Under consideration for publication in Math. Struct. in Comp. Science

A Deterministic Rewrite System for the Probabilistic λ-Calculus

Thomas LEVENTIS

Institut de Mathématiques de Marseille (I2M)
Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France

Institut de Recherche en Informatique Fondamentale (IRIF)
Université Paris Diderot, Paris, France

E-mail: thomas.leventis@ens-lyon.org

Received November 8, 2018

In this paper we present an operational semantics for the "call-by-name" probabilistic λ-calculus, whose main feature is to use only deterministic relations and to have no constraint on the reduction strategy. The calculus enjoys similar properties to the usual λ-calculus. In particular we prove it to be confluent, and we prove a standardisation theorem.

Introduction

Probabilistic λ-calculus Randomness appeared naturally in the world of programming languages. It is very useful to have some probabilistic primitive, for instance to generate random data on which to run a computation, or to implement some non deterministic heuristic. To study the role of this non determinism from a theoretical point of view one need to enrich the usual models of computation, such as Turing machines or λ-calculus, with probabilistic behaviours.

In this article we consider the λ-calculus with an additional family of constructors $+_p$ for $p \in [0,1]$ representing binary probabilistic choices. Given any two terms M and N there is a term $M +_p N$ which will behave as M with probability p and as N with probability $1-p$. A problem arises immediately when we consider such constructors: we have to choose a reduction strategy for the calculus to actually describe a probabilistic behaviour. Indeed let us consider the term $(\lambda b.\underline{equiv}\ b\ b)\ (\mathbb{T} +_p \mathbb{F})$ where $p \in]0,1[$, $\mathbb{T}$ and $\mathbb{F}$ are some encoding of the booleans and $\underline{equiv}$ is a binary function encoding boolean equivalence. If we directly reduce the sum this term reduces to $(\lambda b.\underline{equiv}\ b\ b)\ \mathbb{T}$ with probability p and to $(\lambda b.\underline{equiv}\ b\ b)\ \mathbb{F}$ with probability $1-p$, and in both cases this will reduce to $\mathbb{T}$. But if we decide to perform the β-reduction first, we will be able to reduce $\underline{equiv}\ (\mathbb{T} +_p \mathbb{F})\ (\mathbb{T} +_p \mathbb{F})$ to false. So there exist two probabilistic λ-calculi, the call-by-name one and the call-by-value one. In the literature those calculi are always studied with a very constrained reduction strategy, such that there is always at most one available redex (see for instance (Ehrhard and Danos 2011) and (Ehrhard, Pagani, and Tasson 2011) about the call-by-name semantics, or (Dal Lago and Zorzi 2012) about both call-by-value and call-by-name). Then from every term there is either a unique β-reduction, or two reductions of single sum with probabilities p and $1-p$, or no reduction at all, and we can easily build the semantics of a term with probability distributions.

Reductions and equations In the deterministic λ-calculus β-reduction can be turned into an equality on terms. The reduction can be performed under any context so the reflexive symmetric transitive closure $=_\beta$ of β-reduction is a congruence, i.e. it is an equivalence and if $M =_\beta N$ then $C[M] =_\beta C[N]$ for any context C. Besides reduction is confluent so two terms are equivalent if and only if they have a common reduct. More generally any relation $=_\mathcal{T}$ on terms is an *equational theory* if it is a congruence and it respects the β-rule, i.e. $(\lambda x.M)\ N =_\mathcal{T} M\left[{}^{N}/_{x}\right]$ for any M and N.

None of this works with the usual presentations of the probabilistic λ-calculus. The reduction is not confluent (as we have $M +_p N \xrightarrow{p} M$ and $M +_p N \xrightarrow{1-p} N$, and in general M and N have no common reduct), it can not be performed under arbitrary context, and it is not even a pure relation between terms, as it is indexed by probabilities. Yet it makes sense to think about equational theories for a probabilistic λ-calculus. An example of this is the model defined by Ehrhard, Pagani and Tasson (Ehrhard, Pagani, and Tasson 2011). They give a denotational semantics $[\![M]\!]$ for every probabilistic term M, and although they restrict their calculus to the head reduction we can observe that we have $[\![C[M]]\!] = [\![C[N]]\!]$ whenever $[\![M]\!] = [\![N]\!]$, and $[\![(\lambda x.M)\ N]\!] = [\![M\left[{}^{N}/_{x}\right]]\!]$ for any M and N.

The purpose of this article is to give a suitable description of the call-by-name probabilistic λ-calculus from which we can derive a notion of probabilistic equational theory. In other words we define a rewrite system based on deterministic reductions performed under arbitrary contexts. We also prove that this reduction is confluent, hence the probabilistic version of β-equivalence is characterised by the existence of a common reduct, just like the deterministic one. In addition we prove a standardisation theorem, which is necessary to link the general notion of reduction to the head reduction. Moreover we show that these properties are preserved when one consider the probabilistic choice modulo barycentric equivalence (Stone 1949): commutativity, associativity, idempotence and simplification of the trivial choices $+_1$ and $+_0$.

Related work The presentation of a quantitative λ-calculus with a deterministic and contextual rewrite system already exists in the literature. De'Liguoro and Piperno (Liguoro and Piperno 1995) studied such a system for the non-deterministic λ-calculus, i.e. the λ-calculus with a simple unlabelled choice constructor $+$ such that $M+N$ behaves as M or as N, and they proved a standardisation theorem. But their calculus is not confluent as they still consider a non-deterministic reduction where $M+N$ actually reduces into either M or N. Unlike in the probabilistic λ-calculus, it makes sense to say that the term $(\lambda b.\underline{equiv}\ b\ b)\ (\mathbb{T} +_p \mathbb{F})$ has six possible reductions (two if we reduce the sum first, four is we duplicate it) and that it reduces non-deterministically either to $\mathbb{T}$ or to $\mathbb{F}$.

The work presented in this paper is closer to what exists in the algebraic λ-calculus, i.e.

the λ-calculus with arbitrary linear combinations instead of just probability distributions. In particular a confluence result similar to ours has been proven by Alberti (Alberti 2014). Actually his techniques could likely be used to prove a confluence property in the probabilistic case, and conversely our proofs can be adapted to a more general setting of terms with labelled sums $+_l$. The main difference in our approaches is that he considers as algebraic terms some equivalence classes of syntactic terms, whereas we introduce the same equivalences but as an additional structure. This may be why he could not prove any standardisation result, which we do by playing on the equivalences we use.

Layout of the paper In Section 1 of the paper we describe the probabilistic λ-terms, and we sketch a proof of confluence of β-reduction (Proposition 1.2.1) to introduce the tools and technique we will use for more complex theorems, and in particular the notions of residuals and parallel reductions.

In Section 2 we detail the reduction $\to_+$ associated to sums (Definition 2.1). We show that every term has a unique normal form for this reduction (Theorem 2.2.3), which we call its *canonical form.*

We prove the confluence of the union of β-reduction and reduction of sums in Section 3 (Proposition 3.1.7), which is preserved when we consider the whole equivalence induced by $\to_+$ (Theorem 3.2.2). We also show that the calculus can be described with a notion of β-reduction between canonical forms (Theorem 3.2.3).

In Section 4 we introduce a weaker reduction for sums corresponding to weak head reduction, which has a more convenient structure than $\to_+$ and which is in particular orthogonal to β-reduction. We show a relation between this new reduction and $\to_+$ (Theorem 4.2.5).

Section 5 is dedicated to the proof of some standardisation theorems. We first prove a standardisation result for the calculus with weak reduction of sums (Theorem 5.1.9) using well-known proof techniques based on the orthogonality of the reduction system. Then we use the correspondence proven in the previous section to get a standardisation theorem for $\to_+$ (Theorem 5.2.3).

In Sections 6 we introduce an additional equivalence on terms. This equivalence expresses the commutativity, the associativity and the idempotence of probabilistic choice, as well as the triviality of choices with probability 0 or 1. We show that the sums in terms modulo this equivalence describe precisely probability distributions (Theorem 6.1.6). We then show that idempotence does play a role in the computations, but the rest of this equivalence has little influence on the reductions (Proposition 6.2.7).

In Section 7 we further study the role of idempotence, and we achieve new confluence and standardisation results with our additional equivalence (Theorems 7.1.10 and 7.2.3).

In conclusion in Section 8 we use our contextual operational semantics to justify the definition of a notion of probabilistic theory, and we define the observational equivalence of terms as an example.

Notation: If $\mathcal{R}$ and $\mathcal{S}$ are two relations we note $\mathcal{R} \cdot \mathcal{S}$ their composition, such that $M\ \mathcal{R} \cdot \mathcal{S}\ N$ iff there exists P with $M\ \mathcal{R}\ P$ and $P\ \mathcal{S}\ N$.

For any relation $\mathcal{R}$, we use the following notations:

— $\mathcal{R}^?$ is its reflexive closure: $M\ \mathcal{R}^?\ N$ iff $M\ \mathcal{R}\ N$ or $M = N$;
— $\mathcal{R}^+$ is its transitive closure;
— $\mathcal{R}^*$ is its reflexive transitive closure;
— $\mathcal{R}^n$ for $n \in \mathbb{N}$ is its n-iteration: $\mathcal{R}^0$ is the identity relation and $\mathcal{R}^{n+1} = \mathcal{R} \cdot \mathcal{R}^n$;
— for a reduction $\to$ we write $\leftarrow$ for the anti-reduction: $M \leftarrow N$ iff $N \to M$.

1. β-Reduction with Sums

1.1. *Probabilistic Terms*

We extend the syntax of the λ-calculus with probabilistic sums.

Definition 1.1. The set Λ^+ of probabilistic λ-terms is given inductively by:

$$M, N \in \Lambda^+ := x \mid \lambda x.M \mid M\ N \mid M +_p N,\ \ p \in [0,1].$$

Terms in Λ^+ are considered modulo α-equivalence, which is induced by $\lambda x.M =_\alpha \lambda y.N$ whenever y is not free in M and N is obtained by substituting all free occurrences of x in M by y.

Notation: When reading terms, we will always consider sums first: for instance the term $\lambda x.M +_p N\ P$ is to be read $(\lambda x.M) +_p (N\ P)$. We also consider abstractions before applications and we read the application as left associative: typically a head normal form $\lambda x_1...x_n.y\ P_1\ ...\ P_m$ is to be read $\lambda x_1...x_n.((x\ P_1)\ ...\ P_n)$.

Definition 1.2. The substitution $M\left[{}^P/_x\right]$ of P for x in M is defined by:
— $x\left[{}^P/_x\right] = P$;
— $y\left[{}^P/_x\right] = y$ if $y \neq x$;
— $(\lambda y.M)\left[{}^P/_x\right] = \lambda y.\left(M\left[{}^P/_x\right]\right)$ if $y \neq x$ and y is not free in P;
— $(M\ N)\left[{}^P/_x\right]) = \left(M\left[{}^P/_x\right]\right)\ \left(N\left[{}^P/_x\right]\right)$;
— $(M +_p M')\left[{}^P/_x\right]) = \left(M\left[{}^P/_x\right]\right) +_p \left(M'\left[{}^P/_x\right]\right)$.

Definition 1.3. We define contexts as terms with exactly one occurrence of a hole [ ]

$$C := [\ ] \mid \lambda x.C \mid C\ N \mid M\ C \mid C +_p N \mid M +_p C$$

and we note $C[P]$ the substitution of P for [ ] in C with variable capture:
— $([\ ])[P] = P$;
— $(\lambda y.C)[P] = \lambda y.C[P]$;
— $(C\ N)[P] = C[P]\ N$;
— $(M\ C)[P] = M\ C[P]$;
— $(C +_p N)[P] = C[P] +_p N$;
— $(M +_p C)[P] = M +_p C[P]$;

If C and C' are contexts then we define a context $C[C']$ in a similar way.

1.2. *β-Reduction*

Our goal is to give a meaningful operational semantics for this syntax using contextual relations, i.e. relations $\mathcal{R}$ such that if $M \mathrel{\mathcal{R}} N$ then $C[M] \mathrel{\mathcal{R}} C[N]$ for every context C, exclusively. Such an operational semantics necessarily includes β-reduction under arbitrary context.

Definition 1.4. The β-reduction $\to_\beta$ is:

$$(\lambda x.M)\ N \to_\beta M\left[{}^{N}/_{x}\right]$$

extended to arbitrary contexts:

$$\text{if } M \to_\beta M' \text{ then } C[M] \to_\beta C[M'].$$

Before giving any rule for probabilistic sums, we can check that the presence of sums does not hinder the standard proof of the confluence of β-reduction. The original proof of confluence for the deterministic λ-calculus uses labelled redexes to define a notion of parallel reduction which enjoys the diamond property, hence is confluent, and whose transitive closure matches the transitive closure of the β-reduction, hence the latter is also confluent ((Barendregt 1981), Section 11.2). Another technique is to define the same notion of parallel reduction directly by induction on terms (Takahashi 1995). The latter allows for more concise definitions and proofs to prove a confluence property. But the parallel reduction is also useful to prove the standardisation, and its inductive definition becomes far less convenient when we want to decompose the parallel reduction itself, especially in our probabilistic setting. So in this paper we will use labelled terms, and we will further discuss this choice in Subsection 5.1 before proving our standardisation theorem.

Here we will just sketch the proof for the confluence of β-reduction alone, and we will detail a similar but slightly more complicated proof involving some reduction rules for the sums in Section 3.

Subterms and Redexes. Given a term M we call *subterm* of M a term P in a given position in M, i.e. a pair (C, P) such that $M = C[P]$. Remark that a subterm of M is entirely defined by the context C, as for a fixed C there is at most one P such that $C[P] = M$, but it is often more convenient to describe the term P. For instance if we consider a term of the form $M_1 +_p M_2$, we sometimes write "the subterm M_1" for the subterm $([\ ] +_p M_2, M_1)$.

A *β-redex* of M is a subterm of M of the form $(C, (\lambda x.P)\ Q)$. If $\Delta = (C, (\lambda x.P)\ Q)$ is a redex of M we write $M \xrightarrow{\Delta}_\beta C\left[P\left[{}^{Q}/_{x}\right]\right]$ the reduction of Δ in M.

Labelled Terms. We consider terms with labelled β-redexes, which are given by:

$$\mathcal{M}, \mathcal{N} \in \Lambda_l^+ := x \mid \lambda x.\mathcal{M} \mid \mathcal{M}\ \mathcal{N} \mid \mathcal{M} +_p \mathcal{N} \mid (\lambda_0 x.\mathcal{M})\ \mathcal{N}$$

We generalise the definitions of variable substitution and context to this labelled setting. We also define β-reduction:

$$(\lambda x.\mathcal{M})\ \mathcal{N} \to_\beta \mathcal{M}\left[{}^{\mathcal{N}}/_{x}\right]$$

as well as β_0-reduction of labelled redexes:

$$(\lambda_0 x.\mathcal{M})\ \mathcal{N} \to_{\beta_0} \mathcal{M}\left[{}^{\mathcal{N}}/x\right]$$

both extended under arbitrary context.

Given a term M and a set $\mathcal{F}$ of β-redexes in M, for each subterm (C, P) of M we defined a labelled term $(C, P)_{\mathcal{F}}$, by induction on P:

— $(C, x)_{\mathcal{F}} = x$;
— $(C, \lambda x.P)_{\mathcal{F}} = \lambda x.\mathcal{P}$ with $\mathcal{P} = (C[\lambda x.[\ ]], P)_{\mathcal{F}}$;
— $(C, P\ Q)_{\mathcal{F}} = \mathcal{P}\ \mathcal{Q}$ with $\mathcal{P} = (C[[\ ]\ Q], P)_{\mathcal{F}}$ and $\mathcal{Q} = (C[P\ [\ ]], Q)_{\mathcal{F}}$ if $(C, P\ Q) \notin \mathcal{F}$;
— $(C, (\lambda x.P)\ Q)_{\mathcal{F}} = (\lambda_0 x.\mathcal{P})\ \mathcal{Q}$ with $\mathcal{P} = (C[(\lambda x.[\ ])\ Q], P)_{\mathcal{F}}$ and $\mathcal{Q} = (C[(\lambda x.P)\ [\ ]], Q)_{\mathcal{F}}$ if $(C, (\lambda x.P)\ Q) \in \mathcal{F}$;
— $(C, P +_p Q)_{\mathcal{F}} = \mathcal{P} +_p \mathcal{Q}$ with $\mathcal{P} = (C[[\ ] +_p Q], P)_{\mathcal{F}}$ and $\mathcal{Q} = (C[P +_p [\ ]], Q)_{\mathcal{F}}$.

Then the *labelling of* M *by* $\mathcal{F}$ is defined as $M_{\mathcal{F}} = ([\ ], M)_{\mathcal{F}}$. Conversely if $\mathcal{M} \in \Lambda_l^+$ we write $|\mathcal{M}|$ the term obtained by erasing all the labels in $\mathcal{M}$. In particular we always have $|M_{\mathcal{F}}| = M$. One can also easily check that given any $\mathcal{M} \in \Lambda_l^+$, there is a unique term M and a unique set $\mathcal{F}$ such that $\mathcal{M} = M_{\mathcal{F}}$.

Residual of β-redexes. Given a set $\mathcal{F}$ of β-redexes in a term M, every reduction $M \xrightarrow{\Delta}_{\beta} N$ can be lifted to Λ_l^+ along the labelling of M by $\mathcal{F}$. If $\Delta \notin \mathcal{F}$ then one can prove there is a unique labelled term $\mathcal{N}$ such that $M_{\mathcal{F}} \to_{\beta} \mathcal{N}$ and $|\mathcal{N}| = N$. If $\Delta \in \mathcal{F}$ then let Δ' be the labelled redex of $M_{\mathcal{F}}$ corresponding to Δ, there is a unique $\mathcal{N}$ such that $M_{\mathcal{F}} \xrightarrow{\Delta'} \mathcal{N}$ and we have $|\mathcal{N}| = N$. In both cases there is a unique set $\mathcal{G}$ of β-redexes in N such that $N_{\mathcal{G}} = \mathcal{N}$: this set is called the *set of residuals of* $\mathcal{F}$ for the reduction $M \xrightarrow{\Delta}_{\beta} N$.

Remark that the residuals are defined for the reduction *of a particular redex*. For instance let $I = \lambda x.x$, the term $I\ (I\ I)$ can only β-reduce into $I\ I$, but there are two different β-redexes which make this reduction possible. If we consider for instance the outermost β-redex $([\ ], I\ (I\ I))$, it may have one residual or none depending on the reduced redex:

$$(\lambda_0 x.x)\ (\underline{(\lambda x.x)\ (\lambda x.x)}) \to_{\beta} (\lambda_0 x.x)\ (\lambda x.x)$$
$$\underline{(\lambda_0 x.x)\ ((\lambda x.x)\ (\lambda x.x))} \to_{\beta_0} (\lambda x.x)\ (\lambda x.x)$$

Parallel Reduction. Reducing a set $\mathcal{F}$ of β-redexes in a term M means β_0-normalising $M_{\mathcal{F}}$. For deterministic terms the reduction β_0 is known to be weakly confluent and strongly normalising (Barendregt 1981), and the proof easily generalises to terms with sums, hence every labelled term has a unique normal form. Besides a normal form for $\to_{\beta_0}$ is precisely a term without label. So the result of the parallel reduction of $\mathcal{F}$ in M is well defined, and it is the unique term N such that $M_{\mathcal{F}} \to_{\beta_0}^* N$.

Definition 1.5. We write $\mathrm{NF}_0(M, \mathcal{F})$ for the unique β_0-normal form of $M_{\mathcal{F}}$, $M \xrightarrow{\mathcal{F}}_{\beta_{/\!/}} N$ if $N = \mathrm{NF}_0(M, \mathcal{F})$, and more generally $M \to_{\beta_{/\!/}} N$ if there exists $\mathcal{F}$ such that $M \xrightarrow{\mathcal{F}}_{\beta_{/\!/}} N$.

Confluence. The parallel reduction has the diamond property. Indeed if $\mathcal{F}$ and $\mathcal{F}'$ are sets of β-redexes in M and $\mathcal{G}'$ is the set of residuals of $\mathcal{F}'$ in $\mathrm{NF}_0(M, \mathcal{F})$ then $\mathrm{NF}_0(\mathrm{NF}_0(M, \mathcal{F}), \mathcal{G}') = \mathrm{NF}_0(M, \mathcal{F} \cup \mathcal{F}')$. This means that if $M \xrightarrow{\mathcal{F}}_{\beta_{/\!/}} N$ and $M \xrightarrow{\mathcal{F}'}_{\beta_{/\!/}} N'$ then let $P = \mathrm{NF}_0(M, \mathcal{F} \cup \mathcal{F}')$, we have $N \to_{\beta_{/\!/}} P$ and $N' \to_{\beta_{/\!/}} P$.

$$\begin{tikzcd}
& M \arrow[dl, "\mathcal{F}"'] \arrow[dr, "\mathcal{F}'"] \arrow[dd, dashed, "\mathcal{F} \cup \mathcal{F}'"] & \\
N \arrow[dr, dashed, "\mathcal{G}'"'] & & N' \arrow[dl, dashed, "\mathcal{G}"] \\
& P &
\end{tikzcd}$$

Moreover we have obviously $\to_\beta \subset \to_{\beta_{/\!/}}$ and $\to_{\beta_{/\!/}} \subset \to_\beta^*$, so $\to_{\beta_{/\!/}}^* = \to_\beta^*$. Consequently the confluence of the β-reduction is equivalent to the confluence of $\to_{\beta_{/\!/}}$, which is an immediate corollary of its diamond property.

