.
3. All the final semantics that we have introduced do not yield "standard" denotational models for λ-calculus. Nevertheless, they can count as *compositional*, in that they induce observational equivalences, which are congruences w.r.t. the syntactical operators of the language. They can be seen to provide, in effect, alternative presentations of the, obviously *fully abstract*, term model. To this end it is useful to extend the equivalences considered to open λ-terms:
   Let $\sigma \in \{v, l, h, n, i, e\}$ and let $P, P' \in \Lambda$ be s.t. $FV(P, P') \subseteq \{x_1, \ldots, x_n\}$. We say that $P \approx_\sigma^{app} P'$ if and only if, for all $P_1, \ldots, P_n \in \Lambda^0$, $P[P_i/x_i] \approx_\sigma^{app} P'[P_i/x_i]$.
   Models could have been defined also using the technique of *processes as terms* introduced by J.J.M.M.Rutten (see [17, 19]).
4. We conjecture that both the *syntactical induction-coinduction* principle and the alternative final description of $\approx_v$ could be defined and shown to hold for all the observational equivalences discussed in the paper. For instance, the definitions and proofs, presented here for $\approx_v$, can be readily adapted to the case of $\approx_e$. The crucial fact is that both equivalences have "computationally adequate" inverse limit models with λ-definable projections. And hence one can obtain $\approx_\sigma$ (for $\sigma \in \{v, e\}$) by defining inductively on approximations a quotient of the interior of the model. This is shown for $\approx_v$ in [6]; the appropriate denotational model for $\approx_e$ is the one discussed in [11].
5. Purely set theoretic models in $ZF^- X_1$, where values are modeled by set-theoretic functions, can be readily obtained for $\approx_v$ and $\approx_l$ if we modify the definition of the functors F_v, F_l, G_v by replacing the disjoint union by the set-theoretic union and the atom $*$ by the empty set.

References

1. S.Abramsky, L.Ong, *Full Abstraction in the Lazy Lambda Calculus,* Information and Computation, 105(2):159–267, 1993.
2. P.Aczel, *Non-wellfounded sets,* Number 14, Lecture Notes CSLI, 1988.
3. P.Aczel, N.Mendler, *A final coalgebra theorem* Category Thepry and Computer Science Proceedings, D.Pitt et al. eds., Springer LNCS n.389:357-365, 1989.
4. H.Barendregt, *The Lambda Calculus, its Syntax and Semantics,* North Holland, Amsterdam, 1984.
5. M.Coppo, M.Dezani-Ciancaglini, M.Zacchi, *Type Theories, Normal Forms and D_∞-Lambda-Models,* Information and Computation, 72(2):85–116, 1987.
6. L.Egidi, F.Honsell, S.Ronchi Della Rocca, *Operational, denotational and logical Descriptions: a Case Study,* Fundamenta Informaticae, 16(2):149–169, 1992.
7. M.Fiore, *A Coinduction Principle for Recursive Data Types Based on Bisimulation,* 8th LICS Conference Proceedings, IEEE Computer Society Press:110-119, 1993.
8. M.Forti, F.Honsell, *Set Theory with Free Construction Principles,* Annali Scuola Normale Sup. Pisa, Cl. Sci., (IV), 10:493–522, 1983.
9. M.Forti, F.Honsell, M.Lenisa, *Processes and Hyperuniverses,* MFCS'94 Conference Proceedings, I.Privara et al. eds., Springer LNCS n.841:352-363, 1994.
10. F.Honsell, M.Lenisa, *Some Results on Restricted λ-calculi,* MFCS'93 Conference Proceedings, A.Borzyszkowski et al. eds., Springer LNCS n.711:84-104, 1993.
11. F.Honsell, S.Ronchi Della Rocca, *An approximation theorem for topological lambda models and the topological incompleteness of lambda calculus*, J. of Computer and System Sciences (45) 1:49-75, 1992.
12. R.Milner, *Operational and Algebraic Semantics of Concurrent Processes,* Handbook of Theoretical Computer Science, Ch.19, 1990.
13. C.H.L.Ong, *The lazy lambda calculus: an investigation into the foundations of functional programming,* Ph.D. thesis, Imperial College of Science and Technology, University of London, 1988.
14. A.M.Pitts, *Relational Properties of Recursively Defined Domains,* 8th LICS Conference Proceedings, IEEE Computer Society Press:86-97, 1993.
15. G.D.Plotkin, *Call-by-name, Call-by-value and the λ-calculus,* Theoretical Computer Science (1):125-159, 1975.
16. S.Ronchi Della Rocca, *International Summer School in Logic for Computer Science,* Chambery 28/6 – 9/7 1993, lecture notes.
17. J.J.M.M.Rutten, *Processes as terms: non-wellfounded models for bisimulation,* Math.Struct.Comp.Sci., 2(3):257–275, 1992.
18. J.J.M.M.Rutten, D.Turi, *On the Foundations of Final Semantics: Non-Standard Sets, Metric Spaces, Partial Orders,* REX Conference Proceedings, J.deBakker et al. eds., Springer LNCS n.666:477-530, 1993.
19. D.Turi, B.Jacobs, *On final Semantics for applicative and non-deterministic languages,* Fifth Biennial Meeting on Category Theory and Computer Science, Amsterdam, 1993.
20. C.P.Wadsworth, *The relation between computational and denotational properties for Scott's D_∞-models of the λ-calculus,* SIAM J. of Computing, 5(3):488-521 ,1976.

