Solvability in a Probabilistic Setting

Simona Ronchi Della Rocca
Dipartimento di Informatica, Università di Torino, Italy

Ugo Dal Lago
Dipartimento di Informatica, Università di Bologna, Italy

Claudia Faggian
Université de Paris, IRIF, CNRS, France

Abstract

The notion of solvability, crucial in the λ-calculus, is conservatively extended to a probabilistic setting, and a complete characterization of it is given. The employed technical tool is a type assignment system, based on non-idempotent intersection types, whose typable terms turn out to be precisely the terms which are solvable with nonnull probability. We also supply an operational characterization of solvable terms, through the notion of head normal form, and a denotational model of $\Lambda_\oplus$, itself induced by the type system, which equates all the unsolvable terms.

2012 ACM Subject Classification Theory of computation → Lambda calculus; Theory of computation → Models of computation

Keywords and phrases Probabilistic Computation, Lambda Calculus, Solvability, Intersection Types

Digital Object Identifier 10.4230/LIPIcs.FSCD.2020.1

Category Invited Talk

1 Introduction

In *probabilistic computation*, the current state of the underlying program or machine can evolve in different ways depending on the outcome of probabilistic choices, this way turning an essentially deterministic process into a stochastic one. This computing paradigm has proved useful, in particular, in the area of cryptography [20] or in so-called randomized algorithmics [25]. From a theoretical point of view, all evolutions of a computation possibly contribute to the final result, according to the laws of probability. As a consequence, the result of a computation, if formalized using rewriting, is not *a* normal form with respect to some set of reduction rules, but a *probabilistic distribution* on all the possible outcomes. If the languages one has in mind are higher-order probabilistic languages, a natural model to consider is the λ-calculus, of course enriched with one or more probabilistic constructs.

The simplest approach, followed in [11, 8, 13] consists in endowing the λ-calculus with an operator $\oplus$ modeling fair coin flipping. This suffices to reach universality [8]: the mere presence of binary fair probabilistic choice allows to get all *computable* probability distributions on the natural numbers. The resulting calculus, called $\Lambda_\oplus$, is however well-known to be non-confluent, as recalled in Example 3.2 below. In the literature, such a problem has been handled by fixing deterministic reduction strategies. In [15], a foundational investigation of all this has been initiated following a principle stated by Plotkin [28], where a clear distinction is made between *calculi* and *programming languages*: the former consist of reduction rules (and are thus independent of any reduction strategy), enjoy confluence and standardization, while the latter are implementations of calculi, obtained by fixing a deterministic standard strategy. The aforementioned reference [28] is the first considering $\Lambda_\oplus$ as a calculus in the former sense, endowing it with the β-rule in its full generality, and

5th International Conference on Formal Structures for Computation and Deduction (FSCD 2020).
Editor: Zena M. Ariola; Article No. 1; pp. 1:1–1:17
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

a rule dealing with the probabilistic operator $\oplus$. The main results were that, under mild conditions on the probabilistic rule, confluence and standardization hold. This is done both in presence of call-by-name and of call-by-value evaluation.

In this paper, we continue this foundational investigation, from a semantic perspective. We restrict ourselves to consider the call-by-name version of $\Lambda_\oplus$, and we study in it the notion of *solvability*. Solvability is a central notion in λ-calculus theory: *solvable* programs are those which are meaningful, i.e. those that can produce any desired result when applied to a suitable sequence of arguments. More formally, a closed term M of the usual λ-calculus is solvable if there is a sequence of terms $\vec{P}$ such that $M\vec{P}$ reduces to the identity. The importance of this notion is witnessed by the fact that it is sound to equate all unsolvable terms in any denotational semantics. We define solvability in $\Lambda_\oplus$, in a conservative way with respect to λ-calculus, as follows. A closed term M of $\Lambda_\oplus$ is said to be *p-solvable*, where $p \in \mathbb{R}_{[0,1]}$ if p is the least upper bound on the probabilities of observing the identity in the distributions obtained reducing $M\vec{P}$ for any sequence of terms $\vec{P}$. In order to study solvability, we use a type assignment system, based on non-idempotent intersection types, where types are multisets (so intersections) of simple types, weighted by probabilities. The result we obtain is a complete characterization of probabilistic solvability: an operational characterization, through the notion of head normal form, a logical one, through typing, and a denotational one. In fact, the type assignment system we define supplies a model for $\Lambda_\oplus$, giving not trivial denotation to all and only the p-solvable terms, for a strictly positive real.

Related Work. The idea of endowing the λ-calculus with a form of probabilistic choice is not at all new (see, e.g., [31, 17, 27, 29, 11, 13, 8]). Most of the introduced idioms, however, come with a fixed reduction strategy, i.e. they are indeed languages, not calculi, according to Plotkin's distinction. To the authors' knowledge, the only proposals of a probabilistic λ-calculus in which reduction is studied independently on a *specific strategy* are the call-by-name calculus introduced in [22], which stems from the line of work of differential [14] and algebraic [32] λ-calculi, and the already mentioned work by the first and third authors [15].

The study of semantical properties of probabilistic λ-calculi has itself a long tradition, starting from the pioneering contributions which introduced and studied the so-called probabilistic powerdomain [31, 17], down to some deep observations about the technical problems one inevitably encounters along this route [18], until, e.g., a very recent contribution about how probabilistic higher-order computation can be reconciled with domain-theoretic semantics based on continuous functions, through call-by-push-value [16].

An alternative way of giving a denotational semantics to probabilistic λ-calculi based on coherent spaces has also been investigated [9], and the obtained model has been proved fully abstract for a probabilistic variation on Plotkin's PCF [12]. A model (itself based on coherent spaces) for a calculus very similar in spirit to $\Lambda_\oplus$ has been given [13], and proved fully abstract [24]. A full abstraction result has also been proved [5] for a model which is based on the weighted relational semantics [21].

Observational equivalence for a probabilistic λ-calculus has been further studied by Leventis [23] who proved it to coincide with the equivalence induced by Nakajima trees, i.e. Böhm trees quotiented by infinitary extensionality. A probabilistic variation on Abramsky's applicative bisimilarity has been proved sound for contextual equivalence in an untyped λ-calculus with weak-head evaluation [7], and fully abstract in presence of sequencing [19] or head evaluation.

An intersection type assignment system for $\Lambda_\oplus$ has been designed in [4], but with an aim different than the one we have in the present paper: the authors show how *weak*-head termination can be characterized by typability in idempotent intersection types. Type disciplines of other kinds, sized types [6] and linear dependent types [1] in particular, have been shown to be sound for termination of probabilistic higher-order programs.

Outline. In Section 2, we give basic notions about probability theory. In particular, we define the central notions of distribution and multidistribution. In Section 3, we introduce the calculus $\Lambda_\oplus$, a λ-calculus endowed with a probabilistic choice operator, indicated as $\oplus$. The calculus' operational semantics is given in terms of multidistributions, following [2, 15]. Then, we define the semantic notion of p-solvability, and the notion of having head normal form with probability p. In Section 4, we give a type assignment system, based on non-idempotent intersection types, where types are multidistributions of simple types, and we prove that the system enjoys the good properties of subject reduction and expansion. In Section 5 we give, through the type assignment system, a threefold characterization of solvability: operational, logical and denotational. In fact, the type assignment system induces a model for $\Lambda_\oplus$ in which terms are interpreted by *sets* of typings. Section 6 contains some concluding remarks and hints for future work. Some technical proofs are in the Appendix.

2 Preliminaries

A *discrete probability space* is given by a pair (Ω,μ), where Ω is a *countable* set, and μ is a *discrete probability distribution* on Ω, *i.e.* is a function from Ω to $[0,1]\subset\mathbb{R}$ such that $\|\mu\| := \sum_{\omega\in\Omega}\mu(\omega)=1$. In this case, a probability measure is assigned to any subset $\mathcal{A}\subseteq\Omega$ as $\mu(\mathcal{A}) = \sum_{\omega\in\mathcal{A}}\mu(\omega)$. Given a countable set Ω, a function $\mu:\Omega\to[0,1]$ is a probability *subdistribution* if $\|\mu\|\leq 1$. We write $\mathtt{DST}(\Omega)$ for the set of subdistributions on Ω. Subdistributions allow us to deal with partial results and non-successful computations. Slightly abusively, we often use the term distribution also when referring to subdistributions.

Let (Ω,μ) be as above. Any *function* $F:\Omega\to\Delta$, where Δ is another countable set, *induces a probability distribution* μ^F on Δ by composition: $\mu^F(d\in\Delta) := \mu(F^{-1}(d))$ *i.e.* $\mu\{\omega\in\Omega : F(\omega)=d\}$. The *support* of μ is $Supp(\mu)=\{\omega : \mu(\omega)>0\}$. We represent a distribution by explicitly indicating the support, and (as superscript) the probability μ assigns to each element. We write $\mu=\{a_1^{p_1},\ldots,a_n^{p_n}\}$ (where the a_is are pairwise distinct) if $\mu(a_i)=p_i$ for every $1\leq i\leq n$ and $\mu(b)=0$ for every $b\in\{a_1,\ldots,a_n\}$.

3 The Calculus

Terms and Contexts. *Terms* of $\Lambda_\oplus$ are generated by the grammar

$$M,N,P,Q ::= x \mid \lambda x.M \mid MM \mid M\oplus M \qquad (\textbf{Terms})$$

where x ranges over a countable set of *variables* (indicated as $x,y,\ldots$). As usual, $\lambda xy.PQR$ abbreviates $\lambda x.(\lambda y.(PQ)R)$, $\vec{x}$ and $\vec{M}$ denote respectively a sequence of variables and a sequence of terms, and $|\vec{x}|$ and $|\vec{M}|$ denote their lenghts. Free variables are defined as usual. $M[N/x]$ denotes the term obtained by the capture-avoiding substitution of N for each free occurrence of x in M. Terms we use frequently in our examples are $\mathsf{I}=\lambda x.x$, $\Delta=\lambda x.xx$, $\mathsf{K}=\lambda xy.x$, $\mathsf{O}=\lambda xy.y$ and $\Omega=(\lambda x.xx)(\lambda x.xx)$.