Proposition 1.2.1. The β-reduction $\to_\beta$ is confluent on Λ^+.

2. Probabilistic Choice

2.1. *Operational Semantics of Sums*

The sum $M +_p N$ is meant to describe a probabilistic choice between M (with probability p) and N (with probability $1 - p$). So its operational semantics is usually given in terms of probabilistic reductions:

$$M +_p N \xrightarrow{p} M$$

$$M +_p N \xrightarrow{1-p} N$$

where $\xrightarrow{p}$ is to be understood as a reduction which happens with probability p.

For this semantics to actually describe a probabilistic behaviour we cannot allow reductions under arbitrary contexts and must choose a strategy. A simple example of that is the term $\delta\ (x +_{\frac{1}{2}} y)$ with $\delta = \lambda x.x\ x$. If we allow reduction under arbitrary context, we have the following reduction paths:

$$\delta\ (x +_{\frac{1}{2}} y) \begin{cases} \xrightarrow{1/2} & \delta\ x \ \to_\beta \ x\ x \\ \xrightarrow{1/2} & \delta\ y \ \to_\beta \ y\ y \end{cases}$$

$$\delta\ (x +_{\frac{1}{2}} y) \to_\beta (x +_{\frac{1}{2}} y)\ (x +_{\frac{1}{2}} y) \begin{cases} \xrightarrow{1/2} & x\ (x +_{\frac{1}{2}} y) \begin{cases} \xrightarrow{1/2} & x\ x \\ \xrightarrow{1/2} & x\ y \end{cases} \\ \xrightarrow{1/2} & y\ (x +_{\frac{1}{2}} y) \begin{cases} \xrightarrow{1/2} & y\ x \\ \xrightarrow{1/2} & y\ y \end{cases} \end{cases}$$

Intuitively one could say that if we reduce the sum first we get a first probability distribution over normal forms ($x\ x$ and $y\ y$ with probability $\frac{1}{2}$) and if we β-reduce first we get a second distribution ($x\ x$, $x\ y$, $y\ x$ and $y\ y$ with probability $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$). But we expect each probabilistic term to describe a single probability distribution.

When defining such a probabilistic calculus, one has to choose between call-by-name and call-by-value strategies. In this article we are interested in the call-by-name variant of this calculus. In this case the reduction of sums is usually restricted to head contexts as in (Dal Lago and Zorzi 2012) or (Ehrhard, Pagani, and Tasson 2011):

$$\lambda x_1 ... x_n.(M +_p N)\ P_1\ ...\ P_m \xrightarrow{p} \lambda x_1 ... x_n.M\ P_1\ ...\ P_m$$

$$\lambda x_1 ... x_n.(M +_p N)\ P_1\ ...\ P_m \xrightarrow{1-p} \lambda x_1 ... x_n.N\ P_1\ ...\ P_m$$

Then how to transpose such reductions into a contextual calculus? The trick is simple and is used in other quantitative calculi such as the algebraic λ-calculus (Vaux 2009),(Alberti 2014), although such a presentation has not been studied before in the probabilistic case. The key idea is not to let the probabilities leave the realm of terms. The fact that the term $\lambda x_1 ... x_n.(M +_p N)\ P_1\ ...\ P_m$ reduces into both $\lambda x_1 ... x_n.M\ P_1\ ...\ P_m$ and $\lambda x_1 ... x_n.N\ P_1\ ...\ P_m$ should not be described by two reductions labelled with probabilities. We can express this by a single reduction on terms without any additional information:

$$\lambda x_1 ... x_n.(M +_p N)\ P_1\ ...\ P_m \to \lambda x_1 ... x_n.M\ P_1\ ...\ P_m +_p \lambda x_1 ... x_n.N\ P_1\ ...\ P_m.$$

Then we obtain a calculus which describes the head reduction of sums, but in which it is perfectly safe to reduce under any context:

$$C\left[\lambda x_1 ... x_n.(M +_p N)\ P_1\ ...\ P_m\right] \to C\left[\lambda x_1 ... x_n.M\ P_1\ ...\ P_m +_p \lambda x_1 ... x_n.N\ P_1\ ...\ P_m\right].$$

It is actually more convenient to decompose such a reduction in elementary steps.

Definition 2.1. The reduction of sums $\to_+$ on Λ^+ is given by

$$\lambda x.(M +_p N) \to_+ \lambda x.M +_p \lambda x.N$$

$$(M +_p N)\ P \to_+ M\ P +_p N\ P$$

extended to arbitrary contexts.

Definition 2.2. The relation $\equiv_+$ is the reflexive symmetric transitive closures of $\to_+$.

2.2. *Canonical Terms*

The reduction $\to_+$ does not have an important computational meaning. Its role is to move the sums out of the terms in order to let the β-reduction do the actual computation. For instance if we have a term $((\lambda x.M) +_p N)\ P$, we need the reduction of sums to write $((\lambda x.M) +_p N)\ P \to_+ (\lambda x.M)\ P +_p N\ P$, but the main point of this reduction is to allow the β-reduction $(\lambda x.M)\ P +_p N\ P \to_\beta M\left[{}^{P}/_{x}\right] +_p N\ P$. For this reason the reduction $\to_+$ is very simple: it is confluent and strongly normalising, hence every term has a unique normal form for $\to_+$.

Proposition 2.2.1. The reduction $\to_+$ is weakly confluent.

Proof. If $M \to_+ N_1$ and $M \to_+ N_2$ we show by induction on M that there is P such that $N_i \to_+^* P$ for $i \in \{1,2\}$.

— If $(\lambda x.(Q_1 +_p Q_2)) \to_+ \lambda x.Q_1 +_p \lambda x.Q_2$ and $(\lambda x.(Q_1 +_p Q_2)) \to_+ (\lambda x.(Q_1' +_p Q_2'))$ with $Q_i \to_+^? Q_i'$ for $i \in \{1,2\}$ then both terms reduce to $\lambda x.Q_1' +_p \lambda x.Q_2'$.
— If $(Q_1 +_p Q_2)\ R \to_+ Q_1\ R +_p Q_2\ R$ and $(Q_1 +_p Q_2)\ R \to_+ (Q_1' +_p Q_2')\ R'$ with $Q_i \to_+^? Q_i'$ for $i \in \{1,2\}$ and $R \to_+^? R'$ then both terms reduce to $Q_1'\ R' +_p Q_2'\ R'$.
— Otherwise M is not a reduced redex and the result is immediate. For instance if $Q\ R \to_+ Q_1\ R_1$ and $Q\ RQ_2\ R_2$ with $Q \to_+^? Q_i$ and $R \sum^? R_i$ for $i \in \{1,2\}$ then either trivially or by induction hypothesis there are Q' and R' such that $Q_i \to_+^* Q'$ and $R_i \to_+^* R'$, hence $Q_i\ R_i \to_+^* Q'\ R'$.

□

Proposition 2.2.2. The reduction $\to_+$ is strongly normalising.

Proof. To prove the proposition we define a suitable weight on terms which decreases along the reductions. We define a family of weight functions $w_d : \Lambda^+ \to \mathbb{N}$ for $d \in \mathbb{N}$ by induction on terms:

— $\forall d \in \mathbb{N}, w_d(x) = 0$;
— $\forall d \in \mathbb{N}, w_d(\lambda x.M) = w_{d+1}(M)$;
— $\forall d \in \mathbb{N}, w_d(M\ N) = (w_{d+1}(M) + 1)(w_0(N) + 1)$;
— $\forall d \in \mathbb{N}, w_d(M +_p N) = w_d(M) + w_d(N) + d$.

Then whenever we have a reduction $M \to_+ N$ we have $w_d(M) > w_d(N)$ for all $d \in \mathbb{N}$. This is a consequence of the three following facts.

(1) $\forall M, N, \forall C,$ $\quad \big(\forall d \in \mathbb{N}, w_d(M) > w_d(N)\big) \Rightarrow \forall d \in \mathbb{N}, w_d(C[M]) > w_d(C[N])$.
(2) $\forall M, N, \forall d \in \mathbb{N},$ $\quad w_d\big(\lambda x.(M +_p N)\big) > w_d\big(\lambda x.M +_p \lambda x.N\big)$.
(3) $\forall M, N, P, \forall d \in \mathbb{N},$ $w_d\big((M +_p N)\ P\big) > w_d\big(M\ P +_p N\ P\big)$.

For instance we have $w_0((x +_p y)\ z) = 1$ as we have a sum on the left side of an application and $w_0(x\ y +_p y\ z) = 0$; $w_0(\lambda x.\lambda y.(x +_p y)) = 2$ as the sum is below two abstractions, $w_0(\lambda x.(\lambda y.x +_p \lambda y.y)) = 1$ and $w_0(\lambda x.\lambda y.x +_p \lambda x.\lambda y.y) = 0$; but $w_0(z\ (x +_p y)) = 0$ as the sum is directly on the right side of an application. □

Theorem 2.2.3. Every term has a unique normal form for $\to_+$. These normal forms, called *canonical terms*, are of the form

$$\begin{aligned} M, N &:= v \mid M +_p N \\ v &:= x \mid \lambda x.v \mid v\ M. \end{aligned}$$

We call *values* the canonical terms which are not sums.

Proof. The reduction $\to_+$ is weakly confluent and strongly normalising so it is also confluent and every term has a unique normal form. Besides it is easy to check that the canonical terms we described are indeed normal forms for $\to_+$. Conversely an induction on the structure of terms gives that every normal form for $\to_+$ is canonical.

— Variables are canonical.

— If $\lambda x.M$ is normal then M is normal, hence by induction hypothesis M is canonical, and M is not a sum.
— If $M\ N$ is normal then M and N are normal, hence by induction hypothesis they are canonical, and M is not a sum.
— If $M +_p N$ is normal then M and N are normal, hence canonical.

$\square$

Definition 2.3. We write $\mathrm{can}(M)$ the canonical form of a term M, i.e. its unique normal form for $\to_+$, and we define the reduction $\to_{\mathrm{can}}$ by $M \to_{\mathrm{can}} \mathrm{can}(M)$ for all M.

Remark that $\to_{\mathrm{can}}$ is an extremely simple reduction. In particular we do not extend it under arbitrary contexts, so from a term M there is always a unique reduction $M \to_{\mathrm{can}} \mathrm{can}(M)$, and $\to_{\mathrm{can}}$ is idempotent (i.e. $\to_{\mathrm{can}} \cdot \to_{\mathrm{can}} = \to_{\mathrm{can}}$), so it is not relevant to consider reduction paths of length more than 1.

3. The Full Calculus

3.1. *Confluence*

We proved some results independently for β-reduction and the reduction of sums, but we are interested in the calculus with both of them. It is interesting to know that $\to_\beta$ and $\to_+$ are both confluent but we want to prove that the reduction $\to_\beta \cup \to_+$ is also confluent. To do so we proceed in the same way as to prove the confluence of β-reduction alone, by defining a parallel reduction enjoying the diamond property. More precisely we will define a canonicalising parallel β-reduction: the reduction of sums being strongly normalising we will define the reduction of a set of β-redexes along with the normalisation by $\to_+$. As in Section 1 we consider terms with labelled β-redexes

$$\mathcal{M}, \mathcal{N} := x \mid \lambda x.\mathcal{M} \mid \mathcal{M}\ \mathcal{N} \mid \mathcal{M} +_p \mathcal{N} \mid (\lambda_0 x.\mathcal{M})\ \mathcal{N}$$

along with the β-reduction

$$(\lambda x.\mathcal{M})\ \mathcal{N} \to_\beta \mathcal{M}\left[{}^{\mathcal{N}}/x\right],$$

the β_0-reduction of labelled redexes

$$(\lambda_0 x.\mathcal{M})\ \mathcal{N} \to_{\beta_0} \mathcal{M}\left[{}^{\mathcal{N}}/x\right],$$

and the reduction of sums

$$\lambda x.(\mathcal{M} +_p \mathcal{N}) \to_+ \lambda x.\mathcal{M} +_p \lambda x.\mathcal{N}$$

$$(\mathcal{M} +_p \mathcal{N})\ \mathcal{P} \to_+ \mathcal{M}\ \mathcal{P} +_p \mathcal{N}\ \mathcal{P}$$

$$(\lambda_0 x.(\mathcal{M} +_p \mathcal{N}))\ \mathcal{P} \to_+ (\lambda_0 x.\mathcal{M})\ \mathcal{P} +_p (\lambda_0 x.\mathcal{N})\ \mathcal{P}$$

all extended to arbitrary contexts. We want to prove that the reduction $\to_{\beta_0} \cup \to_+$ associate a unique normal form to every labelled term.

Lemma 3.1.1. The reduction $\to_+$ is substitutive: if $\mathcal{M} \to_+ \mathcal{M}'$ then $\mathcal{M}\left[{}^{\mathcal{P}}/x\right] \to_+ \mathcal{M}'\left[{}^{\mathcal{P}}/x\right]$.

Proof. By a simple induction on the context of the reduction. □

Proposition 3.1.2. The reduction $\to_{\beta_0} \cup \to_{+}$ is weakly confluent.

Proof. We reason by induction on the structure of the reduced term.

— If $(\lambda_0 x.\mathcal{M})\ \mathcal{P} \to_{\beta_0} \mathcal{M}\left[{}^{\mathcal{P}}/_{x}\right]$ and $(\lambda_0 x.\mathcal{M})\ \mathcal{P} \to_{+} (\lambda_0 x.\mathcal{M}')\ \mathcal{P}$ with $\mathcal{M} \to_{+} \mathcal{M}'$ then the previous lemma gives $\mathcal{M}\left[{}^{\mathcal{P}}/_{x}\right] \to_{+} \mathcal{M}'\left[{}^{\mathcal{P}}/_{x}\right]$ and $(\lambda_0 x.\mathcal{M}')\ \mathcal{P} \to_{\beta_0} \mathcal{M}'\left[{}^{\mathcal{P}}/_{x}\right]$.
— If $(\lambda_0 x.\mathcal{M})\ \mathcal{P} \to_{\beta_0} \mathcal{M}\left[{}^{\mathcal{P}}/_{x}\right]$ and $(\lambda_0 x.\mathcal{M})\ \mathcal{P} \to_{+} (\lambda_0 x.\mathcal{M})\ \mathcal{P}'$ with $\mathcal{P} \to_{+} \mathcal{P}'$ then $\mathcal{M}\left[{}^{\mathcal{P}}/_{x}\right] \to_{+}^{*} \mathcal{M}\left[{}^{\mathcal{P}'}/_{x}\right]$ and $(\lambda_0 x.\mathcal{M})\ \mathcal{P}' \to_{\beta_0} \mathcal{M}'\left[{}^{\mathcal{P}'}/_{x}\right]$.
— If $(\lambda_0 x.(\mathcal{M} +_p \mathcal{N}))\ \mathcal{P} \to_{\beta_0} (\mathcal{M} +_p \mathcal{N})\left[{}^{\mathcal{P}}/_{x}\right] = \mathcal{M}\left[{}^{\mathcal{P}}/_{x}\right] +_p \mathcal{N}\left[{}^{\mathcal{P}}/_{x}\right]$ and $(\lambda_0 x.(\mathcal{M} +_p \mathcal{N}))\ \mathcal{P} \to_{+} (\lambda_0 x.\mathcal{M})\ \mathcal{P} +_p (\lambda_0 x.\mathcal{N})\ \mathcal{P}$ then we have $(\lambda_0 x.\mathcal{M})\ \mathcal{P} +_p (\lambda_0 x.\mathcal{N}) \to_{\beta_0}^{*} \mathcal{M}\left[{}^{\mathcal{P}}/_{x}\right] +_p \mathcal{N}\left[{}^{\mathcal{P}}/_{x}\right]$.
— If $\lambda x.(\mathcal{M}_1 +_p \mathcal{M}_2) \to_{+} \lambda x.\mathcal{M}_1 +_p \lambda x.\mathcal{M}_2$ and $\lambda x.(\mathcal{M}_1 +_p \mathcal{M}_2) \to_{\beta_0} \lambda x.(\mathcal{M}'_1 +_p \mathcal{M}'_2)$ with $\mathcal{M}_i \to_{\beta_0}^{?} \mathcal{M}'_i$ for $i \in \{1,2\}$ then $\lambda x.\mathcal{M}_1 +_p \lambda x.\mathcal{M}_2 \to_{\beta_0} \lambda x.\mathcal{M}'_1 +_p \lambda x.\mathcal{M}'_2$ and $\lambda x.(\mathcal{M}'_1 +_p \mathcal{M}'_2) \to_{+} \lambda x.\mathcal{M}'_1 +_p \lambda x.\mathcal{M}'_2$.
— If $(\mathcal{M}_1 +_p \mathcal{M}_2)\ \mathcal{M}_3 \to_{+} \mathcal{M}_1\ \mathcal{M}_3 +_p \mathcal{M}_2\ \mathcal{M}_3$ and $(\mathcal{M}_1 +_p \mathcal{M}_2)\ \mathcal{M}_3 \to_{\beta_0} (\mathcal{M}'_1 +_p \mathcal{M}'_2)\ \mathcal{M}'_3$ with $\mathcal{M}_i \to_{\beta_0}^{?} \mathcal{M}'_i$ for $i \in \{1,2;3\}$ then $\mathcal{M}_1\ \mathcal{M}_3 +_p \mathcal{M}_2\ \mathcal{M}_3 \to_{\beta_0}^{*} \mathcal{M}'_1\ \mathcal{M}'_3 +_p \mathcal{M}'_2\ \mathcal{M}'_3$ and $(\mathcal{M}'_1 +_p \mathcal{M}'_2)\ \mathcal{M}'_3 \to_{+} \mathcal{M}'_1\ \mathcal{M}'_3 +_p \mathcal{M}'_2\ \mathcal{M}'_3$.
— The other cases are given either by the weak confluence of $\to_{\beta_0}$ or $\to_{+}$, or by induction hypothesis.

□

Proposition 3.1.3. The reduction $\to_{\beta_0} \cup \to_{+}$ is strongly normalising.

Proof. To prove this result we adapt the weight of the corresponding result in (Barendregt 1981) so that the weight of a term decreases along the β_0-reduction and is preserved by $\to_{+}$. Given a term $\mathcal{M}$ we define a weight w on $\mathcal{M}$ by giving a value $w(o) \in \mathbb{N}$, $w(o) \geq 2$ for each variable occurrence o in $\mathcal{M}$ (i.e. each couple $(\mathcal{C}, x)$ such that $\mathcal{M} = \mathcal{C}[x]$). Then we define $w(\mathcal{P})$ for every subterm $\mathcal{P}$ of $\mathcal{M}$ (where the context is implicit) by

— $w(\lambda x.\mathcal{P}) = w(\mathcal{P})$;
— $w(\mathcal{P}\ \mathcal{Q}) = w(\mathcal{P}) \times w(\mathcal{Q})$;
— $w(\mathcal{P} +_p \mathcal{Q}) = w(\mathcal{P}) + w(\mathcal{Q})$;
— $w((\lambda_0 x.\mathcal{P})\ \mathcal{Q}) = w(\mathcal{P}) \times w(\mathcal{Q})$.

Then given a reduction $\mathcal{M} \to_{\beta_0} \mathcal{N}$ a weight on $\mathcal{M}$ induces a weight on $\mathcal{N}$ in a natural way, as every variable occurrence in $\mathcal{N}$ comes from a unique variable occurrence in $\mathcal{M}$. If $\mathcal{M} = \mathcal{C}\left[(\lambda_0 x.\mathcal{P})\ \mathcal{Q}\right] \to_{\beta_0} \mathcal{C}\left[\mathcal{P}\left[{}^{\mathcal{Q}}/_{x}\right]\right] = \mathcal{N}$ then

— the weight of a variable occurrence in $\mathcal{C}$ in $\mathcal{N}$ is the weight of the corresponding variable occurrence in $\mathcal{C}$ in $\mathcal{M}$;
— the weight of a variable occurrence in P in $\mathcal{N}$ is the weight of the corresponding variable occurrence in $\mathcal{P}$ in $\mathcal{M}$;
— the weight of a variable occurrence in one of the copies of $\mathcal{Q}$ in $\mathcal{N}$ is the weight of the corresponding variable occurrence in $\mathcal{Q}$ in $\mathcal{M}$.

Similarly we can derive a weight on $\mathcal{N}$ from a reduction $\mathcal{M} \to_{+} \mathcal{N}$.

