The discriminating power of multiplicities in the λ-calculus

Gérard Boudol, Laneve Cosimo

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

The discriminating power of multiplicities in the λ-calculus

Gérard Boudol and Cosimo Laneve

N° 2441

Décembre 1994

PROGRAMME 2

Calcul symbolique,

programmation

et génie logiciel

Rapport de recherche

1995

ISSN 0249-6399

The discriminating power of multiplicities in the λ-calculus *

Gérard Boudol and Cosimo Laneve**

Programme 2 — Calcul symbolique, programmation et génie logiciel
Projet Meije

Rapport de recherche n°2441 — Décembre 1994 — 41 pages

Abstract: The λ-calculus with multiplicities is a refinement of the lazy λ-calculus where the argument in an application comes with a multiplicity, which is an upper bound to the number of its uses. This introduces potential deadlocks in the evaluation. We study the discriminating power of this calculus over the usual λ-terms. We prove in particular that the observational equivalence induced by contexts with multiplicities coincides with the equality of Lévy-Longo trees associated with λ-terms. This is a consequence of the characterization we give of the corresponding observational precongruence, as an intensional preorder involving η-expansion, namely Ong's lazy Plotkin-Scott-Engeler preorder.

Key-words: functional and concurrent languages, semantics, lambda-calculus

(Résumé : tsvp)

*Partially supported by the ESPRIT Basic Research Project 6454 - CONFER.
**Present address: Università di Bologna, Dipartimento di Matematica, Piazza di Porta San Donato 5, 40126 Bologna, Italy.

Unité de recherche INRIA Sophia-Antipolis
2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex (France)
Téléphone : (33) 93 65 77 77 – Télécopie : (33) 93 65 77 65

Le pouvoir de discrimination des multiplicités dans le λ-calcul

Résumé : Le λ-calcul avec multiplicités est un raffinement du λ-calcul faible, où l'argument dans une application est muni d'une multiplicité qui indique combien de fois on peut l'utiliser. Ceci introduit des blocages potentiels dans l'évaluation. Nous étudions le pouvoir de discrimination de ce calcul sur les λ-termes usuels. Nous montrons en particulier que l'équivalence observationnelle que l'on obtient coïncide avec l'égalité des arbres de Lévy-Longo associés aux λ-termes. Ceci est une conséquence de la caractérisation que nous obtenons pour la précongruence observationnelle, comme un préordre intensionnel qui comporte la η-expansion. Ce préordre a été introduit par Ong sous le nom de préordre de Plotkin-Scott-Engeler.

Mots-clé : langages fonctionnels et parallèles, sémantique, lambda-calcul

1 Introduction

The λ-calculus with multiplicities was introduced in [5] to the purpose of studying the relationship between the λ-calculus and Milner's π-calculus [12]. It is a "resource conscious" refinement of the λ-calculus, based on the following observation: in an application MN the argument N is *infinitely available* for the function M. This appears clearly in the process of β-reduction: when M is an abstraction λxR, the application MN reduces to $R[{}^{N}/_{x}]$ where the argument is copied within R as many times as there are free occurrences of x. One cannot predict the "multiplicity" of x in R, because R could be reduced to another term where this variable is duplicated. For instance if $R = (\mathbf{2}x)$ where $\mathbf{2} = \lambda fy.f(fy)$, then the variable x occurs just once in R, but R reduces to a term where x appears twice.

In our refinement of the λ-calculus, any argument comes with an explicit, finite or infinite *multiplicity*. Namely, we write MN^m where $m \in \mathsf{N} \cup \{\infty\}$, meaning that N is available *at most m times* for the function M. As a particular case, we get the usual λ-terms, where all the multiplicities are infinite – in which case we may omit them, to keep the standard notation. For example, in $(\lambda x(\mathbf{2}x))\mathbf{I}^1$, the argument $\mathbf{I} = \lambda zz$ (the identity) is available only once. We have to be careful in defining the reduction process, since reducing this term to $\mathbf{2I}$ would mean to transform a limited resource into an infinite one, for $\mathbf{2}x$ really stands for $\mathbf{2}x^\infty$. Obviously this is not what we want. Then we have to *delay* in some sense the use of the resource, until something is really needed for the variable it is bound to.

To this purpose, it is convenient to use the notion of *explicit substitution* of Curien et al. [1]. That is, we extend the syntax with the construct $M\langle {}^{N^m}/_{x}\rangle$ meaning that N is substituted for x in M at most m times, and we modify the β-reduction rule in the obvious way:

$$(\lambda xM)N^m \;\rightarrow\; M\langle {}^{N^m}/_{x}\rangle$$

In this paper we adopt the "lazy" regime of reduction of Abramsky and Ong [2], where, in order to compute MN^m, one has first to evaluate M, hopefully to an abstraction. Similarly, to compute a term $M\langle {}^{N^m}/_{x}\rangle$ one first computes M. Then one fetches a sample N of the resource (if any, that is if $m > 0$), leaving the rest for future use, whenever the computation cannot proceed without a value for x, that is whenever x occurs in the *head position* in M. In this case $M = xQ_1 \cdots Q_k$, where the Q_i's are either arguments with multiplicity R^m or substitution items $\langle {}^{R^m}/_{y_i}\rangle$ $(y_i \neq x)$, and the following reduction takes place:

$$xQ_1 \cdots Q_k\langle {}^{N^{m+1}}/_{x}\rangle \;\rightarrow\; NQ_1 \cdots Q_k\langle {}^{N^m}/_{x}\rangle$$

Defined in this way, the reduction process is entirely *deterministic*: for any term there is at most one reduction that can be performed from it in one step.

Clearly what is new, with respect to the usual λ-calculus, is the possibility of *deadlock*: if something is needed for x, but there is no resource available for x, then no reduction is possible. For instance, we have:

$$\begin{array}{rcl}(\lambda x.xx)\mathbf{I}^1 & \rightarrow & (xx)\langle \mathbf{I}^1/x\rangle \\ & \rightarrow & (\mathbf{I}x)\langle \mathbf{I}^0/x\rangle \\ & \rightarrow & z\langle x^\infty/z\rangle\langle \mathbf{I}^0/x\rangle \\ & \rightarrow & x\langle x^\infty/z\rangle\langle \mathbf{I}^0/x\rangle\end{array}$$

However, we do not wish to regard a normal form such as $x\langle x^\infty/z\rangle\langle \mathbf{I}^0/x\rangle$ as a meaningful value. Here, as in the lazy λ-calculus, a *value* is a normal form that waits for an input. In other words, a value is an abstraction, up to the identification of $(\lambda xM)\langle N^m/y\rangle$ with $\lambda x(M\langle N^m/y\rangle)$. This allows us to define the *observational semantics* of the calculus, namely the preorder $M \preceq_{\mathsf{m}} N$, as follows:

> $M \preceq_{\mathsf{m}} N$ if and only if for any context (with multiplicities) C, if $C[M]$ reduces to a value, then $C[N]$ reduces to a value, too.

Using Albert Meyer's terminology, one can read "$C[M]$ reduces to a value" as "$C[M]$ gives back the prompt". Then there are two ways for a term of not "giving back the prompt": either it diverges, having an infinite computation, or its evaluation ends up with a deadlock, that is a normal form which is not observable. In our observational semantics we do not distinguish deadlock from divergence.

Our purpose in this paper is to determine precisely to which extent the λ-calculus with multiplicities is a refinement of the usual λ-calculus. That is, we shall study and characterize the discriminating power of contexts with multiplicities over the λ-terms. In more technical terms, we will study the *restriction to λ-terms* of the preorder $\preceq_{\mathsf{m}}$.

Our results are as follows: foremost, it is immediate that $\preceq_{\mathsf{m}}$ is strictly finer than the observational preorder, denoted $\preceq_\ell$, that we get by restricting the contexts to be λ-calculus contexts, with infinite multiplicities (this is the preorder defined by Abramsky and Ong [2]). For instance, $x(\lambda y.xy) \preceq_\ell xx$, while $x(\lambda y.xy) \not\preceq_{\mathsf{m}} xx$ since $(x(\lambda y.xy))\langle \mathbf{I}^1/x\rangle$ reduces to the value $\lambda y((xy)\langle x^\infty/z\rangle\langle \mathbf{I}^0/x\rangle)$, whereas this is not the case for $xx\langle \mathbf{I}^1/x\rangle$, as we have seen above. In other words, the lazy λ-calculus is sensitive to the lack of resources. This is not very surprising. As we will see, the extra discriminating power of finite multiplicities only shows up when applied to λ-terms exhibiting themselves some multiplicity: if M and N are two *affine* λ-terms, which use resources at most once, that is terms where any variable (free or bound) has at most one occurrence, then

$$M \preceq_{\mathsf{m}} N \Leftrightarrow M \preceq_\ell N$$

We also examine the possible weakenings of the theory $\preceq_{\mathsf{m}}$, by adding new axioms. We show that no such weakening can be as weak as $\preceq_\ell$.

Our main result is the characterization of the preorder $\preceq_{\mathsf{m}}$ over λ-terms. We show that it coincides with the *lazy Plotkin-Scott-Engeler preorder* introduced by Ong in [13]. This is an ordering on an intensional representation of λ-terms, the so-called Lévy-Longo trees. These are like Böhm trees, fitted in with the lazy regime where any divergent term as $\Omega =$

$(\lambda x\, xx)(\lambda x\, xx)$ is different from $\lambda x\, \Omega$. The lazy PSE ordering on these trees is basically the prefix ordering, for which Ω is less than everything, together with the facts that η-expansion is increasing, that is $M \preceq_{\mathsf{m}} \lambda x.Mx$ (x is not free in M), and that anything is less than a term of infinite order – typically, Ξ such that $\Xi =_\beta \lambda x\, \Xi$, for instance $\Xi = (\lambda fx.ff)(\lambda fx.ff)$.

The lazy PSE ordering was introduced by Ong to characterize the "local structure" (following Barendregt's terminology) of some models of the lazy λ-calculus. An immediate consequence of our main result and of results by Ong and Abramsky (namely the Theorem 3.4.1.3 of [13] and Proposition 7.2.10 of [2]) is that finite multiplicities provide us with strictly more discriminating power than the convergence testing combinators, introduced by Abramsky and Ong [2] to make the lazy λ-calculus "complete" in some sense. Regarding the *parallel* convergence testing combinator, this may be surprising because there is no parallel feature in our λ-calculus with multiplicities. The same remark holds as well if, instead of using this combinator, we use a non-deterministic choice, as in [4], or a parallel composition of functions, as in [6]. Therefore the sensitivity of the λ-calculus to the lack of resources is much greater than one could think.

Another consequence of our characterization result is that the observational *equivalence* $M \simeq_{\mathsf{m}} N$ over λ-terms, meaning that for any context C with multiplicities, $C[M]$ has a value if and only if $C[N]$ has a value, coincides with the equality of the associated Lévy-Longo trees. From this and previous results of Sangiorgi [14], we can draw some conclusions. Sangiorgi studied the equivalence $M \simeq_\pi N$ induced by Milner's encoding of the λ-calculus into the π-calculus [12], and he showed in particular that this coincides with the equality of the associated Lévy-Longo trees. We can then conclude that, as far as the λ-calculus is concerned, the π-calculus and the λ-calculus with multiplicities have the same discriminating power:

$$M \simeq_\pi N \;\Leftrightarrow\; M \simeq_{\mathsf{m}} N$$

Again, this may be surprising because the latter is a deterministic calculus, with no parallel facility. We must also point out that Sangiorgi used a kind of bisimulation as the semantic equivalence, while we use the much weaker notion of observational equivalence. Nevertheless, our results show that even if we use an observational equivalence for the π-calculus, we still keep the same induced semantics on λ-terms, namely the equality of Lévy-Longo trees.

Sangiorgi also showed that one cannot go beyond $\simeq_\pi$ by extending the contexts using "well-formed operators", while adding a unary non-deterministic operator $\uplus M$ with evaluation rules

$$\uplus M \;\rightarrow\; M \qquad\qquad \text{and} \qquad\qquad \uplus\, M \;\rightarrow\; \Omega$$

is enough to get the full discriminating power of the π-calculus. Note that this operator has some flavour of introducing potential deadlocks, since $\uplus M \rightarrow \Omega$ means that M, regarded as a resource, may vanish. However, this is only true if we defer a part of the discriminating power to the semantic equivalence itself, using the bisimulation for instance. Sangiorgi then concluded that "*non-determinism is exactly what is necessary to add to the λ-calculus to make it as discriminating as the π-calculus*". As far as one is committed to use "well-formed operators", while being allowed to use a bisimulation semantics, this is true. However, using explicit substitutions, which provide us with a computationally meaningful construct, and still using an observational semantics, which in non-deterministic calculi is usually far less discriminating than bisimulation, we may have a different conclusion: *the possibility of deadlocks is essentially what the π-calculus adds to the lazy λ-calculus.*

2 The λ-calculus with multiplicities

2.1 Syntax

As usual, we assume given a countable set Var of variables, ranged over by x, y, $z\ldots$ The set Λ_{m} of terms of the λ-calculus with multiplicities, or λ_{m}-terms, is generated by the following grammar:

$$E \ ::= \ x \ \mid \ \lambda x E \ \mid \ (EE^m) \ \mid \ (E\langle E^m/x\rangle)$$

where m is a positive integer, or the infinite multiplicity, that is $m \in \mathsf{N} \cup \{\infty\}$. To avoid any confusion with usual λ-terms, denoted by M, $N\ldots$, we use E, $F\ldots$ to range over λ_{m}-terms (here "m" stands for "multiplicities". In [5] we also introduced a λ_{r}-calculus, i.e. a λ-calculus with *resources*). However, we shall most often omit the infinite multiplicity, writing EF and $E\langle F/x\rangle$ for EF^∞ and $E\langle F^\infty/x\rangle$ respectively. That is, we regard ordinary λ-terms as particular cases of terms with multiplicities. We call E^m a *bag*, made of m copies of the term E. The set of bags is denoted by Π. We use P, Q, $R\ldots$ to range over bags, or substitutions $\langle E^m/x\rangle$. As it is standard, we abbreviate $\lambda x_1 \cdots \lambda x_n E$ into $\lambda x_1 \cdots x_n.E$ and use $EP_1 \cdots P_k$ to denote $(\cdots(EP_1)\cdots P_k)$, where the P_i's are bags or substitutions.

The notions of free and bound variables are the standard ones (see [3]), augmented by the following items:

- the free variables of a bag E^m or a substitution $\langle E^m/x\rangle$ are the free variables of E;
- in $E\langle P/x\rangle$ every free occurrence of x in E is bound by the substitution $\langle P/x\rangle$.

We denote by $\mathsf{fv}(E)$ (resp. $\mathsf{bv}(E)$) the set of free (resp. bound) variables of a term E, and by Λ^0_{m} the set of closed terms. The set of closed bags is denoted Π^0.

As usual, a (λ_{m}-)*context* is any term built using the syntax of λ_{m}-terms, plus an additional constant $[]$, the hole. Filling the hole with a term E in a context C results in a term denoted $C[E]$. Note that free variables of E may be bound by the context in $C[E]$. We denote the set of contexts by $\Lambda_{\mathsf{m}}[]$, and use C, $D\ldots$ to range over λ_{m}-contexts.

2.2 Syntactic equality

We shall consider λ_{m}-terms up to α-conversion. To define the syntactic equality of terms, we use the (partial) operation of *renaming* $E[y/x]$ where y is neither free nor bound in E.

This is defined as follows:

$$z[y/x] = \begin{cases} y & \text{if } z = x \\ z & \text{otherwise} \end{cases}$$

$$(\lambda z E)[y/x] = \begin{cases} \lambda z E & \text{if } z = x \\ \lambda z (E[y/x]) & \text{otherwise} \end{cases}$$

$$(EF^m)[y/x] = (E[y/x])(F[y/x])^m$$

$$(E\langle F^m/z\rangle)[y/x] = \begin{cases} E\langle (F[y/x])^m/z\rangle & \text{if } z = x \\ (E[y/x])\langle (F[y/x])^m/z\rangle & \text{otherwise} \end{cases}$$

Now the *syntactic equality* of λ_{m}-terms is the congruence $\equiv$ generated by the following axioms:

$\lambda x E = \lambda y (E[y/x])$	$y \notin \mathsf{fv}(E) \cup \mathsf{bv}(E)$
$E\langle P/x\rangle = (E[y/x])\langle P/y\rangle$	$y \notin \mathsf{fv}(E) \cup \mathsf{bv}(E)$
$(\lambda x E)\langle P/y\rangle = \lambda x (E\langle P/y\rangle)$	$x \notin \mathsf{fv}(P) \cup \{y\}$

The third axiom has been added because it is more convenient to deal with normal forms which are simply abstractions, rather than "closures" $(\lambda x E)\langle P_1/x_1\rangle \cdots \langle P_k/x_k\rangle$. It is immediate, by induction on the definition of $E \equiv F$ (including the implicit axioms of reflexivity, symmetry, transitivity and congruence), that $E \equiv F \Rightarrow \mathsf{fv}(E) = \mathsf{fv}(F)$.

Proposition 2.1

1. *$\lambda x E \equiv (\lambda y F)\langle P/z\rangle$ implies $E = E'\langle Q/z'\rangle$ and $E'[w/x]\langle Q/z'\rangle \equiv F[w/y]\langle P/z\rangle$;*
2. *$\lambda x E \equiv \lambda y F$ implies $E[z/x] \equiv F[z/y]$;*
3. *$E\langle P/x\rangle \equiv F\langle Q/y\rangle$ implies $P \equiv Q$ and $E[z/x] \equiv F[z/y]$;*
4. *$EP \equiv F$ implies $F = F'Q$ and $E \equiv F'$ and $P \equiv Q$; if $E \equiv F$ and $E = \lambda x E'$ or $E = E'\langle P/x\rangle$ then $F = \lambda y F$ or $F = F'\langle Q/y\rangle$.*

PROOF: By induction on the inference of the syntactic equivalence, straightforward. ■

2.3 The reduction relation

Evaluation in the λ-calculus with multiplicities follows the *lazy* strategy of Abramsky and Ong [2], where one does not evaluate the body M of an abstraction $\lambda x M$, nor the argument N in an application MN. However, there are some differences, mainly because we perform the substitutions explicitly, and in a delayed manner. As we said in the introduction, the substitution is performed in the following way:

$$xQ_1 \cdots Q_k\langle N^{m+1}/x\rangle \rightarrow NQ_1 \cdots Q_k\langle N^m/x\rangle$$

provided that no Q_i is a substitution for x, and x is not free in N. To formalize this reduction by means of rules, we introduce an auxiliary evaluation relation $\leadsto$, which is only defined for terms of the form $E\langle F/x\rangle$ where x occurs in the head position in E. Then $E\langle F/x\rangle \leadsto E'$ means that E' is E where F is placed in the head position (provided that no free variable is captured).