Contexts and *surface contexts* are generated by the grammars:

$$\mathbf{C} ::= \Box \mid M\mathbf{C} \mid \mathbf{C}M \mid \lambda x.\mathbf{C} \mid \mathbf{C} \oplus M \mid M \oplus \mathbf{C} \qquad (\textbf{Contexts})$$

$$\mathbf{S},\mathbf{W},\mathbf{T} ::= \Box \mid \lambda x.\mathbf{S} \mid \mathbf{S}M \qquad (\textbf{Surface Contexts})$$

where $\Box$ denotes the *hole* of the context. Given a context $\mathbf{C}$, we denote by $\mathbf{C}(M)$ the term obtained from $\mathbf{C}$ by filling the hole with M, allowing the capture of free variables. Similarly for surface contexts. Since the hole will be filled with a redex, surface contexts formalize the fact that the redex (the hole) is neither in argument position nor in the scope of a $\oplus$.

Multidistributions. To syntactically represent the global evolution of a probabilistic system, we rely on the notion of multidistribution [2].

A *multiset* is a (finite) list of elements, modulo reordering, ranged over by $\mathtt{m}$,$\mathtt{n}$. Let $\mathtt{m}$ be a multiset of pairs of the form pM, with $p \in]0,1]$, and $M \in \Lambda_\oplus$. We call $\mathtt{m} = [p_i M_i \mid i \in I]$ (where the index set I ranges over finite subsets of a countable set a *multidistribution on* $\Lambda_\oplus$ if $\sum_{i \in I} p_i \leq 1$ (think of list concatenation). We denote by $\mathtt{MDST}(\Lambda_\oplus)$ the set of all multidistributions. We write the multidistribution $[1M]$ simply as $[M]$. The sum of multidistributions is denoted by $+$. The product $q \cdot \mathtt{m}$ of a scalar q and a multidistribution $\mathtt{m}$ is defined pointwise: $q \cdot [p_1 M_1, ..., p_n M_n] = [(qp_1)M_1, ..., (qp_n)M_n]$.

Intuitively, a multidistribution $\mathtt{m} \in \mathtt{MDST}(\Lambda_\oplus)$ is a syntactical representation of a discrete probability space where to each element of the space are associated a probability and a term of $\Lambda_\oplus$. To the multidistribution $\mathtt{m} = [p_i M_i \mid i \in I]$, we associate a probability distribution $\mu_{\mathtt{m}} \in \mathtt{DST}(\Lambda_\oplus)$ as follows:

$$\mu_{\mathtt{m}}(M) = \sum_{i \in I} q_i \qquad q_i = \begin{cases} p_i & \text{if } M_i = M \\ 0 & \text{otherwise} \end{cases}$$

(Observe that, $\mathtt{m}$ being a multiset, there are in general *more than one* elements $p_i M_i$ where $M_i = M$, or even multiple copies of the same element). As usual (see Section 2), the distribution $\mu_{\mathtt{m}}$ assigns a probability measure to *every* subset of $\Lambda_\oplus$, namely the sum of the probabilities of its elements. That is, given a set of terms $\mathcal{T} \subseteq \Lambda_\oplus$,

$$\mu_{\mathtt{m}}(\mathcal{T}) = \sum_{M \in \mathcal{T}} \mu_{\mathtt{m}}(M)$$

▶ **Example 3.1** (Distributions vs. Multidistributions)**.** If $\mathtt{m} = [\frac{1}{2}a, \frac{1}{2}a]$, then $\mu_{\mathtt{m}} = \{a^1\}$. Please observe the *different nature* of distributions and multidistributions: if $\mathtt{n} = [1a]$, then $\mathtt{m} \neq \mathtt{n}$, but $\mu_{\mathtt{m}} = \mu_{\mathtt{n}}$.

Reduction Rules. We first define reduction rules on terms (Fig. 1), and one-step reduction from terms to multidistributions (Fig. 2). We then lift the definition of reduction to a binary relation on $\mathtt{MDST}(\Lambda_\oplus)$.

Observe that in the λ-calculus, a reduction step is given by the closure under context of the reduction rules. However, a reduction from terms to terms is not informative enough in a probabilistic setting, because the likelihood of each reduction step needs to be taken into account. The meaning of $M \oplus N$ is that this term reduces to either M or N, *with equal probability* $\frac{1}{2}$. There are various ways to formalize this fact, and here we follow [2, 15] and use multidistributions.

The *reduction rules* on the terms of $\Lambda_\oplus$ are defined in Fig. 1. The *(one-step) reduction*

The β-rule	*Probabilistic Rules*
$(\lambda x.M)N \mapsto_\beta M[N/x]$	$M \oplus N \mapsto_{l\oplus} M \quad M \oplus N \mapsto_{r\oplus} N$

Figure 1 Reduction Rules.

relations $\to_\beta, \to_\oplus \subseteq \Lambda_\oplus \times \mathtt{MDST}(\Lambda_\oplus)$ are defined in Fig. 2. Observe that the probabilistic rules $\mapsto_{r\oplus,l\oplus}$ are closed only under surface contexts, while the reduction rule $\mapsto_\beta$ is closed under general contexts. We denote by $\to$ the union $\to_\beta \cup \to_\oplus$. We lift the reduction relation

$$\frac{(\lambda x.M)N \mapsto_\beta M[N/x]}{\mathbf{C}((\lambda x.M)N) \to_\beta [\mathbf{C}(M[N/x])]} \qquad \frac{M \oplus N \mapsto_{l\oplus} M \quad M \oplus N \mapsto_{r\oplus} N}{\mathbf{S}(M \oplus N) \to_\oplus [\frac{1}{2}\mathbf{S}(M), \frac{1}{2}\mathbf{S}(N)]}$$

Figure 2 Reduction Steps.

$\to \subseteq \Lambda_\oplus \times \mathtt{MDST}(\Lambda_\oplus)$ to a relation $\Rightarrow \subseteq \mathtt{MDST}(\Lambda_\oplus) \times \mathtt{MDST}(\Lambda_\oplus)$, as defined in Fig. 3. Observe that $\Rightarrow$ is a reflexive relation.

We define in the same way the lifting of any relation $\to_r \subseteq \Lambda_\oplus \times \mathtt{MDST}(\Lambda_\oplus)$ to a binary

$$\frac{}{[M] \Rightarrow [M]} \qquad \frac{M \to \mathtt{m}}{[M] \Rightarrow \mathtt{m}} \qquad \frac{([M_i] \Rightarrow \mathtt{m}_i)_{i \in I}}{[p_i M_i \mid i \in I] \Rightarrow +_{i \in I} p_i \cdot \mathtt{m}_i}$$

Figure 3 Lifting $\to$ to $\Rightarrow$.

relation $\Rightarrow_r$ on $\mathtt{MDST}(\Lambda_\oplus)$. In particular, we lift $\to_\beta, \to_\oplus$ to $\Rightarrow_\beta, \Rightarrow_\oplus$. The definition of lifting allows us to apply a reduction step $\to$ to any number of M_i in the multidistribution $\mathtt{m} = [p_i M_i \mid i \in I]$. If no M_i is reduced, then $\mathtt{m} \Rightarrow \mathtt{m}$ (the relation $\Rightarrow$ is reflexive).

Reduction Sequences. A $\Rightarrow$-sequence (or *reduction sequence*) from $\mathtt{m}$ is a sequence $\mathtt{m}_0, \mathtt{m}_1, \mathtt{m}_2, \dots$ such that $\mathtt{m} = \mathtt{m}_0$ and $\mathtt{m}_n \Rightarrow \mathtt{m}_{n+1}$ for every $n \in \mathbb{N}$. We write $\mathtt{m} \Rightarrow^* \mathtt{n}$ to indicate that there is a *finite* sequence from $\mathtt{m}$ to $\mathtt{n}$, and $\langle \mathtt{m}_n \rangle_{n \in \mathbb{N}}$ for an *infinite sequence.*

Confluence. It has been proved [15] that the reduction $\Rightarrow$ enjoys the confluence property. The restriction of $\Rightarrow_\oplus$ to surface contexts is essential to obtain confluence, as the following example shows.

▶ **Example 3.2.** Let M be the term $\Delta(\mathsf{K} \oplus \mathsf{I})$. $[M] \Rightarrow_\beta [(\mathsf{K} \oplus \mathsf{I})(\mathsf{K} \oplus \mathsf{I})] \Rightarrow^* [\frac{1}{4}\mathsf{KK}, \frac{1}{4}\mathsf{KI}, \frac{1}{4}\mathsf{IK}, \frac{1}{4}\mathsf{II}] \Rightarrow^*$ $[\frac{1}{4}\lambda x.\mathsf{K}, \frac{1}{4}\lambda x.\mathsf{I}, \frac{1}{4}\mathsf{K}, \frac{1}{4}\mathsf{I}]$, which is a multidistribution on normal forms. But, if we would allow $\Rightarrow_\oplus$ also in the argument position, the result would be: $[M] \Rightarrow_\oplus [\frac{1}{2}\Delta\mathsf{K}, \frac{1}{2}\Delta\mathsf{I}] \Rightarrow^* [\frac{1}{2}\mathsf{KK}, \frac{1}{2}\mathsf{II}] \Rightarrow_\oplus$ $[\frac{1}{2}\lambda x.\mathsf{K}, \frac{1}{2}\mathsf{I}]$, which is a different multidistribution, again on normal forms!

Head Normal Forms. The notion of head normal form can be extended to $\Lambda_\oplus$. Head normal forms (shortly *hnfs*) are the normal forms of *surface reduction* $\overset{s}{\to}$, i.e., the closure of both β and $\oplus$ reduction rules under surface contexts $\mathbf{S}$. Let us write $\mathcal{H}$ for the set of *head normal forms*, which can be seen as being defined by the following grammar:

$$\mathcal{H} ::= \lambda x.\mathcal{H} \mid \mathcal{K}; \qquad \mathcal{K} ::= x \mid \mathcal{K}M.$$

It is easy to check that any term of $\Lambda_\oplus$ can be written the following form:

$$\lambda x_1...x_n.\zeta M_1...M_m,$$

where $m,n \geq 0$ and ζ (the *head*) is either a variable or a redex. So, as in the λ-calculus, the head normal forms are the terms having a variable in their head position.

