Probabilistic Termination by Monadic Affine Sized Typing

UGO DAL LAGO, University of Bologna & INRIA Sophia Antipolis
CHARLES GRELLOIS, INRIA Sophia Antipolis

We introduce a system of monadic affine sized types, which substantially generalizes usual sized types and allows in this way to capture probabilistic higher-order programs that terminate almost surely. Going beyond plain, strong normalization without losing soundness turns out to be a hard task, which cannot be accomplished without a richer, quantitative notion of types, but also without imposing some affinity constraints. The proposed type system is powerful enough to type classic examples of probabilistically terminating programs such as random walks. The way typable programs are proved to be almost surely terminating is based on reducibility but requires a substantial adaptation of the technique.

CCS Concepts: • **Theory of computation** → **Probabilistic computation**; **Type theory**;

Additional Key Words and Phrases: Probabilistic lambda-calculus, almost-sure termination, sized types, affine types, parametrized reducibility

ACM Reference format:
Ugo Dal Lago and Charles Grellois. 2019. Probabilistic Termination by Monadic Affine Sized Typing. *ACM Trans. Program. Lang. Syst.* 41, 2, Article 10 (March 2019), 65 pages.
https://doi.org/10.1145/3293605

10

1 INTRODUCTION

Probabilistic models are more and more pervasive in computer science [32, 36, 40]. Moreover, the concept of the algorithm originally assuming determinism has been relaxed so as to allow probabilistic evolution since the very early days of theoretical computer science [21]. All this has given impetus to research on probabilistic programming languages, which, however, have been studied at a large scale only in the last 20 years, following advances in randomized computation [35], cryptographic protocol verification [4, 5], and machine learning [25]. Probabilistic programs can be seen as ordinary programs in which specific instructions are provided to make the program evolve probabilistically rather than deterministically. The typical examples are instructions for sampling from a given distribution toolset or for performing probabilistic choice.

One of the most crucial properties a program should satisfy is *termination*: the execution process should be guaranteed to end. In (non)deterministic computation, this is easy to formalize, since any computation path is only considered qualitatively, and termination is a Boolean predicate on programs: any nondeterministic program either terminates—in must or may sense—or does not. In probabilistic programs, on the other hand, any terminating computation path is attributed a

This work is partially supported by the ANR projects 12IS02001 PACE and 14CE250005 ELICA.

Authors' addresses: U. Dal Lago, University of Bologna & INRIA Sophia Antipolis, Dipartimento di Informatica - Scienza e Ingegneria, Mura Anteo Zamboni, 7, 40127 Bologna, Italy; email: ugo.dallago@unibo.it. C. Grellois, INRIA Sophia Antipolis; email: charles.grellois@inria.fr.

0164-0925/2019/03-ART10 $15.00
https://doi.org/10.1145/3293605

probability, and thus termination becomes a *quantitative* property. It is therefore natural to consider a program terminating when its terminating paths form a set of measure one or, equivalently, when it terminates with maximal probability. This is dubbed "almost sure termination" (AST for short) in the literature [9], and many techniques for automatically and semiautomatically checking programs for AST have been introduced in the last years [13, 14, 22, 23]. All of them, however, focus on imperative programs. While probabilistic functional programming languages are nowadays among the most successful ones in the realm of probabilistic programming [25], it is not clear at all whether the existing techniques for imperative languages could be easily applied to functional ones, especially when higher-order functions are involved.

In this article, we introduce a system of monadic affine sized types for a simple probabilistic λ-calculus with recursion and show that it guarantees the AST property for all typable programs. The type system, described in Section 4, can be seen as a nontrivial variation on the sized types of Hughes et al. [27], whose main novelties are the following:

- Types are generalized so as to be *monadic*, this way encapsulating the kind of information we need to type nontrivial examples. This information, in particular, is taken advantage of when typing recursive programs.
- Typing rules are *affine*: higher-order variables cannot be freely duplicated. This is quite similar to what happens when characterizing polynomial-time functions by restricting higher-order languages akin to the λ-calculus [15, 26, 39]. Without affinity, the type system is bound to be unsound for AST, as explained on page 6.