Definition 2.2 *The* reduction relation $\rightarrow_{\mathsf{m}}$ *on* Λ_{m} *is the least one satisfying the following rules:*

$$
\begin{array}{|ccc|}
\hline
(\lambda x E)P \rightarrow E\langle P/x\rangle & \dfrac{E \rightarrow E'}{EP \rightarrow E'P} & \dfrac{E \rightarrow E'}{E\langle P/x\rangle \rightarrow E'\langle P/x\rangle} \\[2ex]
\multicolumn{2}{|c}{\dfrac{E\langle F/x\rangle \leadsto E'}{E\langle F^{m+1}/x\rangle \rightarrow E'\langle F^{m}/x\rangle}\ x \notin \mathsf{fv}(F)} & \dfrac{F \equiv E \qquad E \rightarrow E'}{F \rightarrow E'} \\
\hline
\end{array}
$$

where the fetch relation $\leadsto$ *is defined by:*

$$
\begin{array}{|ccc|}
\hline
x\langle E/x\rangle \leadsto E & \dfrac{E\langle F/x\rangle \leadsto E'}{(EP)\langle F/x\rangle \leadsto E'P} & \dfrac{E\langle F/x\rangle \leadsto E'}{(E\langle P/z\rangle)\langle F/x\rangle \leadsto E'\langle P/z\rangle}\ z \neq x \text{ and } z \notin \mathsf{fv}(F) \\
\hline
\end{array}
$$

One should note that, to infer $E \rightarrow_{\mathsf{m}} E'$, one often has to rename bound variables in E to fulfil the requirements about bound variables in the rules. For instance, we have:

$$(x(xy))\langle x/y\rangle\langle (xy)^2/x\rangle \;\equiv\; (z(zu))\langle z/u\rangle\langle (xy)^2/z\rangle \;\rightarrow_{\mathsf{m}} ((xy)(zu))\langle z/u\rangle\langle (xy)^1/z\rangle$$

The following proposition guarantees that $\rightarrow_{\mathsf{m}}$ is deterministic, up to $\equiv$. That is, if E reduces into F and G then $F \equiv G$.

Proposition 2.3 *The relation $\equiv$ is consistent w.r.t. $\rightarrow_{\mathsf{m}}$. Namely*

$$E \equiv F \;\&\; E \rightarrow_{\mathsf{m}} E' \;\&\; F \rightarrow_{\mathsf{m}} F' \;\Rightarrow\; E' \equiv F'.$$

In particular, if $E \rightarrow_{\mathsf{m}} E'$ and $E \rightarrow_{\mathsf{m}} F'$ then $E' \equiv F'$.

PROOF: The interesting case is when $F \rightarrow_{\mathsf{m}} F'$ is not due to the last rule of Definition 2.2. Under this assumption and by Proposition 2.1, both E and F must have an abstraction or a variable in the head position. We check the determinacy for these two cases, focusing on the cases where E and F are redexes, since the general case follows by congruence.

1. $(\lambda x E)P \equiv (\lambda y F)Q \;\Rightarrow\; E\langle P/x\rangle \equiv F\langle Q/y\rangle$. By Proposition 2.1(4), $(\lambda x E)P \equiv (\lambda y F)Q$ implies $\lambda x E \equiv \lambda y F$ & $P \equiv Q$. Now let $z \notin \mathsf{fv}(E) \cup \mathsf{bv}(E) \cup \mathsf{fv}(F) \cup \mathsf{bv}(F)$. Then $\lambda z E[z/x] \equiv \lambda x E \equiv \lambda y F \equiv \lambda z F[z/y]$. Moreover, by Proposition 2.1(2), $E[z/x] \equiv F[z/y]$. Hence, by congruence, $E[z/x]\langle P/z\rangle \equiv F[z/y]\langle Q/z\rangle$. The result is obtained by noticing that $E\langle P/x\rangle \equiv E[z/x]\langle P/z\rangle$ and $F\langle Q/y\rangle \equiv F[z/y]\langle Q/z\rangle$.

2. $E\langle {}^{F}/_{x}\rangle \leadsto E'$ & $G\langle {}^{H}/_{x}\rangle \leadsto G'$ & $E \equiv G$ & $F \equiv H \;\Rightarrow\; E' \equiv G'$. Actually, besides this statement, we shall also prove that

   $$(*) \qquad E\langle {}^{F}/_{x}\rangle \leadsto E' \;\Rightarrow\; E[{}^{w}/_{z}]\langle {}^{F}/_{x}\rangle \leadsto E'[{}^{w}/_{z}],\;\; w \notin \mathsf{fv}(E)\cup\mathsf{bv}(E)\cup\mathsf{fv}(F)\cup\{x\}$$

   We proceed by induction on the proof of $E\langle {}^{F}/_{x}\rangle \leadsto E'$.

   (a) The basic case is $E = x$ and $E' = F$. Then $G = x$ is an immediate consequence of $E \equiv G$. Hence $E' \equiv G'$ follows from $F \equiv H$. Also $(*)$ is straightforward since $z \notin \mathsf{fv}(F)$ (hence $z \notin \mathsf{fv}(H)$ because $F \equiv H$).

   (b) If $E\langle {}^{F}/_{x}\rangle \leadsto E'$ is proved using the second rule of the fetch relation, then $E' \equiv G'$ follows from to Proposition 2.1(4) and the inductive hypothesis.

   (c) Otherwise $E = E''\langle {}^{P}/_{z}\rangle$, and $G = G''\langle {}^{P'}/_{z'}\rangle$, by Proposition 2.1(4) and the fact that the fetch relation is not defined on abstractions. Let $w \notin \mathsf{fv}(F)\cup\{x\}\cup \mathsf{fv}(E)\cup\mathsf{bv}(E)\cup\mathsf{fv}(G)\cup\mathsf{bv}(G)$. By Proposition 2.1(3) $E''[{}^{w}/_{z}] \equiv G''[{}^{w}/_{z'}]$; hence since by hypothesis $E''\langle {}^{F}/_{x}\rangle \leadsto E^{+}$ and $G''\langle {}^{H}/_{x}\rangle \leadsto G^{+}$, we have by $(*)$, $E''[{}^{w}/_{z}]\langle {}^{F}/_{x}\rangle \leadsto E^{+}[{}^{w}/_{z}]$ and $G''[{}^{w}/_{z'}]\langle {}^{H}/_{x}\rangle \leadsto G^{+}[{}^{w}/_{z}]$. Again, by inductive hypothesis, $E^{+}[{}^{w}/_{z}] \equiv G^{+}[{}^{w}/_{z}]$. The equivalence $E' \equiv G'$ follows easily by congruence.

   We let the reader check that $(*)$ holds. ■

3 The observational semantics

3.1 Observational preorder and the context lemma

In this section we introduce the observational semantics of our λ-calculus with multiplicities. It is an instance of the standard Morris' "extensional operational semantics" (see [3], Exercise 16.5.5, and [9], Chapter IV). The idea is to use the syntactic machinery to "test" the expressions by plugging them into contexts, and looking for an observable result, that is, a value. Then an expression is better than another one if it passes successfully more tests.

A term E is a *value* if there exist x and F such that $E \equiv \lambda x F$. Let $\xrightarrow{*}_{\mathsf{m}}$ be the reflexive and transitive closure of $\rightarrow_{\mathsf{m}}$. Then the convergence and divergence predicates on closed terms are defined, as usual, by:

$$\begin{array}{lll} E \Downarrow_{\mathsf{m}} \Leftrightarrow \exists E' \text{ value. } E \xrightarrow{*}_{\mathsf{m}} E' & & (\text{“}E \text{ converges”}) \\ E \Uparrow_{\mathsf{m}} \Leftrightarrow \neg(E \Downarrow_{\mathsf{m}}) & & (\text{“}E \text{ diverges”}) \end{array}$$

Notice that $E \Uparrow_{\mathsf{m}}$ does not necessarily mean, as in the lazy λ-calculus, that the evaluation of E does not terminate. Indeed, E "diverges" if its evaluation ends up with a *deadlock*, typically a term of the form $xQ_1 \cdots Q_k\langle {}^{E^0}/_{x}\rangle R_1 \cdots R_n$ (where no Q_i is a substitution for x). However, a deadlock is regarded as semantically the same as a truly divergent term, such as $\Omega = (\lambda x\, xx^{\infty})(\lambda x\, xx^{\infty})^{\infty}$. It should be obvious that

$$F \equiv E \;\Rightarrow\; (E \Downarrow_{\mathsf{m}} \Leftrightarrow\; F \Downarrow_{\mathsf{m}})$$

Definition 3.1 *The* observational preorder *is the relation* $\preceq_{\mathsf{m}}$ *on* Λ_{m} *defined as follows:*

$$E \preceq_{\mathsf{m}} F \overset{\text{def}}{\iff} \forall C.\ C[E], C[F] \in \Lambda^0_{\mathsf{m}} \ \Rightarrow\ (C[E] \Downarrow_{\mathsf{m}} \Rightarrow\ C[F] \Downarrow_{\mathsf{m}})$$

Two terms E and F are observationally equivalent*, in notation $E \simeq_{\mathsf{m}} F$, whenever $E \preceq_{\mathsf{m}} F$ and $F \preceq_{\mathsf{m}} E$.*

Due to the universal quantification over contexts, the definition of the observational preorder is not very manageable: it is usually quite difficult to prove or disprove an inequation $E \preceq_{\mathsf{m}} F$. In [11], Milner stated and proved a property called the *context lemma*, which was then generalized to the λ-calculus by Lévy [9], establishing that, in order to "test" a (closed) λ-term, it is enough to apply it. We now show that this also holds in the λ-calculus with multiplicities (see also [5]). To this end, we first introduce a restricted kind of contexts, the *applicative contexts*, ranged over by K, $L \ldots$ These are given by the grammar:

$$K \ ::= \ [] \ \mid \ (KP) \ \mid \ K\langle P/x \rangle$$

where P is any bag. This allows us to define the *applicative testing* preorder, which is the observational preorder restricted to applicative contexts, that is:

$$E \preceq^{\mathcal{A}}_{\mathsf{m}} F \overset{\text{def}}{\iff} \forall K.\ K[E], K[F] \in \Lambda^0_{\mathsf{m}} \ \Rightarrow\ (K[E] \Downarrow_{\mathsf{m}} \Rightarrow\ K[F] \Downarrow_{\mathsf{m}})$$

Before proving the context lemma, establishing that the two preorders $\preceq_{\mathsf{m}}$ and $\preceq^{\mathcal{A}}_{\mathsf{m}}$ coincide, we need some auxiliary results.

Lemma 3.2 *Let K be an applicative context and x a variable, not free in K. Then for any finite set U of variables and any $z \notin U \cup \{x\}$, there exists an applicative K' such that*

$$\mathsf{fv}(E) \cup \mathsf{bv}(E) \subseteq U \ \Rightarrow\ K[E[z/x]] \equiv (K'[E])[z/x]$$

PROOF: By induction on K. This is trivial for $K = []$. If $K = (LP)$ then x is not free in L nor P, and $K[E[z/x]] = (L[E[z/x]])P$. By induction hypothesis,

$$K[E[z/x]] \equiv (L'[E])[z/x]P$$

for some L', therefore $K[E[z/x]] \equiv ((L'[E])P)[z/x]$.

If $K = L\langle P/y \rangle$, there are two cases: if $y = x$, then we let $L' = L[v/x]$ for some fresh v (in particular, v is not in $U \cup \{z\}$). We have $K[E[z/x]] \equiv L'[E[z/x]]\langle P/v \rangle$. By induction hypothesis, there exists L'' such that $L'[E[z/x]] \equiv (L''[E])[z/x]$, therefore

$$K[E[z/x]] \equiv (L''[E]\langle P/v \rangle)[z/x]$$

Otherwise ($y \neq x$), x cannot be free in L (nor in P). Then by induction hypothesis there exists L' such that $K[E[z/x]] \equiv (L'[E][z/x])\langle P/y \rangle$. Again, there are two subcases: if $y = z$ then $K[E[z/x]] \equiv (L'[E])\langle P/x \rangle$, therefore $K[E[z/x]] \equiv (L'[E])\langle P/x \rangle[z/x]$ since x is not free in P. Otherwise ($y \neq z$), we have $K[E[z/x]] \equiv (L'[E])\langle P/y \rangle[z/x]$. ∎

An immediate consequence of this lemma is:

Corollary 3.3 *If $E \preceq^{\mathcal{A}}_{\mathsf{m}} F$ and $z \neq x$ is neither free nor bound in E and F, then*

$$E[z/x] \preceq^{\mathcal{A}}_{\mathsf{m}} F[z/x]$$

Now we prove the context lemma – thanks to the explicit substitutions, one may even prove it for arbitrary terms (that is, not necessarily closed).

Lemma 3.4 (The Context Lemma)

$$E \preceq_{\mathsf{m}} F \;\Leftrightarrow\; E \preceq^{\mathcal{A}}_{\mathsf{m}} F$$

PROOF: The direction "$\Rightarrow$" is obvious. To establish "$\Leftarrow$", we use the notion of a *multiple context*, that is the notion of a context where there may be several kinds of holes, indexed by positive integers, i.e. $[]_i$. For any such context involving only holes whose indexes are less than k, we define $C[E_1, \ldots, E_k]$ in the obvious way, that is by filling the hole $[]_i$ by the corresponding term E_i. We shall also use the notation $C[\widetilde{E}]$ for $C[E_1, \ldots, E_k]$. We show the following:

$$E_1 \preceq^{\mathcal{A}}_{\mathsf{m}} F_1 \;\&\; \ldots \;\&\; E_k \preceq^{\mathcal{A}}_{\mathsf{m}} F_k \;\&\; C[E_1, \ldots, E_k] \Downarrow_{\mathsf{m}} \;\Rightarrow\; C[F_1, \ldots, F_k] \Downarrow_{\mathsf{m}}$$

where $C[\widetilde{E}]$ and $C[\widetilde{F}]$ are closed. Assuming that $C[\widetilde{E}] \xrightarrow{*}_{\mathsf{m}} V$ for some value V, we show that $C[\widetilde{F}] \Downarrow_{\mathsf{m}}$, by induction on (l, h), w.r.t. the lexicographic ordering, where l is the length of the evaluation sequence $C[\widetilde{E}] \xrightarrow{*}_{\mathsf{m}} V$, and h is the number of occurrences of holes in C. We may write $C = C_0 C_1 \cdots C_m$ where C_0 is either a hole $[]_i$, or a variable x, or an abstraction context $\lambda x\, B$, and the C_j's, for $j > 0$, are bags or substitution contexts. We examine the possible cases (this proof technique is directly adapted from Lévy's one [9], with the notable difference that we are dealing with open terms).

$C_0 = \lambda x\, B$.

There are two sub-cases. If $l = 0$ then $C[\widetilde{E}] = V$ is a value, therefore $C_1[\widetilde{E}], \ldots, C_m[\widetilde{E}]$ are substitution items, and $C[\widetilde{F}]$ is a value too.

Otherwise ($l > 0$), there exists j ($1 \leq j \leq m$) such that $C_1, \ldots, C_{j-1}$ are substitution items, say $\langle D_1/x_1 \rangle, \ldots, \langle D_{j-1}/x_{j-1} \rangle$, and $C_j[\widetilde{E}]$ is a bag. Notice that

$$(\lambda x B[\widetilde{E}])\langle D_1[\widetilde{E}]/x_1 \rangle \cdots \langle D_{j-1}[\widetilde{E}]/x_{j-1} \rangle \equiv \lambda z((B[\widetilde{E}])[z/x]\langle D_1[\widetilde{E}]/x_1 \rangle \cdots \langle D_{j-1}[\widetilde{E}]/x_{j-1} \rangle)$$

for some fresh variable z. It should be clear that there exists a context B', obtained from B by renaming x by z, and by replacing some of the holes $[]_i$ by $[]_{i+k}$, such that $(B[\widetilde{E}])[z/x] = B'[\widetilde{E}, \widetilde{E}']$ where $E'_i = E_i[z/x]$. By definition of the reduction relation, the term

$$G = (\lambda z(B'[\widetilde{E}, \widetilde{E}']\langle D_1[\widetilde{E}]/x_1 \rangle \cdots \langle D_{j-1}[\widetilde{E}]/x_{j-1} \rangle))C_j[\widetilde{E}] \cdots C_m[\widetilde{E}]$$

reduces to V by an evaluation sequence of length l. Let

$$C' = B'\langle D_1/x_1\rangle \cdots \langle D_{j-1}/x_{j-1}\rangle\langle C_j/z\rangle C_{j+1}\cdots C_m$$

Then $G \to_{\mathsf{m}} C'[\widetilde{E},\widetilde{E}']$ and $C[\widetilde{F}] \to_{\mathsf{m}} C'[\widetilde{F},\widetilde{F}']$ where $F'_i = F_i[z/x]$, therefore we may apply the induction hypothesis to conclude, since $E_i \preceq^{\mathcal{A}}_{\mathsf{m}} F_i \;\Rightarrow\; E'_i \preceq^{\mathcal{A}}_{\mathsf{m}} F'_i$ by the Corollary 3.3.

$C_0 = x$.

In this case $l > 0$ since $C[\widetilde{E}]$ is not a value. Moreover, there exists j $(1 \leq j \leq m)$ such that $C_j = \langle D^{r+1}/x\rangle$ and no $C_1[\widetilde{E}],\cdots,C_{j-1}[\widetilde{E}]$ is a substitution for x. Without entering into the notational details, one can see that

$$C[\widetilde{E}] \to_{\mathsf{m}} D[\widetilde{E}]Q_1\cdots Q_{j-1}\langle D[\widetilde{E}]^r/z\rangle C_{j+1}[\widetilde{E}]\cdots C_m[\widetilde{E}]$$

where z is fresh, and $Q_1,\ldots,Q_{j-1}$ are built from $C_1[\widetilde{E}],\ldots,C_{j-1}[\widetilde{E}]$ by renaming x into z, and by renaming into fresh variables the variables that are bound by the C_i's (that is, more precisely, the variables $y \in \mathsf{fv}(D[\widetilde{E}])$ such that, for some $i < j$, C_i is a substitution for y). It should be clear that we may find a context C' and renamings $\widetilde{E^1},\ldots,\widetilde{E^s}$ such that

$$C'[\widetilde{E},\widetilde{E^1},\ldots,\widetilde{E^s}] = D[\widetilde{E}]Q_1\cdots Q_{j-1}\langle D[\widetilde{E}]^r/z\rangle C_{j+1}[\widetilde{E}]\cdots C_m[\widetilde{E}]$$

so that we also have

$$C[\widetilde{F}] \to_{\mathsf{m}} C'[\widetilde{F},\widetilde{F^1},\ldots,\widetilde{F^s}]$$

where the $\widetilde{F^i}$'s are obtained from $\widetilde{F}$ by performing the same renamings as in $\widetilde{E^i}$ w.r.t. $\widetilde{E}$. Therefore we can conclude using the induction hypothesis, and the Corollary 3.3.

$C_0 = [\,]_i$.