Given a weight w on $\mathcal{M}$, it is easy to see that if a reduction $\mathcal{M} \to_{+} \mathcal{N}$ induces the weight w' on $\mathcal{N}$ then $w(\mathcal{M}) = w'(\mathcal{N})$: indeed we have $w(\lambda x.(\mathcal{P} +_p \mathcal{Q})) = w'(\lambda x.\mathcal{P} +_p \lambda x.\mathcal{Q})$ and $w((\mathcal{P} +_p \mathcal{Q})\ R) = w'(\mathcal{P}\ R +_p \mathcal{Q}\ R)$. But if $\mathcal{M} \to_{\beta_0} \mathcal{N}$ we do not necessarily have $w(\mathcal{M}) > w'(\mathcal{N})$.

To prove our proposition we need to look at particular weights: a weight w on $\mathcal{M}$ is called *decreasing* if for every β_0-redex $(\lambda_0 x.\mathcal{P})\ \mathcal{Q}$ in $\mathcal{M}$ we have $w(o) \geq w(\mathcal{Q})$ for every occurrence o of x in $\mathcal{P}$. Decreasing weights have the following properties.

(1) For every labelled term $\mathcal{M}$ there exists a decreasing weight on $\mathcal{M}$.
(2) If w is a decreasing weight on $\mathcal{M}$ and $\mathcal{M} \to_{\beta_0} \mathcal{N}$ then the induced weight w' on $\mathcal{N}$ is decreasing and we have $w(\mathcal{M}) > w'(\mathcal{N})$.

Proving that decreasing weights exist and are preserved by reduction is easily done by induction on the terms. Then if $\mathcal{M} = \mathcal{C}\left[(\lambda_0 x.\mathcal{P})\ \mathcal{Q}\right] \to_{\beta_0} \mathcal{C}\left[\mathcal{P}\left[{}^{\mathcal{Q}}/_{x}\right]\right] = \mathcal{N}$, w is decreasing on $\mathcal{M}$ and w' is the induced weight on $\mathcal{N}$ then we have $w(\mathcal{P}) \geq w'\left(\mathcal{P}\left[{}^{\mathcal{Q}}/_{x}\right]\right)$ and $w(\mathcal{Q}) \geq 2$ so $w\left((\lambda_0 x.\mathcal{P})\ \mathcal{Q}\right) > w'\left(\mathcal{P}\left[{}^{\mathcal{Q}}/_{x}\right]\right)$. See (Barendregt 1981) for a detailed proof in the deterministic case, which easily generalises to terms with sums.

Then for any term $\mathcal{M}$ there is a decreasing weight w on $\mathcal{M}$, and every reduction $\mathcal{M}\ (\to_{\beta_0} \cup \to_+)^*\ \mathcal{N}$ has at most $w(\mathcal{M})$ β_0-reduction steps. But we also know that $\to_+$ is strongly normalising so $\to_{\beta_0} \cup \to_+$ is strongly normalising. ☐

Definition 3.1. Given a term M and a set $\mathcal{F}$ of β-redexes in M we write $\mathrm{NF}_0^{\mathrm{c}}(M, \mathcal{F})$ for the unique normal form of $M_{\mathcal{F}}$ for $\to_{\beta_0} \cup \to_+$. We write $M \xrightarrow{\mathcal{F}}_{\beta^{\mathrm{c}}_{/\!/}} N$ if $N = \mathrm{NF}_0^{\mathrm{c}}(M, \mathcal{F})$, and more generally $M \to_{\beta^{\mathrm{c}}_{/\!/}} N$ if there exists $\mathcal{F}$ such that $M \xrightarrow{\mathcal{F}}_{\beta^{\mathrm{c}}_{/\!/}} N$.

Proposition 3.1.4. The canonicalising parallel reduction is the parallel reduction followed by a canonicalisation:

$$\to_{\beta^{\mathrm{c}}_{/\!/}} = \to_{\beta_{/\!/}} \cdot \to_{\mathrm{can}} .$$

Proof. Given a term M and a set $\mathcal{F}$ of redexes in M, we have $M_{\mathcal{F}} \to_{\beta_0}^* \mathrm{NF}_0(M, \mathcal{F}) \to_+^* \mathrm{can}\left(\mathrm{NF}_0(M, \mathcal{F})\right)$, hence $M\ (\to_{\beta_0} \cup \to_+)^*\ \mathrm{can}\left(\mathrm{NF}_0(M, \mathcal{F})\right)$. But $\mathrm{can}\left(\mathrm{NF}_0(M, \mathcal{F})\right)$ is by construction a normal form for $\to_+$, and $\mathrm{NF}_0(M, \mathcal{F})$ is a β_0-normal form so it is clear that $\mathrm{can}\left(\mathrm{NF}_0(M, \mathcal{F})\right)$ is also β_0-normal. Then necessarily $\mathrm{can}\left(\mathrm{NF}_0(M, \mathcal{F})\right) = \mathrm{NF}_0^{\mathrm{c}}(M, \mathcal{F})$. ☐

Proposition 3.1.5. The reduction $\to_{\beta^{\mathrm{c}}_{/\!/}}$ has the diamond property.

Proof. If $\mathcal{F}$ and $\mathcal{F}'$ are sets of β-redexes in M, $M \xrightarrow{\mathcal{F}}_{\beta^{\mathrm{c}}_{/\!/}} N$, $\mathcal{G}'$ is the set of residuals of $\mathcal{F}'$ in N and $N \xrightarrow{\mathcal{G}'}_{\beta^{\mathrm{c}}_{/\!/}} P$ then $M \xrightarrow{\mathcal{F} \cup \mathcal{F}'}_{\beta^{\mathrm{c}}_{/\!/}} P$. ☐

As our parallel reductions do not describe in detail the reductions of the sums we do not have $\to_{\beta^{\mathrm{c}}_{/\!/}}^* = (\to_\beta \cup \to_+)^*$ and we can not deduce directly the confluence of $\to_\beta \cup \to_+$ from the confluence of $\to_{\beta^{\mathrm{c}}_{/\!/}}$. But this construction of the parallel reductions gives a very interesting result about the β-reduction and the canonicalisation.

Definition 3.2. The *canonicalising β-reduction* $\to_{\beta^{\mathrm{c}}}$ is defined between canonical terms by

$$\to_{\beta^{\mathrm{c}}} = \to_\beta \cdot \to_{\mathrm{can}}$$

Proposition 3.1.6. For any reduction $M \to_\beta N$ we have $\mathrm{can}(M) \to^*_{\beta^c} \mathrm{can}(N)$.

Proof. Let Δ be a β-redex such that $M \xrightarrow{\Delta}_\beta N$, we have $M_{\{\Delta\}} \to_{\beta_0} N \to^*_+ \mathrm{can}(N)$. But we also have $M_{\{\Delta\}} \to^*_+ \mathrm{can}(M_{\{\Delta\}})$. Then as $\to_{\beta_0} \cup \to_+$ is strongly normalising we can find a sequence of reductions $\mathrm{can}(M_{\{\Delta\}}) \to_{\beta_0} \cdot \to^*_+ \mathcal{M}_1 \to_{\beta_0} \cdot \to^*_+ ... \to_{\beta_0} \cdot \to^*_+ \mathcal{M}_n$ such that the $\mathcal{M}_i$s are canonical and $\mathcal{M}_n$ is β_0-normal. As $\to_{\beta_0} \cup \to_+$ is confluent we necessarily have $\mathcal{M}_n = \mathrm{can}(N)$.

Now if we erase the labels we get $|\mathrm{can}(M_{\{\Delta\}})| \to_{\beta^c_{/\!/}} |\mathcal{M}_1| \to_{\beta^c_{/\!/}} \ldots \to_{\beta^c_{/\!/}} |\mathcal{M}_n| = \mathrm{can}(N)$, and we can check that $|\mathrm{can}(M_{\{\Delta\}})| = \mathrm{can}(M)$. □

This result and the observation that $\to_\beta \cup \to_{\mathrm{can}} \subset \to_{\beta^c_{/\!/}}$ allow us to conclude.

Proposition 3.1.7. The reduction $\to_\beta \cup \to_+$ is confluent.

Proof. Whenever $M \to_\beta N$ we have $\mathrm{can}(M) \to^*_{\beta^c_{/\!/}} \mathrm{can}(N)$, and whenever $M \to_+ N$ we have $\mathrm{can}(M) = \mathrm{can}(N)$. Then if $M\ (\to_\beta \cup \to_+)^*\ N_i$ for $i \in \{1,2\}$ we have $\mathrm{can}(M) \to^*_{\beta^c_{/\!/}} \mathrm{can}(N_i)$, and the confluence of $\to_{\beta^c_{/\!/}}$ gives a (canonical) term P such that $\mathrm{can}(N_i) \to^*_{\beta^c_{/\!/}} P$. In particular we have $N_i\ (\to_\beta \cup \to_+)^*\ P$ for $i \in \{1,2\}$. □

3.2. *β-Reduction Modulo Equivalences*

We can observe that the proof of confluence of $\to_\beta \cup \to_+$ only uses two facts about the reductions of sums: they preserve β_0-normal forms and with β-reduction they are sufficient to recover the parallel reduction, i.e. $\to_{\beta^c_{/\!/}} \subset (\to_\beta \cup \to_+)^*$. So we can do the same proof with the whole equivalence $\equiv_+$.

Proposition 3.2.1. The reduction $\to_\beta \cup \equiv_+$ is confluent.

Of course it is not natural to treat $\equiv_+$ as a reduction. It would be more intuitive to say that β-reduction modulo $\equiv_+$ is confluent. Reduction systems modulo an equivalence relation have been studied in (Bezem, Klop, and Vrijer 2003). The basic idea is that β-reduction modulo an equivalence relation $\equiv$ is β-reduction between equivalences classes of terms. Given $\mathcal{M}, \mathcal{N} \subset \Lambda^+$ two equivalence classes of terms for $\equiv$ we consider that $\mathcal{M}$ β-reduces into $\mathcal{N}$ if there are $M \in \mathcal{M}$ and $N \in \mathcal{N}$ such that $M \to_\beta N$. This induces the following definition on terms.

Definition 3.3. For any equivalence relation $\equiv$ on terms the reduction $\to_{\beta/\equiv}$ is defined by

$$M \to_{\beta/\equiv} N \Leftrightarrow \exists M', N' : M \equiv M' \to_\beta N' \equiv N.$$

A reduction for $\to^*_{\beta/\equiv}$ is then given by a path of the form

$$M_0 \equiv M'_0 \to_\beta M_1 \equiv M'_1 \to_\beta ... \to_\beta M_n \equiv M'_n.$$

In particular by a slight abuse of notations we consider that $M \to^*_{\beta/\equiv} M'$ whenever $M \equiv M'$.

In such a framework two main properties are associated to the notion of confluence.

We say that the β-reduction is *confluent modulo* $\equiv$ if for any terms M_1, M_2, N_1 and N_2 we have:

$$N_1 \leftarrow_\beta^* M_1 \equiv M_2 \rightarrow_\beta^* N_2 \Rightarrow \exists P_1, P_2 \in \Lambda^+ : N_1 \rightarrow_\beta^* P_1 \equiv P_2 \leftarrow_\beta^* N_2.$$

Furthermore the stronger property of being *Church-Rosser modulo* $\equiv$ is defined as:

$$M_1\ (\rightarrow_\beta \cup \leftarrow_\beta \cup \equiv)^*\ M_2 \Rightarrow \exists P_1, P_2 \in \Lambda^+ : M_1 \rightarrow_\beta^* P_1 \equiv P_2 \leftarrow_\beta^* M_2$$

for all M_1 and M_2. In both cases we can see that the equivalence $\equiv$ is not supposed to play any role in the reduction. Two related terms are supposed to reduce into a common result, provided this result and only this result is considered modulo equivalence. It is easy to see that the equivalence $\equiv_+$ does not fit in this framework. For example we have

$$(\lambda x.x +_p y)\ z \equiv_+ (\lambda x.x)\ z +_p y\ z \rightarrow_\beta z +_p y\ z$$

but there are no terms P_1 and P_2 such that

$$(\lambda x.x +_p y)\ z \rightarrow_\beta^* P_1 \equiv_+ P_2 \leftarrow_\beta^* z +_p y\ z.$$

It is not necessary to view the commutation of sums with abstractions and applications as a reduction, but it definitely plays an important role in the reduction. So it is not true that β-reduction is confluent modulo $\equiv_+$. What is true is that β-reduction modulo $\equiv_+$ is confluent.

Theorem 3.2.2. $\rightarrow_{\beta/\equiv_+}$ is confluent.

Another interesting result on β-reduction modulo $\equiv_+$ is given by Proposition 3.1.6.

Theorem 3.2.3. We have $M \rightarrow_{\beta/\equiv_+}^* N$ iff $\mathrm{can}(M) \rightarrow_{\beta^c}^* \mathrm{can}(N)$.

4. Weak and Strong Reductions of Sums

4.1. *Weak Head Reduction of Sums*

We have based our reduction rules for the probabilistic sums on head reduction: our goal was to replace the usual probabilistic reduction rules:

$$\lambda x_1...x_n.(M +_p N)\ P_1\ ...\ P_m \xrightarrow{p} \lambda x_1...x_n.M\ P_1\ ...\ P_m$$

$$\lambda x_1...x_n.(M +_p N)\ P_1\ ...\ P_m \xrightarrow{1-p} \lambda x_1...x_n.N\ P_1\ ...\ P_m$$

by a deterministic reduction:

$$\lambda x_1...x_n.(M +_p N)\ P_1\ ...\ P_m \rightarrow_+^* \lambda x_1...x_n.M\ P_1\ ...\ P_m +_p \lambda x_1...x_n.N\ P_1\ ...\ P_m.$$

But as we mentioned in Subsection 2.2 the purpose of this reduction is mostly to make explicit potential β-redexes: the term $((\lambda x.M) +_p N)\ P$ is a priori β-normal (if M, N and P are normal), but it reduces into $(\lambda x.M)\ P +_p N\ P$ which has a β-redex. To achieve this purpose, we actually need a single reduction rule:

$$(M +_p N)\ P \rightarrow_+ M\ P +_p N\ P.$$

Definition 4.1. The weak reduction of sums $\to_{+^{\mathrm{w}}}$ on Λ^+ is given by

$$(M +_p N)\ P \to_{+^{\mathrm{w}}} M\ P +_p N\ P$$

extended to arbitrary contexts.

This reduction corresponds to a weak head reduction strategy for the sums. One can check that all the properties of $\to_+$ proven in Sections 2 and 3 also hold for $\to_{+^{\mathrm{w}}}$.

Proposition 4.1.1.

(1) The reduction $\to_{+^{\mathrm{w}}}$ is confluent and strongly normalising. We call *weak canonical form* of M and we note $\mathrm{can}^{\mathrm{w}}(M)$ the unique normal form of M for $\to_{+^{\mathrm{w}}}$. We define a reduction $\to_{\mathrm{can}^{\mathrm{w}}}$ by $M \to_{\mathrm{can}^{\mathrm{w}}} \mathrm{can}^{\mathrm{w}}(M)$.
(2) The reduction $\to_{\beta_0} \cup \to_{+^{\mathrm{w}}}$ is confluent and strongly normalising on labelled terms. Given a term M and a set $\mathcal{F}$ of redexes in M we write $\mathrm{NF}_0^{\mathrm{w}}(M, \mathcal{F})$ the unique normal form of $M_{\mathcal{F}}$. We write $M \xrightarrow{\mathcal{F}}_{\beta_{/\!/}^{\mathrm{w}}} N$ if $N = \mathrm{NF}_0^{\mathrm{w}}(M, \mathcal{F})$, and more generally $M \to_{\beta_{/\!/}^{\mathrm{w}}} N$ if there exists $\mathcal{F}$ such that $M \xrightarrow{\mathcal{F}}_{\beta_{/\!/}^{\mathrm{w}}} N$.
(3) The weakly canonicalising parallel reduction $\to_{\beta_{/\!/}^{\mathrm{w}}}$ is equal to $\to_{\beta_{/\!/}} \cdot \to_{\mathrm{can}^{\mathrm{w}}}$.
(4) The reduction $\to_{\beta_{/\!/}^{\mathrm{w}}}$ has the diamond property.
(5) We define the *weakly canonicalising β-reduction* by $\to_{\beta^{\mathrm{w}}} = \to_\beta \cdot \to_{\mathrm{can}^{\mathrm{w}}}$. If $M \to_\beta N$ then $\mathrm{can}^{\mathrm{w}}(M) \to_{\beta^{\mathrm{w}}}^* \mathrm{can}^{\mathrm{w}}(N)$.
(6) The reductions $\to_\beta \cup \to_{+^{\mathrm{w}}}$ and $\to_{\beta/\equiv_{+^{\mathrm{w}}}}$ are confluent.
(7) We have $M \to_{\beta/\equiv_{+^{\mathrm{w}}}}^* N$ iff $\mathrm{can}^{\mathrm{w}}(M) \to_{\beta^{\mathrm{w}}}^* \mathrm{can}^{\mathrm{w}}(N)$.

Proof. The proofs of the corresponding results for $\to_+$ are easily adapted to $\to_{+^{\mathrm{w}}}$. □

This reduction has one major property which $\to_+$ lacks: it is orthogonal to β-reduction. In any term of the form $(\lambda x.(M_1 +_p M_2))\ N$ there are two overlapping redexes, one for $\to_+$ and one for $\to_\beta$. This situation does not occur with $\to_{+^{\mathrm{w}}}$, which makes studying the reduction $\to_\beta \cup \to_{+^{\mathrm{w}}}$ much more convenient. In particular, we will see in Section 5 that we can prove a standardisation theorem for $\to_\beta \cup \to_{+^{\mathrm{w}}}$ using well-known techniques, whereas such an attempt fails for $\to_\beta \cup \to_+$.

Another convenient property, although of much lesser importance, is that weakly canonical forms are easier to manipulate. One can prove that those weakly canonical forms are given inductively by:

$$\begin{aligned} M, N &:= v \mid M +_p N \\ v &:= x \mid \lambda x.M \mid v\ M. \end{aligned}$$

but we can also characterise them without defining values, by:

$$M, N := x\ N_1\ \ldots\ N_m \mid (\lambda x.M)\ N_1\ \ldots\ N_m \mid M +_p N.$$

This makes it much easier to reason by induction on weakly canonical forms than on canonical forms.

4.2. *From Strong Reduction to Weak Reduction*

We have two reductions for the sum, one corresponding to the head reduction usually considered in the literature and the other which we claim is more simple and enjoys more interesting properties than the first one. What is the relation between these two reductions? On one side we obviously have $\to_{+^{\mathrm{w}}} \subset \to_{+}$. On the other side we claimed that the reduction rule $\lambda x.(M +_p N) \to_+ \lambda x.M +_p \lambda x.N$ does not really influence β-reduction, so we could expect that whenever $M \to^*_{\beta/\equiv_+} N$ then $M \to^*_{\beta/\equiv_{+^{\mathrm{w}}}} \cdot \equiv_+ N$. But this clearly does not hold: we have for instance a reduction

$$(\lambda x.(x +_p y))\ z \to_+ (\lambda x.x +_p \lambda x.y)\ z \to_+ (\lambda x.x)\ z +_p (\lambda x.y)\ z \to_\beta z +_p (\lambda x.y)\ z$$

whereas the only possible reduction of this term for $\to_\beta \cup \to_{+^{\mathrm{w}}}$ is

$$(\lambda x.(x +_p y))\ z \to_\beta z +_p y.$$

The reduction rule $\lambda x.(M +_p N) \to_+ \lambda x.M +_p \lambda x.N$ may be used to duplicate β-redexes and we can then choose to reduce only some of the resulting copies, which is impossible with the weak reduction. On the other hand, this example seems to indicate that we can reduce the remaining copies at the end of our strong reduction to reach a result which can be attained by a weak reduction: there is no reduction $\to_\beta \cup \to_{+^{\mathrm{w}}}$ from $(\lambda x.(x +_p y))\ z$ to $z +_p (\lambda x.y)\ z$ but we do have $z +_p (\lambda x.y)\ z \to_\beta z +_p y$ and $(\lambda x.(x +_p y))\ z \to_\beta z +_p y$.

We will prove that every reduction $M \to^*_{\beta/\equiv_+} N$ can be extended with a reduction $N \to^*_{\beta/\equiv_+} N'$ such that $M \to^*_{\beta/\equiv_{+^{\mathrm{w}}}} N'$ (Theorem 4.2.5). More precisely we will prove that every reduction $M\ (\to_\beta \cup \to_+)^*\ N$ can be extended with a reduction $N\ (\to_\beta \cup \to_+)^*\ \mathrm{can}(N')$ (hence $N \to^*_{\beta/\equiv_+} N'$) such that $M\ (\to_\beta \cup \to_{+^{\mathrm{w}}})^*\ N'$, the correspondence between these two results being given by Theorem 3.2.3 and Proposition 4.1.1, (7).