The necessity of both these variations is discussed in Section 2 below. The main result of this article is that typability in monadic affine sized types entails AST, a property that is proved using an adaptation of the Girard-Tait reducibility technique [24]. This adaptation is technically involved, as it needs substantial modifications to deal with possibly infinite and probabilistic computations. In particular, every reducibility set is parameterized by a real number p, and terms belonging to this set are guaranteed to terminate, but only with probability p. The idea of parameterizing such sets already appears in work by the first author and Hofmann [18], in which a notion of realizability parameterized by resource monoids is considered. These realizability models are, however, studied in relation to linear logic and to the complexity of normalization and do not fit as such to our setting, even if they provided some inspiration. In our approach, the fact that recursively defined terms are AST comes from a continuity argument on this parameter: we can prove, by unfolding such terms, that they terminate with probability p for every $p < 1$, and continuity then allows one to take the limit and deduce that they are AST. This soundness result is technically speaking the main contribution of this article and is described in Section 6.

Versions of This Article. This article extends the ESOP 2017 conference version [30] by the same authors.

1.1 Related Work

Sized types have been originally introduced by Hughes, Pareto, and Sabry [27] in the context of reactive programming. A series of papers by Barthe and colleagues [3, 6, 7] presents sized types in a way similar to the one we will adopt here, although still for a deterministic functional language. Contrary to the other works on sized types, their type system is proved to admit a decidable type inference; see the unpublished tutorial [6]. Abel developed independently of Barthe and colleagues a similar type system featuring size information [1]. These three lines of work allow polymorphism, arbitrary inductive data constructors, and ordinal sizes, so that data such as infinite trees can be manipulated. These three features will be absent from our system in order to focus on the main challenge, i.e., the treatment of probabilistic recursive programs. Another interesting approach is the one of Xi's Dependent ML [41], in which a system of lightweight dependent types allows a more liberal treatment of the notion of size, over which arithmetic or conditional operations may in particular be applied. Termination is ensured by checking during typing that a given metric decreases during recursive calls. This system is well adapted for practical termination checking and can be extended with mutual recursion, inductive types, and polymorphism but does not feature ordinal sizes. See [1] for a detailed comparison of the previously cited systems. Some works along these lines are able to deal with coinductive data, as well as inductive ones [1, 3, 27]. They are related to Amadio and Coupet-Grimal's work on guarded types ensuring productivity of infinite structures such as streams [2]. None of these works deal with probabilistic computation, in particular with almost sure termination.

There has been a lot of interest recently about probabilistic termination as a verification problem in the context of imperative programming [13, 14, 22, 23]. Most of the literature deals, invariably, with some form of while-style language without higher-order functions. A possible approach is to reduce AST for probabilistic programs to the termination of nondeterministic programs [22]. Another one is to extend the concept of the ranking function to the probabilistic case [13, 14, 23]. Bournez and Garnier obtained in this way the notion of the Lyapunov ranking function [8], but such functions capture a notion more restrictive than AST: *positive* almost sure termination, meaning that the program is AST and terminates in expected finite time. To capture AST, the notion of ranking supermartingale [12] has been used. Note that the use of ranking supermartingales allows one to deal with programs that are both probabilistic and nondeterministic [13, 23] and even to reason about programs with real-valued variables [14]. Another but related line of work deals with program logics and about how the latter can be useful in analyzing the termination behavior and expected runtime of imperative probabilistic programs [29, 33, 34].

From a recursion-theoretic point of view, checking (positive) almost-sure termination is harder than checking termination of nonprobabilistic programs, where termination is at least recursively enumerable, although undecidable: in a universal probabilistic imperative programming language, almost-sure termination is Π_2^0 complete, while positive almost-sure termination is Σ_2^0 complete [28].

Some recent works by Cappai, the first author, and Parisen Toldin [11, 19] introduce type systems ensuring that all typable programs can be evaluated in probabilistic *polynomial time*. This is too restrictive for our purposes. On the one hand, we aim at termination, and restricting to polynomial-time algorithms would be an overkill. On the other hand, the above-mentioned type systems guarantee that the length of *all* probabilistic branches are uniformly bounded (by the same polynomial). This would limit the focus to terms in which infinite computations are forbidden, while we want the set of such computations to have probability 0. In fact, the results we present in this article can be seen as a first step toward a type system characterizing average polynomial time, in the style of implicit computational complexity [16].