Let C' be the context $E_iC_1\cdots C_m$. It has $h-1$ holes, and obviously $C'[\widetilde{E}] = C[\widetilde{E}]$, therefore by induction hypothesis $C'[\widetilde{F}] \Downarrow_{\mathsf{m}}$. Since $C'[\widetilde{F}] = E_iC_1[\widetilde{F}]\cdots C_m[\widetilde{F}]$ and $E_i \preceq^{\mathcal{A}}_{\mathsf{m}} F_i$, we conclude $F_iC_1[\widetilde{F}]\cdots C_m[\widetilde{F}] \Downarrow_{\mathsf{m}}$, that is $C[\widetilde{F}] \Downarrow_{\mathsf{m}}$. ■

The context lemma is important, for several reasons. For instance, it shows that the λ_{m}-terms may be regarded as functions, since their semantics is determined by applying them to arguments (for a formalization of this point, see the "functionality theory" of [5]). Moreover, by restricting the universal quantification over contexts, the context lemma allows us to prove more easily some semantic relations. For instance, one can use it to check the following:

Proposition 3.5 *If $E \to_{\mathsf{m}} F$ then $E \simeq_{\mathsf{m}} F$. Moreover, $\Omega \preceq_{\mathsf{m}} E \preceq_{\mathsf{m}} \Xi$ for any E.*

We recall that $\Xi = (\lambda fx.ff)(\lambda fx.ff)$. This term is such that $K[\Xi] \Downarrow_{\mathsf{m}}$ for any applicative context K, while $K[\Omega] \Uparrow_{\mathsf{m}}$.

Using the context lemma, one can also prove that η-expansion is increasing with respect to the observational preorder, as we indicated in the introduction:

Lemma 3.6 $x \notin \mathsf{fv}(E) \;\Rightarrow\; E \preceq_{\mathsf{m}} \lambda x(Ex^\infty)$

To prove this property, one first proves the following auxiliary statements:

1. $E\langle{}^{P}\!/_{x}\rangle \simeq_{\mathsf{m}} E\langle{}^{y^\infty}\!/_{x}\rangle\langle{}^{P}\!/_{y}\rangle$, provided that $y \notin \mathsf{fv}(E) \cup \mathsf{fv}(P)$;
2. $EP \simeq_{\mathsf{m}} Ex^\infty\langle{}^{P}\!/_{x}\rangle$, provided that $x \notin \mathsf{fv}(E) \cup \mathsf{fv}(P)$.

We refer to [7] for the details. In this paper, we shall give an easy proof of this lemma, using a result established in [5]; this proof is deferred to the Appendix A.

3.2 Encoding the lazy λ-calculus

In this section we recall from [5] some facts concerning the embedding of the (lazy) λ-calculus of Abramsky and Ong [2] into the λ-calculus with multiplicities, and we begin to study the semantics induced by the latter over λ-terms. It was shown in [5] that the context with multiplicities are strictly more discriminating than ordinary contexts (that is, with implicit infinite multiplicities). Here we show that this happens in a way that cannot be softened by adding further equations over Λ_{m}. The set Λ of λ-terms is given by the usual syntax:

$$M \;::=\; x \;\mid\; \lambda x M \;\mid\; (MM)$$

As we said, to avoid any confusion with terms with multiplicities, we use the standard symbols, that is $M, N \ldots$ to range over Λ. The set of closed λ-terms is Λ^0. The lazy evaluation $M \rightarrow_\ell N$ is defined by the following rules:

$$(\lambda x M)N \rightarrow M[{}^{N}\!/_{x}] \qquad\qquad \frac{M \rightarrow M'}{MN \rightarrow M'N}$$

We shall also write $M \rightarrow_\beta N$ for the standard β-reduction (that is, $M = C[(\lambda x R)S]$ and $N = C[R[{}^{S}\!/_{x}]]$ for some λ-context C with one hole), and $M =_\beta N$ for the β-conversion. We let $\preceq_\ell$ and $\simeq_\ell$ be the observational preorder and equivalence of the lazy λ-calculus.

There is an obvious encoding of Λ into Λ_{m}, consisting in imposing an explicit infinite multiplicity on the argument of an application, that is:

$$\begin{array}{rcl} [\![x]\!] & = & x \\ [\![\lambda x M]\!] & = & \lambda x[\![M]\!] \\ [\![MN]\!] & = & [\![M]\!][\![N]\!]^\infty \end{array}$$

In [5] it is proved that this encoding is correct, i.e.:

$$[\![M]\!] \preceq_{\mathsf{m}} [\![N]\!] \;\Rightarrow\; M \preceq_\ell N$$

The converse implication holds if one restricts the observational preorder of λ_{m} to using only terms with infinite multiplicities. This means that we may regard the lazy λ-calculus as a subcalculus of the λ_{m}-calculus. For this reason infinite multiplicities may be omitted, writing EF instead of EF^{∞}. Similarly, we also omit the explicit mention of the encoding, writing M for $[\![M]\!]$. We note also that the equivalence $\simeq_{\mathsf{m}}$ on λ-terms induced by the λ-calculus with multiplicities contains the standard β-conversion (see [5], and the Appendix A):

Proposition 3.7 $M =_{\beta} N \;\Rightarrow\; M \simeq_{\mathsf{m}} N.$

It was shown in [5] that the encoding of the λ-calculus into the λ_{m}-calculus is not fully abstract. In other words, the contexts with multiplicities provide us with more discriminating power. A typical example is this:

Example 3.8 *The two λ-terms $\Delta = \lambda x\, xx$ and $\Delta' = \lambda x\, x(\lambda y\, xy)$ are equated in the lazy λ-calculus, that is $\Delta \simeq_{\ell} \Delta'$, but their encodings $\lambda x\, xx^{\infty}$ and $\lambda x\, x(\lambda y\, xy^{\infty})^{\infty}$ are not λ_{m}-observationally equivalent, since $(\lambda x\, xx^{\infty})\mathbf{I}^1 \Uparrow_{\mathsf{m}}$ while $(\lambda x\, x(\lambda y\, xy^{\infty})^{\infty})\mathbf{I}^1 \Downarrow_{\mathsf{m}}$, where $\mathbf{I} = \lambda z z$ is the identity (the pair of terms Δ, Δ' has been used by Lévy in [9], Proposition 5.3.5, for a similar purpose).*

In other words, the preorder *on λ-terms* induced by the λ-calculus with multiplicities, that is $M \preceq_{\mathsf{m}} N$, or more accurately $[\![M]\!] \preceq_{\mathsf{m}} [\![N]\!]$, is strictly finer than $M \preceq_{\ell} N$. We shall see several other examples in the Section 6.

To conclude this section, we show that the λ_{m}-calculus cannot be weakened by adding new inequations to coincide with the lazy λ-calculus (over Λ). To this end, we adapt the notion of *functionality order* of a term (see [10, 2]). Roughly, the functionality order of E is the number of nested abstraction E has or is convertible to.

Definition 3.9 *The* order *of a λ_{m}-term is inductively defined as follows;*

i. E has order 0, denoted $E \in \mathsf{O}^{\mathsf{m}}_0$, if $\neg(\,\exists F.\; E \xrightarrow{*}_{\mathsf{m}} \equiv \lambda x F\,)$;

ii. E has order $n+1$, denoted $E \in \mathsf{O}^{\mathsf{m}}_{n+1}$, if $\exists F \in \mathsf{O}^{\mathsf{m}}_n.\; M \xrightarrow{*}_{\mathsf{m}} \equiv \lambda x F$;

iii. E has order ∞, denoted $E \in \mathsf{O}^{\mathsf{m}}_{\infty}$, if $E \notin \mathsf{O}^{\mathsf{m}}_n$, for any n.

The terms of proper order n *are defined as follows:*

1. E has proper order 0, denoted $E \in \mathsf{PO}^{\mathsf{m}}_0$, if $E \in \mathsf{O}^{\mathsf{m}}_0$ and $\neg(\exists \vec{P}\; \exists x \in \mathsf{fv}(E).\; E \xrightarrow{*}_{\mathsf{m}} x\vec{P})$. Every term in $\mathsf{PO}^{\mathsf{m}}_0$ is said strongly unsolvable*;*

2. E has proper order $n+1$, denoted $E \in \mathsf{PO}^{\mathsf{m}}_{n+1}$, if $\exists F \in \mathsf{PO}^{\mathsf{m}}_n.\; E \xrightarrow{*}_{\mathsf{m}} \equiv \lambda x F$;

3. E has proper order ∞, denoted $E \in \mathsf{PO}^{\mathsf{m}}_{\infty}$, iff $E \in \mathsf{O}^{\mathsf{m}}_{\infty}$.

For example $\lambda x\, x \in \mathsf{O}^{\mathsf{m}}_1$ but $\lambda x\, x \notin \mathsf{PO}^{\mathsf{m}}_1$, while $\lambda x\, \Omega \in \mathsf{PO}^{\mathsf{m}}_1$. Notice that $x\langle F^0/x\rangle \in \mathsf{PO}^{\mathsf{m}}_0$ because x does not occur free in $x\langle F^0/x\rangle$. An example of term of infinite order is Ξ. It is easy to see, using the Context Lemma, that the following property holds:

Remark 3.10 *If $O \in \mathsf{PO}^{\mathsf{m}}_0$ and $T \in \mathsf{PO}^{\mathsf{m}}_{\infty}$ then $O \preceq_{\mathsf{m}} E \preceq_{\mathsf{m}} T$ for any E.*

To state our next results, it is convenient to use Barendregt's notations [3] regarding the inequational and equational theories, that is:

$$\lambda\ell_{\mathsf{m}} \vdash E \sqsubseteq F \stackrel{\text{def}}{\Longleftrightarrow} E \preceq_{\mathsf{m}} F$$
$$\lambda\ell_{\mathsf{m}} \vdash E = F \stackrel{\text{def}}{\Longleftrightarrow} E \simeq_{\mathsf{m}} F$$

and similarly for the lazy λ-calculus, using $\lambda\ell$ in this case. More generally, when we write $\mathcal{T} \vdash E \sqsubseteq F$, we mean that the pair (E, F) is in the precongruence generated by the set $\mathcal{T}$ of pairs of λ_{m}-terms. We now show that $\lambda\ell_{\mathsf{m}}$ is a (maximal) "fully lazy theory", as is $\lambda\ell$ (see [2]).

Proposition 3.11

i. $\lambda\ell_{\mathsf{m}}$ *is a* fully lazy *theory, that is for every* $m, n \in \mathsf{N} \cup \{\infty\}$*:*

$$\forall E \in \mathsf{PO}^{\mathsf{m}}_{m}\ \forall F \in \mathsf{PO}^{\mathsf{m}}_{n}.\ (\lambda\ell_{\mathsf{m}} \vdash E \sqsubseteq F \iff m \leq n)\,;$$

ii. $\lambda\ell_{\mathsf{m}}$ *is a* maximal *fully lazy theory, i.e. for* $E, F \in \Lambda_{\mathsf{m}}$ *such that* $\lambda\ell_{\mathsf{m}} \nvdash E \sqsubseteq F$*, then either* $\lambda\ell_{\mathsf{m}} + E \sqsubseteq F$ *is inconsistent or* $\lambda\ell_{\mathsf{m}} + E \sqsubseteq F$ *is not fully lazy.*

Proof: *i.* It is easy to check that if $E \in \mathsf{PO}^{\mathsf{m}}_{m}$ and $F \in \mathsf{PO}^{\mathsf{m}}_{n}$ with $m \leq n$ then $\lambda\ell_{\mathsf{m}} \vdash E \sqsubseteq F$. Now if $E \in \mathsf{PO}^{\mathsf{m}}_{m}$ and $F \in \mathsf{PO}^{\mathsf{m}}_{n}$ are such that $\lambda\ell_{\mathsf{m}} \vdash E \sqsubseteq F$ then $m \leq n$, because otherwise we would have $F\ \underbrace{\mathbf{I}^1 \cdots \mathbf{I}^1}_{m-1\text{ times}} \Uparrow_{\mathsf{m}}$ while $E\ \underbrace{\mathbf{I}^1 \cdots \mathbf{I}^1}_{m-1\text{ times}} \Downarrow_{\mathsf{m}}$.

ii. If $\lambda\ell_{\mathsf{m}} \nvdash E \sqsubseteq F$ then there exists a context C such that $C[E] \Downarrow_{\mathsf{m}}$ and $C[F] \Uparrow_{\mathsf{m}}$. That is $C[F] \in \mathsf{PO}^{\mathsf{m}}_{0}$ and $C[E] \in \mathsf{O}^{\mathsf{m}}_{n}$, $n \neq 0$. There are two cases:

1. $C[E] \in \mathsf{PO}^{\mathsf{m}}_{n}$. Then the theory $\lambda\ell_{\mathsf{m}} + E \sqsubseteq F$ it is not fully lazy, since $C[E] \sqsubseteq C[F]$ is a consequence of $E \sqsubseteq F$.

2. $C[E] \in \mathsf{O}^{\mathsf{m}}_{n} \setminus \mathsf{PO}^{\mathsf{m}}_{n}$. Then $C[E] \simeq_{\mathsf{m}} \lambda x_1 \cdots x_n.\, x_i \vec{P}$. Let k be the number of bags in the vector $\vec{P}$. It is clear that the context

$$C' = C\, \underbrace{\mathbf{I}^1 \cdots \mathbf{I}^1}_{i-1\text{ times}} (\lambda z_1 \cdots z_k.\, \mathbf{I})^1\ \underbrace{\mathbf{I}^1 \cdots \mathbf{I}^1}_{n-i\text{ times}}$$

is such that $C'[E] \simeq_{\mathsf{m}} \mathbf{I}$ and $C'[F] \simeq_{\mathsf{m}} \Omega$. So

$$\begin{array}{lll} \lambda\ell_{\mathsf{m}} + E \sqsubseteq F \vdash & \mathbf{I} \sqsubseteq \Omega & \text{hence} \\ & \mathbf{I}X^1 \sqsubseteq \Omega X^1 & \text{since } \sqsubseteq \text{ is a precongruence, hence} \\ & X \sqsubseteq \Omega & \end{array}$$

Since $\lambda\ell_{\mathsf{m}} \vdash \Omega \sqsubseteq X$ for any X, we have $\lambda\ell_{\mathsf{m}} + E \sqsubseteq F \vdash X = \Omega$, therefore $\lambda\ell_{\mathsf{m}} + E \sqsubseteq F \vdash X = Y$ for any X and Y, that is, the theory is inconsistent. ∎

An immediate consequence of this proposition is the following:

Corollary 3.12 *There exists no set $\mathcal{T}$ of pairs of λ_{m}-terms such that, for every $M, N \in \Lambda$, $\lambda\ell_{\mathsf{m}} + \mathcal{T} \vdash M \sqsubseteq N$ if and only if $\lambda\ell \vdash M \sqsubseteq N$.*

PROOF: Assume that there is such a $\mathcal{T}$. Then $\lambda\ell_{\mathsf{m}} + \mathcal{T}$ is consistent, since $\mathbf{I} \not\preceq_{\ell} \Omega$. On the other hand, $\lambda\ell_{\mathsf{m}} + \mathcal{T} \vdash \lambda x\, xx = \lambda x\, x\lambda y\, xy$ (see Example 3.8), therefore $\lambda\ell_{\mathsf{m}} + \mathcal{T} \vdash (\lambda x\, xx)\mathbf{I}^1 = (\lambda x\, x\lambda y\, xy)\mathbf{I}^1$. By the Proposition 3.5 $\lambda\ell_{\mathsf{m}} + \mathcal{T} \vdash (\lambda x\, xx)\mathbf{I}^1 = x\langle \mathbf{I}^0/x\rangle$ and $\lambda\ell_{\mathsf{m}} + \mathcal{T} \vdash (\lambda x\, x\lambda y\, xy)\mathbf{I}^1 = \lambda y(xy\langle \mathbf{I}^0/x\rangle)$. By the previous proposition $\lambda\ell_{\mathsf{m}} + \mathcal{T} \vdash x\langle \mathbf{I}^0/x\rangle = \Omega$ and $\lambda\ell_{\mathsf{m}} + \mathcal{T} \vdash \lambda y(xy\langle \mathbf{I}^0/x\rangle) = \lambda y\Omega$ therefore $\lambda\ell_{\mathsf{m}} + \mathcal{T}$ is not fully lazy. But this contradicts a result of Ong (see [2]) that $\lambda\ell$ is fully lazy. ∎

4 Lazy approximants and Lévy-Longo trees

In this section we start relating the observational preorder $\preceq_{\mathsf{m}}$ with an intensional representation ([1]) of λ-terms due to Lévy, and studied by Longo [10] and Ong [13]. In [8], Lévy introduced a refinement of Wadsworth's notion of syntactic approximant [15], suited for the lazy λ-calculus where $\lambda x\,\Omega$ must be distinguished from Ω. Then Lévy defined an interpretation of the λ-calculus, based on this notion of "lazy approximant", and he showed that the induced preorder is a precongruence. Lévy's interpretation may be defined as follows:

Definition 4.1 *The set $\mathcal{N}$ of (lazy) approximants, ranged over by $A, B, \cdots$, is the least subset of Λ containing $\lambda x_1 \ldots x_n.\,\Omega$, and $\lambda x_1 \ldots x_n.\, xA_1 \cdots A_m$ whenever $A_i \in \mathcal{N}$. For $M \in \Lambda$, the* direct approximation *of M is the term $\varpi(M)$ of $\mathcal{N}$ inductively defined by:*

$$\begin{array}{rcl} \varpi(\lambda\vec{x}.\,(\lambda y.\,M)NM_1\cdots M_k) & = & \lambda\vec{x}.\,\Omega \\ \varpi(\lambda\vec{x}.\,yM_1\cdots M_k) & = & \lambda\vec{x}.\,y\,\varpi(M_1)\cdots\varpi(M_k) \end{array}$$

The interpretation *of $M \in \Lambda$ is $\mathcal{A}(M) = \{\varpi(N) \mid M =_{\beta} N\}$.* Lévy's preorder *on λ-terms, denoted $M \sqsubseteq_{\mathcal{L}} N$, is the inclusion of sets of approximants $\mathcal{A}(M) \subseteq \mathcal{A}(N)$. The equality $M =_{\mathcal{L}} N$ is $\mathcal{A}(M) = \mathcal{A}(N)$.*

One may characterize Lévy's preorder on approximants: $A \sqsubseteq_{\mathcal{L}} B$ if and only if A is a *prefix* of B, where the prefix ordering is the precongruence $\leq$ on approximants generated by $\Omega \leq A$. Using the Proposition 3.5, it is easy to see that

$$A \leq B \;\Rightarrow\; A \preceq_{\mathsf{m}} B$$

The Church-Rosser property has the following consequence (see [8]):

Lemma 4.2 *For any $M \in \Lambda$, the set $\mathcal{A}(M)$ is directed with respect to the prefix preorder, namely*

$$\forall A', A'' \in \mathcal{A}(M)\ \exists A \in \mathcal{A}(M).\ A' \leq A\ \&\ A'' \leq A$$

[1] by "intensional" we mean "relying on the syntactic shape of the terms, up to β-conversion".

Moreover, it is easy to see that $\mathcal{A}(M)$ is in fact an *ideal*, that is it is downward-closed with respect to the prefix ordering:

$$A \in \mathcal{A}(M) \;\&\; B \leq A \;\Rightarrow\; B \in \mathcal{A}(M)$$

This is because any term M is β-convertible to a redex, namely $(\mathbf{I}\,M)$, whose direct approximation is Ω. In particular, we have $\Omega \in \mathcal{A}(M)$ for any M.