Appendix A

We recall some categorical definitions (for more details see [18]).

Definition 21 (F-coalgebra). Let C be a category and $F : C \to C$ an endofunctor.
i) A *F-coalgebra* is a pair (A, α), where A is an object of C and $\alpha : C \to F(C)$ is a morphism of C.
ii) Let C_F be the category whose objects are F-coalgebras and whose morphisms are F-coalgebra morphisms. A F-coalgebra morphism $f : (A, \alpha) \to (B, \beta)$ is an arrow $f : A \to B$ in the category C such that the following diagram commutes:

$$\begin{tikzcd} A \arrow[r, "\alpha"] \arrow[d, "f"'] & F(A) \arrow[d, "F(f)"] \\ B \arrow[r, "\beta"] & F(B) \end{tikzcd}$$

Definition 22 (F-bisimulation). Let C be a category with products. Let $F : C \to C$ be an endofunctor and (A, α) a F-coalgebra. A *F-bisimulation on the F-coalgebra* (A, α) is a relation $R \subseteq A \times A$ such there exists an arrow $\gamma : R \to F(R)$ which makes the following diagram commutes:

$$\begin{tikzcd} A \arrow[d, "\alpha"'] & R \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] \arrow[d, "\gamma"] & A \arrow[d, "\alpha"] \\ F(A) & F(R) \arrow[l, "F(\pi_1)"'] \arrow[r, "F(\pi_2)"] & F(A) \end{tikzcd}$$

In the following definition we recall the Antifoundation Axiom FCU and the Super Strong Extensionality Axiom up to the set of atoms U ($SSExtU$), which is a consequence of FCU and gives an interesting characterization of the equality between sets (for more details see [8]):

Definition 23. i) Unique Free Construction with respect to a set U of atoms FCU:
Let X be a set such that $X \cap U = \emptyset$. For every function $f : X \to \mathcal{P}(X \cup U)$ there is a unique function $g : X \to V$ verifying

$$g(x) = (f(x) \cap U) \cup \{g(y) \mid y \in f(x) \cap X\}, \quad \forall x \in X.$$

ii) Super Strong Extensionality Axiom up to the set of atoms U $SSExtU$:
Let V be the set theoretic Universe. Two sets X, Y are equal if and only if there exists a relation $R \subseteq V \times V$ such that $(X, Y) \in R$ and $R \subseteq (R)^+$, where $(\ )^+$ is the operator on relations defined by

$$\begin{aligned}(R)^+ = \{(X, Y) \mid\ & X \cap U = Y \cap U \quad \text{and} \\ & \forall W \in (f(X) \setminus U). \exists Z \in (Y \setminus U).\ (W, Z) \in R \quad \text{and} \\ & \forall W \in (Y \setminus U). \exists Z \in (X \setminus U).\ (W, Z) \in R\}.\end{aligned}$$

(Categorically, R is a $\mathcal{P}(\ ) + U$-bisimulation on the coalgebra $(V,\ id)$.)

This article was processed using the LaTeX macro package with LLNCS style