We will denote by $\to_{+^\lambda}$ the relation $\to_+ \setminus \to_{+^{\mathrm{w}}}$, i.e. the reduction given by

$$\lambda x.(M +_p N) \to_{+^\lambda} \lambda x.M +_p \lambda x.N$$

extended to arbitrary context.

Lemma 4.2.1. If M is weakly canonical and $M \to_{+^\lambda} \cdot \to_{\mathrm{can}^{\mathrm{w}}} \cdot \to_{\beta_{/\!/}} N$ then there is N' such that $N \to_{\beta_{/\!/}} N'$ and $M \to_{\beta_{/\!/}} \cdot \to^*_+ N'$.

$$\begin{tikzcd}
M \arrow[r, "+^\lambda"] \arrow[dd, dashed, "\beta_{/\!/}"'] & {} \arrow[r, "\mathrm{can}^{\mathrm{w}}"] & {} \arrow[d, "\beta_{/\!/}"] \\
& & N \arrow[d, dashed, "\beta_{/\!/}"] \\
{} \arrow[rr, dashed, "+^*"'] & & N'
\end{tikzcd}$$

Proof. We reason by induction on M as a weakly canonical term, as described at the end of Subsection 4.1. We detail two cases, where we have an abstraction in weak head position.

— If

$$\begin{aligned}(\lambda x.M)\ P\ R_1\ ...\ R_m \to_{+^\lambda} \cdot \to_{\mathrm{can^w}} (\lambda x.M')\ P'\ R'_1\ ...\ R'_m \\ \to_{\beta_{/\!/}} N\left[{}^{Q}/_{x}\right]\ S_1\ ...\ S_m\end{aligned}$$

with $M \to^?_{+^\lambda} \cdot \to_{\mathrm{can^w}} M' \to_{\beta_{/\!/}} N$, $P \to^?_{+^\lambda} \cdot \to_{\mathrm{can^w}} P' \to_{\beta_{/\!/}} Q$ and $R_i \to^?_{+^\lambda} \cdot \to_{\mathrm{can^w}} R'_i \to_{\beta_{/\!/}} S_i$ for $i \le m$ then applying the induction hypothesis to these reductions gives terms N', Q' and S'_i and we get

$$\begin{aligned}N\left[{}^{Q}/_{x}\right]\ S_1\ ...\ S_m &\to_{\beta_{/\!/}} N'\left[{}^{Q'}/_{x}\right]\ S'_1\ ...\ S'_m \\ (\lambda x.M)\ P\ R_1\ ...\ R_m &\to_{\beta_{/\!/}} \cdot \to^*_{+} N'\left[{}^{Q'}/_{x}\right]\ S'_1\ ...\ S'_m.\end{aligned}$$

— If

$$\begin{aligned}(\lambda x.(M_1 +_p M_2))\ P\ R_1\ ...\ R_m &\to_{+^\lambda} (\lambda x.M_1 +_p \lambda x.M_2)\ P\ R_1\ ...\ R_m \\ &\to_{\mathrm{can^w}} (\lambda x.M_1)\ P\ R_1\ ...\ R_m +_p (\lambda x.M_2)\ P\ R_1\ ...\ R_m \\ &\to_{\beta_{/\!/}} T_1 +_p T_2\end{aligned}$$

with $(\lambda x.M_1)\ P\ R_1\ ...\ R_m \to_{\beta_{/\!/}} T_1$ and $(\lambda x.M_2)\ P\ R_1\ ...\ R_m \to_{\beta_{/\!/}} T_2$ then for each $j \in \{1,2\}$ there are reductions $M_j \to_{\beta_{/\!/}} N_j$, $P \to_{\beta_{/\!/}} Q_j$ and $R_i \to_{\beta_{/\!/}} S_{i,j}$ such that $T_j = (\lambda x.N_j)\ Q_j\ S_{1,j}\ ...\ S_{m,j}$ or $T_j = N_j\left[{}^{Q_j}/_{x}\right]\ S_{1,j}\ ...\ S_{m,j}$. In either case as $P \to_{\beta_{/\!/}} Q_j$ for $j \in \{1,2\}$ there is Q' such that $P \to_{\beta_{/\!/}} Q'$ and $Q_j \to_{\beta_{/\!/}} Q'$ for $j \in \{1,2\}$, and similarly there are terms S'_i such that $R_i \to_{\beta_{/\!/}} S'_i$ and for $j \in \{1,2\}$, $S_{i,j} \to_{\beta_{/\!/}} S'_i$. We get

$$\begin{aligned}T_1 +_p T_2 &\to_{\beta_{/\!/}} N_1\left[{}^{Q'}/_{x}\right]\ S'_1\ ...\ S'_m +_p N_2\left[{}^{Q'}/_{x}\right]\ S'_1\ ...\ S'_m \\ (\lambda x.(M_1 +_p M_2))\ P\ R_1\ ...\ R_m &\to_{\beta_{/\!/}} \left(N_1\left[{}^{Q'}/_{x}\right] +_p N_2\left[{}^{Q'}/_{x}\right]\right)\ S'_1\ ...\ S'_m \\ &\to^*_{+} N_1\left[{}^{Q'}/_{x}\right]\ S'_1\ ...\ S'_m +_p N_2\left[{}^{Q'}/_{x}\right]\ S'_1\ ...\ S'_m.\end{aligned}$$

The other cases are immediate. □

Lemma 4.2.2. If M is weakly canonical then there is a reduction $M\ (\to_{+^\lambda} \cdot \to_{\mathrm{can^w}})^*\ \mathrm{can}(M)$.

Proof. We know that the length of the reductions $M \to^*_{+} \mathrm{can}(M)$ is bounded (Proposition 2.2.2), so we reason by induction on this bound. If M is canonical the result is immediate. Otherwise M is weakly canonical but not canonical, so there is a reduction $M \to_{+^\lambda} M'$, and by induction hypothesis we have $\mathrm{can^w}(M')\ (\to_{+^\lambda} \cdot \to_{\mathrm{can^w}})^*\ \mathrm{can}(M')$. But $\mathrm{can}(M') = \mathrm{can}(M)$ so $M \to_{+^\lambda} M' \to_{\mathrm{can^w}} \mathrm{can^w}(M')\ (\to_{+^\lambda} \cdot \to_{\mathrm{can^w}})^*\ \mathrm{can}(M)$. □

Lemma 4.2.3. If M is weakly canonical and $M \to_{\mathrm{can}} \cdot \to_{\beta^{\mathrm{c}}_{/\!/}} N$ then there is N' such that $N \to^*_{\beta^{\mathrm{c}}_{/\!/}} N'$ and $M \to_{\beta^{\mathrm{c}}_{/\!/}} N'$. In particular we have $M \to_{\beta^{\mathrm{w}}_{/\!/}} \cdot \to_{\mathrm{can}} N'$.

Proof. If M is a weakly canonical term, according to the previous lemma we have

$M\ (\to_{+^\lambda}\cdot\to_{\mathrm{can}^{\mathrm{w}}})^*\ \mathrm{can}(M)$. We reason by induction on the length of such a reduction. If M is canonical then we just choose $N' = N$. Otherwise the induction step is given by the following diagram:

$$\begin{tikzcd}
M \arrow[r, "+^\lambda\cdot\mathrm{can}^{\mathrm{w}}"] \arrow[ddd, dashed, "\beta_{/\!/}"'] \arrow[ddr, phantom, "(4.2.1)"] & P \arrow[r, "(+^\lambda\cdot\mathrm{can}^{\mathrm{w}})^*"] \arrow[dd, dashed, "\beta_{/\!/}"'] \arrow[dr, phantom, "(\mathrm{IH})"] & \mathrm{can}(M) \arrow[d, "\beta^{\mathrm{c}}_{/\!/}"] \\
 & & N \arrow[d, dashed, "(\beta^{\mathrm{c}}_{/\!/})^*"] \\
 & Q \arrow[r, dashed, "\mathrm{can}"'] \arrow[d, dashed, "\beta_{/\!/}"] & N' \arrow[d, dashed, "\beta^{\mathrm{c}}_{/\!/}"] \\
{} \arrow[r, dashed, "+^*"'] & Q' \arrow[r, dashed, "\mathrm{can}"'] & \mathrm{can}(Q')
\end{tikzcd}$$

If $M \to_{+^\lambda}\cdot\to_{\mathrm{can}^{\mathrm{w}}} P\ (\to_{+^\lambda}\cdot\to_{\mathrm{can}^{\mathrm{w}}})^*\ \mathrm{can}(M) \to_{\beta^{\mathrm{c}}_{/\!/}} N$ then the induction hypothesis gives the terms Q and N'. Next we apply Lemma 4.2.1 to M, P and Q to get the term Q'. Now let $\mathcal{F}$ be the set of β-redexes in Q such that $Q \xrightarrow{\mathcal{F}}_{\beta_{/\!/}} Q'$, we have $\mathrm{can}(Q') = \mathrm{NF}^{\mathrm{c}}_0(Q,\mathcal{F})$ (see the proof of Proposition 3.1.4). But as $Q \to^*_{+} N'$, let $\mathcal{F}'$ be the set of residuals of $\mathcal{F}$ in N, we have $\mathrm{NF}^{\mathrm{c}}_0(Q,\mathcal{F}) = \mathrm{NF}^{\mathrm{c}}_0(N',\mathcal{F}')$. In other words, $N' \to_{\beta^{\mathrm{c}}_{/\!/}} \mathrm{can}(Q')$.

Finally we can remark that $\to_{\beta^{\mathrm{w}}_{/\!/}}\cdot\to_{\mathrm{can}} = \to_{\beta^{\mathrm{c}}_{/\!/}}$. $\square$

Lemma 4.2.4. If M is weakly canonical and $M \to_{\mathrm{can}}\cdot\to^*_{\beta^{\mathrm{c}}_{/\!/}} N$ then there is N' such that $N \to^*_{\beta^{\mathrm{c}}_{/\!/}} N'$ and $M \to^*_{\beta^{\mathrm{w}}_{/\!/}}\cdot\to_{\mathrm{can}} N'$.

Proof. We reason by induction on the length of the reduction $\to^*_{\beta^{\mathrm{c}}_{/\!/}}$. We recall that for $n \in \mathbb{N}$ we write $\to^n_{\beta^{\mathrm{c}}_{/\!/}}$ for n iterations of the reduction $\to_{\beta^{\mathrm{c}}_{/\!/}}$. If $M \to_{\mathrm{can}} N$ then we just choose $N' = N$. Otherwise the induction step is given by the following diagram:

$$\begin{tikzcd}
M \arrow[r, "\mathrm{can}"] \arrow[d, dashed, "\beta^{\mathrm{w}}_{/\!/}"'] \arrow[dr, phantom, "(4.2.3)"] & \mathrm{can}(M) \arrow[r, "\beta^{\mathrm{c}}_{/\!/}"] & P \arrow[r, "(\beta^{\mathrm{c}}_{/\!/})^n"] \arrow[d, dashed, "(\beta^{\mathrm{c}}_{/\!/})^*"'] \arrow[dr, phantom, "(3.1.5)"] & N \arrow[d, dashed, "(\beta^{\mathrm{c}}_{/\!/})^*"] \\
M' \arrow[rr, dashed, "\mathrm{can}"'] \arrow[d, dashed, "(\beta^{\mathrm{w}}_{/\!/})^*"'] \arrow[drr, phantom, "(\mathrm{IH})"] & & P' \arrow[r, dashed, "(\beta^{\mathrm{c}}_{/\!/})^n"'] & Q \arrow[d, dashed, "(\beta^{\mathrm{c}}_{/\!/})^*"] \\
{} \arrow[rrr, dashed, "\mathrm{can}"'] & & & N'
\end{tikzcd}$$

If $M \to_{\mathrm{can}}\cdot\to_{\beta^{\mathrm{c}}_{/\!/}} P \to^n_{\beta^{\mathrm{c}}_{/\!/}} N$ then the previous lemma gives the terms M' and P'. The diamond property of $\to_{\beta^{\mathrm{c}}_{/\!/}}$ (Proposition 3.1.5) gives the existence of Q with $P' \to^n_{\beta^{\mathrm{c}}_{/\!/}} Q$.

Then we simply conclude by induction hypothesis on the reduction $M' \to_{\mathrm{can}} P' \to^n_{\beta^{\mathrm{c}}_{/\!/}} Q$. □

Theorem 4.2.5. If $M \to^*_{\beta/\equiv_+} N$ then there is N' such that $N \to^*_{\beta/\equiv_+} N'$ and $M \to^*_{\beta/\equiv_{+^{\mathrm{w}}}} N'$.

Proof. If $M \to^*_{\beta/\equiv_+} N$ then according to Theorem 3.2.3 we have $\mathrm{can}(M) \to^*_{\beta^{\mathrm{c}}} \mathrm{can}(N)$, thus $\mathrm{can}^{\mathrm{w}}(M) \to_{\mathrm{can}} \cdot \to^*_{\beta^{\mathrm{c}}_{/\!/}} \mathrm{can}(N)$. The previous lemma states the existence of terms M' and N' such that $\mathrm{can}(N) \to^*_{\beta^{\mathrm{c}}_{/\!/}} N'$ and $\mathrm{can}^{\mathrm{w}}(M) \to^*_{\beta^{\mathrm{w}}_{/\!/}} M' \to_{\mathrm{can}} N'$, hence $N \to^*_{\beta/\equiv_+} M'$ and $M \to^*_{\beta/\equiv_{+^{\mathrm{w}}}} M'$. □

5. Standardisation

The purpose of this article is to present a probabilistic λ-calculus which does not require a specific reduction strategy and allows reduction under arbitrary context. In this setting there may be multiple ways to reduce a term, hence it was important to prove the confluence of the reduction. But another interesting result we would like to obtain is a standardisation theorem: it is very convenient to have that a term M reduces into N if and only if there is a standard reduction from M to N, i.e. a reduction where redexes are reduced from left to right and from top to bottom. Unfortunately this does not hold for the reduction $\to_{\beta/\equiv_+}$ (see (Alberti 2014), example 3.2.10): if you consider the reduction

$$(\lambda x.I\ (y +_p z))\ u \to_\beta (\lambda x.(y +_p z))\ u \equiv_+ (\lambda x.y)\ u +_p (\lambda x.z)\ u \to_\beta y +_p (\lambda x.z)\ u$$

where $I = \lambda x.x$, there cannot be any standard reduction from $(\lambda x.I\ (y +_p z))\ u$ to $y +_p (\lambda x.z)\ u$. Indeed $(\lambda x.I\ (y +_p z))\ u$ is only equivalent to itself with respect to $\equiv_+$, and the only possible β-reductions are $(\lambda x.I\ (y +_p z))\ u \to_\beta I\ (y +_p z)$, which does not reduce into $y +_p (\lambda x.z)\ u$, and $(\lambda x.I\ (y +_p z))\ u \to_\beta (\lambda x.(y +_p z))\ u$, which cannot be further β-reduced if we want to respect the intuition of what a standard reduction should be. But just like the comparison between the weak and strong reductions of sums, we can see that our reduction can be extended and then turned into a standard one. We have $y +_p (\lambda x.z)\ u \to_\beta y +_p z$ and there is a standard reduction

$$(\lambda x.I\ (y +_p z))\ u \to_\beta I\ (y +_p z) \to_\beta y +_p z.$$

As we mentioned in Subsection 4.1 the main reason we considered the weak reduction of sums is that the reduction system $\to_\beta \cup \to_{+^{\mathrm{w}}}$ does not have any critical pair, thus this problem of redex duplication does not occur and the usual proof techniques for the standardisation are known to work. So we will first prove a standardisation property for $\to_{\beta/\equiv_{+^{\mathrm{w}}}}$. Then using the theorem 4.2.5 we can prove that every β-reduction modulo $\equiv_+$ can be extended then turned into a standard one.

5.1. *Strong Standardisation of the Weak Reduction*

The key property behind the notion of standardisation is that every reduction can be transformed into a sequence of head reductions followed by internal reductions. In an orthogonal system system such as $\to_\beta \cup \to_{+^{\mathrm{w}}}$ there is a natural way to define head and internal reductions: head redexes are simply outermost redexes and the orthogonality ensures that those are preserved by reduction. For instance in our case we could say that terms of the form $(\lambda x.M)\ N$ or $(M +_p N)\ P$ are directly head redexes, and otherwise:

— x has no head redex,
— the head redexes of $\lambda x.M$ are the head redexes of M,
— the head redexes of $M\ N$ are the head redexes of M
— and the head redexes of $M +_p N$ are the head redexes of M or N.

But once again we are not really interested in $\to_{+^{\mathrm{w}}}$ as a reduction, and we rather want to prove a standardisation property for $\to_{\beta/\equiv_{+^{\mathrm{w}}}}$. This reduction being characterised by weakly canonicalising reduction $\to_{\beta^{\mathrm{w}}}$ between weakly canonical terms (by Proposition 4.1.1, (7)), we will not give a general notion of $\beta{+}^{\mathrm{w}}$-head redex but we will only define weakly canonicalising head β-reduction and internal β-reduction on weakly canonical terms.

Definition 5.1. The *head β-redexes* of the weakly canonical terms are defined as follows:

— the head redexes of $M +_p N$ are the head redexes of M and the head redexes of N;
— $x\ P_1\ ...\ P_m$ has no head redex;
— the head redexes of $\lambda x.M$ are the head redexes of M;
— the only head redex of $(\lambda x.M)\ N\ P_1\ ...\ P_m$ is $(\lambda x.M)\ N$.

β-redexes which are not head β-redexes are called *internal*.

Definition 5.2.

(1) Given a weakly canonical term M and a reduction $M \xrightarrow{\Delta}_\beta \cdot \to_{\mathrm{can}^{\mathrm{w}}} N$ we note $M \to_{\mathrm{h}^{\mathrm{w}}} N$ if Δ is a head redex of M, and $M \to_{\mathrm{i}^{\mathrm{w}}} N$ if Δ is internal.
(2) Given a set $\mathcal{F}$ of β-redexes in a weakly canonical term M and a reduction $M \xrightarrow{\mathcal{F}}_{\beta^{\mathrm{w}}_{/\!/}} N$ we note $M \to_{\mathrm{h}^{\mathrm{w}}_{/\!/}} N$ if $\mathcal{F}$ is a set of head redexes, and $M \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}} N$ if it contains only internal redexes.

Proposition 5.1.1. $\to_{\mathrm{h}^{\mathrm{w}}} \subset \to_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \subset \to^*_{\mathrm{h}^{\mathrm{w}}}$ and $\to_{\mathrm{i}^{\mathrm{w}}} \subset \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}} \subset \to^*_{\mathrm{i}^{\mathrm{w}}}$.

Proof. We have immediately $\to_{\mathrm{h}^{\mathrm{w}}} \subset \to_{\mathrm{h}^{\mathrm{w}}_{/\!/}}$ and $\to_{\mathrm{i}^{\mathrm{w}}} \subset \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}}$. Moreover an easy induction on M shows that if $\mathcal{F}$ is a set of head redexes in M and $M \to_{\mathrm{h}^{\mathrm{w}}} N$ then the residuals of $\mathcal{F}$ in N are also head redexes, and similarly the residuals of internal redexes by an internal reduction are internal redexes. □

This definition of the head and internal reductions is actually perfectly suited for the strong reduction of sums as well. The problem is that the following property only holds for the weak reduction.

Lemma 5.1.2. If $M \to_{\mathrm{i}^{\mathrm{w}}} N$ and $\mathcal{H}$ is a set of head β-redexes in N then $\mathcal{H}$ is exactly the set of residuals of a set $\mathcal{H}'$ of head redexes in M.

Proof. By a simple induction on N. The key idea is that if $N = (\lambda x.N_0)\ Q\ S_1\ ...\ S_m$ then necessarily $M = (\lambda x.M_0)\ P\ R_1\ ...\ R_m$ with $M_0 \to_{\beta^{\mathrm{w}}} N_0$, $P \to_{\beta^{\mathrm{w}}} Q$ and $R_i \to_{\beta^{\mathrm{w}}} S_i$ for $i \leq m$. □

In our counterexample to the standardisation of $\to_\beta \cup \to_+$ we precisely chose a reduction

$$(\lambda x.I\ (y +_p z))\ u \to_{\mathrm{i}^{\mathrm{c}}} (\lambda x.y)\ u +_p (\lambda x.z)\ u$$

such that the redex $(\lambda x.y)\ u$ is not the unique residual of a head redex in the initial term, so we could not transform our reduction into a reduction $\to^*_{\mathrm{h}^{\mathrm{c}}} \cdot \to^*_{\mathrm{i}^{\mathrm{c}}}$, and our solution to this problem was exactly to further reduce all the other residuals of the original head redex $(\lambda x.I\ (y +_p z))\ u$. But with the weak reduction the internal reduction does not interact with the head redexes and we can prove that $\to^*_{\beta^{\mathrm{w}}} \subset \to^*_{\mathrm{h}^{\mathrm{w}}} \cdot \to^*_{\mathrm{i}^{\mathrm{w}}}$.