In [10], Longo gave a suggestive presentation of Lévy's interpretation, by means of what is now called Lévy-Longo trees. These are refinements of the well-known Böhm trees (see [3]), that deal with the fact that, in the lazy λ-calculus, the unsolvable terms cannot be all identified to Ω. The Lévy-Longo trees are possibly infinite, node-labelled trees, where the labels are either Υ, representing terms of infinite order as Ξ, or $\lambda x_1 \ldots x_n.\perp$, representing terms as $\lambda x_1 \cdots x_n.\Omega$, or $\lambda x_1 \ldots x_n.x$, representing the "head" of a solvable term, as in Böhm trees. To define these trees, let us first recall the notion of a λ-term of *proper order* n, with $n \in \mathsf{N} \cup \{\infty\}$ (see [2]).

1. $M \in \mathsf{PO}_0 \;\Leftrightarrow\; \forall M'.\, M \xrightarrow{*}_\ell M' \;\Rightarrow\; \exists M''.\, M' \rightarrow_\ell M''$
2. $M \in \mathsf{PO}_{n+1} \;\Leftrightarrow\; \exists x\, \exists M' \in \mathsf{PO}_n.\, M \xrightarrow{*}_\ell \lambda x M'$
3. $M \in \mathsf{PO}_\infty \;\Leftrightarrow\; \forall n\, \exists x_1, \ldots, x_n\, \exists M'.\, M =_\beta \lambda x_1 \cdots x_n.\, M'$

Definition 4.3 *The* Lévy-Longo tree *of a λ-term M, $\mathsf{LT}(M)$, is defined inductively as follows:*

1. $\mathsf{LT}(M) = \Upsilon$, *if* $M \in \mathsf{PO}_\infty$*;*
2. $\mathsf{LT}(M) = \lambda x_1 \cdots x_n.\perp$, *if* $M \in \mathsf{PO}_n$ *with* $n \in \mathsf{N}$*;*
3. $\mathsf{LT}(M) = \begin{array}{ccc} & \lambda x_1 \cdots x_n.\, x & \\ \mathsf{LT}(M_1) & \cdots & \mathsf{LT}(M_k) \end{array}$, if $M =_\beta \lambda x_1 \cdots x_n.\, x M_1 \cdots M_k$.

To recover Lévy's ordering $M \sqsubseteq_{\mathcal{L}} N$ on the tree representation, one defines an operation $\lambda x T$ on trees, consisting in prefixing the label of the root of T by λx, with the rule that $\lambda x \Upsilon = \Upsilon$. Then a tree T is less than T' whenever T' is obtained from T by replacing some leaves labelled $\lambda x_1 \ldots x_n.\perp$ in T by trees $\lambda x_1 \ldots x_n.T''$. Obviously, $M =_{\mathcal{L}} N$ means that M and N have the same Lévy-Longo tree, $\mathsf{LT}(M) = \mathsf{LT}(N)$.

One may observe that if a term M has a β-normal form, then $\mathsf{LT}(M)$ is finite. On the other hand, an example of infinite Lévy-Longo tree is provided by Wadsworth's combinator $\mathbf{J}$, satisfying $\mathbf{J} =_\beta \lambda xy.\, x(\mathbf{J}y)$, which may be given by $\mathbf{J} = (\lambda f \lambda xy.\, x(ffy))(\lambda f \lambda xy.\, x(ffy))$.

The tree for this term is:

$$\begin{array}{rcl} \mathsf{LT}(\mathbf{J}) & = & \lambda x y_0 . x \\ & & | \\ & & \lambda y_1 . y_0 \\ & & | \\ & & \lambda y_2 . y_1 \\ & & \vdots \end{array}$$

Regarding Lévy's interpretation, a natural question is: is there any observational semantics on λ-terms that coincides with $\sqsubseteq_{\mathcal{L}}$? As a matter of fact, neither $\preceq_\ell$ nor even $\preceq_{\mathsf{m}}$ does provide a right answer, for two reasons:

1. Lévy's preorder does not deal with the fact that terms of infinite order are greater than any other one (Remark 3.10). For instance, $\mathbf{I} \not\sqsubseteq_{\mathcal{L}} \Xi$ since $\mathcal{A}(\mathbf{I}) = \{\Omega,\ \lambda x\,\Omega, \lambda x\,x\}$ while $\mathcal{A}(\Xi) = \{\Omega, \lambda x_1\,\Omega, \ldots, \lambda x_1 \ldots x_n.\,\Omega, \ldots\}$.

2. Lévy's preorder does not deal with the fact that η-expansion is increasing (Lemma 3.6). Typically, $x \not\sqsubseteq_{\mathcal{L}} \lambda y.\,xy$ since $\mathcal{A}(x) = \{\Omega, x\}$ while $\mathcal{A}(\lambda y\,xy) = \{\Omega, \lambda y\,\Omega, \lambda y\,x\Omega, \lambda y\,xy\}$.

In other words, Lévy's interpretation is not fully abstract with respect to $\preceq_{\mathsf{m}}$. However, we will see that it is adequate, that is $M \sqsubseteq_{\mathcal{L}} N \ \Rightarrow\ M \preceq_{\mathsf{m}} N$. Moreover, we will prove in the next section that Lévy's interpretation is *equationally* fully abstract with respect to $\preceq_{\mathsf{m}}$, that is $M =_{\mathcal{L}} N \ \Leftrightarrow M \simeq_{\mathsf{m}} N$.

As a first step towards these results, we establish a property that we call the *approximation lemma* (*cf.* [9]). It states that, in order for $C[M]$ to converge (where M is a λ-term and C a λ_{m}-context), only a finite amount of information about M is needed to know. Intuitively, this should be clear, because M can only participate by a finite number of reduction steps in the convergent computation of $C[M]$. Moreover, it is only whenever M shows up in the head position, as a function applied to a series of arguments, that it has to exhibit some specific finite intensional content, like beginning with a series of abstractions. Then any term having at least the same intensional content is as good as M, as far as the convergence within the context C is concerned. The appropriate formalization of "finite intensional content" is given by approximants.

Lemma 4.4 (The Approximation Lemma) *For any λ_{m}-context C and for every $M \in \Lambda$ with $C[M]$ closed:*

$$C[M] \Downarrow_{\mathsf{m}} \Leftrightarrow\ \exists A.\ A \in \mathcal{A}(M)\ \&\ C[A] \Downarrow_{\mathsf{m}}$$

Proof: The implication "$\Leftarrow$" is easy, since $M =_\beta N$ implies $M \simeq_{\mathsf{m}} N$ (Proposition 3.7), and $\varpi(N) \preceq_{\mathsf{m}} N$ since $\preceq_{\mathsf{m}}$ is a precongruence such that $\Omega \preceq_{\mathsf{m}} X$ for any X.

To establish "$\Rightarrow$" we use multiple contexts, as in the context lemma (again, the explicit substitution construct is very convenient for this proof). We recall that $C[\,\widetilde{M}\,]$ stands for $C[M_1, \ldots, M_p]$. Let us denote by $C[\,\widetilde{M}\,] \Downarrow^l_{\mathsf{m}}$ the fact that $C[\,\widetilde{M}\,]$ converges to an abstraction in l steps. Then we show the following:

if $C[\,M_1,\ldots,M_p\,]$ is closed and $C[\,M_1,\ldots,M_p\,] \Downarrow^{l}_{\mathsf{m}}$ then there exist $A_1,\ldots,A_p$ such that $A_i \in \mathcal{A}(M_i)$ and $C[\,A_1,\ldots,A_p\,] \Downarrow_{\mathsf{m}}$.

We proceed by induction on l, observing that the context C may be written $C_0C_1\cdots C_n$ where C_0 is either a hole $[]_i$, or a variable x, or an abstraction context $\lambda x\, D$, and the C_j's, for $j>0$, are either D^m or $\langle D^m/x\rangle$ for some context D. We examine the possible cases, the interesting one being $C_0 = []_i$ (case 2 below). We do not investigate the case $C_0 = x$, since we may let $C' = []_{p+1}C_1\cdots C_n$, so that $C[E_1,\ldots,E_p] = C'[E_1,\ldots,E_p,x]$ for any $E_1,\ldots,E_p$, and we are in case 2 (or more precisely case 2.b, where one can see that the approximant taken for x in order for $C'[E_1,\ldots,E_p,x]$ to converge is x itself).

1. $C_0 = \lambda x\, D$.

   There are two cases. If $l=0$ then $C[\,\widetilde{M}\,]$ is, possibly up to $\equiv$, an abstraction and we can let $A_i = \Omega$ for any i. Otherwise $(l>0)$, there exists i such that C_i is a bag D'^m, and C_j are substitutions for any $j<i$. Then we have

   $$C[\,\widetilde{M}\,] \;\equiv\!\rightarrow_{\mathsf{m}}\; D''[\,\widetilde{M},\widetilde{M'}\,]C_1[\,\widetilde{M}\,]\cdots C_{i-1}[\,\widetilde{M}\,]\langle D'[\,\widetilde{M}\,]^m/z\rangle C_{i+1}[\,\widetilde{M}\,]\cdots C_n[\,\widetilde{M}\,]\Downarrow^{l-1}_{\mathsf{m}}$$

   where z is fresh and D'' is obtained from D by renaming x by z, and by replacing some of the holes $[]_i$ by $[]_{i+p}$, so that $(D[\widetilde{M}])[z/x] = D''[\widetilde{M},\widetilde{M'}]$ where $M'_i = M_i[z/x]$. Then we let $C' = D''C_1\cdots C_{i-1}\langle D'^m/z\rangle C_{i+1}\cdots C_n$, so that $C[\,\widetilde{E}\,] \rightarrow_{\mathsf{m}} C'[\,\widetilde{E},\widetilde{E}[z/x]\,]$ for any $E_1,\ldots,E_p$ such that $C[\,\widetilde{E}\,]$ is closed (note that the choice of z only depends on the variables that are bound by C). By induction hypothesis there exist $A'_1,\ldots,A'_p$ and $B_1,\ldots,B_p$ such that $A'_i \in \mathcal{A}(M_i)$ and $B_i \in \mathcal{A}(M'_i)$, and $C'[A'_1,\ldots,A'_p,B_1,\ldots,B_p] \Downarrow_{\mathsf{m}}$. It is easy to check that there exist $A''_1,\ldots,A''_p$ such that $A''_i \in \mathcal{A}(M_i)$ and $B_i = A''_i[z/x]$. By Lemma 4.2, there exist $A_1,\ldots,A_p$, respectively approximants of $M_1,\ldots,M_p$ such that $A'_i \leq A_i$ and $A''_i \leq A_i$, therefore $C[\,\widetilde{A}\,]\Downarrow_{\mathsf{m}}$ since $C[\,\widetilde{A}\,] \rightarrow_{\mathsf{m}} C'[\,\widetilde{A},\widetilde{A}[z/x]\,]$, and $B_i \leq A_i[z/x]$.

2. $C_0 = []_i$.

   Here $C[\,\widetilde{M}\,] = M_iC_1[\,\widetilde{M}\,]\cdots C_n[\,\widetilde{M}\,]$. If $l=0$ then M_i is an abstraction, and $C_1[\,\widetilde{M}\,]$, …, $C_n[\,\widetilde{M}\,]$ are susbtitution items. In this case we let $A_i = \varpi(M_i)$, and $A_j = \Omega$ for any $j \neq i$. Otherwise $(l>0)$, there are three cases, according as M_i is a normal form (with respect to $\rightarrow_\ell$) or not:

   (a) $M_i = \lambda x\, M'$.

   Let us assume, for simplicity, that C_1 is an argument context D^m (otherwise we have to use $\equiv$ to push the context C_1 under the abstraction λx, as in the case (1)). Then the normalizing derivation of $C[\,\widetilde{M}\,]$ is of the form

   $$C[\,\widetilde{M}\,] \rightarrow_{\mathsf{m}} M'\langle D[\,\widetilde{M}\,]^m/x\rangle C_2[\,\widetilde{M}\,]\cdots C_n[\,\widetilde{M}\,]\Downarrow^{l-1}_{\mathsf{m}}$$

   We let $C' = []_{p+1}\langle D^m/x\rangle C_2\cdots C_n$, so that for any $E_1,\ldots,E_p$ where $E_i = \lambda x\, E$ such that $C[E_1,\ldots,E_p]$ is closed, we have $C[E_1,\ldots,E_p] \rightarrow_{\mathsf{m}} C'[E_1,\ldots,E_p,E]$.

By induction hypothesis, there exist approximants $A'_1, \dots, A'_{p+1}$ of $M_1, \dots, M_p, M'$ such that $C'[\widetilde{A}'] \Downarrow_{\mathsf{m}}$. We let $A_j = A'_i$ for $j \neq i$, and let $A_i \in \mathcal{A}(M_i)$ be such that $A'_i \leq A_i$ and $\lambda x A_{p+1} \leq A_i$. Such an A_i exists by the Lemma 4.2, since $\lambda x A_{p+1} \in \mathcal{A}(M_i)$, and we have $A_i = \lambda x B$ for some $B \geq A_{p+1}$. Then $C[\widetilde{A}] \rightarrow_{\mathsf{m}} B\langle D[\widetilde{A}]^m/x\rangle C_2[\widetilde{A}] \cdots C_n[\widetilde{A}]$, therefore $C[\widetilde{A}] \Downarrow_{\mathsf{m}}$.

(b) $M_i = xL_1 \cdots L_q$.

In this case, there exists $j \leq n$ such that $C_j = \langle D^{m+1}/x\rangle$. Without entering into the notational details, one can see that the normalizing derivation of $C[\widetilde{M}]$ is of the form

$$C[\widetilde{M}] \rightarrow_{\mathsf{m}} (D[\widetilde{M}]L'^{\infty}_1 \cdots L'^{\infty}_q)Q_1 \cdots Q_{j-1}\langle D[\widetilde{M}]^m/z\rangle \cdots C_n[\widetilde{M}] \Downarrow_{\mathsf{m}}^{l-1}$$

where z is fresh and $L'_1, \dots, L'_q$ and $Q_1, \dots, Q_{j-1}$ are respectively obtained from $L_1, \dots, L_q$ and $C_1[\widetilde{M}], \dots, C_{j-1}[\widetilde{M}]$ by renaming some variables (which are bound in $C[\widetilde{M}]$) with fresh ones. It should be clear that we may find a context C' and renamings $\widetilde{M^1}, \dots, \widetilde{M^s}$ such that

$$C'[\widetilde{M}, \widetilde{L}, \widetilde{M^1}, \dots, \widetilde{M^s}] = (D[\widetilde{M}]L'^{\infty}_1 \cdots L'^{\infty}_q)Q_1 \cdots Q_{j-1}\langle D[\widetilde{M}]^m/z\rangle \cdots C_n[\widetilde{M}]$$

By induction hypothesis, there exist approximants

$$A'_1, \dots, A'_p, \dots, A^h_1, \dots, A^h_p, \dots, B_1, \dots, B_q$$

of $M_1, \dots, M_p, \dots, M^h_1, \dots, M^h_p, \dots, L'_1, \dots, L'_q$, with $1 \leq h < s$, such that

$$C'[\widetilde{A}', \widetilde{B}, \widetilde{A}^1, \dots, \widetilde{A}^s] \Downarrow_{\mathsf{m}}$$

Obviously, for some appropriate renamings $B'_1, \dots, B'_q$ of $B_1, \dots, B_q$, the term $xB_1 \cdots B_q$ is an approximant of M_i. We use again the Lemma 4.2 to conclude.

(c) If M_i is not a lazy normal form, then the normalizing derivation of $C[\widetilde{M}]$ must start with a reduction of M_i. We have shown in [5] that in this case there exists $M' \in \Lambda$ such that $M_i \rightarrow_{\ell} M'$ and $M'C_1[\widetilde{M}] \cdots C_n[\widetilde{M}] \Downarrow_{\mathsf{m}}^{l'}$ with $l' \leq l - 1$. Then we let $C' = []_{p+1}C_1 \cdots C_n$ and we use the induction hypothesis and the Lemma 4.2. ■

The Approximation Lemma holds in particular for any λ-context C (that is, more accurately, with infinite multiplicities) and in this case the lemma is a particular case of Lévy's Lemma 5.7 in [8] that his preordering $\sqsubseteq_{\mathcal{L}}$ is a precongruence, which he proved using the idea of Welch that inside-out reductions are complete in some sense (see [3]). However, the proof we give is, to our view, simpler. One should remark that we make an essential use of the construct of *explicit substitutions*, in the cases (1) and (2.a).

As a corollary of the Approximation Lemma, we can now prove the adequacy result mentioned above, relating Lévy's interpretation to the observational preorder:

Theorem 4.5 *For any λ-terms M and N*

$$M \sqsubseteq_{\mathcal{L}} N \;\Rightarrow\; M \preceq_{\mathsf{m}} N$$

PROOF: Assume that $M \sqsubseteq_{\mathcal{L}} N$ and $C[M] \Downarrow_{\mathsf{m}}$, with both $C[M]$ and $C[N]$ closed. Then, by the Approximation Lemma, there exists $A \in \mathcal{A}(M)$ such that $C[A] \Downarrow_{\mathsf{m}}$. By definition of $\sqsubseteq_{\mathcal{L}}$, we also have $A \in \mathcal{A}(N)$, therefore $C[N] \Downarrow_{\mathsf{m}}$ by the Approximation Lemma again. ∎

5 The characterization of the discriminating power of multiplicities

We have seen that Lévy's preorder cannot coincide with the restriction of $\preceq_{\mathsf{m}}$ to λ-terms. In this section we aim at characterizing this observational preorder as an ordering on Lévy-Longo trees, or more accurately as an intensional ordering on λ-terms. We already indicated that the reasons for the above discrepancy are the absence of a top element in $\sqsubseteq_{\mathcal{L}}$ and that η-expansion is not $\sqsubseteq_{\mathcal{L}}$-increasing. Therefore it is necessary to weaken the intensional preorder $\sqsubseteq_{\mathcal{L}}$ into another one, taking into account these features. This was done by Ong in [13] (Definition 3.4.1.1), who called *lazy Plotkin-Scott-Engeler preorder* the following preorder:

Definition 5.1 *The preorder $M \sqsubseteq^{\eta}_{\mathcal{L}} N$ on λ-terms is given by*

$$M \sqsubseteq^{\eta}_{\mathcal{L}} N \;\stackrel{\text{def}}{\Longleftrightarrow}\; \forall k \in \mathsf{N}.\; M \leq^{\eta}_{k} N$$

where $\leq^{\eta}_{k}$ is defined as follows:

1. *$M \leq^{\eta}_{0} N$ for any M and N;*
2. *$M \leq^{\eta}_{k+1} N$ if and only if*
   (a) *$N \in \mathsf{PO}_{\infty}$ or*
   (b) *$M \in \mathsf{PO}_{n}$ and $N =_{\beta} \lambda x_1 \cdots x_m.N'$ with $m \geq n$, or*
   (c) *$M =_{\beta} \lambda x_1 \ldots x_n.xM_1 \cdots M_s$ and $N =_{\beta} \lambda x_1 \ldots x_n y_1 \ldots y_t.xN_1 \cdots N_s Y_1 \cdots Y_t$ for some $N_1, \ldots, N_s$ and $Y_1, \ldots, Y_t$ such that $M_i \leq^{\eta}_{k} N_i$ and $y_j \leq^{\eta}_{k} Y_j$, with $y_j \notin \mathsf{fv}(xM_1 \cdots M_s)$.*

In the clause 2(c) of the definition we implicitly assume that the variables $x_1, \ldots, x_n, y_1, \ldots, y_t$ are distinct. It is easy to check that $\leq^{\eta}_{k}$ and $\sqsubseteq^{\eta}_{\mathcal{L}}$ are indeed preorders. Moreover, it should be clear that

$$M =_{\mathcal{L}} N \;\Leftrightarrow\; M \sqsubseteq^{\eta}_{\mathcal{L}} N \;\&\; N \sqsubseteq^{\eta}_{\mathcal{L}} M$$

since $M \sqsubseteq_{\mathcal{L}} N \;\Rightarrow\; M \leq^{\eta}_{k} N$ for any k. A remark on the notation: we should have used $\sqsubseteq^{\eta,\top}_{\mathcal{L}}$ rather than $\sqsubseteq^{\eta}_{\mathcal{L}}$ (and similarly for $\leq^{\eta}_{k}$) to indicate the two ingredients added to Lévy's preorder. We will content ourselves here using the simpler $\sqsubseteq^{\eta}_{\mathcal{L}}$ – as a matter of fact, the main difficulties come from η-expansion.