In order to generalize the notion of *having a head normal form* to $\Lambda_\oplus$ we need to take into account probabilities. Recall that $\mathcal{H}$ is the set of head normal forms, and therefore $\mu_{\mathtt{m}}(\mathcal{H})$ is the probability assigned by $\mu_{\mathtt{m}}$ to the event "a term is in head normal form". Let then p be any strictly positive real number. Then:

- A term M *has head normal form with probability at least* p (notation $\geq p$-*hnf*) if there is $\mathtt{m}$ such that $[M] \Rightarrow^* \mathtt{m}$ and $\mu_{\mathtt{m}}(\mathcal{H}) \geq p$.
- A term M *has head normal form with probability* p (notation p-*hnf*) if $p = sup\{q \mid [M] \Rightarrow^* \mathtt{m}$ and $\mu_{\mathtt{m}}(\mathcal{H}) = q\}$.
- A term M *has not head normal form* if for every $\mathtt{m}$ it holds that $[M] \Rightarrow^* \mathtt{m}$ implies $\mu_{\mathtt{m}}(\mathcal{H}) = 0$.

Note that even if a term has head normal form with probability 1, that degree of certitude is not necessarily reached in any finite number of steps, as the following Example (point 2) shows:

▶ **Example 3.3.**
1. $M = \lambda yz.(y\mathsf{I} \oplus y)\Omega$ has $\geq\frac{1}{2}$-*hnf* and ≥ 1-*hnf* so it has 1-*hnf*. In fact $[M] \Rightarrow \mathtt{m} = [\frac{1}{2}\lambda yz.y\mathsf{I}\Omega, \frac{1}{2}\lambda yz.y\Omega]$, and both components of $\mathtt{m}$ are in *hnf*.
2. Let $N = \lambda x.xx \oplus \mathsf{I}$, and let $M = NN$. It is easy to check that $[NN] \Rightarrow^* [\frac{1}{2}NN, \frac{1}{2}\mathsf{I}]$, which, for every n, reduces to $\mathtt{m}$ such that $\mu_{\mathtt{m}}(\mathcal{H}) = \frac{1}{2} + \frac{1}{4} + .. + \frac{1}{2^n}$. So M has $\geq \sum_1^n \frac{1}{2^n}$-*hnf*, for all $n \geq 1$, and it thus has 1-*hnf*.
3. $\mathsf{I} \oplus \Omega$ has $\frac{1}{2}$-*hnf*.

3.1 Solvability

The notion of solvability is a central semantic notion, capturing the property of a term being meaningful (or a program being meaningful, if we consider closed terms). In λ-calculus, the semantic notion of solvability has its operational counterpart in that of head normal form, moreover it can be characterized by suitable intersection type assignment systems. We will show that similar properties hold for $\Lambda_\oplus$.

Let us first recall this notion for λ-calculus [3]. A λ-term M is solvable if there is a surface context such that $\mathbf{S}(M)$ β-reduces to the identity I (obviously considering surface contexts restricted to λ-calculus[1]). Closed solvable terms represent meaningful programs: if M is closed and solvable, then M can produce any desired result when applied to a suitable sequence of arguments. The importance of this notion is certified by the fact that it is sound to equate all unsolvable terms in any denotational semantics.

Extending this notion to $\Lambda_\oplus$ is indeed possible by way of the following definition, where p is any strictly positive real:

- A term M is *solvable with probability at least* p (notation $\geq p$-solvable) if there is a surface context $\mathbf{S}$ such that $[\mathbf{S}(M)] \Rightarrow^* \mathtt{m}$, and $\mu_{\mathtt{m}}(\mathsf{I}) \geq p$.

[1] To be precise, the definition of solvability for λ-calculus uses the notion of head context, which is a restriction of that of surface context. But since the two induced β-reductions have the same normal forms, using one or the other in the definition is equivalent. Here, the use of surface contexts is motivated by the fact that the surface reduction is standard, while head reduction is not (see [15]).

- A term M is *p-solvable* if $p = sup\{q \mid M$ is $\geq q$-solvable$\}$.
- A term is unsolvable if for every $\mathbf{S}$ it holds that $[\mathbf{S}(M)] \Rightarrow^* \mathtt{m}$ implies $\mu_{\mathtt{m}}(\mathsf{I}) = 0$.

This definition, when restricted to the syntax of λ-calculus, coincides with the standard one. Note that, while proving a term $\geq p$-solvable requires to exhibit just *one* context, proving that a term is p-solvable may require to exhibit *an infinite* number of contexts. Point 3 of the following example is an instance of this fact.

▶ **Example 3.4.**

1. $M = \lambda yz.(y\mathsf{I} \oplus y)\Omega$ is 1-solvable and $\geq\frac{1}{2}$-solvable and $\geq$1-solvable. A context playing the job is $\Box(\lambda xt.\mathsf{I})(\lambda xt.\mathsf{I})\mathsf{I}$. In fact, $[M(\lambda xt.\mathsf{I})(\lambda xt.\mathsf{I})\mathsf{I}] \Rightarrow^*_\beta [((\lambda xt.\mathsf{I})\mathsf{I} \oplus (\lambda xt.\mathsf{I}))\Omega\mathsf{I}] \Rightarrow_\oplus [\frac{1}{2}(\lambda xt.\mathsf{I})\mathsf{I}\Omega\mathsf{I}, \frac{1}{2}(\lambda xt.\mathsf{I})\Omega\mathsf{I}] \Rightarrow^*_\beta [\frac{1}{2}\mathsf{I}\mathsf{I}, \frac{1}{2}(\lambda t.\mathsf{I})\mathsf{I}] \Rightarrow_\beta [\frac{1}{2}\mathsf{I}, \frac{1}{2}\mathsf{I}]$.
2. Consider the term NN, defined in Example 3.3.2. Clearly NN is $\geq\sum_1^n \frac{1}{2^n}$-solvable, for every $n > 0$. To prove that M is 1-solvable, the context $\Box$ suffices.
3. Let Y be a fixed-point operator, whose behavior is $[YM] \Rightarrow^* [M(YM)]$. Then $[Y(\mathsf{K} \oplus \mathsf{O})] \Rightarrow^* [\frac{1}{2}\lambda x.Y(\mathsf{K} \oplus \mathsf{O}), \frac{1}{2}\mathsf{I}] \Rightarrow^* [\frac{1}{2}\mathsf{I}, \frac{1}{4}\lambda x.\mathsf{I}, \frac{1}{8}\lambda x_1x_2.\mathsf{I}, \ldots, \frac{1}{2^n}\lambda x_1..x_{n+1}.\mathsf{I}]$, $(n \geq 0)$. Then the context $\mathbf{S}_n = \Box \underbrace{\mathsf{I}..\mathsf{I}}_{n+1}$ is a witness that M is $\geq\sum_1^n \frac{1}{2^n}$-solvable, for every $n \geq 0$. Taking the supremum, this term is 1-solvable.

Three Characterizations of Solvability. In the λ-calculus, solvability can be characterized [3, 30] in three different ways:

- *operationally*, through the notion of head normal form;
- *logically*, through suitable type assignment systems, based on intersection types;
- *denotationally*, though some denotational models.

To be more precise, the operational characterization says that a term is solvable if and only if it has head normal form, the logical characterization says that there are type assignment systems assigning types to all and only the solvable terms, and the denotational one says that there are λ-models which are sensible, i.e., which assign a significant denotation to all and only the solvable terms. The aim of this paper is to show that similar characterizations hold also for $\Lambda_\oplus$, taking into account the differences between the two calculi. For proving all three characterizations, one tool is sufficient, namely an intersection type assignment system.

4 A Type Assignment System for $\Lambda_\oplus$

In this section we will present a type assignment system, based on non idempotent intersection types, which is the technical tool we will use to characterize the solvability property of $\Lambda_\oplus$.

Types. Types are defined by the following grammars:

$$\mathtt{A},\mathtt{B} ::= \alpha \mid \mathcal{A} \to \mathtt{A} \qquad \textbf{(Simple Types)}$$

$$\mathsf{a},\mathsf{b},\mathsf{c} ::= \langle p_1\mathtt{A}_1, \ldots, p_n\mathtt{A}_n \rangle \qquad \textbf{(Types)}$$

$$\mathcal{A},\mathcal{B} ::= [\mathsf{a}_1, \ldots, \mathsf{a}_n] \qquad \textbf{(Context Types)}$$

where $n \geq 0$, α ranges over a countable set of constants, types are multidistributions on simple types, and context types are multisets of types. Note that the definition of types as multidistributions means that, syntactically, $\mathsf{a} = \langle p_1\mathtt{A}_1, \ldots, p_n\mathtt{A}_n \rangle$ implies $p_i > 0$ for every $1 \leq i \leq n$ and $\Sigma_{1\leq i\leq n} p_i \leq 1$. If $\mathsf{a} = \langle p_i A_i \mid i \in I \rangle$ then its *norm* is $\|\mathsf{a}\| = \sum_i p_i$. As usual, in simple types, the type constructor $\to$ associates to the right.

Type Contexts. Type contexts, ranged over by Γ,Δ,Φ,Ψ are partial functions from variables to context types, with finite domain. $\Gamma\uplus\Delta$ denotes the function such that $(\Gamma\uplus\Delta)(x)=\Gamma(x)+\Delta(x)$. We denote by $\leq$ the set-theoretical order between partial functions.

The Type Assignment System $\mathcal{S}$. The type assignment $\mathcal{S}$ is given in Fig. 4; it proves judgements of the shape $\Gamma\vdash M:\mathtt{a}$, or $\Gamma\vdash M:\mathcal{A}$, where Γ is a type context, M is a term, $\mathtt{a}$ is a type and $\mathcal{A}$ is a context type.