Lemma 5.1.3. If $M \to_{\beta^{\mathrm{w}}_{/\!/}} N$ then $M \to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cdot \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}} N$.

Proof. Let $\mathcal{F}$ be the set of β-redexes of M such that $M \xrightarrow{\mathcal{F}}_{\beta^{\mathrm{w}}_{/\!/}} N$. The reduction $(\to_{\beta_0} \cup \to_{+^{\mathrm{w}}})$ is strongly normalising (Proposition 3.1.3) so we reason by induction on the maximal length of the reductions $M_{\mathcal{F}}\ (\to_{\beta_0} \cup \to_{+^{\mathrm{w}}})^*\ N$. If $\mathcal{F}$ contains no head redex then $M \xrightarrow{\mathcal{F}}_{\mathrm{i}^{\mathrm{w}}_{/\!/}} N$. Otherwise there is a nonempty set $\mathcal{H} \subset \mathcal{F}$ of head redexes and we have a reduction $M \xrightarrow{\mathcal{H}}_{\mathrm{h}^{\mathrm{w}}_{/\!/}} P$. Let $\mathcal{F}'$ be the set of residuals of $\mathcal{F} \setminus \mathcal{H}$ in P, we have $P \xrightarrow{\mathcal{F}'}_{\beta^{\mathrm{w}}_{/\!/}} N$ and by induction hypothesis $P \to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cdot \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}} N$. □

Lemma 5.1.4. If $M \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}} \cdot \to_{\mathrm{h}^{\mathrm{w}}_{/\!/}} N$ then $M \to_{\beta^{\mathrm{w}}_{/\!/}} N$ (hence $M \to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cdot \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}} N$).

Proof. By definition there are a term P, a set $\mathcal{F}$ of internal β-redexes in M and a set $\mathcal{H}$ of head β-redexes in P such that $M \xrightarrow{\mathcal{F}}_{\mathrm{i}^{\mathrm{w}}_{/\!/}} P$ and $P \xrightarrow{\mathcal{H}}_{\mathrm{h}^{\mathrm{w}}_{/\!/}} N$. According to Proposition 5.1.1 we have $M \to^*_{\mathrm{i}^{\mathrm{w}}} P$, and Proposition 5.1.2 gives a set $\mathcal{H}'$ of (head) redexes in M such that $\mathcal{H}$ is exactly the set of residuals of $\mathcal{H}'$. Thus there is a direct reduction $M \xrightarrow{\mathcal{F} \cup \mathcal{H}'}_{\beta^{\mathrm{w}}_{/\!/}} N$. □

Lemma 5.1.5. If $M \to^*_{\mathrm{i}^{\mathrm{w}}_{/\!/}} \cdot \to_{\mathrm{h}^{\mathrm{w}}_{/\!/}} N$ then $M \to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cdot \to^*_{\mathrm{i}^{\mathrm{w}}_{/\!/}} N$.

Proof. By induction on the length of the reduction $\to^*_{\mathrm{i}^{\mathrm{w}}_{/\!/}}$, using the previous lemma. □

Proposition 5.1.6. If $M \to^*_{\beta^{\mathrm{w}}} N$ then $M \to^*_{\mathrm{h}^{\mathrm{w}}} \cdot \to^*_{\mathrm{i}^{\mathrm{w}}} N$.

Proof. We have $\to^*_{\beta^{\mathrm{w}}} = \to^*_{\beta^{\mathrm{w}}_{/\!/}}$, $\to^*_{\mathrm{h}^{\mathrm{w}}} = \to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}}$ and $\to^*_{\mathrm{i}^{\mathrm{w}}} = \to^*_{\mathrm{i}^{\mathrm{w}}_{/\!/}}$ so we rather reason with parallel reductions.

If $M \to^*_{\beta^{\mathrm{w}}_{/\!/}} N$ then Lemma 5.1.3 gives $M\ (\to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cdot \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}})^*\ N$, which corresponds exactly to $M\ (\to_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cup \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}})^*\ N$. We reason by induction on the length of such a reduction. If $M = N$ the result is immediate. If $M\ (\to_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cup \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}})^*\ P \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}} N$ then by induction hypothesis $M \to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cdot \to^*_{\mathrm{i}^{\mathrm{w}}_{/\!/}} P$ and the result is immediate. If $M\ (\to_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cup \to_{\mathrm{i}^{\mathrm{w}}_{/\!/}})^*\ P \to_{\mathrm{h}^{\mathrm{w}}_{/\!/}} N$ then by induction hypothesis $M \to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cdot \to^*_{\mathrm{i}^{\mathrm{w}}_{/\!/}} P$, and with the previous lemma we get $M \to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cdot (\to^*_{\mathrm{h}^{\mathrm{w}}_{/\!/}} \cdot \to^*_{\mathrm{i}^{\mathrm{w}}_{/\!/}}) N$. □

Labelled reduction versus inductive parallel reduction. As we mentioned in Subsection 1.2 one can define parallel β-reduction using labels as we did, following the constructions in (Barendregt 1981), or one can define it inductively as in (Takahashi 1995).

The inductive definition is less elaborated and may be better suited to prove confluence properties, but when one wants to prove that β-reduction decomposes into head reductions followed by internal reductions, not being able to reason with sets of β-redexes and residuals is an important drawback. Furthermore in our case the situation is made even more complicated by the fact that we are not interested in head and internal β-reductions, but in *weakly canonicalising* head and internal β-reductions, and those are not easily defined by induction. We believe a proof of our standardisation theorem using inductively defined parallel reductions is possible, but we doubt it would be more concise than the one using labels.

The standardisation is an extension of Proposition 5.1.6: a reduction is standard if it is a sequence of head reductions followed by some internal reductions, and those internal reductions are locally standard. So once again we will not try to define a general notion of standard reduction for $\to_\beta \cup \to_{+^{\mathrm{w}}}$ but we will only consider weakly canonicalising β-reductions between weakly canonical terms. More precisely we define inductively a relation on weakly canonical terms characterising the existence of a standard reduction.

Definition 5.3. The relation $M \twoheadrightarrow_{S^{\mathrm{w}}} N$ is defined between weakly canonical terms by:

$$\frac{M_1 \twoheadrightarrow_{S^{\mathrm{w}}} N_1 \qquad M_2 \twoheadrightarrow_{S^{\mathrm{w}}} N_2}{M_1 +_p M_2 \twoheadrightarrow_{S^{\mathrm{w}}} N_1 +_p N_2} \qquad\qquad \frac{M \twoheadrightarrow_{S^{\mathrm{w}}} N}{\lambda x.M \twoheadrightarrow_{S^{\mathrm{w}}} \lambda x.N}$$

$$\frac{\forall i \leq m, P_i \twoheadrightarrow_{S^{\mathrm{w}}} Q_i}{y\ P_1\ ...\ P_m \twoheadrightarrow_{S^{\mathrm{w}}} y\ Q_1\ ...\ Q_m}$$

$$\frac{\mathrm{can}^{\mathrm{w}}\left(M\left[{}^{P}/{}_{y}\right]\ Q_1\ ...\ Q_m\right) \twoheadrightarrow_{S^{\mathrm{w}}} N}{(\lambda y.M)\ P\ Q_1\ ...\ Q_m \twoheadrightarrow_{S^{\mathrm{w}}} N}$$

$$\frac{M \twoheadrightarrow_{S^{\mathrm{w}}} N \qquad \forall i \leq m, P_i \twoheadrightarrow_{S^{\mathrm{w}}} Q_i \qquad m > 0}{(\lambda y.M)\ P_1\ ...\ P_m \twoheadrightarrow_{S^{\mathrm{w}}} (\lambda y.N)\ Q_1\ ...\ Q_m}$$

Remark that the rule for $(\lambda y.M)\ P_1\ ...\ P_m \twoheadrightarrow_{S^{\mathrm{w}}} (\lambda y.N)\ Q_1\ ...\ Q_m$ with $m = 0$ corresponds exactly to the rule for $\lambda x.M \twoheadrightarrow_{S^{\mathrm{w}}} \lambda x.N$, so it may seem awkward to distinguish the two cases. We chose to do so in order to match Definition 5.1 of head redexes: one rule applies only to abstractions which are not part of a β-redex, while the other only applies to β-redexes.

We want to prove that $\twoheadrightarrow_{S^{\mathrm{w}}}$ corresponds to $\to^*_{\beta^{\mathrm{w}}}$. First it is easy to check that the relation $\twoheadrightarrow_{S^{\mathrm{w}}}$ is included in the canonicalising β-reduction.

Proposition 5.1.7. If $M \twoheadrightarrow_{S^{\mathrm{w}}} N$ then $M \to^*_{\beta^{\mathrm{w}}} N$.

Proof. By induction on $\twoheadrightarrow_{S^{\mathrm{w}}}$. □

The converse is easily proven using Proposition 5.1.6.

Lemma 5.1.8. If $M \to^*_{\mathrm{h}^{\mathrm{w}}} \cdot \twoheadrightarrow_{S^{\mathrm{w}}} N$ then $M \twoheadrightarrow_{S^{\mathrm{w}}} N$.

Proof. By induction on the length of the reduction $\to^*_{\mathrm{h}^{\mathrm{w}}}$. If $M \twoheadrightarrow_{S^{\mathrm{w}}} N$ the result is immediate, and if $M \to_{\mathrm{h}^{\mathrm{w}}} \cdot \to^*_{\mathrm{h}^{\mathrm{w}}} \cdot \twoheadrightarrow_{S^{\mathrm{w}}} N$ then by induction hypothesis $M \to_{\mathrm{h}^{\mathrm{w}}} \cdot \twoheadrightarrow_{S^{\mathrm{w}}} N$ and we reason by induction on M and $\twoheadrightarrow_{S^{\mathrm{w}}}$. If M is a sum or an abstraction then the result is immediate by induction hypothesis. The only other possible case is $M = (\lambda y.M_0)\ P\ Q_1\ ...\ Q_m$ and $M \to_{\mathrm{h}^{\mathrm{w}}} \mathrm{can}^{\mathrm{w}}\left(M_0\left[{}^{P}/_{y}\right]\ Q_1\ ...\ Q_m\right) \twoheadrightarrow_{S^{\mathrm{w}}} N$ so we have $M \twoheadrightarrow_{S^{\mathrm{w}}} N$. □

Theorem 5.1.9. If $M \to^*_{\beta/\equiv_{+^{\mathrm{w}}}} N$ then $\mathrm{can}^{\mathrm{w}}(M) \twoheadrightarrow_{S^{\mathrm{w}}} \mathrm{can}^{\mathrm{w}}(N)$.

Proof. If $M \to^*_{\beta/\equiv_{+^{\mathrm{w}}}} N$ Proposition 4.1.1, (7) gives $\mathrm{can}^{\mathrm{w}}(M) \to^*_{\beta^{\mathrm{w}}} \mathrm{can}^{\mathrm{w}}(N)$, so we want to prove that if M and N are weakly canonical and $M \to^*_{\beta^{\mathrm{w}}} N$ then $M \twoheadrightarrow_{S^{\mathrm{w}}} N$. Moreover Proposition 5.1.6 gives that if $M \to^*_{\beta^{\mathrm{w}}} N$ then $M \to^*_{\mathrm{h}^{\mathrm{w}}} P \to^*_{\mathrm{i}^{\mathrm{w}}} N$, and according to the previous lemma if $P \twoheadrightarrow_{S^{\mathrm{w}}} N$ then $M \twoheadrightarrow_{S^{\mathrm{w}}} N$. In the end we want to prove that if $M \to^*_{\mathrm{i}^{\mathrm{w}}} N$ then $M \twoheadrightarrow_{S^{\mathrm{w}}} N$. We reason by induction on N.

- If $N = N_1 +_p N_2$ and $M \to^*_{\mathrm{i}^{\mathrm{w}}} N$ then necessarily $M = M_1 +_p N_2$ with $M_i \to^*_{\mathrm{i}^{\mathrm{w}}} N_i$ for $i \in \{1, 2\}$ so by induction hypothesis $M_i \twoheadrightarrow_{S^{\mathrm{w}}} N_i$, and $M \twoheadrightarrow_{S^{\mathrm{w}}} N$.
- The case $N = \lambda x.N_0$ is similar.
- If $N = (\lambda y.N_0)\ Q_1\ ...\ Q_m$ with $m > 0$ and $M \to^*_{\mathrm{i}^{\mathrm{w}}} N$ then $M = (\lambda y.M_0)\ P_1\ ...\ P_m$ with $M_0 \to^*_{\beta^{\mathrm{w}}} N_0$ and $P_i \to^*_{\beta^{\mathrm{w}}} Q_i$ for $i \leq m$. By induction hypothesis we know that if $S \to^*_{\mathrm{i}^{\mathrm{w}}} N_0$ then $S \twoheadrightarrow_{S^{\mathrm{w}}} N_0$, and from there we know how to deduce $M_0 \twoheadrightarrow_{S^{\mathrm{w}}} N_0$. Similarly we get $P_i \twoheadrightarrow_{S^{\mathrm{w}}} Q_i$ for $i \leq m$, and we have $M \twoheadrightarrow_{S^{\mathrm{w}}} N$.
- The case $N = y\ Q_1\ ...\ Q_m$ is similar.

□

5.2. *Weak Standardisation of the Strong Reduction*

As we mentioned in the previous Subsection, Definition 5.1 for the head β-redexes of a weakly canonical term is easily transposed to the strongly canonical terms:

- the head redexes of $M +_p N$ are the head redexes of M and the head redexes of N;
- $\lambda x_1...x_n.y\ P_1\ ...\ P_m$ has no head redex;
- the only head redex of $\lambda x_1...x_n.(\lambda y.v)\ N\ P_1\ ...\ P_m$ is $(\lambda y.v)\ N$.

We obtain the following notion of standard reduction.

Definition 5.4. The relation $M \twoheadrightarrow_S N$ between canonical terms is defined by:

$$\frac{M_1 \twoheadrightarrow_S N_1 \qquad M_2 \twoheadrightarrow_S N_2}{M_1 +_p M_2 \twoheadrightarrow_S N_1 +_p N_2} \qquad \frac{\forall i \leq m, P_i \twoheadrightarrow_S Q_i}{\lambda x_1...x_n.y\ P_1\ ...\ P_m \twoheadrightarrow_S \lambda x_1...x_n.y\ Q_1\ ...\ Q_m}$$

$$\frac{\mathrm{can}\left(\lambda x_1...x_n.v\left[{}^{P}/_{y}\right]\ Q_1\ ...\ Q_m\right) \twoheadrightarrow_S N}{\lambda x_1...x_n.(\lambda y.v)\ P\ Q_1\ ...\ Q_m \twoheadrightarrow_S N}$$

$$\frac{v \twoheadrightarrow_S N \qquad \forall i \leq m, P_i \twoheadrightarrow_S Q_i \qquad m > 0}{\lambda x_1...x_n.(\lambda y.v)\ P_1\ ...\ P_m \twoheadrightarrow_S \mathrm{can}\left(\lambda x_1...x_n.(\lambda y.N)\ Q_1\ ...\ Q_m\right)}$$

Remark that in this case an abstraction $\lambda x.M$ is canonical only if M is a value, and it is not sufficient to know that M is canonical. So in the last rule when we perform some internal reductions under a β-redex we need to canonicalise the whole result to be sure we get a canonical term.

Proposition 5.2.1. If $M \twoheadrightarrow_S N$ then $M \to^*_{\beta^c} N$.

Proof. By induction on $\twoheadrightarrow_S$. $\square$

Now we want to prove that every weak standard reduction gives a strong standard one.

Proposition 5.2.2. If $M \twoheadrightarrow_{S^w} N$ then $\mathrm{can}(M) \twoheadrightarrow_S \mathrm{can}(N)$.

Proof. We reason by induction on $\twoheadrightarrow_{S^w}$. We need to show that the following rules are admissible for $\twoheadrightarrow_S$:

$$\frac{\mathrm{can}(M_1) \twoheadrightarrow_S \mathrm{can}(N_1) \qquad \mathrm{can}(M_2) \twoheadrightarrow_S \mathrm{can}(N_2)}{\mathrm{can}(M_1 +_p M_2) \twoheadrightarrow_S \mathrm{can}(N_1 +_p N_2)} \qquad \frac{\mathrm{can}(M) \twoheadrightarrow_S \mathrm{can}(N)}{\mathrm{can}(\lambda x.M) \twoheadrightarrow_S \mathrm{can}(\lambda x.N)}$$

$$\frac{\forall i \leq m, \mathrm{can}(P_i) \twoheadrightarrow_S \mathrm{can}(Q_i)}{\mathrm{can}(y\ P_1\ ...\ P_m) \twoheadrightarrow_S \mathrm{can}(y\ Q_1\ ...\ Q_m)}$$

$$\frac{\mathrm{can}\left(M\left[{}^{P}/{}_{y}\right]\ Q_1\ ...\ Q_m\right) \twoheadrightarrow_S \mathrm{can}(N)}{\mathrm{can}((\lambda y.M)\ P\ Q_1\ ...\ Q_m) \twoheadrightarrow_S \mathrm{can}(N)}$$

$$\frac{\mathrm{can}(M) \twoheadrightarrow_S \mathrm{can}(N) \qquad \forall i \leq m, \mathrm{can}(P_i) \twoheadrightarrow_S \mathrm{can}(Q_i) \qquad m > 0}{\mathrm{can}((\lambda y.M)\ P_1\ ...\ P_m) \twoheadrightarrow_S \mathrm{can}((\lambda y.N)\ Q_1\ ...\ Q_m)}$$

Some of these are immediately given by the rules of $\twoheadrightarrow_S$. We have $\mathrm{can}(M_1 +_p M_2) = \mathrm{can}(M_1) +_p \mathrm{can}(M_2)$ and $\mathrm{can}(x\ P_1\ ...\ P_m) = x\ \mathrm{can}(P_1)\ ...\ \mathrm{can}(P_m)$ so the first and third rules are trivially admissible. The others can be simplified. For instance we have $\mathrm{can}(\lambda x.M) = \mathrm{can}(\lambda x.\mathrm{can}(M))$, so $\mathrm{can}(M) \twoheadrightarrow_S \mathrm{can}(N)$ implies $\mathrm{can}(\lambda x.M) \twoheadrightarrow_S \mathrm{can}(\lambda x.N)$ for all terms M and N if and only if $M \twoheadrightarrow_S N$ implies $\mathrm{can}(\lambda x.M) \twoheadrightarrow_S \mathrm{can}(\lambda x.N)$ for all canonical terms M and N. So we want to prove the three following rules with canonical terms:

$$\frac{M \twoheadrightarrow_S N}{\mathrm{can}(\lambda x.M) \twoheadrightarrow_S \mathrm{can}(\lambda x.N)}$$

$$\frac{\mathrm{can}\left(M\left[{}^{P}/{}_{y}\right]\ Q_1\ ...\ Q_m\right) \twoheadrightarrow_S N}{\mathrm{can}((\lambda y.M)\ P\ Q_1\ ...\ Q_m) \twoheadrightarrow_S N}$$

$$\frac{M \twoheadrightarrow_S N \qquad \forall i \leq m, P_i \twoheadrightarrow_S Q_i \qquad m > 0}{\mathrm{can}((\lambda y.M)\ P_1\ ...\ P_m) \twoheadrightarrow_S \mathrm{can}((\lambda y.N)\ Q_1\ ...\ Q_m)}$$

The last two rules are very close to actual rules of $\twoheadrightarrow_S$ and they are easily proven by induction on the canonical structure of M. For instance for the second one:

— if $M = v$ is a value then $\mathrm{can}((\lambda y.v)\ P\ Q_1\ ...\ Q_m) = (\lambda y.v)\ P\ Q_1\ ...\ Q_m$ (as we assume w.l.o.g. all the subterms P and Q_i canonical) and we get exactly a rule of $\twoheadrightarrow_S$;

— if $M = M_1 +_p M_2$ then

$$\begin{aligned}\mathrm{can}\left((M_1 +_p M_2)\left[{}^{P}/_{y}\right]\ Q_1\ ...\ Q_m\right) = &\quad \mathrm{can}\left(M_1\left[{}^{P}/_{y}\right]\ Q_1\ ...\ Q_m\right)\\ & +_p \mathrm{can}\left(M_2\left[{}^{P}/_{y}\right]\ Q_1\ ...\ Q_m\right)\\ \mathrm{can}((\lambda y.(M_1 +_p M_2))\ P\ Q_1\ ...\ Q_m) = &\quad \mathrm{can}((\lambda y.M_1)\ P\ Q_1\ ...\ Q_m)\\ & +_p \mathrm{can}((\lambda y.M_2)\ P\ Q_1\ ...\ Q_m)\end{aligned}$$

and we have necessarily $N = N_1 +_p N_2$ with $\mathrm{can}\left(M_i\left[{}^{P}/_{y}\right]\ Q_1\ ...\ Q_m\right) \twoheadrightarrow_S N_i$ for $i \in \{1, 2\}$ so we conclude by induction hypothesis.