Example 5.2 *Recall that Wadsworth's combinator* $\mathbf{J}$ *satisfies* $\mathbf{J} =_\beta \lambda xy.x(\mathbf{J}y)$. *It is equal to the identity in Scott's* D_∞ *model of the* λ*-calculus. One can check that* $z \preceq^\eta_k (\mathbf{J}z)$ *for any* k, *therefore* $\mathbf{I} \sqsubseteq^\eta_{\mathcal{L}} \mathbf{J}$. *On the other hand* $\mathbf{J} \not\sqsubseteq^\eta_{\mathcal{L}} \mathbf{I}$, *since* $\mathbf{J}z \not\preceq^\eta_1 z$. *Note that these two terms are distinguished in the lazy* λ*-calculus, since* $\mathbf{J}\Omega \Downarrow_\ell$ *while* $\mathbf{I}\Omega \Uparrow_\ell$. *We let the reader check that* $\Delta' \not\sqsubseteq^\eta_{\mathcal{L}} \Delta$ *(see the Example 3.8), and that the least index* k *such that* $\Delta' \not\preceq^\eta_k \Delta$ *is* 2. *This is the level at which their Lévy-Longo trees differ.*

The following properties should be obvious:

Remark 5.3

1. $M \preceq^\eta_{k+1} N \;\Rightarrow\; M \preceq^\eta_k N$;
2. $M' =_\beta M \sqsubseteq^\eta_{\mathcal{L}} N =_\beta N' \;\Rightarrow\; M' \sqsubseteq^\eta_{\mathcal{L}} N'$;
3. $M \preceq^\eta_k N \;\Rightarrow\; \lambda x\, M \preceq^\eta_k \lambda x\, N$.

We shall use the fact that the preorder $\sqsubseteq^\eta_{\mathcal{L}}$ is in some sense a fixpoint. More precisely, we will use the following:

Proposition 5.4 *If* $M \sqsubseteq^\eta_{\mathcal{L}} N$ *and* $M =_\beta \lambda z_1 \ldots z_m.xM_1 \cdots M_s$ *then either* $N \in \mathsf{PO}_\infty$ *or* $N =_\beta \lambda z_1 \ldots z_m y_1 \ldots y_t.xN_1 \cdots N_s Y_1 \cdots Y_t$ *with* $y_j \notin \mathsf{fv}(xM_1 \cdots M_s)$, $M_i \sqsubseteq^\eta_{\mathcal{L}} N_i$ *and* $y_j \sqsubseteq^\eta_{\mathcal{L}} Y_j$.

PROOF: Assume that $M \sqsubseteq^\eta_{\mathcal{L}} N$ and $M =_\beta \lambda z_1 \ldots z_m.xM_1 \cdots M_s$ and $N \notin \mathsf{PO}_\infty$. Then for any k there exist $N^k_1, \ldots, N^k_s$, $y_1, \ldots, y_{t_k}$ not in $\mathsf{fv}(xM_1 \cdots M_s)$ and $Y^k_1, \ldots, Y^k_{t_k}$ such that $N =_\beta \lambda z_1 \ldots z_m y_1 \ldots y_{t_k}.xN^k_1 \cdots N^k_s Y^k_1 \cdots Y^k_{t_k}$ with $M_i \preceq^\eta_k N^k_i$ and $y_j \preceq^\eta_k Y^k_j$. By the Church-Rosser property, $t_k = t_h$, $N^k_i =_\beta N^h_i$ and $Y^k_j =_\beta Y^h_j$ for any k, h. Then by the remark above, we may let $N_i = N^0_i$ and $Y_j = Y^0_j$. ■

Before proving our characterization result, establishing that $\preceq_{\mathsf{m}}$ and $\sqsubseteq^\eta_{\mathcal{L}}$ coincide over λ-terms, let us first point out a source of trouble. It is not very difficult to see that, if both A and B are approximants, then $A \sqsubseteq^\eta_{\mathcal{L}} B$ if and only if there exists an η-expansion A' of A (2) such that $A' \leq B$ (the "only if" part may be proved by induction on B). However, it is **not** true in general that if $M \sqsubseteq^\eta_{\mathcal{L}} N$, then for any approximant A of M, there exists an η-expansion of A which is an approximant of N. For instance, we have $\mathcal{A}(x) = \{\, \Omega, x \,\}$ and

$$\mathcal{A}(\mathbf{J}x) = \{\, \Omega\,,\ \lambda y_0\Omega\,,\ \lambda y_0.x\Omega\,,\ \lambda y_0.x\lambda y_1\Omega\,,\ \lambda y_0.x\lambda y_1.y_0\Omega\,, \ldots \}$$

therefore, although $x \sqsubseteq^\eta_{\mathcal{L}} \mathbf{J}x$, there is no $A \in \mathcal{A}(\mathbf{J}x)$ such that $x \sqsubseteq^\eta_{\mathcal{L}} A$. The fact that we cannot express $\sqsubseteq^\eta_{\mathcal{L}}$ in terms of approximants is one of the main difficulties in proving our characterization theorem. There are two main points in the proof: the *crux lemma*, solving the afore-mentioned difficulty, and the *separation lemma*.

[2] M' is an η-expansion of M if it results from M by a sequence of rewriting steps $N \to \lambda x(Nx)$, where x is not free in N, performed in any context.

Lemma 5.5 (the Crux Lemma) $x \sqsubseteq^{\eta}_{\mathcal{L}} X \;\Rightarrow\; x \preceq_{\mathsf{m}} X$.

It is possible to give a direct proof of this lemma, showing that $C[X] \Downarrow_{\mathsf{m}}$ if $x \sqsubseteq^{\eta}_{\mathcal{L}} X$ and $C[x] \Downarrow_{\mathsf{m}}$, see [7]. However, this proof is quite difficult and technical (hence the name of the lemma). In this paper we give an easy proof, using a result established in [5], namely that there is a functional interpretation of the λ-calculus with multiplicities which is adequate with respect to the observational semantics. Since we do not deal with this functional interpretation here, we defer the proof of the Crux Lemma to the Appendix A. We note an immediate consequence of this lemma:

Corollary 5.6 *For any $A \in \mathcal{N}$, $A \sqsubseteq^{\eta}_{\mathcal{L}} M \;\Rightarrow\; A \preceq_{\mathsf{m}} M$.*

PROOF: By induction on A.

1. If $A = \lambda x_1 \ldots x_n.\,\Omega$, that is $A \in \mathsf{PO}_n$, then $M \in \mathsf{O}_m$ for some $m \geq n$, and it is easy to see that $A \preceq_{\mathsf{m}} M$ in this case, since $\Omega \preceq_{\mathsf{m}} X$ for any X (see the Proposition 3.5).

2. If $A = \lambda x_1 \ldots x_m.\, x A_1 \cdots A_s$, then , by the Proposition 5.4, there are two cases: either $M \in \mathsf{PO}_{\infty}$, in which case $A \preceq_{\mathsf{m}} M$ is obvious (see the Remark 3.10), or $M =_{\beta} \lambda x_1 \ldots x_m y_1 \ldots y_t.\, x M_1 \cdots M_s Y_1 \cdots Y_t$ with $A_i \sqsubseteq^{\eta}_{\mathcal{L}} M_i$ and $y_j \sqsubseteq^{\eta}_{\mathcal{L}} Y_j$. By induction hypothesis $A_i \preceq_{\mathsf{m}} M_i$. Moreover, by the Lemma 3.6 (η-expansion is increasing):

$$A \preceq_{\mathsf{m}} \lambda x_1 \ldots x_m y_1 \ldots y_t.\, x A_1 \cdots A_s y_1 \cdots y_t$$

   Then we use the Crux Lemma, that is $y_i \preceq_{\mathsf{m}} Y_i$ for $1 \leq i \leq t$, and the fact that $\preceq_{\mathsf{m}}$ is a precongruence to conclude. ∎

The second key lemma, the *separation lemma*, states that if M and N intensionally differ at some finite order, that is $M \not\preceq^{\eta}_{k} N$, then one can test the difference in the λ_{m}-calculus. That is, there is a λ_{m}-context C *separating* these two terms, in the sense that $C[M] \Downarrow_{\mathsf{m}}$ while $C[N] \Uparrow_{\mathsf{m}}$. The proof, which is long and technical, uses a refinement of the classical "Böhm-out technique" (see [3]). As such, it uses the *tupling* combinators

$$\mathbf{P}_n = \lambda x_1 \cdots x_{n+1}.\, x_{n+1} x_1 \cdots x_n$$

and the associated *projections*

$$\mathbf{U}^n_i = \lambda x_1 \cdots x_n.\, x_i$$

Then we show that if $M \not\preceq^{\eta}_{k} N$ then M and N may be separated by means of a context of the form

$$[]\, \langle {\mathbf{P}_{q_1}}^{m_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{m_n}/x_n\rangle P_1 \cdots P_r$$

where the bags P_j's are either ${\mathbf{P}_q}^m$, where m is finite, or $(\mathbf{U}^n_i)^{\infty}$, or Ω^{∞}, and the m_j's are finite. Let us call *canonical* a context of this form.

Lemma 5.7 (The Separation Lemma) *Let $M, N \in \Lambda$ be such that $M \not\preceq^{\eta}_{k} N$ for some k. Then there exists a canonical context $C \in \Lambda_{\mathsf{m}}[]$ closing both M and N such that $C[M] \Downarrow_{\mathsf{m}}$ and $C[N] \Uparrow_{\mathsf{m}}$.*

The proof is given in the Appendix B. Let us just recall here Böhm's extraction technique: assume for instance that

where Y is "obviously" not greater than V – for instance $Y = \Omega$ while V is a value (we assume that X, Y, Z and V are closed). Then one would like to find a context C extracting the difference, showing up Y from $C[M]$, and V from $C[N]$. Since Y and V are the second arguments of x in (xXY) and (xXV), one should substitute $\mathbf{U}_2^2$ for the second occurrence of x in M and N. However, we cannot simply use the context $[]\mathbf{U}_2^2$, since x occurs in the head position in M and N. In this position, the difference is in the first argument, so that we should use here $\mathbf{U}_1^2$ for x. The solution is to first replace x by a tupling combinator, namely $\mathbf{P}_2$ (because x is of arity 2 in M and N), since then the abstraction on x will be replaced by a series of new abstractions, one for each occurrence of x, as it is clear from:

$$\begin{array}{lcl} M\mathbf{P}_2 & =_\beta & \lambda z.z(\lambda y.yXY)Z \\ N\mathbf{P}_2 & =_\beta & \lambda z.z(\lambda y.yXV)Z \end{array}$$

Now we can apply these terms to the sequence $\mathbf{U}_1^2\mathbf{U}_2^2$ to achieve the desired extraction. Note that a complication arises whenever x also occurs free in the pair of subterms that we want to extract, because this variable will be replaced by a tupling combinator. For instance, as we will see in the Proposition 6.1, using λ-contexts cannot allow us to separate the two terms $\mathbf{G}_1 = \lambda x.x(x\Omega\Omega)(x(x\Omega\Omega)\Omega)$ and $\mathbf{G}_2 = \lambda x.x(x\Omega\Omega)(x(\lambda y\, x\Omega\Omega y)\Omega)$ – the reader is invited to draw the trees: both $\mathbf{G}_1\mathbf{P}_2\mathbf{U}_2^2\mathbf{U}_1^2$ and $\mathbf{G}_2\mathbf{P}_2\mathbf{U}_2^2\mathbf{U}_1^2$ converge, respectively to $(x\Omega\Omega)[\mathbf{P}_2/x] =_\beta \lambda z.z\Omega\Omega$ and $(\lambda y.x\Omega\Omega y)[\mathbf{P}_2/x]$. A similar, simpler example is provided by the pair $\Delta = \lambda x\, xx$ and $\Delta' = \lambda x\, x(\lambda y\, xy)$ of λ-terms of Example 3.8. This is where we use the multiplicities, basically by allowing only a finite amount of tupling combinators.

We must point out that the separation lemma is the only occasion where we really need the power of finite multiplicities. In the basic case where $M \not\preceq_1^\eta N$, it turns out that M and N are already separable in the lazy λ-calculus. However, it is important to be able to give in this case a resource of multiplicity 1 for the head variable (if any). Then to prove by induction the separation property for $M \not\preceq_{k+1}^\eta N$, we just increase the multiplicity of the resource for the head variable by one. The discussion above should also indicate that, as regards affine λ-terms, the multiplicities are harmless: for such terms, any bag E^{m+1} is like E^∞ since E will never be used more than once (E^0 is always like Ω^∞, since deadlock and divergence are identified). That is, one may prove (see [7], and the Appendix B):

Lemma 5.8 (The Separation Lemma, Affine case) *Let M and N be affine λ-terms such that $M \not\preceq_k^\eta N$ for some k. Then there exists a λ-context C closing both M and N such that $C[M] \Downarrow_{\mathsf{m}}$ and $C[N] \Uparrow_{\mathsf{m}}$. That is, $M \preceq_\ell N \;\Rightarrow\; M \sqsubseteq_{\mathcal{L}}^\eta N$ for M and N affine λ-terms.*

Now we can establish our main result:

Theorem 5.9 *Let M and N be λ-terms. Then $M \sqsubseteq_{\mathcal{L}}^{\eta} N$ if and only if $M \preceq_{\mathsf{m}} N$.*

PROOF: The implication "$\Leftarrow$" is the Separation Lemma. For the converse, assume that $M \sqsubseteq_{\mathcal{L}}^{\eta} N$ and $C[M] \Downarrow_{\mathsf{m}}$. Then by the Approximation Lemma there exists $A \in \mathcal{A}(M)$ such that $C[A] \Downarrow_{\mathsf{m}}$. Clearly $A \sqsubseteq_{\mathcal{L}}^{\eta} M$, therefore $A \sqsubseteq_{\mathcal{L}}^{\eta} N$, hence $C[N] \Downarrow_{\mathsf{m}}$ by the Corollary 5.6. ∎

An obvious corollary of this theorem is:

Corollary 5.10 $M =_{\mathcal{L}} N \Leftrightarrow M \simeq_{\mathsf{m}} N$.

We mentioned in the introduction that Sangiorgi showed in [14] that the equality of Lévy-Longo trees $=_{\mathcal{L}}$ also coincides with the equality $\simeq_{\pi}$ induced by the π-calculus over λ-terms. We can then conclude that, as far as equality of λ-terms is concerned, the π-calculus and the λ_{m}-calculus have the same discriminating power:

$$(*) \qquad M \simeq_{\pi} N \;\Leftrightarrow\; M \simeq_{\mathsf{m}} N \;\Leftrightarrow\; M =_{\mathcal{L}} N$$

for $M, N \in \Lambda$. At first sight, this may be surprising, since the latter is a deterministic calculus. Then one could interpret this result as meaning that parallelism is of little use in separating λ-term: what is important is to be able to distinguish the successive appearances of a given variable in the head position. As shown by Böhm, the λ-calculus provides part of this ability. But it generally fails distinguishing subterms like $xM_1 \cdots M_k$ and $\lambda y.xM_1 \cdots M_k y$, and this is where limited resources are useful.

Regarding the affine λ-terms, it is easy to see that the various preorders we considered (except Lévy's one) collapse down to $\preceq_{\ell}$. Obviously one has $\preceq_{\mathsf{m}} \subseteq \preceq_{\ell}$, therefore $\sqsubseteq_{\mathcal{L}}^{\eta} \subseteq \preceq_{\ell}$ by the Theorem 5.9, while the other inclusion is the Lemma 5.8. Therefore, we have:

Theorem 5.11 *Let M and N be affine λ-terms. Then*

$$M \sqsubseteq_{\mathcal{L}}^{\eta} N \;\Leftrightarrow\; M \preceq_{\mathsf{m}} N \;\Leftrightarrow\; M \preceq_{\ell} N$$

In particular, for affine λ-terms, the π-calculus, as well as the λ_{m}-calculus, is not more discriminating than the usual lazy λ-calculus, that is:

$$(**) \qquad M \simeq_{\pi} N \;\Leftrightarrow\; M \simeq_{\ell} N$$

for M and N affine. The results $(*)$ and $(**)$ provide a justification for our initial idea, which was to introduce multiplicities as a means to study the relationship between the π-calculus and the λ-calculus. Moreover, it looks somewhat easier to use applicative λ_{m}-contexts to show the equality or difference of λ-terms, rather than π-calculus contexts – this is not quite fair, however: the simplest way is to draw Lévy-Longo trees!

6 Convergence testing and parallel features

Our sentence that "parallelism is of little use in separating λ-term" sounds contradicting the idea that, according to Plotkin and Abramsky, some parallel features are missing from the λ-calculus. In this section we discuss this point.

In investigating the *full abstraction problem* for the lazy λ-calculus, Abramsky found out that this calculus is not expressive enough. The problem is the following (see [2]): there is a canonical denotational semantics, in a domain satisfying $D = (D \to D)_\perp$, which is adequate for the lazy λ-calculus, but the semantic preorder it induces is strictly finer than the observational preorder $\preceq_\ell$. In other words, lazy λ-contexts are not powerful enough. Abramsky showed that it is necessary to add some "convergence testing" and parallel facilities to gain the same discriminating power as the denotational semantics.