▶ **Notation 4.1.** *$p\mathtt{a}$ denotes $\langle pq\mathtt{A}\,|\,q\mathtt{A}\in\mathtt{a}\rangle$; recall that $+$ denotes the multiset union, so $\mathtt{a}+\mathtt{b}$ denotes the concatenation of $\mathtt{a}$ and $\mathtt{b}$. The type $\langle 1\mathtt{A}\rangle$ is abbreviated by $\langle\mathtt{A}\rangle$. $\Gamma\vdash M:\mathtt{a}$ denotes the existence of a derivation proving this judgment, while $\Pi\triangleright\Gamma\vdash M:\mathtt{a}$ denotes a particular type derivation proving the judgment $\Gamma\vdash M:\mathtt{a}$. If $\Pi\triangleright\Gamma\vdash M:\mathtt{a}$, then M and $\mathtt{a}$ are respectively the subject and the object of Π. $\vdash M:\mathtt{a}$ abbreviates $\emptyset\vdash M:\mathtt{a}$, where $\emptyset$ is the empty function. If $\Pi\triangleright\Gamma\vdash M:\mathtt{a}$ we say that M is typed in Π with probability $\|\mathtt{a}\|$.*

$$\frac{\mathtt{a}\in\mathcal{A}}{\Gamma,x:\mathcal{A}\vdash x:\mathtt{a}}\ var$$

$$\frac{\Gamma,x:\mathcal{A}\vdash M:\langle p_i\mathtt{A}_i\,|\,i\in I\rangle}{\Gamma\vdash\lambda x.M:\langle p_i(\mathcal{A}\to\mathtt{A}_i)\,|\,i\in I\rangle}\ \to I \qquad \frac{\Gamma\vdash M:\langle p_i(\mathcal{A}_i\to\mathtt{B}_i)\,|\,i\in I\rangle\quad(\Delta_i\vdash N:\mathcal{A}_i)_{i\in I}}{\Gamma\uplus_{i\in I}\Delta_i\vdash MN:\langle p_i\mathtt{B}_i\,|\,i\in I\rangle}\ \to E$$

$$\frac{(\Gamma_i\vdash M:\mathtt{a}_i)_{i\in I}}{\uplus_i\Gamma_i\vdash M:[\mathtt{a}_i\,|\,i\in I]}\ !$$

$$\frac{\Gamma\vdash M:\mathtt{a}\quad\Gamma\vdash N:\mathtt{b}}{\Gamma\vdash M\oplus N:\frac{1}{2}\mathtt{a}+\frac{1}{2}\mathtt{b}}\ \oplus \qquad \frac{\Gamma\vdash M:\mathtt{a}}{\Gamma\vdash M\oplus N:\frac{1}{2}\mathtt{a}}\ \oplus_l \qquad \frac{\Gamma\vdash N:\mathtt{a}}{\Gamma\vdash M\oplus N:\frac{1}{2}\mathtt{a}}\ \oplus_r$$

Figure 4 The Type Assignment System $\mathcal{S}$.

The *size* of a derivation Π, denoted by $|\Pi|$, is defined as follows. Note that the size of Π is not the number of its rule applications (the dimension of the derivation tree) because of the cases of rules (!) and ($\oplus$).

- If Π is an application of the rule (var), then $|\Pi|=1$;
- If Π ends with an application of rule ($\to I$), with premise Φ, then $|\Pi|=|\Phi|+1$;
- If Π ends with an application of rule ($\to E$), with premise Φ and $(\Psi_i)_{i\in I}$, then $|\Pi|=|\Phi|+\sum_{i\in I}|\Psi_i|+1$;
- If Π ends with an application of rule (!), with premises $(\Phi_i)_{i\in I}$, then $|\Pi|=\sum_{i\in I}|\Phi_i|$;
- If Π ends with an application of rule ($\oplus$), with premises Φ,Ψ, then $|\Pi|=max\{|\Phi|,|\Psi|\}+1$;
- If Π ends with an application of rule ($\oplus_r$) (resp. ($\oplus_l$)), with premise Φ, then $|\Pi|=|\Phi|+1$;

The size of a derivation is a key notion here, since one of the characterizations, namely the proof of (1⇒2) in Theorem 5.1 is by induction on it. A benefit of using non idempotent intersection consists in the fact that it is possible to define a measure of derivations that decreases while reducing the subject. In case of λ-calculus, the size corresponds to the dimension of the derivation tree, i.e., the number of rule applications in it. Here the additive behavior of the ($\oplus$) rule obliges us to a different choice.

Some comments about the rules of $\mathcal{S}$ are in order. The rule (var) uses implicitly a weakening property. Rules ($\to I$) and ($\to E$) are similar to the usual rules for λ-calculus. Note that to the subject of the major premise is assigned a type, while to the subject of the minor premise is assigned a context type; this is possible through rule (!). The three rules for the constructor $\oplus$ are as expected. Notice that these last rules treat type environments addively, while the rule $(\to E)$ treats them multiplicatively. The use of an additive presentation for $\oplus$ rules is justified by the fact that that, in order to have the subject reduction property with respect to $\Rightarrow_\oplus$, we need weakening, as the following example shows.

▶ **Example 4.2.** Let $M=\lambda x.(x\oplus \mathsf{I})$. The following (incomplete) derivation can be built:

$$\dfrac{\dfrac{\dfrac{}{x:[\langle \mathtt{A}\rangle]\vdash x:\langle \mathtt{A}\rangle}\ var \qquad \dfrac{}{\vdash \mathsf{I}:\langle[\langle \mathtt{B}\rangle]\to \mathtt{B}\rangle}\ \to I}{x:[\langle \mathtt{A}\rangle]\vdash x\oplus \mathsf{I}:\langle \frac{1}{2}\mathtt{A},\frac{1}{2}([\langle \mathtt{B}\rangle]\to \mathtt{B})\rangle}\ \oplus}{\vdash \lambda x.(x\oplus \mathsf{I}):\langle \frac{1}{2}([\langle \mathtt{A}\rangle]\to \mathtt{A},\frac{1}{2}([\langle \mathtt{A}\rangle]\to[\langle \mathtt{B}\rangle]\to \mathtt{B}\rangle}\ \to I$$

Note that $[M]\Rightarrow_\oplus[\frac{1}{2}\mathsf{I},\frac{1}{2}\lambda x.\mathsf{I}]$; while $\vdash \mathsf{I}:\langle[\langle \mathtt{A}\rangle]\to \mathtt{A}\rangle$, it is necessary to have weakening in order to built a derivation proving $\vdash \lambda x.\mathsf{I}:\langle[\langle \mathtt{A}\rangle]\to([\langle \mathtt{B}\rangle]\to \mathtt{B})\rangle$.

In fact the weakening rule is derivable, as the following property formalizes.

▶ **Property 4.3.** $\Pi\triangleright\Gamma\vdash M:\mathtt{a}$ *implies there is* Φ *such that* $\Phi\triangleright\Gamma\uplus\Delta\vdash M:\mathtt{a}$ *and* $|\Phi|=|\Pi|$.

On the other hand, the multiplicative presentation for the rule $(\to E)$ comes naturally from the use of non idempotent intersection, which avoid the use of difficult tools to prove termination, like computability or reducibility candidates.

Rule (!) allows to assign context types to terms, and it can assign the type context $[]$ to any term, in case I is the empty set. It is not strictly necessary, a system without it could be easily designed, but it allows for an easy presentation of the $(\to E)$ rule. Note that the rule cannot be iterated.

The system can assign type also to terms with untyped subterms, through rules $(\oplus_l)$,$(\oplus r)$ and $(\to E)$, in case $I=\emptyset$. Consider the following examples:

▶ **Example 4.4.**

$$\dfrac{\dfrac{\dfrac{}{x:[\langle \mathtt{A}\rangle]\vdash x:\langle \mathtt{A}\rangle}\ var}{\vdash \mathsf{I}:\langle[\langle \mathtt{A}\rangle]\to \mathtt{A}\rangle}\ \to I}{\vdash \mathsf{I}\oplus\Omega:\langle \frac{1}{2}([\langle \mathtt{A}\rangle]\to \mathtt{A})\rangle}\ \oplus_l \qquad \dfrac{\dfrac{\dfrac{}{x:[\langle[]\to \mathtt{B}\rangle]\vdash x:\langle[]\to \mathtt{B}\rangle}\ var}{x:[\langle[]\to \mathtt{B}\rangle]\vdash x\Omega:\langle \mathtt{B}\rangle}\ \to E}{\vdash \lambda x.x\Omega:\langle[\langle[]\to \mathtt{B}\rangle]\to \mathtt{B}\rangle}\ \to I$$

Properties of the Type Assignment System $\mathcal{S}$. The system $\mathcal{S}$ enjoys the good properties we expect, namely subject reduction and expansion. Before going into that, we need to prove an important property of surface contexts, namely that terms filling their hole positions inherit from them both the typability and the norm, as expressed by the following lemma.

▶ **Lemma 4.5.** *If* $\Pi\triangleright\Gamma\vdash \mathbf{S}(M):\mathtt{a}$*, then there are* Δ *and* $\mathtt{b}$ *such that* $\Delta\vdash M:\mathtt{b}$*, where* $\|\mathtt{a}\|=\|\mathtt{b}\|$.

Proof. By induction on $\mathbf{S}$. If $\mathbf{S}=\square$ the proof is obvious. If $\mathbf{S}=\mathbf{T}N$, then Π is of the shape:

$$\dfrac{\Pi'\triangleright\Gamma'\vdash \mathbf{T}(M):\langle p_i(\mathcal{A}_i\to \mathtt{A}_i)\,|\,i\in I\rangle \qquad (\Delta_i\vdash N:\mathcal{A}_i)_{i\in I}}{\Gamma'\uplus_{i\in I}\Delta_i\vdash \mathbf{T}(M)N:\langle p_i\mathtt{A}_i\,|\,i\in I\rangle}\ \to E$$

We conclude by induction. If $\mathbf{S}=\lambda x.\mathbf{T}$, then the claim follows by induction, too. ◀

Typing is preserved by both reduction and expansion, but these properties, which are standard in intersection type assignment systems, must be adapted to the probabilistic setting.

▶ **Lemma 4.6** (One-Step Subject Reduction)**.** *Let* $\Pi\triangleright\Gamma\vdash M:\mathtt{a}$.

1. *If* $M \to_\beta [N]$ *then there is* $\Psi \triangleright \Gamma \vdash N : \mathtt{a}$.
2. *If* $M \to_\oplus [\frac{1}{2}N_1, \frac{1}{2}N_2]$*, then one of the two following cases happens:*
   - $\mathtt{a} = \frac{1}{2}\mathtt{a}_1 + \frac{1}{2}\mathtt{a}_2$ *and* $\Psi_1 \triangleright \Gamma \vdash N_1 : \mathtt{a}_1$, $\Psi_2 \triangleright \Gamma \vdash N_2 : \mathtt{a}_2$*;*
   - $\mathtt{a} = \frac{1}{2}\mathtt{b}$ *and* $\Psi \triangleright \Gamma \vdash N_i : \mathtt{b}$*, for some* $1 \leq i \leq 2$.