The only remaining case is the rule specific to $\twoheadrightarrow_{S^{\mathrm{w}}}$, i.e. the abstraction rule.

$$\frac{M \twoheadrightarrow_S N}{\mathrm{can}(\lambda x.M) \twoheadrightarrow_S \mathrm{can}(\lambda x.N)}$$

We prove this by a simple induction on $M \twoheadrightarrow_S N$. □

With this result we obtain our weak standardisation property for the calculus with the strong reduction of sums.

Theorem 5.2.3. If $M \to^*_{\beta/\equiv_+} N$ then there exists N' such that $N \to^*_{\beta/\equiv_+} N'$ and $\mathrm{can}(M) \twoheadrightarrow_S N'$.

Proof. If $M \to^*_{\beta/\equiv_+} N$ then according to Theorem 4.2.5 there is N' such that $N \to^*_{\beta/\equiv_+} N'$ and $M \to^*_{\beta/\equiv_{+^{\mathrm{w}}}} N'$. Then the standardisation theorem for the weak reduction (Theorem 5.1.9) states that $\mathrm{can}^{\mathrm{w}}(M) \twoheadrightarrow_{S^{\mathrm{w}}} \mathrm{can}^{\mathrm{w}}(N')$, and the previous result gives a standard reduction $\mathrm{can}(\mathrm{can}^{\mathrm{w}}(M)) \twoheadrightarrow_S \mathrm{can}(\mathrm{can}^{\mathrm{w}}(N'))$, i.e. $\mathrm{can}(M) \twoheadrightarrow_S \mathrm{can}(N')$. Besides $N \to^*_{\beta/\equiv_+} N'$ so we obviously have $N \to^*_{\beta/\equiv_+} \mathrm{can}(N')$. □

Corollary 5.2.4. If N is a canonical β-normal form then $M \to^*_{\beta/\equiv_+} N$ if and only if $\mathrm{can}(M) \twoheadrightarrow_S N$.

6. Probabilities and Barycentric Equivalences

In Section 4 we claimed that weak reduction $\to_{+^{\mathrm{w}}}$ has better properties than $\to_+$ and Theorem 4.2.5 states that $\to_+$ is not much more expressive than $\to_{+^{\mathrm{w}}}$. Thus it would make sense to forget about $\to_+$ and work with $\to_{+^{\mathrm{w}}}$ exclusively. However $\to_{+^{\mathrm{w}}}$ is not particularly better behaved than $\to_+$ with respect to the issues addressed in the rest of the paper, so we will keep working with $\to_+$. The following results, and in particular Theorems 7.1.10 and 7.2.3, also hold for $\to_{+^{\mathrm{w}}}$, and the proofs are easily derived from those for $\to_+$.

6.1. *Terms with Probability Distributions*

So far we have proven some important results about our reduction systems. But something is missing: the probabilities do not play any role. We could actually replace the probabilistic choices $+_p$ for $p \in [0,1]$ by any family of labelled sums $+_l$ for $l \in \mathcal{L}$ an arbitrary set of labels, and all our previous results would still hold. This situation is directly inherited from the presentation of the probabilistic λ-calculus with probabilistic reductions, where the computation on the probabilities is defined at a higher level. For instance in our example at the beginning of Section 2.1 we described the probabilistic reductions of the term $\delta\ (x +_{\frac{1}{2}} y)$ and then we claimed that in the call-by-name case this term reduces in $x\ x$, $x\ y$, $y\ x$ or $y\ y$, each with probability $\frac{1}{4}$, but how one can compute this distribution is not explained in the calculus: it is a higher level construction based on all possible reductions. In this work we want to internalise as much as possible the semantics of the probabilistic terms by syntactic means. To turn the sums into probability distributions, we quotient them by the following equivalences, which correspond to the equations of *barycentric algebras* (Stone 1949).

Definition 6.1. We define four equivalences on terms by:

$$M +_p N \equiv_\gamma N +_{1-p} M \qquad\qquad (M +_p N) +_q P \equiv_\alpha M +_{pq} (N +_{\frac{q-pq}{1-pq}} P)$$

$$M +_p M \equiv_\mathcal{I} M \qquad\qquad M +_1 N \equiv_\mathcal{Z} M$$

extended under arbitrary contexts. Together these define the equivalence $\equiv_{\text{bar}}$:

$$\equiv_{\text{bar}} = \ (\equiv_\gamma \cup \equiv_\alpha \cup \equiv_\mathcal{I} \cup \equiv_\mathcal{Z})^*.$$

Remark: The associativity rule $(M +_p N) +_q P \equiv_\alpha M +_{pq} (N +_{\frac{q-pq}{1-pq}} P)$ is well defined only when $pq \neq 1$. When $pq = 1$, i.e. $p = q = 1$, we can use the other rules to complete the associativity. For any $r \in [0,1]$ we can write

$$\begin{aligned} M +_1 (N +_r P) &\equiv_\gamma (P +_{1-r} N) +_0 M \\ &\equiv_\alpha P +_0 (N +_0 M) \\ &\equiv_\gamma (M +_1 N) +_1 P \end{aligned}$$

or

$$\begin{aligned} M +_1 (N +_r P) &\equiv_\mathcal{Z} M \\ &\equiv_\mathcal{Z} M +_1 N \\ &\equiv_\mathcal{Z} (M +_1 N) +_1 P. \end{aligned}$$

Considering terms modulo $\equiv_{\text{bar}}$ amounts to viewing trees of sums as probability distributions. For instance if we consider terms with only variables and sums, there is a simple way to associate with any term M a probability distribution over variables $\mathcal{V}_M : \mathit{Var} \to [0,1]$. We just define $\mathcal{V}_x(x) = 1$ and $\mathcal{V}_x(y) = 0$ if $y \neq x$, and $\mathcal{V}_{M+_pN} = p\mathcal{V}_M + (1-p)\mathcal{V}_N$. Then we can prove that $M \equiv_{\text{bar}} N$ if and only if $\mathcal{V}_M = \mathcal{V}_N$.

But $\equiv_{\text{bar}}$ is not the first equivalence on terms we are interested in. We already looked at the β-reduction modulo $\equiv_+$, so if we further quotient by $\equiv_{\text{bar}}$ we obtain the following equivalence on terms.

Definition 6.2. The equivalence $\equiv$ is given by

$$\equiv \; = \; (\equiv_+ \cup \equiv_{\text{bar}})^*$$

As usual when considering terms modulo $\equiv_+$ we are mostly interested in canonical terms. So first we will prove that $\equiv$ is entirely described by the relation $\equiv_{\text{bar}}$ on canonical terms (Corollary 6.1.5). To do so we consider the four following reductions:

$$M +_p N \to_\gamma N +_{1-p} M \qquad (M +_p N) +_q P \to_\alpha M +_{pq} (N +_{\frac{q-pq}{1-pq}} P)$$

$$M \to_{\mathcal{I}} M +_p M \qquad M +_1 N \to_{\mathcal{Z}} M$$

extended under arbitrary context. We obviously have $\equiv_\gamma \; = \; \to_\gamma^*$ (as $\leftarrow_\gamma \; = \; \to_\gamma$), $\equiv_\alpha \; = (\to_\alpha \cup \leftarrow_\alpha)^*$, $\equiv_{\mathcal{I}} = (\to_{\mathcal{I}} \cup \leftarrow_{\mathcal{I}})^*$ and $\equiv_{\mathcal{Z}} = (\to_{\mathcal{Z}} \cup \leftarrow_{\mathcal{Z}})^*$. We will also prove that some reductions commute so it will be useful to consider some parallel reduction. We will only need to consider an independent parallel reduction: given a reduction $\to$ we define $\xrightarrow{i}$ by

$$\frac{}{M \xrightarrow{i} M} \qquad \frac{M \to N}{M \xrightarrow{i} N}$$

$$\frac{M \xrightarrow{i} N}{\lambda x.M \xrightarrow{i} \lambda x.N} \qquad \frac{M \xrightarrow{i} N \qquad M' \xrightarrow{i} N'}{M\ M' \xrightarrow{i} N\ N'} \qquad \frac{M_1 \xrightarrow{i} N_1 \qquad M_2 \xrightarrow{i} N_2}{M_1 +_l M_2 \xrightarrow{i} N_1 +_l N_2}$$

Lemma 6.1.1. For any reduction $\to \in \{\to_\gamma, \to_\alpha, \leftarrow_\alpha, \to_{\mathcal{I}}, \leftarrow_{\mathcal{I}}, \to_{\mathcal{Z}}, \leftarrow_{\mathcal{Z}}\}$, if $M \xrightarrow{i} N_1 +_p N_2$ then

(1) there is N' such that $\lambda x.N_1 +_p \lambda x.N_2 \to_+^* N'$ and $\lambda x.M \to_+^* \cdot \xrightarrow{i} N'$

(2) there is N' such that $N_1\ P +_p N_2\ P \to_+^* N'$ and $M\ P \to_+^* \cdot \xrightarrow{i} N'$.

Proof. We reason by induction on $M \xrightarrow{i} N_1 +_p N_2$. Either $M = M_1 +_p M_2$ with $M_i \xrightarrow{i} N_i$ for $i \in \{1, 2\}$ and the result is immediate, or $M \to N_1 +_p N_2$. In the second case we can assume w.l.o.g. the reduction context to be empty. Then we need to detail the different possible reductions for $\to$. There are three non trivial cases, which involve two reductions of sums. For instance with $\to_\alpha$, if $M = (N_1 +_p N_2) +_q N_3 \to_\alpha N_1 +_{pq} (N_2 +_{\frac{q-pq}{1-pq}} N_3)$ then we have $\lambda x.M \to_+^* (\lambda x.N_1 +_p \lambda x.N_2) +_q \lambda x.N_3 \to_\alpha \lambda x.N_1 +_{pq} (\lambda x.N_2 +_{\frac{q-pq}{1-pq}} \lambda x.N_3)$ and $\lambda x.N_1 +_{pq} \lambda x.(N_2 +_{\frac{q-pq}{1-pq}} N_3) \to_+ \lambda x.N_1 +_{pq} (\lambda x.N_2 +_{\frac{q-pq}{1-pq}} \lambda x.N_3)$. $\square$

Lemma 6.1.2. For any reduction $\to \in \{\to_\gamma, \to_\alpha, \leftarrow_\alpha, \to_{\mathcal{I}}, \leftarrow_{\mathcal{I}}, \to_{\mathcal{Z}}, \leftarrow_{\mathcal{Z}}\}$, if $M \xrightarrow{i} N \to_+ N'$ then there exist M' and N'' such that $M \to_+^* M' \xrightarrow{i} N''$ and $N' \to_+^* N''$.

$$\begin{tikzcd}
M \arrow[rr, dashed, "+^*"] \arrow[d, "i"'] & & M' \arrow[d, dashed, "i"] \\
N \arrow[r, "+"'] & N' \arrow[r, dashed, "+^*"'] & N''
\end{tikzcd}$$

Proof. We reason by induction on the reduction $M \xrightarrow{i} N$.

— If $M \xrightarrow{i} M$ the result is immediate.

— If $M \to N$ (with an empty reduction context) then we need to detail the cases of each reduction $\to$. They are all simple as the reduction $\to_+$ does not interact with the reduction $\to$. For instance if $M \to_{\mathcal{I}} M +_p M \to_+ N +_p M$ then $N +_p M \to_+ N +_p N$ and $M \to_+ N \to_{\mathcal{I}} N +_p N$.

— If $\lambda x.M \xrightarrow{i} \lambda x.N$ with $M \xrightarrow{i} N$ then we have either $\lambda x.N \to_+ \lambda x.N'$ with $N \to_+ N'$ and the result is immediate by induction hypothesis, or $N = N_1 +_p N_2$ and $\lambda x.N \to_+ \lambda x.N_1 +_p \lambda x.N_2$ and the result is given by the previous lemma.

— The case $M\ M' \xrightarrow{i} N\ N'$ is similar: it is immediate by induction hypothesis or given by the previous lemma.

— If $M_1 +_p M_2 \xrightarrow{i} N_1 +_p N_2$ with $M_i \xrightarrow{i} N_i$ for $i \in \{1, 2\}$ then necessarily the reduction $\to_+$ is of the form $N_1 +_p N_2 \to_+ N_1' +_p N_2'$ with $N_i \to_+ N_i'$, so the result is immediate by induction hypothesis.

□

Proposition 6.1.3. For any reduction $\to \in \{\to_\gamma, \to_\alpha, \leftarrow_\alpha, \to_{\mathcal{I}}, \leftarrow_{\mathcal{I}}, \to_{\mathcal{Z}}, \leftarrow_{\mathcal{Z}}\}$, if $M \xrightarrow{i} N$ then $\mathrm{can}(M) \xrightarrow{i} \mathrm{can}(N)$.

Proof. The reduction $\to_+$ is strongly normalising (Proposition 2.2.2), so if $M \xrightarrow{i} N$ we can use the previous lemma in an induction on the maximal length of the reductions of N to prove that there is M' such that $M \to_+^* M' \xrightarrow{i} \mathrm{can}(N)$. But we proved this for all our barycentric reductions and their anti-reductions, so if $\mathrm{can}(N) \xleftarrow{i} M'$ then there is N' such that $\mathrm{can}(N) \to_+^* N' \xleftarrow{i} \mathrm{can}(M')$, with necessarily $N' = \mathrm{can}(N)$ and $\mathrm{can}(M') = \mathrm{can}(M)$, hence $\mathrm{can}(M) \xrightarrow{i} \mathrm{can}(N)$. □

Corollary 6.1.4. The equivalences $\equiv_\gamma$, $\equiv_\alpha$, $\equiv_{\mathcal{I}}$, $\equiv_{\mathcal{Z}}$ and $\equiv_{\mathrm{bar}}$ are stable by canonicalisation.

Proof. By definition of these equivalences as closures of the corresponding reductions.

□

Corollary 6.1.5. We have $M \equiv N$ if and only if $\mathrm{can}(M) \equiv_{\mathrm{bar}} \mathrm{can}(N)$.

According to Theorem 2.2.3 the canonical terms are trees of sums with values at the leaves. As trees of sums modulo $\equiv_{\mathrm{bar}}$ describe probability distributions, classes of canonical terms modulo $\equiv_{\mathrm{bar}}$ are distributions over values.

Theorem 6.1.6. The classes of terms modulo $\equiv$ are:

$$\overline{M}, \overline{N} := \text{finite probability distributions over values } \overline{v}$$
$$\overline{v} := x \mid \lambda x.\overline{v} \mid \overline{v}\ \overline{M}.$$

A detailed proof of this theorem can be found in (Leventis 2016), with a slightly different barycentric equivalence: the rule $M +_1 N \equiv_{\mathcal{Z}} M$ is replaced by $M +_1 N \equiv_{\mathcal{Z}} M +_1 P$, which is equivalent in the presence of the idempotence $M +_p M \equiv_{\mathcal{I}} M$.

6.2. *Reduction Modulo* $\equiv_{\mathrm{bar}}$

In Sections 3 and 5 we proved some confluence and standardisation results for the β-reduction modulo $\equiv_+$. Now we want to show that these properties also hold for the β-reduction modulo $\equiv$. In Subsection 3.2 we discussed more specifically the relation between β-reduction modulo $\equiv_+$ and the usual notion of reduction modulo equivalence. We claimed that the equivalence $\equiv_+$ has too much computational content for the confluence and Church-Rosser properties modulo $\equiv_+$ to be relevant, but here the equivalence $\equiv_{\mathrm{bar}}$ is not supposed to play any role in the reduction. We could expect a very strong result stating that if $M \to^*_{\beta/\equiv} N$ then $M \to^*_{\beta/\equiv_+} \cdot \equiv_{\mathrm{bar}} N$. With this we could deduce the confluence of the reduction $\to_{\beta/\equiv}$ from the confluence of $\to_{\beta/\equiv_+}$, and the Church-Rosser property modulo $\equiv_{\mathrm{bar}}$ would follow. Unfortunately one rule of $\equiv_{\mathrm{bar}}$ makes this property fail. Indeed if $M \to_\beta N$ we have

$$M \to_{\mathcal{I}} M +_p M \to_\beta N +_p M$$

and in general there is no way to write a reduction $M \to^*_{\beta/\equiv_+} \cdot \equiv_{\mathrm{bar}} N +_p M$. However if we consider $\to_{\mathcal{I}}$ as part of the reduction, we can prove that if $M \to^*_{\beta/\equiv} N$ then $M\ (\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}})^* \cdot \equiv_{\mathrm{bar}} N$.

Definition 6.3. The reduction $\to_{\mathrm{bar}}$ is defined as

$$\to_{\mathrm{bar}} = \to_\gamma \cup \to_\alpha \cup \leftarrow_\alpha \cup \leftarrow_{\mathcal{I}} \cup \to_{\mathcal{Z}} \cup \leftarrow_{\mathcal{Z}} .$$

We obviously have $\equiv_{\mathrm{bar}} = (\to_{\mathcal{I}} \cup \to_{\mathrm{bar}})^*$. Our goal is to prove:

$$\to^*_{\beta/\equiv} \subset (\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}})^* \cdot \to^*_{\mathrm{bar}}$$

More precisely we will prove this for canonical terms, in order for $\to^*_{\beta/\equiv_+}$ and $\to^*_{\beta^c}$ to coincide.

Lemma 6.2.1. If $M \to^*_{\mathrm{bar}} N$ then $\mathrm{can}(M) \to^*_{\mathrm{bar}} \mathrm{can}(N)$. In other words if $M \to^*_{\mathrm{bar}} \cdot \to_{\mathrm{can}} N$ then $M \to_{\mathrm{can}} \cdot \to^*_{\mathrm{bar}} N$.

Proof. This is a consequence of Proposition 6.1.3. □

Lemma 6.2.2. If M is canonical and $M \to_{\mathrm{bar}} \cdot \xrightarrow{i}_\beta N$ then $M \xrightarrow{i}_\beta \cdot \to^*_{\mathrm{bar}} N$.

Proof. By a simple induction on the reductions. The key idea is that as M is canonical $\to_{\mathrm{bar}}$ can not create any β-redex: we don't have any reduction of the form $(\lambda x.P +_p \lambda x.P)\ Q \to_{\mathrm{bar}} (\lambda x.P)\ Q \to_\beta P\left[{}^{Q}/_{x}\right]$ or $(\lambda x.P +_1 R)\ Q \to_{\mathrm{bar}} (\lambda x.P)\ Q \to_\beta P\left[{}^{Q}/_{x}\right]$. □

Proposition 6.2.3. If $M\ (\to_{\beta/\equiv_+} \cup \to_{\mathrm{bar}})^*\ N$ then $\mathrm{can}(M) \to^*_{\beta^c} \cdot \to^*_{\mathrm{bar}} \mathrm{can}(N)$.

Proof. First observe that if M is canonical and $M \to^*_{\mathrm{bar}} \cdot \to_{\beta^c} N$ then the previous lemma gives $M \xrightarrow{i}_\beta \cdot \to^*_{\mathrm{bar}} \cdot \to_{\mathrm{can}} N$ hence $M \xrightarrow{i}_\beta \cdot \to_{\mathrm{can}} \cdot \to^*_{\mathrm{bar}} N$ by Lemma 6.2.1 and $M \to^*_{\beta^c} \cdot \to^*_{\mathrm{bar}} N$ by Theorem 3.2.3. Now if $M\ (\to_{\beta/\equiv_+} \cup \to_{\mathrm{bar}})^*\ N$ then Lemma 6.2.1 and Theorem 3.2.3 give

$$\mathrm{can}(M) = M_0 \to^*_{\mathrm{bar}} M'_0 \to_{\beta^c} M_1 \to^*_{\mathrm{bar}} M'_1 \to_{\beta^c} \ldots \to_{\beta^c} M_n \to^*_{\mathrm{bar}} M'_n = \mathrm{can}(N)$$

where all the terms M_i and M'_i are canonical. By an easy induction on n we get that $\mathrm{can}(M) \to_{\beta^c}^* \cdot \to_{\mathrm{bar}}^* \mathrm{can}(N)$. □

To deal with $\to_{\mathcal{I}}$ it is convenient to observe that the reduction $\to_{\mathcal{I}}^*$ has an inductive characterisation.