Let us recall some definitions and facts regarding the convergence testing combinators. Abramsky and Ong established the non full abstraction result by showing that the two λ-terms $\mathbf{A}_1 = \lambda x\, x(x(\lambda y\Omega)\Omega)(\lambda y\Omega)$ and $\mathbf{A}_2 = \lambda x\, x(\lambda z\, x(\lambda y\Omega)\Omega z)(\lambda y\Omega)$ are observationally indistinguishable, whereas they are denotationally different. Their difference shows off once one adds to the calculus a (sequential) convergence testing combinator. That is, we enrich the syntax of λ-calculus with the following clause:

$$M \ ::= \ \cdots \ \mid \ (\mathsf{c}M)$$

(this is not exactly the way taken in [2], where a constant C is considered; this may be defined by $\mathsf{C} = \lambda x(\mathsf{c}x)$). The extended set of terms is Λ_{c}, and the reduction rules are those of the lazy λ-calculus plus the following ones:

$$\mathsf{c}(\lambda x M) \to \mathbf{I} \qquad\qquad \frac{M \to M'}{\mathsf{c}M \to \mathsf{c}M'}$$

In this calculus, a value is still any abstraction. Then we can rephrase the notion of convergence, denoted $\Downarrow_{\mathsf{c}}$, and those of observational preorder and equivalence, denoted $\preceq_{\mathsf{c}}$ and $\simeq_{\mathsf{c}}$, respectively. In this theory the two terms $\mathbf{A}_1$ and $\mathbf{A}_2$ above are distinguished, since $\mathbf{A}_1\mathsf{C} \Uparrow_{\mathsf{c}}$ while $\mathbf{A}_2\mathsf{C} \Downarrow_{\mathsf{c}}$. These two terms are also distinguished in the λ-calculus with multiplicities, since $\mathbf{A}_1(\mathbf{U}_1^2)^1 \Uparrow_{\mathsf{m}}$ while $\mathbf{A}_2(\mathbf{U}_1^2)^1 \Downarrow_{\mathsf{m}}$.

Extending the lazy λ-calculus with a sequential convergence testing is not enough, however. Abramsky showed that one needs to add some parallel facility. That is, the syntax of λ-calculus is now enriched with the following clause:

$$M \ ::= \ \cdots \ \mid \ (\mathsf{p}MM)$$

Again, Abramsky and Ong considered a combinator P, which may be defined by $\mathsf{P} = \lambda xy.(\mathsf{p}xy)$. The reduction rules are those of the lazy λ-calculus plus the following ones:

$$\mathsf{p}(\lambda x M)N \to \mathbf{I} \qquad\qquad \mathsf{p}M(\lambda x N) \to \mathbf{I} \qquad\qquad \frac{M \to M' \ , \ N \to N'}{\mathsf{p}MN \to \mathsf{p}M'N'}$$

The function p is parallel because it looks at its two arguments simultaneously – and the fastest to converge will win the race. We denote by $\Downarrow_{\mathsf{p}}$ the convergence predicate in the lazy λ_{p}-calculus, and by $\preceq_{\mathsf{p}}$ and $\simeq_{\mathsf{p}}$ the corresponding observational preorder and equivalence. Abramsky showed (see [2], Proposition 7.2.10) that $\preceq_{\mathsf{p}}$ coincides with the semantical preorder induced by the interpretation into the canonical domain $D = (D \to D)_{\perp}$. In this interpretation, the two terms $\Delta = \lambda x\, xx$ and $\Delta' = \lambda x\, x(\lambda y\, xy)$ of Example 3.8 are equal, therefore $\Delta \simeq_{\mathsf{p}} \Delta'$. Note that one may let $(\mathsf{c}M) = (\mathsf{p}MM)$, so that λ_{p} is stronger than λ_{c}. Indeed, Abramsky and Ong give a pair of denotationally distinct λ_{c}-terms (thus distinct in λ_{p}) that are not distinguished in the λ_{c}-calculus.

Since we are studying various preorders over pure λ-terms, one may wonder whether there exists a pair of λ-terms M and N satisfying the same property, that is $M \preceq_{\mathsf{c}} N$ while $M \not\preceq_{\mathsf{p}} N$. The answer is positive. That is, one can show that the theory $\preceq_{\mathsf{p}}$, *restricted to λ-terms*, is strictly stronger than $\preceq_{\mathsf{c}}$.

Proposition 6.1 *Let* $\mathbf{G}_1 = \lambda x.x(x\Omega\Omega)(x(x\Omega\Omega)\Omega)$ *and* $\mathbf{G}_2 = \lambda x.x(x\Omega\Omega)(x(\lambda y\, x\Omega\Omega y)\Omega)$. *Then* $\mathbf{G}_1 \simeq_{\mathsf{c}} \mathbf{G}_2$, *while* $\mathbf{G}_1\mathsf{P} \Uparrow_{\mathsf{p}}$ *and* $\mathbf{G}_2\mathsf{P} \Downarrow_{\mathsf{p}}$.

PROOF: Firstly it is easy to see that $\mathbf{G}_1\mathsf{P} \Uparrow_{\mathsf{p}}$ and $\mathbf{G}_2\mathsf{P} \Downarrow_{\mathsf{p}}$. Now we show that $\mathbf{G}_1 \simeq_{\mathsf{c}} \mathbf{G}_2$. This requires extending the notion of functionality order to the λ_{c}-calculus. In fact, one just has to modify the definition of proper order 0, as follows: M has proper order 0, denoted $M \in \mathsf{PO}_0^{\mathsf{c}}$, iff

$$M \in \mathsf{O}_0^{\mathsf{c}} \text{ and } \neg[\ \exists k \exists x.\ \exists \vec{N}_1, \cdots, \vec{N}_k.\ M \xrightarrow{*}_{\mathsf{c}} \mathsf{c}(\cdots \mathsf{c}(x\vec{N}_1)\vec{N}_2 \cdots)\vec{N}_k\ ]$$

Moreover, we use the property of λ_{c} that $M \Downarrow_{\mathsf{c}} \Rightarrow\ M \simeq_{\mathsf{c}} \lambda x\, Mx$ for any closed M. By this property, we just have to prove that, whenever $M\Omega\Omega \Uparrow_{\mathsf{c}}$, the terms $\mathbf{G}_1 M$ and $\mathbf{G}_2 M$ behave in the same way, as far as convergence is concerned. The interesting cases are when $M \in \mathsf{O}_i^{\mathsf{c}}$ with $i \leq 2$.

($M \in \mathsf{O}_0^{\mathsf{c}}$) Both $\mathbf{G}_1 M \Uparrow_{\mathsf{c}}$ and $\mathbf{G}_2 M \Uparrow_{\mathsf{c}}$.

($M \in \mathsf{O}_1^{\mathsf{c}}$)

If $M \in \mathsf{PO}_1^{\mathsf{c}}$ then it is immediate that $\mathbf{G}_1 M \Uparrow_{\mathsf{c}}$ and $\mathbf{G}_2 M \Uparrow_{\mathsf{c}}$. Otherwise we have $M \simeq_{\mathsf{c}} \lambda x\,(\,\mathsf{c}(\cdots \mathsf{c}(x\vec{N}_1)\vec{N}_2 \cdots)\vec{N}_k\,)$. Then observe that

$$\begin{array}{rcl} \mathbf{G}_2 M & \simeq_{\mathsf{c}} & M(M\Omega\Omega)(M(\lambda y.M\Omega\Omega y)\Omega) \\ & \simeq_{\mathsf{c}} & (\,\mathsf{c}(\cdots \mathsf{c}((M\Omega\Omega)\vec{N}_1)\vec{N}_2 \cdots)\vec{N}_k)(M(\lambda y\, M\Omega\Omega y)\Omega) \Uparrow_{\mathsf{c}} \end{array}$$

and

$$\begin{array}{rcl} \mathbf{G}_1 M & \simeq_{\mathsf{c}} & M(M\Omega\Omega)(M(M\Omega\Omega)\Omega) \\ & \simeq_{\mathsf{c}} & (\,\mathsf{c}(\cdots \mathsf{c}((M\Omega\Omega)\vec{N}_1)\vec{N}_2 \cdots)\vec{N}_k)(M(M\Omega\Omega)\Omega) \Uparrow_{\mathsf{c}} \end{array}$$

($M \in \mathsf{O}_2^{\mathsf{c}}$)

The case where $M \in \mathsf{PO}_2^{\mathsf{c}}$ is again immediate. Otherwise, if $M \in \mathsf{O}_2^{\mathsf{c}} \setminus \mathsf{PO}_2^{\mathsf{c}}$ then we have $M \simeq_{\mathsf{c}} \lambda x_1 x_2.\,(\,\mathsf{c}(\cdots \mathsf{c}(y\vec{N}_1)\vec{N}_2 \cdots)\vec{N}_k\,)$. So we have two subcases, according to $y = x_1$ and $y = x_2$. We leave to the reader to check that, in any case, $\mathbf{G}_1 M \Uparrow_{\mathsf{c}}$ and $\mathbf{G}_2 M \Uparrow_{\mathsf{c}}$. ■

The two λ-terms $\mathbf{G}_1$ and $\mathbf{G}_2$ are also differentiated using contexts with multiplicities. Indeed, it is easy to check that:

$$\mathbf{G}_2\mathbf{P}_2{}^2(\mathbf{U}_2^2)^1(\mathbf{U}_1^2)^1 \Downarrow_{\mathsf{m}} \qquad \text{while} \qquad \mathbf{G}_1\mathbf{P}_2{}^2(\mathbf{U}_2^2)^1(\mathbf{U}_1^2)^1 \Uparrow_{\mathsf{m}}$$

This example shows that contexts with multiplicities may simulate the use of parallel convergence testing. This is a general fact, as we will see. This would be obvious if parallel convergence testing were definable in the λ-calculus with multiplicities. However, this is not the case; in fact, sequential convergence testing is not even λ_{m}-definable, as we now show.

Lemma 6.2 *There is no closed λ_{m}-term T such that for every E:*

$$TE^1 \simeq_{\mathsf{m}} \begin{cases} \mathbf{I} & \text{if } E \Downarrow_{\mathsf{m}} \\ \Omega & \text{otherwise} \end{cases}$$

PROOF: We proceed as Abramsky and Ong [2], showing that for every $E \in \Lambda^0_{\mathsf{m}}$:

$$E \simeq_{\mathsf{m}} \mathbf{I} \;\text{ or }\; [E\Omega^1 \Downarrow_{\mathsf{m}} \Leftrightarrow E(\lambda y\Omega)^1 \Downarrow_{\mathsf{m}}]$$

This is obvious if $E \in \mathsf{O}^{\mathsf{m}}_n$ with $n \neq 1$, or $E \in \mathsf{PO}^{\mathsf{m}}_1$. Now assume that $E \in \mathsf{O}^{\mathsf{m}}_1 \setminus \mathsf{PO}^{\mathsf{m}}_1$. Then $E \xrightarrow{*}_{\mathsf{m}} \equiv \lambda xF$ with $F \xrightarrow{*}_{\mathsf{m}} xQ_1 \cdots Q_k$, where no Q_i is a substitution of the shape $\langle P/x \rangle$. If all the Q_i's are substitutions or $k = 0$ then $E \simeq_{\mathsf{m}} \mathbf{I}$. Otherwise, at least one Q_i is not a substitution and $(\lambda xF)\Omega^1 \Uparrow_{\mathsf{m}}$ and $(\lambda xF)(\lambda y\Omega)^1 \Uparrow_{\mathsf{m}}$. ∎

It is indeed possible to show that the preorder $\preceq_{\mathsf{p}}$, restricted to λ-terms, is weaker than $\preceq_{\mathsf{m}}$, see [7]. However, this turns out to be a consequence of our main result and of some results by Ong. He introduced the lazy PSE-preorder to characterize the local structure of some models of the lazy λ-calculus. In particular, he showed in [13] a soundness result (Theorem 3.4.1.3), and a consequence of this is that the lazy PSE-preorder $\sqsubseteq^{\eta}_{\mathcal{L}}$ is adequate with respect to the denotational semantics. In other words

$$M \sqsubseteq^{\eta}_{\mathcal{L}} N \;\Rightarrow\; M \preceq_{\mathsf{p}} N$$

Then an obvious corollary of this and our characterization theorem (together with previously mentioned facts) is:

Theorem 6.3 *Let M and N be λ-terms. Then*

$$M \preceq_{\mathsf{m}} N \;\Rightarrow\; M \preceq_{\mathsf{p}} N \;\Rightarrow\; M \preceq_{\mathsf{c}} N \;\Rightarrow\; M \preceq_{\ell} N$$

Moreover, none of these implications can be reversed. The counterexamples are:

$$\Delta' \not\preceq_{\mathsf{m}} \Delta \quad \text{while} \quad \Delta' \simeq_{\mathsf{p}} \Delta$$

where $\Delta = \lambda x\, xx$ and $\Delta' = \lambda x\, x(\lambda y\, xy)$,

$$\mathbf{G}_2 \not\preceq_{\mathsf{p}} \mathbf{G}_1 \quad \text{while} \quad \mathbf{G}_2 \simeq_{\mathsf{c}} \mathbf{G}_1$$

where $\mathbf{G}_1 = \lambda x.x(x\Omega\Omega)(x(x\Omega\Omega)\Omega)$ *and* $\mathbf{G}_2 = \lambda x.x(x\Omega\Omega)(x(\lambda y\, x\Omega\Omega y)\Omega)$,

$$\mathbf{A}_2 \not\preceq_{\mathsf{c}} \mathbf{A}_1 \quad \text{while} \quad \mathbf{A}_2 \simeq_{\ell} \mathbf{A}_1$$

where $\mathbf{A}_1 = \lambda x\, x(x(\lambda y\Omega)\Omega)(\lambda y\Omega)$ *and* $\mathbf{A}_2 = \lambda x\, x(\lambda z\, x(\lambda y\Omega)\Omega z)(\lambda y\Omega)$. *However, for* M *and* N *affine* λ*-terms, we have:*

$$M \preceq_{\mathsf{m}} N \;\Leftrightarrow\; M \preceq_{\mathsf{p}} N \;\Leftrightarrow\; M \preceq_{\mathsf{c}} N \;\Leftrightarrow\; M \preceq_{\ell} N$$

We may then conclude that parallelism is indeed useful in separating λ-term, as far as the purpose is to recover the canonical denotational semantics. For this purpose, introducing finite multiplicities is not appropriate: they provide a discriminating power which is far too strong.

We may also note another consequence of Theorem 6.3. In [4] we have shown that the semantical preorder, that is, equivalently, the preorder $\preceq_{\mathsf{p}}$ coincides with the observational preorder on the λ-calculus extended with convergence testing ($\mathsf{c}M$) and non-deterministic choice ($M \oplus N$) (or parallel composition ($M \parallel N$), see [6]). For instance, the two terms $\mathbf{G}_1$ and $\mathbf{G}_2$ of Proposition 6.1 are distinguished using non-deterministic choice: $\mathbf{G}_2(\mathbf{U}_1^2 \oplus \mathbf{U}_2^2)$ has a value, namely $\lambda y(\mathbf{U}_1^2 \oplus \mathbf{U}_2^2)\Omega\Omega y$, while $\mathbf{G}_1(\mathbf{U}_1^2 \oplus \mathbf{U}_2^2)$ always diverges (one does not really need the convergence testing ability in this case, and this is not surprising since these two terms are equated in λ_{c}). Then we may conclude that the λ-calculus with multiplicities is more discriminating over λ-terms than non-deterministic choice.

A η-expansion and the Crux Lemma

In this appendix we prove the "η-expansion lemma" 3.6 and the "crux lemma" 5.5. To this end we use a result established in [5], showing the adequacy of a *functionality theory* with respect to the observational semantics. The functionality theory is an adaptation of the "intersection type discipline" of Coppo *et al.* We refer the reader to [5] for the details, and just recall here the necessary definitions and facts. There are two kinds of functional characters, one for terms of the calculus, and another for the bags. These are given by the grammar:

$$\begin{array}{rcl} \phi & ::= & \omega \mid (\pi \to \phi) \\ \pi & ::= & \phi \mid (\pi \times \pi) \end{array}$$

The functionality theory is an inference system for proving sequents of the form $x_1{:}\,\pi_1, \ldots, x_k{:}\,\pi_k \vdash E{:}\,\phi$ and $x_1{:}\,\pi_1, \ldots, x_k{:}\,\pi_k \vdash P{:}\,\pi$. We use Γ, $\Delta \ldots$ to denote the assumptions, that is sequences $x_1{:}\,\pi_1, \ldots, x_k{:}\,\pi_k$. There is a first group of rules, related to the constructions of the calculus:

$$(\mathrm{var}):\ x{:}\,\phi \vdash x{:}\,\phi \qquad\qquad (\mathrm{abs}):\ \frac{x{:}\,\pi\,,\,\Gamma \vdash E{:}\,\phi}{\Gamma \vdash \lambda x E{:}\,\pi \to \phi}\quad (x \text{ not in } \Gamma)$$

$$(\mathrm{app}):\ \frac{\Gamma \vdash E{:}\,\pi \to \phi\ ,\ \Delta \vdash P{:}\,\pi}{\Gamma\,,\,\Delta \vdash (EP){:}\,\phi} \qquad\qquad (\mathrm{subs}):\ \frac{x{:}\,\pi\,,\,\Gamma \vdash E{:}\,\phi\ ,\ \Delta \vdash P{:}\,\pi}{\Gamma\,,\,\Delta \vdash E\langle P/x\rangle{:}\,\phi}\quad (x \text{ not in } \Gamma)$$

$$(\mathrm{singl}):\ \frac{\Gamma \vdash E{:}\,\phi}{\Gamma \vdash E^1{:}\,\phi} \qquad\qquad (\mathrm{bag}):\ \frac{\Gamma \vdash E^m{:}\,\pi_0\ ,\ \Delta \vdash E^n{:}\,\pi_1}{\Gamma\,,\,\Delta \vdash E^{m+n}{:}\,\pi_0 \times \pi_1}$$

Then there are rules independent from the structure of the terms, which hold for both λ_{m}-terms and bags. In particular, there is a rule subsuming all the structural manipulations on can make on an assumption. To state this rule, we write $\Gamma \gg \Delta$ whenever Δ results from Γ by a sequence of structural manipulations. That is, $\gg$ is the least preorder on assumptions satisfying:

$$\begin{array}{rcll} \Gamma\,,\,x{:}\,\pi\,,\,y{:}\,\pi'\,,\,\Delta & \gg & \Gamma\,,\,y{:}\,\pi'\,,\,x{:}\,\pi\,,\,\Delta & \qquad(\text{exchange}) \\ \Gamma & \gg & x{:}\,\pi\,,\,\Gamma & \qquad(\text{weakening}) \\ x{:}\,\pi_0\,,\,x{:}\,\pi_1\,,\,\Gamma & \gg & x{:}\,\pi_0 \times \pi_1\,,\,\Gamma & \qquad(\text{product}) \end{array}$$

Moreover, we use a congruence $\sim$ over formulae, given as the least one for which the product $\pi_0 \times \pi_1$ is associative, commutative and has ω as a unit. Then, using T to denote either a term or a bag, and τ, σ to denote formulae of any kind, our last rules are:

$$(\mathrm{triv}):\ \vdash T{:}\,\omega \qquad\qquad (\mathrm{struc}):\ \frac{\Gamma \vdash T{:}\,\tau}{\Delta \vdash T{:}\,\sigma}\quad \Gamma \gg \Delta\ \&\ \tau \sim \sigma$$

We shall use $\Gamma \vdash_{\mathcal{F}} T{:}\,\tau$ to mean that the sequent $\Gamma \vdash T{:}\,\tau$ is provable in the functionality system. For any term E, let us denote by $\mathcal{F}(E)$ the set of pairs (Γ, ϕ) such that $\Gamma \vdash_{\mathcal{F}} E{:}\,\phi$. In [5] we proved the following *Adequacy Theorem*:

$$\mathcal{F}(E) \subseteq \mathcal{F}(F) \ \Rightarrow\ E \preceq_{\mathsf{m}} F$$

In the proof we established a property known as the "subject expansion" property. Regarding λ-terms, one may show a stronger property, namely that β-conversion preserves the functional characters:

Lemma A.1 $\quad M =_\beta N \ \Rightarrow\ \mathcal{F}(M) = \mathcal{F}(N)$

PROOF OUTLINE: The implication $M \to_\beta N \ \Rightarrow\ \mathcal{F}(N) \subseteq \mathcal{F}(M)$ is an immediate consequence of the subject expansion property (Proposition 3.7 of [5]), and of the correspondence between β-reduction and λ_{m}-evaluation (see the proof of Proposition 2.8 in [5]). The inclusion $\mathcal{F}((\lambda x M)N) \subseteq \mathcal{F}(M[N/x])$ is an easy consequence of the Lemma 3.4 of [5] (and, again, of the correspondence between the notions of reduction). This implies the "subject reduction" property $M \to_\beta N \ \Rightarrow\ \mathcal{F}(M) \subseteq \mathcal{F}(N)$ since the relation $\mathcal{F}(E) \subseteq \mathcal{F}(F)$ is a precongruence. The lemma follows by the Church-Rosser property. ∎

Note that this lemma, combined with the Adequacy Theorem, shows the Proposition 3.7. We shall also use another property proved in [5] (Lemma 3.3):

Lemma A.2 $\Gamma \vdash_{\mathcal{F}} E\colon \phi$ *if and only if there exists an assumption* Δ *on free variables of* E *such that* $\Delta \gg \Gamma$ *and* $\Delta \vdash_{\mathcal{F}} E\colon \phi$.