Moreover, if the redex is typed in Π*, then* $|\Psi| < |\Pi|$ *(resp.* $|\Psi_i| < |\Pi|$*).*

Proof. The proof is in the Appendix. ◀

▶ **Lemma 4.7** (Subject Reduction). $\Pi \triangleright \Gamma \vdash M : \mathtt{a}$ *and* $[M] \Rightarrow^* \mathtt{m} = [p_i N_i \mid i \in I]$ *imply there is* $J \subseteq I$, $\mathtt{a} = \Sigma_{j \in J} p_j \mathtt{a}_j$ *and* $\Pi_j \triangleright \Gamma \vdash N_j : \mathtt{a}_j$.

Proof. By induction on the lenght of the reduction, using Lemma 4.6. ◀

▶ **Lemma 4.8** (Subject Expansion). $[M] \Rightarrow^* [p_i N_i \mid i \in I]$ *and* $\Gamma \vdash N_j : \mathtt{a}_j$ *for some* $j \in J \subseteq I$ *imply* $\Delta \vdash M : \Sigma_{j \in J} p_j \mathtt{a}_j$*, for some* Δ, $\Gamma \leq \Delta$.

Proof. By induction on the length of the reduction, see the Appendix. ◀

The system can assign type to every *hnf*.

▶ **Property 4.9.** *Let* $M \in \mathcal{H}$*. Then for every* $p \in]0,1]$ *there are* Γ,$\mathtt{a}$ *such that* $\Gamma \vdash M : \mathtt{a}$ *and* $\|\mathtt{a}\| = p$.

Proof. Let $M \in \mathcal{H}$. The proof is by induction on the grammar defining $\mathcal{H}$. Let $M \in \mathcal{K}$: we will prove that, for every $\mathtt{a}$ there is Γ such that $\Gamma \vdash M : \mathtt{a}$. Let $\mathtt{a} = \langle p_i \mathtt{A}_i \mid i \in I \rangle$. If $M = x$, then choose $\Gamma = x : [\mathtt{a}]$, if $M = NP$, where $N \in \mathcal{K}$, then by induction there is Γ such that $\Gamma \vdash N : \langle p_i([\,] \to \mathtt{A}_i) \mid i \in I \rangle$ and the proof follows by rule $(\to E)$. If $M = \lambda x.N$, with $N \in \mathcal{H}$ the proof comes by induction and rule $(\to I)$. ◀

Note that the previous property says, in particular, that if a term is in *hnf*, then it is always possible to assign it a type with norm 1.

5 Characterizing Solvability

The next theorem shows the key result of this paper.

▶ **Theorem 5.1** (Finitary Characterization). *The three following statements are equivalent:*
1. $\Gamma \vdash M : \mathtt{a}$*, with* $\|\mathtt{a}\| = p$.
2. M *has* $\geq p$-hnf.
3. M *is* $\geq p$*-solvable.*

Proof. 1⇒2 Let $\Pi \triangleright \Gamma \vdash M : \mathtt{a}$; we prove that $[M] \Rightarrow^* \mathtt{m}$ with $\mu_{\mathtt{m}}(\mathcal{H}) \geq \|\mathtt{a}\|$, by induction on $|\Pi|$. Note that if M is in head normal form, the claim holds. If $|\Pi| = 1$, then $M = x$ is in *hnf*. Let $|\Pi| > 1$. If M is in *hnf*, the claim holds; if M not in *hnf*, then, according to Lemma 4.6, three cases can happen:

a $M \overset{s}{\to}_\beta [N]$, and $\Psi \triangleright \Gamma \vdash N : \mathtt{a}$;
b $M \overset{s}{\to}_\oplus [\frac{1}{2}N_1, \frac{1}{2}N_2]$, $\mathtt{a} = \frac{1}{2}\mathtt{a}_i$, and $\Psi \triangleright \Gamma \vdash N_i : \mathtt{a}_i$, for some $i \in \{1,2\}$;
c $M \overset{s}{\to}_\oplus [\frac{1}{2}N_1, \frac{1}{2}N_2]$, $\mathtt{a} = \frac{1}{2}\mathtt{a}_1 + \frac{1}{2}\mathtt{a}_2$, and $\Psi_i \triangleright \Gamma \vdash N_i : \mathtt{a}_i$, for all $i \in \{1,2\}$;

and in all cases $|\Psi| < |\Pi|$, $|\Psi_i| < |\Pi|$ by Lemma 4.5. In case **a**, the result follows by induction on the structure of Ψ. In the case **b**, by induction it holds that $[N_i] \Rightarrow^* \mathtt{n}_i$ with $\mu_{\mathtt{n}_i}(\mathcal{H}) \geq \|\mathtt{a}_i\|$, for some $1 \leq i \leq 2$. Assume $i = 1$. Then $[M] \Rightarrow^* \frac{1}{2}\mathtt{n}_1 + \frac{1}{2}[N_2] = \mathtt{m}$, so $\mu_{\mathtt{m}}(\mathcal{H}) \geq \frac{1}{2}\mu_{\mathtt{n}_1}(\mathcal{H}) \geq \|\mathtt{a}\|$. Let us consider case **c**. By i.h., it holds that $[N_i] \Rightarrow^* \mathtt{n}_i$ with $\mu_{\mathtt{n}_i}(\mathcal{H}) \geq \|\mathtt{a}_i\|$ for every $i \in \{1,2\}$. Hence we have that $M \overset{s}{\to} [\frac{1}{2}N_1, \frac{1}{2}N_2] \Rightarrow^* (\frac{1}{2}\mathtt{n}_1 + \frac{1}{2}\mathtt{n}_2) = \mathtt{m}$, where $\mu_{\mathtt{m}}(\mathcal{H}) = \frac{1}{2}\mu_{n_1}(\mathcal{H}) + \frac{1}{2}\mu_{n_2}(\mathcal{H}) \geq_{i.h.} \frac{1}{2}\|\mathtt{a}_1\| + \frac{1}{2}\|\mathtt{a}_2\| = \|\mathtt{a}\|$.

2⇒3 Assume M in $\mathcal{H}$; we prove that M is 1-solvable, by showing how to build a particular head context $\mathbf{S}$ such that $[\mathbf{S}(M)] \Rightarrow^* [\mathsf{I}]$. Let $M = \lambda x_1...x_n.zM_1...M_m$. If $z = x_i$, for some i, then the context $\Box\underbrace{(\lambda z_1...z_m.\mathsf{I})}_{n}$ does the job. If z is free, then use the context $\lambda z.\Box(\lambda z_1...z_m.\mathsf{I})$. For the general case, let $[M] \Rightarrow^* \mathtt{m}$, and $\mu_{\mathtt{m}}(\mathcal{H}) = q \geq p$. Let $\mathtt{m} = [q_1N_1,...,q_nN_n] + \mathtt{n}$, where $\mu_{\mathtt{n}}(\mathcal{H}) = 0$ and $\Sigma_{1\leq i\leq n}q_i \geq p$. W.l.o.g., we assume M closed. Every N_i is of the shape $\lambda\vec{x_i}.z^i\vec{P_i}$, where $z^i \in \vec{x_i}$; let $r = max_{1\leq i\leq n}|\vec{x_i}|$ and $s = max_{1\leq i\leq n}|\vec{P_i}|$. Choose $w_1,..,w_r$ fresh variables. Then the desired context is:

$$\mathbf{H} = (\lambda w_1...w_r.\Box w_1...w_r)\underbrace{(\lambda t_1...t_{r+s}.\mathsf{I})...(\lambda t_1...t_{r+s}.\mathsf{I})}_{r}\underbrace{\mathsf{I}...\mathsf{I}}_{r+s}$$

$[\mathbf{S}(M)] \Rightarrow^* [q_1\mathbf{S}(N_1),...,q_n\mathbf{S}(N_n)] + \mathtt{p}$. It is sufficient to prove that $[\mathbf{S}(N_i)] \Rightarrow^* [\mathsf{I}]$. Now,

$$[\mathbf{S}(N_i)] \Rightarrow^*_\beta [(\lambda w_1...w_r.w^i\vec{P'_i})\underbrace{(\lambda t_1...t_{r+s}.\mathsf{I})...(\lambda t_1...t_{r+s}.\mathsf{I})}_{r}\underbrace{\mathsf{I}...\mathsf{I}}_{r+s}],$$

where $\vec{P'_i} = \vec{Q}_i w_{|P_i|+1}...w_r$, $w^i \in \{w_1,...,w_r\}$, $Q_i = P_i[\vec{w_i}/\vec{x_i}]$ and $|\vec{P'_i}| \leq |\vec{P_i}| + r - 1 \leq s + r - 1$. Then, after r reduction steps, we obtain: $[(\lambda t_1...t_{r+s}.\mathsf{I})\vec{P''_i}\underbrace{\mathsf{I}...\mathsf{I}}_{r+s}] \Rightarrow^*_\beta [(\lambda\vec{t}.\mathsf{I})\underbrace{\mathsf{I}...\mathsf{I}}_{l}]$, where $l \leq r+s$. Since s $|\vec{t}| \leq r+s$, $[(\lambda\vec{t}.\mathsf{I})\underbrace{\mathsf{I}...\mathsf{I}}_{l}] \Rightarrow^*_\beta [\underbrace{\mathsf{I}...\mathsf{I}}_{o}] \Rightarrow^*_\beta [I]$ $(o \leq r+s-1)$.