Proposition 6.2.4. The reduction $\to_{\mathcal{I}}^*$ is inductively given by

$$\frac{}{x \to_{\mathcal{I}}^* x} \qquad \frac{M \to_{\mathcal{I}}^* N}{\lambda x.M \to_{\mathcal{I}}^* \lambda x.N} \qquad \frac{M \to_{\mathcal{I}}^* N \qquad M' \to_{\mathcal{I}}^* N'}{M\ M' \to_{\mathcal{I}}^* N\ N'}$$

$$\frac{M_1 \to_{\mathcal{I}}^* N_1 \qquad M_2 \to_{\mathcal{I}}^* N_2}{M_1 +_p M_2 \to_{\mathcal{I}}^* N_1 +_p N_2} \qquad \frac{M \to_{\mathcal{I}}^* N_1 \qquad M \to_{\mathcal{I}}^* N_2}{M \to_{\mathcal{I}}^* N_1 +_p N_2}$$

Proof. Let us write $\twoheadrightarrow$ the reduction defined by these rules, we want to prove that it is equal to the reduction $\to_{\mathcal{I}}^*$. First an easy induction gives that $\twoheadrightarrow\ \subset\ \to_{\mathcal{I}}^*$: the cases of the context rules are immediate and for the last rule if $M \to_{\mathcal{I}}^* N_1$ and $M \to_{\mathcal{I}}^* N_2$ then $M \to_{\mathcal{I}} M +_p M \to_{\mathcal{I}}^* M_1 +_p M \to_{\mathcal{I}}^* M_1 +_p M_2$. Conversely to prove $\to_{\mathcal{I}}^*\ \subset\ \twoheadrightarrow$ we reason by induction on the length of the reduction $\to_{\mathcal{I}}^*$. The reduction $\twoheadrightarrow$ is reflexive, and we need to show that if $M \twoheadrightarrow \cdot \to_{\mathcal{I}} N$ then $M \twoheadrightarrow N$, which is again easily achieved by induction on $\twoheadrightarrow$. □

We want to show that $\to_{\mathcal{I}}$ can be performed before the other barycentric transformations: if $M \to_{\mathrm{bar}}^* \cdot \to_{\mathcal{I}}^* N$ then $M \to_{\mathcal{I}}^* \cdot \to_{\mathrm{bar}}^* N$. This is not entirely straightforward. If we have for instance

$$(M +_p N) +_q P \to_{\mathrm{bar}} M +_{pq} (N +_{\frac{q-pq}{1-pq}} P) \to_{\mathcal{I}} M +_{pq} ((N +_{\frac{q-pq}{1-pq}} P) +_r (N +_{\frac{q-pq}{1-pq}} P))$$

we can not directly duplicate the sum $N +_{\frac{q-pq}{1-pq}} P$ in the original term. We need to duplicate N and P separately:

$$(M +_p N) +_q P \to_{\mathcal{I}}^* (M +_p (N +_r N)) +_q (P +_r P) \to_{\mathrm{bar}}^* M +_{pq} ((N +_{\frac{q-pq}{1-pq}} P) +_r (N +_{\frac{q-pq}{1-pq}} P)).$$

Lemma 6.2.5. If $M_1 +_p M_2 \to_{\mathcal{I}}^* N$ then there are terms N_1 and N_2 such that $M_i \to_{\mathcal{I}}^* N_i$ for $i \in \{1,2\}$ and $N_1 +_p N_2 \to_{\mathrm{bar}}^* N$.

Proof. We reason by induction on $\to_{\mathcal{I}}^*$, and there are two possible cases. We may have directly $N = N_1 +_p N_2$ with $M_i \to_{\mathcal{I}}^* N_i$ for $i \in \{1,2\}$. Otherwise we have $N = N_1 +_q N_2$ with $M_1 +_p M_2 \to_{\mathcal{I}}^* N_j$ for $j \in \{1,2\}$ and by induction hypothesis there are terms $N_{1,1}$, $N_{1,2}$, $N_{2,1}$ and $N_{2,2}$ such that $M_i \to_{\mathcal{I}}^* N_{i,j}$ and $N_{1,j} +_p N_{2,j} \to_{\mathrm{bar}}^* N_j$. Then we have $M_i \to_{\mathcal{I}}^* N_{i,1} +_q N_{i,2}$ and

$$(N_{1,1} +_q N_{1,2}) +_p (N_{2,1} +_q N_{2,2}) \to_{\mathrm{bar}}^* (N_{1,1} +_p N_{2,1}) +_q (N_{1,2} +_p N_{2,2}) \to_{\mathrm{bar}}^* N_1 +_q N_2.$$

□

Lemma 6.2.6. If $M \to_{\mathrm{bar}}^* \cdot \to_{\mathcal{I}}^* N$ then $M \to_{\mathcal{I}}^* \cdot \to_{\mathrm{bar}}^* N$.

Proof. If we prove that whenever $M \to_{\mathrm{bar}} \cdot \to_{\mathcal{I}}^* N$ we have $M \to_{\mathcal{I}}^* \cdot \to_{\mathrm{bar}}^* N$ then we can get the proposition by induction on the length of the reduction $\to_{\mathrm{bar}}^*$. This can be done by a simple induction on the reduction $\to_{\mathrm{bar}}$. If the context of the reduction is not empty the result is immediate by induction on $\to_{\mathcal{I}}^*$. Otherwise there are six cases corresponding to the six reductions $\to_\gamma$, $\to_\alpha$, $\leftarrow_\alpha$, $\leftarrow_{\mathcal{I}}$, $\to_{\mathcal{Z}}$ and $\leftarrow_{\mathcal{Z}}$. They are all immediate thanks to the previous lemma (remark that most cases are actually easy by induction hypothesis, and the lemma is mostly useful for the associativity cases). For instance if $M_2 +_p M_1 \to_{\text{bar}} M_1 +_{1-p} M_2 \to_{\mathcal{I}}^* N$ then the lemma gives terms N_1 and N_2 and we have $M_2 +_p M_1 \to_{\mathcal{I}}^* N_2 +_p N_1 \to_{\text{bar}} N_1 +_{1-p} N_2 \to_{\text{bar}}^* N$. □

Together these results ensure that the use of $\to_{\text{bar}}$ can be postponed until the end of the reduction.

Proposition 6.2.7. If $M \to_{\beta/\equiv}^* N$ then $M\ (\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}})^* \cdot \to_{\text{bar}}^* \text{can}(N)$.

Proof. If $M \to_{\beta/\equiv}^* N$ then $M\ (\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}} \cup \to_{\text{bar}})^*\ N$ and using Theorem 3.2.3, Proposition 6.1.3 for $\to_{\mathcal{I}}^*$ and its consequence Proposition 6.2.1 we can canonicalise this reduction to get $\text{can}(M)\ (\to_{\beta^c} \cup \to_{\mathcal{I}} \cup \to_{\text{bar}})^*\ \text{can}(N)$. From Proposition 6.2.3 and Lemma 6.2.6 we deduce $\text{can}(M)\ (\to_{\beta^c} \cup \to_{\mathcal{I}})^* \cdot \to_{\text{bar}}^* \text{can}(N)$. □

We proved that except for $\to_{\mathcal{I}}$, the equivalence $\equiv_{\text{bar}}$ does not influence the reductions. If we prove that $\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}}$ is confluent then we also get that $\to_{\beta/\equiv}$ is confluent, and even that $\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}}$ is Church-Rosser modulo $\equiv_{\text{bar}}$. Yet we will not directly prove this confluence. Indeed the reduction $\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}}$ can still be simplified, and its confluence will be a consequence of the confluence of $\to_{\beta/\equiv_+}$.

7. The Role of Idempotence

In the previous section we showed that among the barycentric reductions $\to_{\mathcal{I}}$ is the only one which interfere with β+-reductions. Yet its role is minimal: in this section we will show that the reduction $\to_{\beta/\equiv_+}$ is actually at the core of the reduction $\to_{\beta/\equiv}$, and that even the role of $\to_{\mathcal{I}}$ is a marginal one.

7.1. *Confluence Modulo* $\equiv_{\text{bar}}$

Unlike the other barycentric transformations, the idempotence reduction $\to_{\mathcal{I}}$ cannot be delayed to the end of the β-reduction. But it can still be isolated from the rest of the reduction. Indeed no particular syntactic structure is required to split a term, so all the splitting can be done at the very beginning of the reduction. Of course we may duplicate too many terms: in the reduction

$$(\lambda x.x\ x)\ M \to_\beta M\ M \to_{\mathcal{I}} M\ (M +_p M)$$

we cannot duplicate M first and then β-reduce to obtain the same result. But we can extend the reduction, or use the symmetrical reduction $\leftarrow_{\mathcal{I}}$ to write

$$(\lambda x.x\ x)\ M \to_{\mathcal{I}} (\lambda x.x\ x)\ (M +_p M) \to_\beta (M +_p M)\ (M +_p M) \leftarrow_{\mathcal{I}} M\ (M +_p M).$$

We will show that if $M\ (\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}})^*\ N$ then $M \to_{\mathcal{I}}^* \cdot \to_{\beta/\equiv_+}^* \cdot \leftarrow_{\mathcal{I}}^* \text{can}(N)$. The confluence of $\to_{\beta/\equiv}$ will follow. To achieve this result we prove that $\to_{\mathcal{I}}$ commutes with the other reductions (Propositions 7.1.1, 7.1.4 and 7.1.6).

Proposition 7.1.1. $\to_{\mathcal{I}}$ is confluent. In other words if $M \leftarrow^*_{\mathcal{I}} \cdot \to^*_{\mathcal{I}} N$ then $M \to^*_{\mathcal{I}} \cdot \leftarrow^*_{\mathcal{I}} N$.

Proof. By a simple induction on $\to^*_{\mathcal{I}}$ using Proposition 6.2.4. □

Lemma 7.1.2.

(1) If $\lambda x.M_0 \to^*_{\mathcal{I}} N$ then there is N_0 such that $M_0 \to^*_{\mathcal{I}} N_0$ and $\lambda x.N_0 \to^*_+ N$.
(2) If $M_1\ M_2 \to^*_{\mathcal{I}} N$ then there are N_1 and N_2 such that $M_i \to^*_{\mathcal{I}} N_i$ for $i \in \{1,2\}$ and $N_1\ N_2\ (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N$.

Proof. We reason by induction on $\to^*_{\mathcal{I}}$ using Proposition 6.2.4, the base cases are immediate.

Otherwise if $\lambda x.M_0 \to^*_{\mathcal{I}} N_i$ for $i \in \{1,2\}$ and $N = N_1 +_p N_2$ then by induction hypothesis we have $M_0 \to^*_{\mathcal{I}} N_{i,0}$ and $\lambda x.N_{i,0} \to^*_+ N_i$ for $i \in \{1,2\}$, so $M_0 \to^*_{\mathcal{I}} N_{1,0} +_p N_{2,0}$ and $\lambda x.(N_{1,0} +_p N_{2,0}) \to_+ \lambda x.N_{1,0} +_p \lambda x.N_{2,0} \to^*_+ N_1 +_p N_2$.

Similarly if $M_1\ M_2 \to^*_{\mathcal{I}} N_i$ for $i \in \{1,2\}$ and $N = N_1 +_p N_2$ then by induction hypothesis we have terms $N_{i,j}$ with $M_j \to^*_{\mathcal{I}} N_{i,j}$ and $N_{i,1}\ N_{i,2}\ (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N_i$. The confluence of $\to_{\mathcal{I}}$ (Proposition 7.1.1) gives a term N'_2 such that $N_{i,2} \to^*_{\mathcal{I}} N'_2$ and we have $M_1 \to^*_{\mathcal{I}} N_{1,1} +_p N_{2,1}$, $M_2 \to^*_{\mathcal{I}} N'_2$ and $(N_{1,1} +_p N_{2,1})\ N'_2 \to_+ N_{1,1}\ N'_2 +_p N_{2,1}\ N'_2 \leftarrow^*_{\mathcal{I}} N_{1,1}\ N_{1,2} +_p N_{2,1}\ N_{2,2}$ with $N_{1,1}\ N_{1,2} +_p N_{2,1}\ N_{2,2}\ (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N_1 +_p N_2$. □

Lemma 7.1.3.

(1) If $\lambda x.M_1 +_p \lambda x.M_2 \to^*_{\mathcal{I}} N$ then $\lambda x.(M_1 +_p M_2) \to^*_{\mathcal{I}} \cdot \to^*_+ N$.
(2) If $M_1\ P +_p M_2\ P \to^*_{\mathcal{I}} N$ then $(M_1 +_p M_2)\ P \to^*_{\mathcal{I}} \cdot (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N$.

Proof. We only detail the case of the application, which is the most complicated. We reason by induction on $\to^*_{\mathcal{I}}$ using Proposition 6.2.4. There are two possible cases.

If $M_1\ P +_p M_2\ P \to^*_{\mathcal{I}} N_1 +_p N_2$ with $M_i\ P \to^*_{\mathcal{I}} N_i$ for $i \in \{1,2\}$ then according to the previous lemma there are N'_1, N'_2, Q_1 and Q_2 such that $M_i \to^*_{\mathcal{I}} N'_i$, $P \to^*_{\mathcal{I}} Q_i$ and $N'_i\ Q_i\ (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N_i$ for $i \in \{1,2\}$. By the confluence of $\to_{\mathcal{I}}$ there is Q such that $Q_i \to^*_{\mathcal{I}} Q$, hence $Q \leftarrow^*_{\mathcal{I}} Q_i$ and $N'_i\ Q\ (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N_i$. Then we have $(M_1 +_p M_2)\ P \to^*_{\mathcal{I}} (N'_1 +_p N'_2)\ Q \to_+ N'_1\ Q +_p N'_2\ Q\ (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N_1 +_p N_2$.

If $M_1\ P +_p M_2\ P \to^*_{\mathcal{I}} N_1 +_q N_2$ with $M_1\ P +_p M_2\ P \to^*_{\mathcal{I}} N_i$ for $i \in \{1,2\}$ then by induction hypothesis $(M_1 +_p M_2)\ P \to^*_{\mathcal{I}} R_i\ (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N_i$, and $(M_1 +_p M_2)\ P \to^*_{\mathcal{I}} R_1 +_q R_2\ (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N_1 +_p N_2$. □

Proposition 7.1.4. If $M \to_+ \cdot \to^*_{\mathcal{I}} N$ then $M \to^*_{\mathcal{I}} \cdot (\to_+ \cup \leftarrow_{\mathcal{I}})^*\ N$.

Proof. We reason by induction on $\to^*_{\mathcal{I}}$ (using Proposition 6.2.4). If the context of the reduction $\to_+$ is empty then the result is immediately given by the previous lemma. Otherwise the result is given by induction hypothesis. □

We have a similar result for the β-reduction.

Lemma 7.1.5. If $M_0 \left[{}^{P}/_{x}\right] \to^*_{\mathcal{I}} N$ then there are terms N_0 and Q such that $M_0 \to^*_{\mathcal{I}} N_0$, $P \to^*_{\mathcal{I}} Q$ and $N_0 \left[{}^{Q}/_{x}\right] \leftarrow^*_{\mathcal{I}} N$.

Proof. We reason by induction on $\to_{\mathcal{I}}^*$. In any case remark that if $M_0 = x$ then $P \to_{\mathcal{I}}^* N$ and we choose $N_0 = x$ and $Q = N$, and if $M_0 = y \neq x$ then $y \to_{\mathcal{I}}^* N$ and we choose $N_0 = N$ (and necessarily x is not free in N_0). Otherwise we detail the different cases of the induction.

— If $M_0 \left[{}^{P}/_{x}\right] \to_{\mathcal{I}}^* N_i$ for $i \in \{1,2\}$ and $N = N_1 +_p N_2$ then by induction hypothesis we have terms $N_{i,0}$ and Q_i such that $M_0 \to_{\mathcal{I}}^* N_{i,0}$, $P \to_{\mathcal{I}}^* Q_i$ and $N_{i,0} \left[{}^{Q_i}/_{x}\right] \leftarrow_{\mathcal{I}}^* N_i$. As $\to_{\mathcal{I}}$ is confluent (Proposition 7.1.1) there is Q such that $Q_i \to_{\mathcal{I}}^* Q$ so $M_0 \to_{\mathcal{I}}^* N_{1,0} +_p N_{2,0}$, $P \to_{\mathcal{I}}^* Q$ and $(N_{1,0} +_p N_{2,0}) \left[{}^{Q}/_{x}\right] \leftarrow_{\mathcal{I}}^* N_{1,0} \left[{}^{Q_1}/_{x}\right] +_p N_{2,0} \left[{}^{Q_2}/_{x}\right] \leftarrow_{\mathcal{I}}^* N_1 +_p N_2$.

— The other cases are immediate when M_0 is not a variable. For instance if $M_0 \left[{}^{P}/_{x}\right] = \lambda y.R$ and $N = \lambda y.N'$ with $R \to_{\mathcal{I}}^* N'$ then necessarily $M_0 = \lambda y.M_0'$ and $R = M_0' \left[{}^{P}/_{x}\right]$, so by induction hypothesis we have $M_0' \to_{\mathcal{I}}^* N_0'$, $P \to_{\mathcal{I}}^* Q$ and $N_0' \left[{}^{Q}/_{x}\right] \leftarrow_{\mathcal{I}}^* N'$, hence $M_0 \to_{\mathcal{I}}^* \lambda y.N_0'$ and $(\lambda y.N_0) \left[{}^{Q}/_{x}\right] \leftarrow_{\mathcal{I}}^* N$.

□

Proposition 7.1.6. If $M \to_\beta \cdot \to_{\mathcal{I}}^* N$ then $M \to_{\mathcal{I}}^* \cdot (\to_\beta \cup \leftarrow_{\mathcal{I}})^* N$.

Proof. We reason by induction on $\to_{\mathcal{I}}^*$. In any case if the context of the reduction $\to_\beta$ is empty, i.e. if we have $(\lambda x.M)\ P \to_\beta M \left[{}^{P}/_{x}\right] \to_{\mathcal{I}}^* N$, then the previous lemma gives terms N_0 and Q and we have $(\lambda x.M_0)\ P \to_{\mathcal{I}}^* (\lambda x.N_0)\ Q \to_\beta N_0 \left[{}^{Q}/_{x}\right] \leftarrow_{\mathcal{I}}^* N$. Otherwise the result is immediate by induction hypothesis. □

We proved that the reduction $\to_{\mathcal{I}}$ commutes with all the reductions we are interested in, but we need to use the reduction $\leftarrow_{\mathcal{I}}$ to get these commutations. For now given a reduction $M\ (\to_\beta \cup \to_+ \cup \to_{\mathcal{I}})^*\ N$ we can get a reduction $M \to_{\mathcal{I}}^* \cdot (\to_\beta \cup \to_+ \cup \leftarrow_{\mathcal{I}})^*$ N. We still need the following property.

Proposition 7.1.7. If $M \leftarrow_{\mathcal{I}}^* \cdot \to_{\beta^c}^* N$ with M canonical then $M \to_{\beta^c}^* \cdot \leftarrow_{\mathcal{I}}^* N$.

Proof. The proposition is given by the equivalent of Lemmas 6.2.1 and 6.2.2 for the single reduction $\leftarrow_{\mathcal{I}}$. □

With these results we can give the expected characterisation of the reduction with idempotence.

Proposition 7.1.8. If $M\ (\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}})^*\ N$ then $M \to_{\mathcal{I}}^* \cdot \to_{\beta/\equiv_+}^* \cdot \leftarrow_{\mathcal{I}}^* \mathrm{can}(N)$.

Proof. If $M\ (\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}})^*\ N$ then Theorem 3.2.3 and Proposition 6.1.3 for $\to_{\mathcal{I}}^*$ give that $M \to_+^* \mathrm{can}(M)\ (\to_\beta \cup \to_+ \cup \to_{\mathcal{I}})^*\ \mathrm{can}(N)$. Then Propositions 7.1.1, 7.1.4 and 7.1.6 give that $M \to_{\mathcal{I}}^* \cdot (\to_\beta \cup \to_+ \cup \leftarrow_{\mathcal{I}})^*\ \mathrm{can}(N)$, thus $M \to_{\mathcal{I}}^* \cdot \to_{\mathrm{can}} \cdot (\to_{\beta^c} \cup \leftarrow_{\mathcal{I}})^*$ $\mathrm{can}(N)$ and with the previous proposition we have $M \to_{\mathcal{I}}^* \cdot \to_{\mathrm{can}} \cdot \to_{\beta^c}^* \cdot \leftarrow_{\mathcal{I}}^* \mathrm{can}(N)$. □

Corollary 7.1.9. If $M \to_{\beta/\equiv}^* N$ then $M \to_{\mathcal{I}}^* \cdot \to_{\beta/\equiv_+}^* \cdot \to_{\mathrm{bar}}^* \mathrm{can}(N)$.

Proof. This is a consequence of the previous proposition and Proposition 6.2.7. □

Now that we have shown that the reductions modulo $\equiv_{\mathrm{bar}}$ are just reductions where the barycentric rules are applied at the beginning and at the end, we can easily prove some confluence properties.

Theorem 7.1.10. The reduction $\to_{\beta/\equiv}$ is confluent.