Before proving the η-expansion lemma, we first observe the following:

Remark A.3 $x\colon \pi \vdash_{\mathcal{F}} x^\infty\colon \pi$ *for any bag formula* π.

The proof, by induction on π, consists in noting that, for $\pi = \phi$, one uses (var), (singl), (triv), (bag) and (struc), as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{x\colon \phi \vdash x\colon \phi}}{x\colon \phi \vdash x^1\colon \phi} \qquad \dfrac{}{\vdash x^\infty\colon \omega}}{x\colon \phi \vdash x^\infty\colon \phi \times \omega}}{x\colon \phi \vdash x^\infty\colon \phi}$$

while for $\pi = \pi_0 \times \pi_1$, we have, by induction on the formulae, and using (bag) and (struc):

$$\dfrac{\dfrac{\dfrac{\vdots}{x\colon \pi_0 \vdash x^\infty\colon \pi_0} \qquad \dfrac{\vdots}{x\colon \pi_1 \vdash x^\infty\colon \pi_1}}{x\colon \pi_0\,,\; x\colon \pi_1 \vdash x^\infty\colon \pi}}{x\colon \pi \vdash x^\infty\colon \pi}$$

Now the Lemma 3.6 is established, thanks to the Adequacy Theorem above, once we have shown:

Lemma A.4 $\Gamma \vdash_{\mathcal{F}} E\colon \phi \;\Rightarrow\; \Gamma \vdash_{\mathcal{F}} \lambda x(Ex^\infty)\colon \phi$ *for any* ϕ*, provided* x *is not free in* E.

PROOF: By the Lemma A.2, one may assume that x is not in Γ. Then one proceeds by induction on ϕ. The statement is trivial for $\phi = \omega$, by (triv) and (struc). Otherwise, $\phi =$

$\pi \to \psi$, and, using the Remark A.3 and the rules (app) and (abs), we have:

$$\dfrac{\dfrac{\dfrac{\vdots}{x{:}\,\pi \vdash x^\infty{:}\,\pi} \qquad \dfrac{\vdots}{\Gamma \vdash E{:}\,\pi \to \psi}}{x{:}\,\pi\,,\, \Gamma \vdash Ex^\infty{:}\,\psi}}{\Gamma \vdash \lambda(Ex^\infty){:}\,\phi}$$

This completes the proof of $E \preceq_{\mathsf{m}} \lambda(Ex^\infty)$. ■

Now from the Adequacy Theorem above, to prove our Crux Lemma it is enough to show:

$$x \sqsubseteq^\eta_{\mathcal{L}} X \ \&\ \Gamma \vdash_{\mathcal{F}} x{:}\,\phi \ \Rightarrow\ \Gamma \vdash_{\mathcal{F}} X{:}\,\phi$$

As a matter of fact, one can note that $\Gamma \vdash_{\mathcal{F}} x{:}\,\phi$ holds basically as an instance of (triv) or (var):

Lemma A.5 *$\Gamma \vdash_{\mathcal{F}} x{:}\,\phi$ if and only if either $\phi = \omega$ or $x{:}\,\psi \gg \Gamma$ for some ψ such that $\phi \sim \psi$.*

PROOF: The "$\Leftarrow$" direction is obvious, given the rules (triv), (var) and (struc) of the functionality system. The converse direction is easily established by induction on the proof of $\Gamma \vdash x{:}\,\phi$, which can only be inferred using (triv), (var) or (struc). ■

Given this lemma, we may reduce our task to show:

Lemma A.6 $x \sqsubseteq^\eta_{\mathcal{L}} X \ \Rightarrow\ x{:}\,\phi \vdash_{\mathcal{F}} X{:}\,\phi$

We proceed by induction on the *size* $|\phi|$ of the formulae, defined as follows:

$$\begin{array}{rcl} |\omega| & = & 0 \\ |\pi \to \phi| & = & 1 + |\pi| + |\phi| \\ |\pi_0 \times \pi_1| & = & |\pi_0| + |\pi_1| \end{array}$$

First we note the following:

Remark A.7 *If the lemma is true up to size k, that is: $|\phi| < k \ \ \& \ x \sqsubseteq^\eta_{\mathcal{L}} X \ \Rightarrow\ x{:}\,\phi \vdash_{\mathcal{F}} X{:}\,\phi$ then*

$$|\pi| < k \ \&\ x \sqsubseteq^\eta_{\mathcal{L}} X \ \Rightarrow\ x{:}\,\pi \vdash_{\mathcal{F}} X^\infty{:}\,\pi$$

The proof is essentially the same as that of Remark A.3, replacing the use of (var) by a call to the hypothesis $x{:}\,\phi \vdash_{\mathcal{F}} X{:}\,\phi$.

The statement of the lemma is obvious for $|\phi| = 0$, that is $\phi = \omega$, by (triv) and (struc). Otherwise, we have $\phi = (\pi_1 \to (\cdots \to (\pi_n \to \omega)\cdots))$ and $X =_\beta \lambda y_1 \ldots y_m . x Y_1 \cdots Y_m$ with

$y_i \sqsubseteq_{\mathcal{L}}^{\eta} Y_i$, by the Proposition 5.4. By the Lemma A.1 (in fact, the subject expansion property would suffice), it is enough to show that

$$x{:}\,\phi \vdash_{\mathcal{F}} \lambda y_1 \ldots y_m.xY_1 \cdots Y_m{:}\,\phi$$

We distinguish two subcases:

(a) $m \leq n$. Then, using the induction hypotheses, we have $y_i{:}\,\pi_i \vdash_{\mathcal{F}} Y_i^{\infty}{:}\,\pi_i$ for $1 \leq i \leq m$, by the Remark A.7. Therefore, using the rules (app) m times and then (abs) m times:

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{y_m{:}\,\pi_m \vdash Y_m^{\infty}{:}\,\pi_m} \qquad \dfrac{\dfrac{\dfrac{\vdots}{y_2{:}\,\pi_2 \vdash Y_2^{\infty}{:}\,\pi_2} \qquad \dfrac{\dfrac{\dfrac{\vdots}{}}{y_1{:}\,\pi_1 \vdash Y_1^{\infty}{:}\,\pi_1} \qquad \dfrac{}{x{:}\,\phi \vdash x{:}\,\phi}}{y_1{:}\,\pi_1, x{:}\,\phi \vdash xY_1^{\infty}{:}\,\pi_2 \to \cdots \pi_n \to \omega}}{\vdots}}{}}{y_m{:}\,\pi_m, \ldots, y_1{:}\,\pi_1, x{:}\,\phi \vdash xY_1^{\infty} \cdots Y_m^{\infty}{:}\,\pi_{m+1} \to \cdots \pi_n \to \omega}}{\vdots}}{x{:}\,\phi \vdash \lambda y_1 \ldots y_m.xY_1^{\infty} \cdots Y_m^{\infty}{:}\,\phi}$$

(b) $m \geq n$. This case is trivial, since for any M and Γ, we have

$$m \geq n \;\Rightarrow\; \Gamma \vdash_{\mathcal{F}} \lambda y_1 \ldots y_m.M{:}\,\pi_1 \to \cdots \pi_n \to \omega$$

This is because

$$y_n{:}\,\pi_n, \ldots, y_1{:}\,\pi_1, \Gamma \vdash \lambda y_{n+1} \ldots y_m.M{:}\,\omega$$

by (triv), and then one uses (abs) repetitively. ■

B The Separation Lemma

In the proof of the lemma, we shall use Σ, $\Xi \ldots$ to denote sequences of substitutions $\langle P_1/x_1\rangle \cdots \langle P_n/x_n\rangle$. More precisely, when we write $E\Sigma$ where $\Sigma = \langle P_1/x_1\rangle \cdots \langle P_n/x_n\rangle$ this must be read $(\cdots(E\langle P_1/x_1\rangle)\cdots\langle P_n/x_n\rangle)$. We recall that $\mathbf{P}_n \stackrel{\text{def}}{=} \lambda x_1 \cdots x_{n+1}.\, x_{n+1}x_1 \cdots x_n$.

Lemma 5.7 (The Separation Lemma) Let $M, N \in \Lambda$ and $\mathsf{fv}(M) \cup \mathsf{fv}(N) \subseteq \{x_1, \ldots, x_n\}$. If $M \not\leq_k^{\eta} N$ then there exists $p_1, \ldots, p_n$ and $m_1, \ldots, m_n$ in N such that for any $q_1, \ldots, q_n$ with $q_i \geq p_i$ there exist closed bags $P_1, \ldots, P_r$ with $M\langle {\mathbf{P}_{q_1}}^{m_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{m_n}/x_n\rangle P_1 \cdots P_r \Downarrow_{\mathsf{m}}$ and $N\langle {\mathbf{P}_{q_1}}^{m_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{m_n}/x_n\rangle P_1 \cdots P_r \Uparrow_{\mathsf{m}}$.

PROOF: We proceed by induction on k, assuming that k is the least integer such that $M \not\lesssim_k^\eta N$.

1. $k = 1$. We systematically examine the possible cases, according as $M \in \mathsf{PO}_m$ for some m or not (i.e. M has a head normal form), and similarly for N.

   (a) If $M \in \mathsf{PO}_m$ and $N \in \mathsf{O}_h$ with $h < m$, we let $m_i = 0$ and $p_i = 0$ for any i, $r = h$ and $P_j = \Omega$ (that is, more precisely, $P_j = \Omega^\infty$ – recall our convention that the omitted multiplicities are ∞) for any $j \leq r$.

   (b) If $M =_\beta \lambda z_1 \ldots z_m.xM_1 \cdots M_s$ and $N \in \mathsf{PO}_h$ with $h < \infty$, there are two sub-cases. In any of these sub-cases, r will not depend on $q_1, \ldots, q_n$.

   i. If $x = x_i$ for some $i \leq n$ (therefore $x \notin \{z_1, \ldots, z_m\}$), then we let $m_j = 0$ and $p_j = 0$ for $j \neq i$ and $m_i = 1$, and $p_i = s + (r - m)$ where $r = \mathsf{max}\,(m, h)$, and finally $P_j = \Omega$ for $1 \leq j \leq r$. Let $\Sigma = \langle \mathbf{P}_{q_1}{}^0/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^0/x_n\rangle$ and $\Xi = \langle P_1/z_1\rangle \cdots \langle P_m/z_m\rangle$. We have, using the fact that $L\langle K^\infty/v\rangle \simeq_{\mathsf{m}} L[K/v]$, as shown in [5]:

$$(M\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)\underbrace{\Omega\cdots\Omega}_{r}$$

$$\begin{aligned}
&\simeq_{\mathsf{m}} \ (\lambda z_1 \ldots z_m.(\mathbf{P}_{q_i} M_1 \cdots M_s)\Sigma)\Omega\cdots\Omega \\
&\simeq_{\mathsf{m}} \ ((\mathbf{P}_{q_i} M_1 \cdots M_s)\Sigma\Xi)\underbrace{\Omega\cdots\Omega}_{r-m} \\
&\simeq_{\mathsf{m}} \ \lambda y_{p_i+1} \ldots y_{q_i+1}.((y_{q_i+1} M_1 \cdots M_s \Omega \cdots \Omega y_{p_i+1} \ldots y_{q_i})\Sigma\Xi)
\end{aligned}$$

   while

$$(N\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)\underbrace{\Omega\cdots\Omega}_{r} \Uparrow_{\mathsf{m}}$$

   since $N \simeq_{\mathsf{m}} \lambda v_1 \ldots v_h.\Omega$ and $r \geq h$.

   ii. If $x = z_i$ for some $i \leq m$ we let $m_j = 0$ for any j, $r = \mathsf{max}\,(m, h)$ and $P_j = \Omega$ for $j \neq i$, and $P_i = \lambda y_1 \ldots y_l.y_l$ where $l = s + (r - m) + 1$. We have, if we let $\Sigma = \langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle$, and $\Xi = \langle P_1/z_1\rangle \cdots \langle P_m/z_m\rangle$:

$$(M\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)\underbrace{\Omega\cdots\Omega(\lambda y_1 \ldots y_l.y_l)\Omega\cdots\Omega}_{r}$$

$$\begin{aligned}
&\simeq_{\mathsf{m}} \ (\lambda z_1 \ldots z_m.(z_i M_1 \cdots M_s)\Sigma)\Omega\cdots\Omega(\lambda y_1 \ldots y_l.y_l)\Omega\cdots\Omega \\
&\simeq_{\mathsf{m}} \ (((\lambda y_1 \ldots y_l.y_l)M_1 \cdots M_s)\Sigma\Xi)\underbrace{\Omega\cdots\Omega}_{r-m} \ \simeq_{\mathsf{m}} \ \mathbf{I}
\end{aligned}$$

and $(N\langle {\mathbf{P}_{q_1}}^{m_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{m_n}/x_n\rangle)P_1 \cdots P_r \Uparrow_{\mathsf{m}}$ as in the previous case.

(c) If $M =_\beta \lambda z_1 \ldots z_m.xM_1 \cdots M_s$ and $N =_\beta \lambda z_1 \ldots z_h.zN_1 \cdots N_t$, there are again several sub-cases.

i. $h < m$. This case is similar to (1.a) above.

ii. If $h \geq m$, we distinguish again two sub-cases. Notice that, by α-conversion, we may assume $z_j \notin \mathsf{fv}(xM_1 \cdots M_s)$ for $m < j \leq h$.

ii.1 If $z \neq x$, then either x is free in M, that is $x = x_i$ for some i, or $x = z_i$ for some $i \leq m$. These cases are respectively similar to (1.b.i) and (1.b.ii) above (note that if z is free in N then $z = x_j$ with $j \neq i$, and $m_j = 0$).

ii.2. If $z = x$ then the only possibility is $t \neq s + h - m$, and still we may have either $x = x_i$ with $i \leq n$ or $x = z_i$ for some $i \leq m$. Note that if $x = x_i$ we may assume, by α-conversion, that $x \neq z_j$ for $m < j \leq h$ (otherwise we are back to the previous case). Therefore there are four sub-cases:

ii.2.1. If $t < s + h - m$ and $x = x_i$, we let $m_j = 0$ and $p_j = 0$ for $j \neq i$ and $m_i = 1$, $p_i = s + h - m$. Now if $q_1, \ldots, q_n$ are such that $q_j \geq p_j$ for any j, we let $r = h + q_i - t + 1$, $P_j = \Omega$ for $j \neq h + q_i - p_i + 1$ (note that $h + q_i - p_i + 1 < r$ since $t < p_i$) and $P_{h+q_i-p_i+1} = \lambda y_1 \ldots y_l.y_l$ where $l = q_i + p_i - t + 1$. Let Σ and Ξ be as in the case (1.b.i). We have – in the underbraced parts, we only count the number of arguments, not of substitutions:

$$
\begin{aligned}
&(M\langle {\mathbf{P}_{q_1}}^{m_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{m_n}/x_n\rangle)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} (\lambda z_1 \ldots z_m.(\mathbf{P}_{q_i} M_1 \cdots M_s)\Sigma)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} ((\mathbf{P}_{q_i} \underbrace{M_1 \cdots M_s)\Sigma\Xi)P_{m+1} \cdots}_{p_i} \underbrace{P_{h+1} \cdots P_{h+q_i-p_i+1}}_{q_i-p_i+1} \underbrace{\cdots P_r}_{p_i-t} \\
&\qquad \simeq_{\mathsf{m}} ((P_{h+q_i-p_i+1} \underbrace{M_1 \cdots M_s\Omega \cdots \Omega)\Sigma\Xi)\Omega \cdots \Omega}_{q_i+p_i-t} \simeq_{\mathsf{m}} \mathbf{I}
\end{aligned}
$$

while

$$
\begin{aligned}
&(N\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ (\lambda z_1 \dots z_h.(\mathbf{P}_{q_i} N_1 \cdots N_t)\Sigma)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ ((\mathbf{P}_{q_i} \underbrace{N_1 \cdots N_t)\Sigma\Xi')P_{h+1} \cdots P_{h+q_i-p_i+1}}_{q_i-p_i+t+1} \underbrace{\cdots P_r}_{p_i-t} \\
&\qquad \simeq_{\mathsf{m}} \ ((\Omega N_1 \cdots N_t \Omega \cdots \Omega)\Sigma\Xi')\Omega \cdots P_{h+q_i-p_i+1} \cdots \Omega \ \simeq_{\mathsf{m}} \ \Omega
\end{aligned}
$$

where $\Xi' = \langle P_1/z_1\rangle \cdots \langle P_h/z_h\rangle$.