3⇒1 Let M be $\geq p$-solvable. Then there is $\mathbf{S}$ such that $[\mathbf{S}(M)] \Rightarrow^* [p_i\mathsf{I} \mid i \in I] + \mathtt{m}$, where $\Sigma_{i\in I}p_i \geq p$. By Property 4.9, there is $\mathtt{a}$, with $\|\mathtt{a}\| = 1$, such that $\vdash \mathsf{I} : \mathtt{a}$, so, by Subject Expansion, $\Gamma \vdash \mathbf{S}(M) : \Sigma_{i\in I}p_i\mathtt{a}$. By Lemma 4.5, $\Gamma \vdash M : \mathtt{b}$, where $\|\mathtt{b}\| = \Sigma_{i\in I}p_i\|\mathtt{a}\|$. ◀

The results of Theorem 5.1 can be extended to the supremum, this way enabling a complete characterization:

▶ **Theorem 5.2** (Characterization). *The three following statements are equivalent*

1. $p = sup\{q \mid \Gamma \vdash M : \mathtt{a}$, *for some* $\Gamma,\mathtt{a}$, *and* $q = \|\mathtt{a}\|\}$.
2. M *has* p-hnf.
3. M *is* p*-solvable.*

Theorem 5.2 implicitly supplies three different characterizations of solvability, similarly to what happens in the λ-calculus. Namely, the equivalence 2⇔3 corresponds to an *operational characterization* of solvability, and 1⇔3 corresponds to a *logical characterization.* Moreover 1⇔2 gives a logical characterization of *hnf*s.

A Model for $\Lambda_\oplus$. $\mathcal{S}$ is an extension of the basic type assignment system defined in [26], which gives rise to a relational model of λ-calculus. It is possible to reason in a similar way here, and to extract from $\mathcal{S}$ a model of $\Lambda_\oplus$, in the sense specified by Property 5.3. As has been proved in [26], following a seminal observation of [10], the interpretation of a term in a model extracted from a type assignment system with non-idempotent intersections, depends not only on the types derivable for it, but also on the related type contexts. In fact, the context is necessary to preserve the quantitative aspect of types. Let us define the basic ingredients of our model. An *abstract typing* is a pair $(\Gamma;\mathtt{a})$, where Γ is a type context and $\mathtt{a}$ is a type, not necessarily related to each-other. Let $\mathcal{T}$ be the set of abstract typings: the space $\mathcal{D}$ of *denotations* of our model is the power set of $\mathcal{T}$. $\mathcal{D}$ is equipped by two operations

$\circ,\oplus:\mathcal{D}\longrightarrow\mathcal{D}$ which allow to interpret terms of $\Lambda_\oplus$. Their definitions reflect, respectively, the behavior of the typing rules $(\rightarrow E)$ and $(\oplus)$ of $\mathcal{S}$:

$$
\begin{aligned}
t_1\circ t_2 =\; & \{(\Gamma;\langle p_i\mathtt{A}_i \mid i\in I\rangle) \mid (\Gamma;\langle p_i([\,]\rightarrow\mathtt{A}_i)\mid i\in I\rangle\in t_1\}\cup \\
& \left\{(\Gamma\uplus_{i\in I}\Delta_i;\langle p_i\mathtt{A}_i\mid i\in I\rangle \;\middle|\; \begin{array}{l}(\Gamma;\langle p_i(\mathcal{A}_i\rightarrow\mathtt{A}_i)\mid i\in I\rangle\in t_1,\mathcal{A}_i=[\mathtt{a}^i_j\mid j\in J_i],\\ (\Delta^i_j;\mathtt{a}^i_j)\in t_2,\Delta_i=\uplus_{j\in J_i}\Delta^i_j\end{array}\right\};\\
t_1\oplus t_2 =\; & \left\{\left(\Gamma;\frac{1}{2}\mathtt{a}\right)\mid(\Gamma;\mathtt{a})\in t_1\right\}\cup\left\{\left(\Gamma;\frac{1}{2}\mathtt{a}\right)\mid(\Gamma;\mathtt{a})\in t_2\right\}\cup\\
& \left\{\left(\Gamma;\frac{1}{2}\mathtt{a}_1+\frac{1}{2}\mathtt{a}_2\right)\mid(\Gamma;\mathtt{a}_i)\in t_i,i=1,2\right\}.
\end{aligned}
$$

Moreover, if $t\in\mathcal{D}$ and p is a probability, $p\bullet t$ denotes the element of $\mathcal{D}$ such that, if $(\Gamma;\mathtt{a})\in t$ then $(\Gamma;p\mathtt{a})\in p\bullet t$. Let ρ be a denotational environment, assigning an element of $\mathcal{D}$ to every variable: the interpretation of a term under the environment ρ is defined by induction as follows:

$$
\begin{aligned}
[\![x]\!]_\rho &= \rho(x)\\
[\![MN]\!]_\rho &= [\![M]\!]_\rho\circ[\![N]\!]_\rho\\
[\![M\oplus N]\!]_\rho &= [\![M]\!]_\rho\oplus[\![N]\!]_\rho\\
[\![\lambda x.M]\!]_\rho &= \{(\Gamma;\langle p_i(\mathcal{A}\rightarrow\mathtt{A}_i)\mid i\in I\rangle\mid\mathcal{A}=[\mathtt{a}_j\mid j\in J],(\Gamma\uplus_{j\in J}\Delta_j;\langle p_i\mathtt{A}_i\mid i\in I\rangle)\in[\![M]\!]_{\rho[t/x]},\\
&\qquad t=\{(\Delta_j;\mathtt{a}_j)\mid j\in J\}\}
\end{aligned}
$$

It is easy to check that the interpretation of a term is related to its concrete typings in the following way:

$$
\begin{aligned}
[\![M]\!]_\rho=\{(\Gamma;\mathtt{a})\mid\Delta\vdash M:\mathtt{a},\Gamma=\uplus_{i\in I}\Delta_i \text{ such that for every } x,\\
\Delta(x)=[\mathtt{a}_i\mid i\in I] \text{ implies } (\Delta_i;\mathtt{a}_i)\in\rho(x)\}
\end{aligned}
$$

If M is a closed term, then its interpretation is even simpler, namely:

$$[\![M]\!]=\{(\Gamma;\mathtt{a})\mid\exists\Gamma,\mathtt{a}.\Gamma\vdash M:\mathtt{a}\}$$

In particular, since M closed and $\Gamma\vdash M:\mathtt{a}$ together imply that $\vdash M:\mathtt{a}$, the interpretation of a closed term depends only on the types derivable for it. In the following we will restrict ourselves to consider only closed terms: clearly all the properties we prove hold also for the open terms, but are expressed in a more cumbersome way.

The model is correct with respect to the operational behavior of $\Lambda_\oplus$, i.e., the following property holds.

▶ **Property 5.3** (Adequacy). *Let M be closed. $M\Rightarrow^*[p_iM_i\mid i\in I]$ implies $[\![M]\!]=\cup_{i\in I}p_i\bullet[\![M_i]\!]$.*

Finally, the model characterizes solvability, in the following sense:

▶ **Property 5.4.** *Let M be closed. M is p-solvable if and only if $p=sup\{q\mid(\Gamma;\mathtt{a})\in[\![M]\!]\}$ and $\|\mathtt{a}\|=q$.*

Note that M is unsolvable if and only if $[\![M]\!]_\rho=\emptyset$ for every ρ, so, using the terminology of λ-calculus, this model is sensible, since it equates all unsolvable terms.

6 Conclusions and Future Work

We investigated the notion of solvability in the context of the calculus $\Lambda_\oplus$ as introduced in [15], and focusing on the call-by-name parameter passing regime. Solvability being a semantic property, and call-by-name and call-by-value behaving quite differently semantically [30], we leave the task of extending our work to call-by-value to some future work. The definition of solvability we give is a conservative extension of the one from the pure λ-calculus, and explicitly takes probability into account: a term is dubbed p-solvable if, put in a suitable context, it reduces to the identity with probabilities *at most*, but *arbitrary close to*, p. We characterize solvability through a type assignment system based on non-idempotent intersection types. Such a system supplies a logical characterization of solvability, but also induces an operational one in which being p-solvable corresponds to having head normal form with probability p. Finally, the type system induces a model for $\Lambda_\oplus$, in which all unsolvable terms (i.e., terms which are 0-solvable) are equated.

It would be interesting to study the theory induced by our model from a finer point of view, in particular with respect to the equivalence it induces on terms. Certainly, this equivalence cannot coincide with the operational one, characterized in [23], since our model is not extensional. The type assignment system could however be enriched with an equivalence between types, in such a way as to induce an extensional model, thus catching the operational semantics of $\Lambda_\oplus$, in the sense of Plotkin.

Moreover, we intend to give a domain-theoretic account of our model: we believe that it gives a logical description of the category of weighted relational models [21], as conjectured by an anonymous referee.

References

1 Martin Avanzini, Ugo Dal Lago, and Alexis Ghyselen. Type-based complexity analysis of probabilistic functional programs. In *Proc. of LICS 2019*, pages 1–13, 2019.

2 Martin Avanzini, Ugo Dal Lago, and Akihisa Yamada. On probabilistic term rewriting. In *Proc. of FLOPS 2018*, pages 132–148, 2018.

3 H. P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics*, volume 103. North Holland, 1984.

4 Flavien Breuvart and Ugo Dal Lago. On intersection types and probabilistic lambda calculi. In David Sabel and Peter Thiemann, editors, *Proc. of PPDP 2018*, pages 8:1–8:13, 2018.

5 Pierre Clairambault and Hugo Paquet. Fully abstract models of the probabilistic lambda-calculus. In *Proc. of CSL 2018*, pages 16:1–16:17, 2018.

6 Ugo Dal Lago and Charles Grellois. Probabilistic termination by monadic affine sized typing. *ACM Trans. Program. Lang. Syst.*, 41(2):10:1–10:65, 2019.

7 Ugo Dal Lago, Davide Sangiorgi, and Michele Alberti. On coinductive equivalences for higher-order probabilistic functional programs. In *Proc. of POPL 2014*, pages 297–308, 2014.

8 Ugo Dal Lago and Margherita Zorzi. Probabilistic operational semantics for the lambda calculus. *RAIRO - Theor. Inf. and Applic.*, 46(3):413–450, 2012.

9 Vincent Danos and Thomas Ehrhard. Probabilistic coherence spaces as a model of higher-order probabilistic computation. *Inf. Comput.*, 209(6):966–991, 2011.

10 Daniel de Carvalho. Execution time of λ-terms via denotational semantics and intersection types. *Math. Struct. Comput. Sci.*, 28(7):1169–1203, 2018.

11 Alessandra Di Pierro, Chris Hankin, and Herbert Wiklicky. Probabilistic lambda-calculus and quantitative program analysis. *J. Log. Comput.*, 15(2):159–179, 2005.

12 T. Ehrhard, C. Tasson, and M. Pagani. Probabilistic coherence spaces are fully abstract for probabilistic PCF. In *Proc. of POPL 2014*, pages 309–320, 2014.