Proof. If $M \to^*_{\beta/\equiv} N_i$ for $i \in \{1,2\}$, since $N_i \equiv \mathrm{can}(N_i)$ and if there is P such that $N_i \to^*_{\beta/\equiv} P$ then $\mathrm{can}(N_i) \to^*_{\beta/\equiv} P$, we can assume w.l.o.g. that the terms N_i are canonical. Now according to the previous corollary for all $i \in \{1,2\}$ there are terms M_i and N'_i such that $M \to^*_{\mathcal{I}} M_i \to^*_{\beta/\equiv_+} N'_i \to^*_{\mathrm{bar}} N_i$. But according to Proposition 7.1.1 $\to_{\mathcal{I}}$ is confluent so there is M' with $M_i \to^*_{\mathcal{I}} M'$. Then we have $M' \ (\to_{\beta/\equiv_+} \cup \to_{\mathrm{bar}})^* \ N_i$ so according to Proposition 6.2.3 there are terms M'_i such that $M' \to^*_{\beta/\equiv_+} M'_i \to^*_{\mathrm{bar}} \mathrm{can}(N_i) = N_i$. Finally the confluence of $\to^*_{\beta/\equiv_+}$ (Theorem 3.2.2) gives a term P such that $M'_i \to^*_{\beta/\equiv_+} P$, and we have $N_i \to^*_{\beta/\equiv} P$.

$$
\begin{tikzcd}
 & & & M \arrow[dl, "\mathrm{split}^*"'] \arrow[dr, "\mathrm{split}^*"] & & & \\
 & & M_1 \arrow[dl, "(\beta/\equiv_+)^*"'] \arrow[dr, dashed, "\mathrm{split}^*"] \arrow[rr, phantom, "(7.1.1)"] & & M_2 \arrow[dl, dashed, "\mathrm{split}^*"'] \arrow[dr, "(\beta/\equiv_+)^*"] & & \\
 & N'_1 \arrow[dl, "\mathrm{bar}^*"'] \arrow[d, phantom, "(6.2.3)"] & & M' \arrow[dl, dashed, "(\beta/\equiv_+)^*"'] \arrow[dr, dashed, "(\beta/\equiv_+)^*"] & & N'_2 \arrow[dr, "\mathrm{bar}^*"] \arrow[d, phantom, "(6.2.3)"] & \\
N_1 & & M'_1 \arrow[ll, dashed, "\mathrm{bar}^*"] \arrow[dr, dashed, "(\beta/\equiv_+)^*"'] \arrow[rr, phantom, "(3.2.2)"] & & M'_2 \arrow[rr, dashed, "\mathrm{bar}^*"'] \arrow[dl, dashed, "(\beta/\equiv_+)^*"] & & N_2 \\
 & & & P & & &
\end{tikzcd}
$$

□

Corollary 7.1.11. The reduction $\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}}$ is Church-Rosser modulo $\equiv_{\mathrm{bar}}$.

Proof. If M_1 and M_2 are in relation for the reflexive symmetric transitive closure of $\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}} \cup \equiv_{\mathrm{bar}}$ then according to the confluence of $\to_{\beta/\equiv}$ there is N such that $M_i \to^*_{\beta/\equiv} N$, and Proposition 6.2.7 gives terms N_i with $M_i \ (\to_{\beta/\equiv_+} \cup \to_{\mathcal{I}})^* \ N_i \to^*_{\mathrm{bar}} \mathrm{can}(N)$, hence $N_1 \equiv_{\mathrm{bar}} N_2$. □

7.2. *Generalised Standardisation Theorem*

We can use this characterisation of the reduction modulo $\equiv_{\mathrm{bar}}$ given by Corollary 7.1.9 to recover a standardisation result.

Proposition 7.2.1. If $M \to^*_{\beta/\equiv} N$ then there exists N' such that $N \to^*_{\beta/\equiv} N'$ and $\mathrm{can}(M) \to^*_{\mathcal{I}} \cdot \twoheadrightarrow_S N'$.

Proof. If $M \to^*_{\beta/\equiv} N$ then Corollary 7.1.9 gives terms M' and N' such that $M \to^*_{\mathcal{I}}$

$M' \to^*_{\beta/\equiv_+} N' \to^*_{\text{bar}} \text{can}(N)$. According to Theorem 5.2.3 there is N'' such that $N' \to^*_{\beta/\equiv_+} N''$ and $\text{can}(M') \twoheadrightarrow_S N''$, hence $\text{can}(N) \to^*_{\beta/\equiv} N''$. Besides Proposition 6.2.1 gives $\text{can}(M) \to^*_{\mathcal{I}} \text{can}(M')$ so $\text{can}(M) \to^*_{\mathcal{I}} \cdot \twoheadrightarrow_S N''$. □

This is not the best result we can achieve. We claimed that the idempotence reduction $\to_{\mathcal{I}}$, unlike the other barycentric transformations, cannot be pushed at the end of the reductions and our counter-example was that if $M \to_\beta N$ then

$$M \to_{\mathcal{I}} M +_p M \to_\beta N +_p M.$$

This is the same reason why the direct standardisation fails: the idempotence may duplicate β-redexes, and we may choose to reduce only some of them, in which case the duplication necessarily occurs before the β-reduction. Then the same solution works: we can extend the original reduction by reducing the remaining copies of duplicated β-redexes (here $N +_p M \to_\beta N +_p N$) to get a reduction which can be simplified (here $M \to_\beta N \to_{\mathcal{I}} N +_p N$).

Lemma 7.2.2. If $M \to^*_{\mathcal{I}} \cdot \twoheadrightarrow_S N$ then there is N' such that $N \to^*_{\beta/\equiv} N'$ and $M \to^*_{\beta/\equiv_+} N'$.

Proof. We reason by induction on $\twoheadrightarrow_S$.

— If $M \to^*_{\mathcal{I}} M'_1 +_p M'_2 \twoheadrightarrow_S N_1 +_p N_2$ with $M'_i \twoheadrightarrow_S N_i$ then either $M = M_1 +_p M_2$ with $M_i \to^*_{\mathcal{I}} M'_i$ and the result is immediate by induction hypothesis, or $M \to^*_{\mathcal{I}} M_i$ for $i \in \{1, 2\}$. In this second case by induction hypothesis there are terms N'_i such that $N_i \to^*_{\beta/\equiv} N'_i$ and $M \to^*_{\beta/\equiv_+} N'_i$. As $\to^*_{\beta/\equiv_+}$ is confluent there is N' such that $N'_i \to^*_{\beta/\equiv_+} N'$, hence $M \to^*_{\beta/\equiv_+} N'$ and $N_1 +_p N_2 \to^*_{\beta/\equiv} N'_1 +_p N'_2 \to^*_{\beta/\equiv_+} N' +_p N' \leftarrow_{\mathcal{I}} N'$.
— If $M \to^*_{\mathcal{I}} \lambda x_1...x_n.y\ P'_1\ ...\ P'_m \twoheadrightarrow_S \lambda x_1...x_n.y\ Q_1\ ...\ Q_m$ with $P'_i \twoheadrightarrow_S Q_i$ then necessarily $M = \lambda x_1...x_n.y\ P_1\ ...\ P_m$ with $P_i \to^*_{\mathcal{I}} P'_i$ and the result follows by induction hypothesis.
— If $M \to^*_{\mathcal{I}} \lambda x_1...x_n.(\lambda y.v')\ P'\ Q'_1\ ...\ Q'_m \twoheadrightarrow_S N$ with $\text{can}\left(\lambda x_1...x_n.v'\left[{}^{P'}/_x\right]\ Q'_1\ ...\ Q'_m\right) \twoheadrightarrow_S N$ then necessarily $M = \lambda x_1...x_n.(\lambda y.v)\ P\ Q_1\ ...\ Q_m$ with $v \to^*_{\mathcal{I}} v'$, $P \to^*_{\mathcal{I}} P'$ and $Q_i \to^*_{\mathcal{I}} Q'_i$. Then $\lambda x_1...x_n.v\left[{}^{P}/_x\right]\ Q_1\ ...\ Q_m \to^*_{\mathcal{I}} \lambda x_1...x_n.v'\left[{}^{P'}/_x\right]\ Q'_1\ ...\ Q'_m$ so $\text{can}\left(\lambda x_1...x_n.v\left[{}^{P}/_x\right]\ Q_1\ ...\ Q_m\right) \to^*_{\mathcal{I}} \text{can}\left(\lambda x_1...x_n.v'\left[{}^{P'}/_x\right]\ Q'_1\ ...\ Q'_m\right)$ and by induction hypothesis there is N' such that $N \to^*_{\beta/\equiv} N'$ and $\text{can}\left(\lambda x_1...x_n.v\left[{}^{P}/_x\right]\ Q_1\ ...\ Q_m\right) \to^*_{\beta/\equiv_+} N'$, thus $M \to^*_{\beta/\equiv_+} N'$.
— The last case is also immediate by induction hypothesis.

□

Theorem 7.2.3. If $M \to^*_{\beta/\equiv} N$ then there is N' such that $N \to^*_{\beta/\equiv} N'$ and $\text{can}(M) \cdot \twoheadrightarrow_S N'$.

Proof. According to Proposition 7.2.1, if $M \to^*_{\beta/\equiv} N$ then there is N' with $N \to^*_{\beta/\equiv} N'$ and $\text{can}(M) \to^*_{\mathcal{I}} \cdot \twoheadrightarrow_S \cdot N'$, and the previous lemma gives a term N'' such that

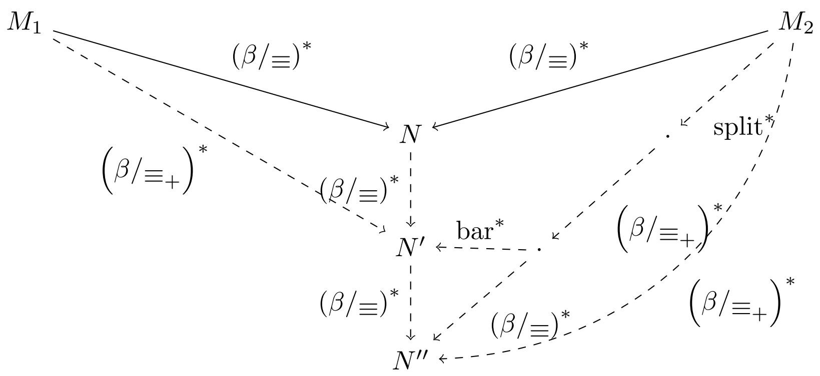
$N' \to^*_{\beta/\equiv} N''$, hence $N \to^*_{\beta/\equiv} N''$, and $\mathrm{can}(M) \to^*_{\beta/\equiv_+} N''$. Then the standardisation theorem (Theorem 5.2.3) gives a term N''' such that $N'' \to^*_{\beta/\equiv} N'''$, hence $N \to^*_{\beta/\equiv} N'''$, and $\mathrm{can}(M) \twoheadrightarrow_S N'''$. □

Corollary 7.2.4. If N is a canonical β-normal form then $M \to^*_{\beta/\equiv} N$ if and only if $\mathrm{can}(M) \twoheadrightarrow_S \cdot \equiv_{\mathrm{bar}} N$.

Proof. First if $\mathrm{can}(M) \twoheadrightarrow_S \cdot \equiv_{\mathrm{bar}} N$ then obviously $M \to^*_{\beta/\equiv} N$.

Conversely if N is a canonical β-normal form and $M \to^*_{\beta/\equiv} N$ then according to the previous theorem there is N' such that $N \to^*_{\beta/\equiv} N'$ and $\mathrm{can}(M) \twoheadrightarrow_S N'$. Using Corollary 7.1.9 we get $N \to^*_{\mathcal{I}} \cdot \to^*_{\beta/\equiv_+} \cdot \to^*_{\mathrm{bar}} N'$ (as N' is necessarily canonical), and by canonicalising all these reductions (Proposition 6.1.3 and Theorem 3.2.3) we get $N \to^*_{\mathcal{I}} N'' \to^*_{\beta^{\mathrm{c}}} \cdot \to^*_{\mathrm{bar}} N'$ with N'' canonical. As N is canonical and β-normal we can check that N'' is also β-normal, hence $N \to^*_{\mathcal{I}} \cdot \to^*_{\mathrm{bar}} N'$, i.e. $N \equiv_{\mathrm{bar}} N'$. □

With this result we can simplify the characterisation of the equivalence induced by $\to_{\beta/\equiv}$. As $\to_{\beta/\equiv}$ is confluent (Theorem 7.1.10) we know that two terms M_1 and M_2 are in relation if and only if there is N such that $M_i \to^*_{\beta/\equiv} N$ for $i \in \{1, 2\}$. With the previous theorem we can find a term N' such that $M_1 \to^*_{\beta/\equiv_+} N'$ and $N \to^*_{\beta/\equiv} N'$, thus $M_2 \to^*_{\beta/\equiv} N'$ and $M_2 \to^*_{\mathcal{I}} \cdot \to^*_{\beta/\equiv_+} \cdot \equiv_{\mathrm{bar}} N'$. Unfortunately we cannot use our results on this last reduction to prove that $\to^*_{\beta/\equiv_+}$ is Church-Rosser modulo $\equiv_{\mathrm{bar}}$. Indeed we can find some term N'' such that $M_2 \to^*_{\beta/\equiv_+} N''$ and $M_1 \to^*_{\beta/\equiv} N''$, but we cannot enforce $M_1 \to^*_{\beta/\equiv_+} \cdot \equiv_{\mathrm{bar}} N''$.



In the end we do not know if the reductions $\to^*_{\beta/\equiv_+}$ and $\to^*_{\beta/\equiv_{+^{\mathrm{w}}}}$ are Church-Rosser modulo $\equiv_{\mathrm{bar}}$ or if the reduction $\to_{\mathcal{I}}$ is really necessary.

8. Conclusion: Equational Theories

8.1. *Definition*

A motivation for this work was to be able to define a notion of equational theory in the probabilistic λ-calculus. To this purpose we defined all our basic relations on terms under arbitrary context, and we can now give a straightforward definition of probabilistic λ-theories.

Definition 8.1. A λ+-*theory* $=_\mathcal{T}$ is a congruence on Λ^+ such that:
— $(\lambda x.M)\ N =_\mathcal{T} M\left[{}^{N}/{}_{x}\right]$;
— $\lambda x.(M +_p N) =_\mathcal{T} \lambda x.M +_p \lambda x.N$;
— $(M +_p N)\ P =_\mathcal{T} M\ P +_p N\ P$;
— $M +_p N =_\mathcal{T} N +_{1-p} M$;
— $(M +_p N) +_q P =_\mathcal{T} M +_{pq} (N +_{\frac{q-pq}{1-pq}} P)$ if $pq \neq 1$;
— $M +_p M =_\mathcal{T} M$;
— $M +_1 N =_\mathcal{T} M$.

Equivalently a λ+-theory is a congruence $=_\mathcal{T}$ such that $\rightarrow_{\beta/\equiv} \subset =_\mathcal{T}$.

Proposition 8.1.1. We note $=_{\beta+}$ the reflexive symmetric transitive closure of $\rightarrow_{\beta/\equiv}$. Then $=_{\beta+}$ is the least λ+-theory for the inclusion, and

$$M =_{\beta+} N \Leftrightarrow \mathrm{can}(M) \rightarrow^*_\mathcal{I} \cdot \rightarrow^*_{\beta^c} \cdot \equiv_{\mathrm{bar}} \cdot \leftarrow^*_{\beta^c} \mathrm{can}(N).$$

Proof. Every theory contains $\rightarrow_{\beta/\equiv}$ and is closed by reflexivity, symmetry and transitivity so it contains $=_{\beta+}$. Besides it is easy to check that $=_{\beta+}$ is a congruence so it is a theory. As for its characterisation, we already stated that $\rightarrow_{\beta/\equiv}$ is confluent (Theorem 7.1.10) so if $M_1 =_{\beta+} M_2$ then there is N such that $M_i \rightarrow^*_{\beta/\equiv} N$ for $i \in \{1,2\}$, and at the end of the previous Section we described how to get a term N' such that $M_1 \rightarrow^*_{\beta/\equiv} N'$ and $\mathrm{can}(M_2) \rightarrow^*_\mathcal{I} \cdot \rightarrow^*_{\beta/\equiv_+} \cdot \rightarrow^*_{\mathrm{bar}} N'$. We can canonicalise these relations using Proposition 6.1.3 and Theorem 3.2.3 to get the expected result. □

8.2. *Example: the Observational Equivalence*

In the deterministic λ-calculus, an important example of theory is the observational equivalence. A term is called *solvable* if its head reduction normalises, and *unsolvable* otherwise; and two terms M and N are *observationally equivalent* if for every context C either $C[M]$ and $C[N]$ are both solvable or they are both unsolvable. The probabilistic notion matching the solvability is the convergence probability. In the usual setting with probabilistic reductions this probability is defined as the sum of all the normalising head reductions:

$$\mathcal{P}_\Downarrow(M) = \sum_{M \xrightarrow{p_1} \ldots \xrightarrow{p_n} h} \prod_{i=1}^{n} p_i$$

where h stands for the head normal forms $\lambda x_1 ... x_n.y\ P_1\ ...\ P_m$. It is interesting to remark that we can give an alternative definition of the same probability in our setting. Indeed we have

$$\mathcal{P}_{\Downarrow}(M) = \sup\{p \mid \exists H, N \in \Lambda^+ : M =_{\beta+} H +_p N, H \text{ hnf}\}$$

where here we say that a term $\lambda x_1 ... x_n.y\ P_1\ ...\ P_m$ is in head normal form, but also that $H_1 +_p H_2$ is in head normal form if H_1 and H_2 are both head normal forms. The equivalence of the two definitions is not trivial and relies on the characterisation of $=_{\beta+}$ and the standardisation theorem. More details can be found in the author's thesis (Leventis 2016). Then a probabilistic observational equivalence can be defined following the deterministic construction.

Definition 8.2. The observational equivalence $=_{obs}$ is given by

$$M =_{obs} N \Leftrightarrow \forall C, \mathcal{P}_{\Downarrow}(C[M]) = \mathcal{P}_{\Downarrow}(C[N]).$$

Proposition 8.2.1. $=_{obs}$ is a λ+-theory.

Proof. If $M =_{\beta+} N$ then from the definition of the convergence probability it is clear that $\mathcal{P}_{\Downarrow}(M) = \mathcal{P}_{\Downarrow}(N)$. Besides if $M =_{\beta+} N$ then $C[M] =_{\beta+} C[N]$ for any context C so two equivalent terms for $=_{\beta+}$ are observationally equivalent. Moreover if $M =_{obs} N$ then we obviously have $C[M] =_{obs} C[N]$ for any context C so $=_{obs}$ is a congruence. □

Our description of the probabilistic λ-calculus is closer to the usual operational semantics of the deterministic calculus, and as such it enjoys many similar properties. It can be used to build many of the usual probabilistic constructions while being based on simple and pure relations on terms, not indexed by any external probability. As such it is also a good starting point to extend some deterministic constructions which were incompatible with the necessity of requiring a particular reduction strategy. Further work can be found in the author's thesis (Leventis 2016). We defined a theory corresponding to the observational equivalence, and it is known that in the deterministic calculus this theory corresponds both to the equality of the infinitely extensional Böhm trees and to the maximal consistent sensible theory. All these notions can be given a probabilistic extension which preserves this correspondence.

References

Alberti, Michele (Dec. 2014). "On operational properties of quantitative extensions of λ-calculus". Theses. Aix Marseille Université ; Università di Bologna. URL: `https://hal.inria.fr/tel-01096067`.

Barendregt, H.P. (1981). *The lambda calculus, its syntax and semantics.* English. Studies in Logic and the Foundations of Mathematics, Vol. 103. Amsterdam - New York - Oxford: North-Holland Publishing Company. XIV, 615 p. \$ 109.75; Dfl. 225.00.

Bezem, Marc, Jan Willem Klop, and Roel de Vrijer (2003). *Term rewriting systems.* Cambridge University Press.

Dal Lago, Ugo and Margherita Zorzi (2012). "Probabilistic operational semantics for the lambda calculus". In: *RAIRO - Theor. Inf. and Applic.* 46.3, pp. 413–450. URL: `http://dx.doi.org/10.1051/ita/2012012`.

Ehrhard, Thomas and Vincent Danos (2011). "Probabilistic coherence spaces as a model of higher-order probabilistic computation". In: *Information and Computation* 152.1, pp. 111–137.

Ehrhard, Thomas, Michele Pagani, and Christine Tasson (2011). "The Computational Meaning of Probabilistic Coherence Spaces". In: *Proceedings of the 26th Annual IEEE Symposium on Logic in Computer Science, LICS 2011, June 21-24, 2011, Toronto, Ontario, Canada*. IEEE Computer Society, pp. 87–96. ISBN: 978-0-7695-4412-0. URL: http://dx.doi.org/10.1109/LICS.2011.29.

Leventis, Thomas (Dec. 2016). "Probabilistic lambda-theories". Theses. Aix-Marseille Université. URL: https://tel.archives-ouvertes.fr/tel-01427279.

Liguoro, Ugo de' and Adolfo Piperno (1995). "Nondeterministic Extensions of Untyped λ-Calculus". In: *Information and Computation* 149-177.122.

Stone, Marshall Harvey (1949). "Postulates for the barycentric calculus". In: *Annali di Matematica Pura ed Applicata* 29.1, pp. 25–30.

Takahashi, Masako (1995). "Parallel reductions in λ-calculus". In: *Information and computation* 118.1, pp. 120–127.

Vaux, Lionel (2009). "The algebraic lambda calculus". In: *Mathematical Structures in Computer Science* 19.5, pp. 1029–1059.