ii.2.2. If $t < s+h-m$ and $x = z_i$, we let $m_j = 0$ for any j, $r = (l-t)+h+1$ where $l = h-m+s$, $P_j = \Omega$ for $j \neq i$ and $j \neq h+1$, $P_i = \lambda y_1 \dots y_{l+1}.y_{l+1}$ and $P_{h+1} = \lambda v_1 \dots v_{l-t+1}.v_{l-t+1}$. We have:

$$
\begin{aligned}
&(M\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ (\lambda z_1 \dots z_m.(z_i M_1 \cdots M_s)\Sigma)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ ((P_i \underbrace{M_1 \cdots M_s)\Sigma\Xi)P_{m+1} \cdots}_{l} P_{h+1} \underbrace{\cdots P_r}_{l-t} \\
&\qquad \simeq_{\mathsf{m}} \ P_{h+1} \cdots P_r \ \simeq_{\mathsf{m}} \ \mathbf{I}
\end{aligned}
$$

while

$$
\begin{aligned}
&(N\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ (\lambda z_1 \dots z_h.(z_i N_1 \cdots N_t)\Sigma)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ ((P_i N_1 \cdots N_t)\Sigma\Xi')\underbrace{P_{h+1} \cdots P_r}_{l-t+1} \ \simeq_{\mathsf{m}} \ P_r = \Omega
\end{aligned}
$$

ii.2.3. If $t > s+h-m$ and $x = x_i$, we let $m_j = 0$ and $p_j = 0$ for $j \neq i$ and $m_i = 1$, $p_i = t$. Now if $q_1, \dots, q_n$ are such that $q_j \geq p_j$ for any j, we let $r = q_i - p_i + h + 1$

and $P_j = \Omega$ for any $j \leq r$. If we let $l = r + s - m$ (note that $l < q_i + 1$), we have:

$$\begin{aligned}
&(M\langle {\mathbf{P}_{q_1}}^{m_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{m_n}/x_n\rangle)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ (\lambda z_1 \ldots z_m.(\mathbf{P}_{q_i} M_1 \cdots M_s)\Sigma)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ ((\mathbf{P}_{q_i} M_1 \cdots M_s)\Sigma\Xi)P_{m+1} \cdots P_h \underbrace{\cdots P_r}_{q_i - p_i + 1} \\
&\qquad \simeq_{\mathsf{m}} \ \lambda y_{l+1} \ldots y_{q_i+1}.((y_{q_i+1} M_1 \cdots M_s P_{m+1} \cdots P_r y_{l+1} \ldots y_{q_i+1})\Sigma\Xi) \Downarrow_{\mathsf{m}}
\end{aligned}$$

while

$$\begin{aligned}
&(N\langle {\mathbf{P}_{q_1}}^{m_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{m_n}/x_n\rangle)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ (\lambda z_1 \ldots z_h.(\mathbf{P}_{q_i} N_1 \cdots N_t)\Sigma)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ ((\mathbf{P}_{q_i} N_1 \cdots N_t)\Sigma\Xi')\underbrace{P_{h+1} \cdots P_r}_{q_i - t + 1} \\
&\qquad \simeq_{\mathsf{m}} \ ((\Omega N_1 \cdots N_t)\Sigma\Xi')P_{h+1} \cdots P_{r-1} \Uparrow_{\mathsf{m}}
\end{aligned}$$

ii.2.4. If $t > s + h - m$ and $x = z_i$ ($i \leq m$), we let $m_j = 0$ and $p_j = 0$ for any j, and $r = h + 1$, $P_j = \Omega$ for $j \neq i$ and $P_i = \lambda y_1 \ldots y_{t+1}.y_{t+1}$. Let $l = s + h - m$. We have:

$$\begin{aligned}
&(M\langle {\mathbf{P}_{q_1}}^{m_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{m_n}/x_n\rangle)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ (\lambda z_1 \ldots z_m.(z_i M_1 \cdots M_s)\Sigma)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ ((P_i M_1 \cdots M_s)\Sigma\Xi)\underbrace{P_{m+1} \cdots P_r}_{h - m + 1} \\
&\qquad \simeq_{\mathsf{m}} \ \lambda y_{l+2} \ldots y_{t+1}.y_{t+1} \Downarrow_{\mathsf{m}}
\end{aligned}$$

while

$$\begin{aligned}
&(N\langle {\mathbf{P}_{q_1}}^{m_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{m_n}/x_n\rangle)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ (\lambda z_1 \ldots z_h.(z_i N_1 \cdots N_t)\Sigma)P_1 \cdots P_r \\
&\qquad \simeq_{\mathsf{m}} \ ((P_i N_1 \cdots N_t \Omega \cdots \Omega)\Sigma\Xi')P_{h+1} \ \simeq_{\mathsf{m}} \ P_{h+1} = \Omega
\end{aligned}$$

2. $k > 1$. Since we assumed that $M \leq^{\eta}_{k'} N$ for any $k' < k$, the only possibility here is $M =_{\beta} \lambda z_1 \ldots z_m.xM_1 \cdots M_s$ and $N =_{\beta} \lambda z_1 \ldots z_m y_1 \ldots y_t.xN_1 \cdots N_s Y_1 \cdots Y_t$, where $y_l \notin \mathsf{fv}(xM_1 \cdots M_s)$, with $M_l \not\leq^{\eta}_{k-1} N_l$ or $y_l \not\leq^{\eta}_{k-1} Y_l$ for some l. Then there are four sub-cases, according as $x = x_i$ or $x = z_i$ for some i. We only examine two of them, the other ones being similar.

   (a) $M_l \not\leq^{\eta}_{k-1} N_l$ and $x = x_i$. We have

   $$\mathsf{fv}(M_l) \cup \mathsf{fv}(N_l) \subseteq \{x_1, \ldots, x_n\} \cup \{z_1, \ldots, z_m\} \cup \{y_1, \ldots, y_t\}$$

   and in fact $\mathsf{fv}(M_l) \subseteq \{x_1, \ldots, x_n\} \cup \{z_1, \ldots, z_m\}$. By induction hypothesis, there exist $\pi_1, \ldots, \pi_n$, $\pi'_1, \ldots, \pi'_m$, $\pi''_1, \ldots, \pi''_m$, and $\mu_1, \ldots, \mu_n$, $\mu'_1, \ldots, \mu'_m$, $\mu''_1, \ldots, \mu''_m$ such that for any $\kappa_1, \ldots, \kappa_n$, $\kappa'_1, \ldots, \kappa'_m$, $\kappa''_1, \ldots, \kappa''_m$ with $\kappa_j \geq \pi_j$, $\kappa'_j \geq \pi'_j$ and $\kappa''_j \geq \pi''_j$ there exist $Q_1, \ldots, Q_\rho$ such that, if we let $\Sigma = \langle {\mathbf{P}_{\kappa_1}}^{\mu_1}/x_1\rangle \cdots \langle {\mathbf{P}_{\kappa_n}}^{\mu_n}/x_n\rangle$, and $\Xi = \langle {\mathbf{P}_{\kappa'_1}}^{\mu'_1}/z_1\rangle \cdots \langle {\mathbf{P}_{\kappa'_m}}^{\mu'_m}/z_m\rangle$ and $\Gamma = \langle {\mathbf{P}_{\kappa''_1}}^{\mu''_1}/y_1\rangle \cdots \langle {\mathbf{P}_{\kappa''_t}}^{\mu''_t}/y_t\rangle$

   $$M_l \Sigma \Xi \Gamma Q_1 \cdots Q_\rho \Downarrow_{\mathsf{m}}$$

   while

   $$N_l \Sigma \Xi \Gamma Q_1 \cdots Q_\rho \Uparrow_{\mathsf{m}}$$

   Then the proof consists in finding a context C such that $C[M]$ is essentially $M_l \Sigma \Xi \Gamma Q_1 \cdots Q_\rho$, while $C[N]$ gives $N_l \Sigma \Xi \Gamma Q_1 \cdots Q_\rho$.

   We let $m_j = \mu_j$ for $j \neq i$ and $m_i = \mu_i + 1$, $p_j = \pi_j$ for $j \neq i$ and $p_i = \mathsf{max}\,(\pi_i, s + t)$. Given $q_1, \ldots, q_n$ such that $q_j \geq p_j$ for any j, let $Q_1, \ldots, Q_\rho$ be a sequence satisfying the property above for $\kappa'_j = \pi'_j$ and $\kappa''_j = \pi''_j$. We let $r = \rho + h$ where $h = q_i + 1 + m - s$, and $P_1, \ldots, P_r$ be the sequence defined as follows:

   $$P_j = \begin{cases} {\mathbf{P}_{\pi'_j}}^{\mu'_j} & \text{if } 1 \leq j \leq m \\ {\mathbf{P}_{\pi''_{j-m}}}^{\mu''_{j-m}} & \text{if } m < j \leq t \\ (\lambda u_1 \ldots u_{q_i}.u_l)^\infty & \text{if } j = h \\ Q_{j-h} & \text{if } h < j \leq r \\ \Omega^\infty & \text{otherwise} \end{cases}$$

   This sequence is thus

   $$P_1 \cdots P_r = {\mathbf{P}_{\pi'_1}}^{\mu'_1} \cdots {\mathbf{P}_{\pi'_m}}^{\mu'_m}\ {\mathbf{P}_{\pi''_1}}^{\mu''_1} \cdots {\mathbf{P}_{\pi''_t}}^{\mu''_t}\ \underbrace{\Omega \cdots \Omega}_{q_i - (s+t)}\ P_h\ Q_1 \cdots Q_\rho$$

   Therefore we have, if $\Sigma = \langle {\mathbf{P}_{q_1}}^{\mu_1}/x_1\rangle \cdots \langle {\mathbf{P}_{q_n}}^{\mu_n}/x_n\rangle$, $\Xi = \langle {\mathbf{P}_{\pi'_1}}^{\mu'_1}/z_1\rangle \cdots \langle {\mathbf{P}_{\pi'_m}}^{\mu'_m}/z_m\rangle$ and $\Gamma = \langle {\mathbf{P}_{\pi''_1}}^{\mu''_1}/y_1\rangle \cdots \langle {\mathbf{P}_{\pi''_t}}^{\mu''_t}/y_t\rangle$, and $\Gamma' = \langle {\mathbf{P}_{\pi''_1}}^{\mu''_1}/v_{s+1}\rangle \cdots \langle {\mathbf{P}_{\pi''_t}}^{\mu''_t}/v_{s+t+1}\rangle$

$$
\begin{aligned}
&(M\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)P_1\cdots P_r\\
&\qquad \simeq_{\mathsf{m}} (\lambda z_1 \ldots z_m.(\mathbf{P}_{q_i}M_1\cdots M_s)\Sigma)P_1\cdots P_r\\
&\qquad \simeq_{\mathsf{m}} ((\mathbf{P}_{q_i}M_1\cdots M_s)\Sigma\Xi)P_{m+1}\cdots P_r\\
&\qquad \simeq_{\mathsf{m}} (\lambda v_{s+t+1}\ldots v_{q_i+1}((v_{q_i+1}M_1\cdots M_s v_{s+1}\cdots v_{q_i})\Sigma\Xi\Gamma'))P_{m+t+1}\cdots P_r\\
&\qquad \simeq_{\mathsf{m}} ((P_h\underbrace{M_1\cdots M_s v_{s+1}\cdots v_{q_i-(s+t)}\Omega\cdots\Omega}_{q_i})\Sigma\Xi\Gamma')Q_1\cdots Q_\rho\\
&\qquad \simeq_{\mathsf{m}} (M_l\Sigma\Xi\Gamma')Q_1\cdots Q_\rho \;\simeq_{\mathsf{m}}\; (M_l\Sigma\Xi\Gamma)Q_1\cdots Q_\rho \Downarrow_{\mathsf{m}}
\end{aligned}
$$

while

$$
\begin{aligned}
&(N\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)P_1\cdots P_r\\
&\qquad \simeq_{\mathsf{m}} (\lambda y_1 \ldots y_t.(\mathbf{P}_{q_i}N_1\cdots N_s Y_1\cdots Y_t)\Sigma\Xi)P_{m+1}\cdots P_r\\
&\qquad \simeq_{\mathsf{m}} ((\mathbf{P}_{q_i}N_1\cdots N_s Y_1\cdots Y_t)\Sigma\Xi\Gamma)P_{m+t+1}\cdots P_r\\
&\simeq_{\mathsf{m}} (\lambda v_{s+t+1}\ldots v_{q_i+1}((v_{q_i+1}N_1\cdots N_s Y_1\cdots Y_t v_{s+t+1}\cdots v_{q_i})\Sigma\Xi\Gamma))P_{m+t+1}\cdots P_r\\
&\qquad \simeq_{\mathsf{m}} ((P_h\underbrace{N_1\cdots N_s Y_1\cdots Y_t\Omega\cdots\Omega}_{q_i})\Sigma\Xi\Gamma)Q_1\cdots Q_\rho\\
&\qquad \simeq_{\mathsf{m}} (N_l\Sigma\Xi\Gamma)Q_1\cdots Q_\rho \Uparrow_{\mathsf{m}}
\end{aligned}
$$

(b) $y_l \not\lesssim^{\eta}_{k-1} Y_l$ and $x = z_i$. We use the induction hypothesis, as in the previous point, that is:

$$y_l\Sigma\Xi\Gamma Q_1\cdots Q_\rho \;\simeq_{\mathsf{m}}\; \mathbf{P}_{\kappa''_l}{}^{\mu''_l}Q_1\cdots Q_\rho \Downarrow_{\mathsf{m}}$$

while

$$N_l\Sigma\Xi\Gamma Q_1\cdots Q_\rho \Uparrow_{\mathsf{m}}$$

We let $m_j = \mu_j$ and $p_j = \pi_j$ for any j. Now given $q_1,\ldots,q_n$ such that $q_j \geq p_j$ for any j, let $Q_1,\ldots,Q_\rho$ be a sequence satisfying the property above for $\kappa'_j = \pi'_j$ and $\kappa''_j = \pi''_j$. We let $q = \mathsf{max}\,(\pi'_i, s+t)$ and $r = \rho + h$ where $h = q+1+m-s$, and let $P_1,\ldots,P_r$ be the following sequence:

$$\mathbf{P}_{\pi'_1}{}^{\mu'_1} \cdots \mathbf{P}_q{}^{\mu'_i+1} \cdots \mathbf{P}_{\pi'_m}{}^{\mu'_m}\ \mathbf{P}_{\pi''_1}{}^{\mu''_1} \cdots \mathbf{P}_{\pi''_t}{}^{\mu''_t}\ \underbrace{\Omega \cdots \Omega}_{q-(s+t)}\ P_h\ Q_1 \cdots Q_\rho$$

where $P_h = (\lambda u_1 \ldots u_{q_i}.u_{s+l})^\infty$. Then it is easy to check (in fact the proof in this case is entirely similar to the previous one) that

$$(M\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)P_1 \cdots P_r\ \simeq_{\mathsf{m}}\ \mathbf{P}_{\pi''_l}{}^{\mu''_l} Q_1 \cdots Q_\rho$$

while

$$(N\langle \mathbf{P}_{q_1}{}^{m_1}/x_1\rangle \cdots \langle \mathbf{P}_{q_n}{}^{m_n}/x_n\rangle)P_1 \cdots P_r\ \simeq_{\mathsf{m}}\ N_l \varSigma \varXi \varGamma Q_1 \cdots Q_\rho$$

The remaining cases, where $M_l \not\lesssim^{\eta}_{k-1} N_l$ and $x = z_i$ or $y_l \not\lesssim^{\eta}_{k-1} Y_l$ and $x = x_i$, are left to the reader. ■

The reader may have noticed that, in the base case ($k = 1$), we could have let $m_i = \infty$ instead of $m_i = 1$. As a matter of fact, M and N are already separated in the lazy λ-calculus whenever $M \not\lesssim^{\eta}_1 N$, since we could replace the bags $\mathbf{P}_{q_i}{}^0$ by Ω^∞. However, we could not prove the induction step ($k > 1$) if we had used λ-calculus contexts for separating M and N such that $M \not\lesssim^{\eta}_1 N$. For instance, we have $\lambda y.xy \not\lesssim^{\eta}_1 x$, which falls into the case (1.(c).i), giving the separating context $[]\langle \mathbf{I}^0/x\rangle$, hence, using the induction step, we find $[]\langle \mathbf{I}^1/x\rangle$ to separate Δ' from Δ.

On the other hand, one may note that if M and N are affine, then in the case ($k > 1$) of the proof, the head variable x cannot occur in $M_1, \ldots, M_s$, $N_1, \ldots, N_s$, $Y_1, \ldots, Y_t$ (and, by the Church-Rosser property, one may assume that these subterms are affine too, since the set of affine λ-terms is closed by β-reduction). This remark may be exploited to modify the proof in order to establish the separation lemma for affine λ-terms, that is the Lemma 5.8. The details are left to the reader (see [7] for a proof).

References

[1] M. Abadi, L. Cardelli, P. L. Curien, and J. J. Lévy. Explicit substitutions. *Journal of Functional Programming*, 1:375 – 416, 1991.

[2] S. Abramsky and L. Ong. Full abstraction in the lazy lambda calculus. *Information and Computation*, 105(2):159 – 267, 1993.

[3] H.P. Barendregt. *The Lambda Calculus.* North-Holland, 1985.

[4] G. Boudol. A Lambda Calculus for Parallel Functions. Technical Report 1231, INRIA Sophia-Antipolis, May 1990.

[5] G. Boudol. The lambda calculus with multiplicities. Technical Report 2025, INRIA Sophia-Antipolis, September 1993. Available as ~ `ftp/pub/papers/boudol/lambadm-prelim.ps.Z` by anonymous ftp at `cma.cma.fr`.

[6] G. Boudol. Lambda-calculi for (strict) parallel functions. *Information and Computation*, 108(1):51–127, 1994.

[7] G. Boudol and C. Laneve. The discriminating power of multiplicities in the λ-calculus. Draft. Available as ~ `ftp/pub/papers/boudol/dpm-draft.ps.Z` by anonymous ftp at `cma.cma.fr`, 1994.

[8] J.J. Lévy. An algebraic interpretation of the $\lambda\beta K$-calculus; and an application of a labelled λ-calculus. *Theoretical Computer Science*, 2(1):97 – 114, 1976.

[9] J.J. Lévy. *Réductions Correctes et Optimales dans le Lambda Calcul.* PhD thesis, Université Paris VII, 1978.

[10] G. Longo. Set theoretical models of lambda calculus: Theories, expansions and isomorphisms. *Annals of Pure and Applied Logic*, 24:153 – 188, 1983.

[11] R. Milner. Fully abstract models of typed λ-calculi. *Theoretical Computer Science*, 4:1 – 22, 1977.

[12] R. Milner. Functions as processes. *Mathematical Structures in Computer Science*, 2:119 – 141, 1992.

[13] L. Ong. *The Lazy Lambda Calculus: an Investigation into the Foundations of Functional Programming*. PhD thesis, Imperial College of Science and Technology, University of London, 1988.

[14] D. Sangiorgi. The lazy lambda calculus in a concurrency scenario. *Information and Computation*, 120(1):120 – 153, 1994.

[15] C. Wadsworth. The relation between computational and denotational properties for Scott D_∞-model of the lambda-calculus. *SIAM Journal on Computing*, 5:488 – 521, 1976.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399