13 Thomas Ehrhard, Michele Pagani, and Christine Tasson. The computational meaning of probabilistic coherence spaces. In *Proc. of LICS 2011*, pages 87–96, 2011.

14 Thomas Ehrhard and Laurent Regnier. The differential lambda-calculus. *Theor. Comput. Sci.*, 309(1-3):1–41, 2003.

15 Claudia Faggian and Simona Ronchi Della Rocca. Lambda calculus and probabilistic computation. In *Proc. of LICS 2019*, pages 1–13, 2019.

16 Jean Goubault-Larrecq. A probabilistic and non-deterministic call-by-push-value language. In *Proc. of LICS 2019*, pages 1–13, 2019.

17 Claire Jones. *Probabilistic non-determinism.* PhD thesis, University of Edinburgh, UK, 1990.

18 Achim Jung and Regina Tix. The troublesome probabilistic powerdomain. *Electron. Notes Theor. Comput. Sci.*, 13:70–91, 1998.

19 Simona Kasterovic and Michele Pagani. The discriminating power of the let-in operator in the lazy call-by-name probabilistic lambda-calculus. In *Proc. of FSCD 2019*, pages 26:1–26:20, 2019.

20 Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography.* Chapman & Hall Cryptography and Network Security Series. Chapman & Hall, 2007.

21 J. Laird, G. Manzonetto, G. McCusker, and M. Pagani. Weighted relational models of typed lambda-calculi. In *proc. of LICS 2013*, pages 301–310, 2013.

22 Thomas Leventis. *Probabilistic lambda-theories.* Phd Thesis, Aix-Marseille Université, 2016. Available at `https://tel.archives-ouvertes.fr/tel-01427279v2/document`.

23 Thomas Leventis. Probabilistic böhm trees and probabilistic separation. In *Proc. of LICS 2018*, pages 649–658, 2018.

24 Thomas Leventis and Michele Pagani. Strong adequacy and untyped full-abstraction for probabilistic coherence spaces. In *Proc. of FoSSaCS 2019*, pages 365–381, 2019.

25 Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms.* Cambridge University Press, 1995.

26 Luca Paolini, Mauro Piccolo, and Simona Ronchi Della Rocca. Essential and relational models. *Math. Struct. Comput. Sci.*, 27(5):626–650, 2017.

27 Sungwoo Park. A calculus for probabilistic languages. In *Proc. of TLDI 2003*, pages 38–49, 2003.

28 Gordon D. Plotkin. Call-by-name, call-by-value and the lambda-calculus. *Theor. Comput. Sci.*, 1(2):125–159, 1975.

29 Norman Ramsey and Avi Pfeffer. Stochastic lambda calculus and monads of probability distributions. In *Proc. of POPL 2002*, pages 154–165, 2002.

30 Simona Ronchi Della Rocca and Luca Paolini. *The Parametric Lambda Calculus - A Metamodel for Computation.* Texts in Theoretical Computer Science. An EATCS Series. Springer, 2004.

31 N. Saheb-Djahromi. Probabilistic LCF. In *Proc. of MFCS 1978*, pages 442–451, 1978.

32 Lionel Vaux. The algebraic lambda calculus. *Mathematical Structures in Computer Science*, 19(5):1029–1059, 2009.

A Some Technical Proofs

Subject Reduction. As usual, the subject reduction property relies on a substitution property.

▶ **Lemma A.1** (Substitution). $\Pi \triangleright \Gamma, x : \mathcal{A} \vdash M : \mathtt{a}$ *(resp.* $\Pi \triangleright \Gamma, x : \mathcal{A} \vdash M : \mathcal{B}$*) and* $\Theta \triangleright \Delta \vdash N : \mathcal{A}$ *imply* $\Pi[\Theta/x] \triangleright \Gamma \uplus \Delta \vdash M[N/x] : \mathtt{a}$ *(resp.* $\Pi[\Theta/x] \triangleright \Gamma \vdash M[N/x] : \mathcal{B}$*). Moreover,* $|\Pi[\Theta/x]| < |\Pi| + |\Theta|$.

Proof. By induction on Π:

- Let Π be:

$$\frac{a \in \mathcal{A}}{\Gamma, y : \mathcal{A} \vdash y : \mathtt{a}}\ var$$

On the other side, the derivation Θ is of the shape:

$$\frac{(\Pi_i \triangleright \Gamma_i \vdash N : \mathtt{a}_i)_{i \in I}}{\uplus_{i \in I} \Gamma_i \vdash N : [\mathtt{a}_i \mid i \in I]}\ !$$

so $\mathtt{a} = \mathtt{a}_i$, for some i. Let $y = x$. By Property 4.3, there is a derivation $\Xi \triangleright \Gamma \uplus_{i \in I} \Gamma_i \vdash N : \mathtt{a}_i$, such that $|\Xi| = |\Pi_i|$. So $\Pi[\Theta/x] = \Xi$. If $y \neq x$, then $\Pi[\Theta/x] = \Pi$. In both cases the condition on the size of $\Pi[\Theta/x]$ is obvious.

- Let Π be:

$$\frac{\Xi \triangleright \Gamma, x : \mathcal{A}' \vdash P : \langle p_i(\mathcal{B}_i \to \mathtt{B}_i) \mid i \in I \rangle \quad (\Pi_i \triangleright \uplus_{i \in I} \Sigma_i, x : \mathcal{A}_i \vdash Q : \mathcal{B}_i)_{i \in I}}{\Gamma \uplus_{i \in I} \Sigma_i \uplus x : (\mathcal{A}' +_{i \in I} \mathcal{A}_i) \vdash PQ : \langle p_i \mathtt{B}_i \mid i \in I \rangle}\ \to E$$

where $\mathcal{A} = \mathcal{A}' +_{i \in I} \mathcal{A}_i$. Then the derivation Θ is:

$$\frac{(\Phi_{\mathtt{b}} \triangleright \Delta_{\mathtt{b}} \vdash N : \mathtt{b})_{\mathtt{b} \in \mathcal{A}'} \quad ((\Phi^i_{\mathtt{c}} \triangleright \Delta^i_{\mathtt{c}} \vdash N : \mathtt{c})_{\mathtt{c} \in \mathcal{A}_i})_{i \in I}}{+_{\mathtt{b} \in \mathcal{A}'} \Delta_{\mathtt{b}} +_{\mathtt{c} \in \mathcal{A}', i \in I} \Delta^i_{\mathtt{c}} \vdash N : \mathcal{A}}\ !$$

Then we can build derivations Ψ and Ψ_i, with subject N, by rule (!) with premises respectively $(\Phi_{\mathtt{b}})_{\mathtt{b} \in \mathcal{A}'}$ and $(\Phi^i_{\mathtt{c}})_{\mathtt{c} \in \mathcal{A}_i}$; by induction there are $\Xi[\Psi/x] \triangleright P[N/x] : \langle p_i(\mathcal{B}_i \to \mathtt{b}_i) \mid i \in I \rangle$ and $(\Pi_i[\Psi_i/x] \triangleright \vdash Q[N/x] : \mathcal{B}_i)$. Since $PQ[N/x] = P[N/x]Q[N/x]$ the result follows by rule $(\to E)$. Moreover by induction $|\Xi[\Psi/x]| < |\Xi| + |\Psi|$, $|\Pi_i[\Psi_i/x]| < |\Pi_i| + |\Psi_i|$, so $|\Pi[\Theta/x]| = |\Xi[\Psi/x]| +_{i \in I} |\Pi_i[\Psi_i/x]| + 1 < |\Xi| + |\Psi| +_{i \in I} (|\Pi_i| + |\Psi_i|) + 1 = |\Pi| + |\Phi'|$

- Let Π be:

$$\frac{\Pi_1 \triangleright \Gamma, x : \mathcal{A} \vdash P : \mathtt{a} \quad \Pi_2 \triangleright \Gamma, x : \mathcal{A} \vdash N : \mathtt{b}}{\Gamma, x : \mathcal{A} \vdash P \oplus Q : \frac{1}{2}\mathtt{a} + \frac{1}{2}\mathtt{b}}\ \oplus$$

By induction there are $\Pi_1[\Theta/x] \triangleright \Gamma \uplus \Delta \vdash P[N/x] : \mathtt{a}$ and $\Pi_2[\Theta/x] \triangleright \Gamma \uplus \Delta \vdash Q[N/x] : \mathtt{b}$, so $\Pi[\Theta/x]$ can be built by rule $(\oplus)$. Moreover $|\Pi[\Theta/x]| = max\{|\Pi_1[\Theta/x]|, |\Pi_2[\Theta/x]|\} + 1 <_{i.h.} max\{(|\Pi_1| + |\Theta|, |\Pi_2| + |\Theta|) + 1 = max\{|\Pi_1|, |\Pi_2|\} + |\Theta| + 1 = |\Pi| + |\Theta|$.

- If the last used rule is $(\to I)$, $(\oplus_r)$, $\oplus_l)$ the proof follows by induction. ◀

Let us identify an occurrence of a term N in a term M by the context $\mathbf{C}$ such that $M = \mathbf{C}(N)$. Then, given a typing derivation $\Pi \triangleright \Gamma \vdash M : \mathtt{a}$, an occurrence of a subterm of M is a typed occurrence of Π if and only if it is the subject of a subderivation of Π.

► **Lemma** (4.6, One-Step Subject Reduction). *Let $\Pi \triangleright \Gamma \vdash M : \mathtt{a}$.*

1. *If $M \to_\beta [M']$ then there is $\Pi' \triangleright \Gamma \vdash M' : \mathtt{a}$.*
2. *If $M \to_\oplus [\frac{1}{2}M_1, \frac{1}{2}M_2]$, then one of the two following cases happens:*
   - $\mathtt{a} = \frac{1}{2}\mathtt{a}_1 + \frac{1}{2}\mathtt{a}_2$ *and $\Pi'_1 \triangleright \Gamma \vdash M_1 : \mathtt{a}_1$, $\Pi'_2 \triangleright \Gamma \vdash M_2 : \mathtt{a}_2$;*
   - $\mathtt{a} = \frac{1}{2}\mathtt{b}$ *and $\Pi' \triangleright \Gamma \vdash M_i : \mathtt{b}$, for some i $(i \in \{1, 2\})$.*

   Moreover, if the redex is typed in Π, then $|\Pi'| < |\Pi|$ (resp. $|\Pi'_i| < |\Pi|$).

Proof.

1. By induction on the context $\mathbf{C}$ such that $M = \mathbf{C}((\lambda x.P)Q)$ and $M' = \mathbf{C}(P[Q/x])$. The base case follows by Lemma A.1, the induction case is easy.

2. By induction on the context $\mathbf{S}$ such that $M = \mathbf{S}(P \oplus Q)$, and then by induction on Π.
   a. In case $\mathbf{S} = \Box$, then the last rule of Π is $\oplus$, and the proof is obvious. Let $\mathbf{S} = \lambda x.\mathbf{S}'$. So $M = \lambda x.N \to_\oplus [\frac{1}{2}\lambda x.N_1, \frac{1}{2}\lambda x.N_2]$, so $N \to_\oplus [\frac{1}{2}N_1, \frac{1}{2}N_2]$. Π is of the shape:

$$\frac{\Pi' \triangleright \Gamma, x:\mathcal{A} \vdash N : \langle p_i \mathtt{A}_i \mid i \in I \rangle}{\Gamma \vdash \lambda x.N : \langle p_i(\mathcal{A} \to \mathtt{A}_i) \mid i \in I \rangle} \to I$$

   by induction $\langle p_i \mathtt{A}_i \mid i \in I \rangle = \langle \frac{q_i}{2}\mathtt{A}_i \mid i \in I_1 \rangle + \langle \frac{q_i}{2}\mathtt{A}_i \mid i \in I_2 \rangle$, and $\Gamma \vdash N_j : \langle q_i \mid i \in I_j \rangle$, where $I = I_1 \cup I_2$, $p_i = \frac{q_i}{2}$ $(j = 1,2)$. Then by rule $(\to I)$, $\Pi_j \triangleright \Gamma \vdash \lambda x.N_j : \langle q_i(\mathcal{A} \to \mathtt{A}_i) \mid i \in I_j \rangle$ $(j = 1,2)$, and then, by rule $(\oplus)$, $\Phi \triangleright \Gamma \vdash \lambda x.N_1 \oplus \lambda x.N_2 : \frac{1}{2}\langle q_i(\mathcal{A} \to \mathtt{A}_i) \mid i \in I_1 \rangle + \frac{1}{2}\langle q_i(\mathcal{A} \to \mathtt{A}_i) \mid i \in I_2 \rangle = \langle p_i(\mathcal{A} \to \mathtt{A}_i) \mid i \in I \rangle$. Moreover, if the redex is typed in Π, then it is typed in Π', so by induction $|\Pi'_i| < |\Pi'|$. Then $|\Phi| = \frac{1}{2}|\Pi_1| + \frac{1}{2}|\Pi_2| = \frac{1}{2}(|\Pi'_1| + |\Pi'_2|) + 1 < |\Pi'| + 1 = |\Pi|$.
   Let $\mathbf{S} = \mathbf{S}'P$. Then $M = NP \to_\oplus [\frac{1}{2}N_1P, \frac{1}{2}N_2P]$ and Π is of the shape:

$$\frac{\Pi' \triangleright \Sigma \vdash N : \langle p_i(\mathcal{A}_i \to \mathtt{A}_i) \mid i \in I \rangle \quad (\Psi_i \triangleright \Delta_i \vdash P : \mathcal{A}_i)_{i \in I}}{\Gamma = \Sigma \uplus_{i \in I} \Delta_i \vdash NP : \langle p_i \mathtt{A}_i \mid i \in I \rangle} \to E$$

   By induction on Π', $I = I_1 \cup I_2$ and $\langle p_i(\mathcal{A}_i \to \mathtt{A}_i) \mid i \in I \rangle = \langle \frac{1}{2}p_i(\mathcal{A}_i \to \mathtt{A}_i) \mid i \in I_1 \rangle + \langle \frac{1}{2}p_i(\mathcal{A}_i \to \mathtt{A}_i) \mid i \in I_2 \rangle$ such that $\Pi_1 \triangleright \Sigma \vdash N_1 : \langle p_i(\mathcal{A}_i \to \mathtt{A}_i) \mid i \in I_1 \rangle$ and $\Pi_2 \triangleright \Sigma \vdash N_2 : \langle p_i(\mathcal{A}_i \to \mathtt{A}_i) \mid i \in I_2 \rangle$, so, by rule $(\to E)$, $\Phi_j \triangleright \Sigma \uplus_{i \in I_j} \Delta_j \vdash N_j P : \langle p_i \mathtt{A}_i \mid i \in I_j \rangle$ $(j \in \{1,2\})$. By Property 4.3, $\Sigma \uplus_{i \in I} \Delta_i \vdash N_j P : \langle p_i \mathtt{A}_i \mid i \in I_j \rangle$. So, by rule $(\oplus)$, we obtain $\Theta \triangleright \Gamma \vdash N_1 P \oplus N_2 P : \frac{1}{2}\langle p_i \mathtt{A}_i \mid i \in I_1 \rangle + \frac{1}{2}\langle p_i \mathtt{A}_i \mid i \in I_2 \rangle$. Moreover, if the redex is typed in Π, then it is typed in Π', so by induction $|\Pi_i| < |\Pi'|$. Since $|\Phi_j| = |\Pi_j| +_{i \in I_j} |\Psi_i| + 1$, we have: $|\Theta| = \frac{1}{2}|\Phi_1| + \frac{1}{2}|\Phi_2| = \frac{1}{2}(|\Pi_1| +_{i \in I_1} |\Psi_i| + 1) + \frac{1}{2}(|\Pi_2| +_{i \in I_2} |\Psi_i| + 1) = \frac{1}{2}(|\Pi_1| + |\Pi_2|) +_{i \in I} |\Psi_i| + 1 < |\Pi'| +_{i \in I} |\Psi_i| + 1 = |\Pi|$.
   b. Similar to the previous case, but easier.

If the redex occurs in an untyped occurence of Π, then, by Lemma 4.5, it is a β-redex. So, if $\Pi \triangleright \Gamma \vdash \mathbf{C}(M) : \mathtt{a}$, and $M \to_\beta M'$, then $\Pi' \triangleright \Gamma \vdash \mathbf{C}(M') : \mathtt{a}$ can be obtained from Π just replacing the occurrence of M by M'. ◀

Subject Expansion.

▶ **Lemma A.2** (Inverse Substitution). *$\Pi \triangleright \Gamma \vdash M[N/x] : \mathtt{a}$ implies there is $\mathcal{A}$ such that $\Sigma, x:\mathcal{A} \vdash M : \mathtt{a}$, $\Delta \vdash N : \mathcal{A}$ and $\Gamma \subseteq \Sigma \uplus \Delta$.*

Proof. By induction on M. All the cases follow easily by induction. Note that, in case all the occurrences of N in M are untyped in Π, then $\mathcal{A} = []$ and $\Sigma = \Gamma$. ◀

▶ **Lemma A.3** (One-Step Subject Expansion).
1. *$\Pi \triangleright \Gamma \vdash M : \mathtt{a}$ and $N \to_\beta [M]$ imply $\Gamma \vdash N : \mathtt{a}$.*
2. *$\Pi \triangleright \Gamma \vdash M : \mathtt{a}$ and $N \to_\oplus [\frac{1}{2}M, \frac{1}{2}P]$ imply $\Gamma \to N : \frac{1}{2}\mathtt{a}$.*
3. *$\Pi \triangleright \Gamma \vdash M_i : \mathtt{a}_i$ for every $1 \leq i \leq 2$ and $N \to_\oplus [\frac{1}{2}M_1, \frac{1}{2}M_2]$ imply $\Gamma \vdash N : \frac{1}{2}\mathtt{a}_1 + \frac{1}{2}\mathtt{a}_2$.*

Proof.
1. By Lemma A.2.
2. By induction on the context $\mathbf{S}$ such that $N = \mathbf{S}(R_1 \oplus R_2) \to_\oplus [\frac{1}{2}\mathbf{S}(R_1), \frac{1}{2}\mathbf{S}(R_2)]$ and either $M = \mathbf{S}(R_1)$ or $P = \mathbf{S}(R_2)$, and then by induction on Π. In case $\mathbf{S} = \Box$, the proof is obvious. Let $\mathbf{S} = \lambda x.\mathbf{S}'$, so $N = \lambda x.Q \to_\oplus [\frac{1}{2}\lambda x.N_1, \frac{1}{2}\lambda x.N_2]$, where $Q \to_\oplus [\frac{1}{2}N_1, \frac{1}{2}N_2]$. Let $\Pi \triangleright \Gamma \vdash \lambda x.N_1 : \mathtt{a}$. Then Π is of the shape:

$$\frac{\Gamma, x:\mathcal{A} \vdash N_1 : \langle p_i \mathtt{B}_i \mid i \in I \rangle}{\Gamma \vdash \lambda x.N_1 : \langle p_i(\mathcal{A} \to \mathtt{B}_i \mid i \in I \rangle} \to I$$

By induction $\Gamma, x:\mathcal{A} \vdash Q : \langle \frac{1}{2} p_i \mathtt{B}_i \mid i \in I \rangle$, and then, by rule $(\to I)$, $\Gamma \vdash \lambda x.Q : \langle \frac{1}{2} p_i (\mathcal{A} \to \mathtt{B}_i) \mid i \in I \rangle$. Let $\mathbf{S} = \mathbf{S}' R$, so $N = QR \to_{\oplus} [\frac{1}{2} N_1 R, \frac{1}{2} N_2 R]$. Then Π is of the shape:

$$\frac{\Gamma \vdash N_1 : \langle p_i (\mathcal{A}_i \to \mathtt{A}_i) \mid i \in I \rangle \quad (\Delta_i \vdash R : \mathcal{A}_i)_{i \in I}}{\Gamma \uplus_{i \in I} \Delta_i \vdash N_1 R : \langle p_i \mathtt{A}_i \mid i \in I \rangle} \to E$$

By induction, $\Gamma \vdash Q : \langle \frac{1}{2} p_i (\mathcal{A}_i \to \mathtt{A}_i) \mid i \in I \rangle$, so the proof follows by rule $(\to E)$. ◀