2 WHY IS MONADIC AFFINE TYPING NECESSARY?

In this section, we justify the design choices that guided us in the development of our type system. As we will see, the nature of AST requires a significant and nontrivial extension of the system of sized types originally introduced to ensure termination in the deterministic case [27].

Sized Types for Deterministic Programs. The simply typed λ-calculus endowed with a typed recursion operator letrec and appropriate constructs for the natural numbers, sometimes called PCF, is already Turing-complete,[1] so that there is no hope to prove it strongly normalizing. Sized types [27]

[1] Indeed, Kleene algebra for partial recursive functions can be easily embedded into PCF and does not even require higher-order recursion nor higher-order copying capabilities.

refine the simply typed system by enriching base types with annotations so as to ensure the termination of any recursive definition. Let us explain the idea of sizes in the simple yet informative case in which the base type is Nat. Sizes are defined by the grammar

$$\mathfrak{s} \quad ::= \quad \mathfrak{i} \mid \infty \mid \widehat{\mathfrak{s}},$$

where $\mathfrak{i}$ is a size variable and $\widehat{\mathfrak{s}}$ is the successor of the size $\mathfrak{s}$—with $\widehat{\infty} = \infty$. These sizes permit one to consider decorations $\mathsf{Nat}^{\mathfrak{s}}$ of the base type Nat, whose elements are natural numbers of size at most $\mathfrak{s}$. The type system ensures that the only constant value of type $\mathsf{Nat}^{\widehat{\mathfrak{i}}}$ is 0, that the only constant values of type $\mathsf{Nat}^{\widehat{\widehat{\mathfrak{i}}}}$ are 0 or $\underline{1} = \mathsf{S}\ 0$, and so on. The type Nat^{∞} is the one of *all* natural numbers and is therefore often denoted as Nat.

The crucial rule of the sized type system, which we present here following Barthe et al. [3], allows one to type recursive definitions as follows:

$$\frac{\Gamma, f\ :\ \mathsf{Nat}^{\mathfrak{i}} \to \sigma \vdash M\ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \sigma\left[\widehat{\mathfrak{i}}/\mathfrak{i}\right] \qquad \mathfrak{i} \text{ pos } \sigma}{\Gamma \vdash \mathsf{letrec}\ f = M\ :\ \mathsf{Nat}^{\mathfrak{s}} \to \sigma\left[\mathfrak{s}/\mathfrak{i}\right]}, \tag{1}$$

where $\mathfrak{i}$ pos σ means that $\mathfrak{i}$ or one of its iterated successors only labels instances of Nat occurring in the positive position (the formal definition is deferred to Figure 4). This typing rule ensures that, to recursively define the function $f = M$, the term M taking an input of size $\widehat{\mathfrak{i}}$ calls f on inputs of *strictly lesser* size $\mathfrak{i}$. This is, for instance, the case when typing the program

$$M_{DBL} \ =\ \mathsf{letrec}\ f = \lambda x.\mathsf{case}\ x\ \mathsf{of}\ \{\, \mathsf{S} \to \lambda y.\mathsf{S}\ \mathsf{S}\ (f\ y) \ \mid\ 0 \to 0\,\},$$

computing recursively the double of an input integer, as the hypothesis of the fix-point rule in a typing derivation of M_{DBL} is

$$f\ :\ \mathsf{Nat}^{\mathfrak{i}} \to \mathsf{Nat} \vdash \lambda x.\mathsf{case}\ x\ \mathsf{of}\ \{\, \mathsf{S} \to \lambda y.\mathsf{S}\ \mathsf{S}\ (f\ y)\ \mid\ 0 \to 0\,\}\ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \mathsf{Nat}.$$

The fact that f is called on an input y of strictly lesser size $\mathfrak{i}$ is ensured by the rule typing the case construction:

$$\frac{\Gamma \vdash x\ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \qquad \Gamma \vdash \lambda y.\mathsf{S}\ \mathsf{S}\ (f\ y)\ :\ \mathsf{Nat}^{\mathfrak{i}} \to \mathsf{Nat} \qquad \Gamma \vdash 0\ :\ \mathsf{Nat}}{\Gamma \vdash \mathsf{case}\ x\ \mathsf{of}\ \{\, \mathsf{S} \to \lambda y.\mathsf{S}\ \mathsf{S}\ (f\ y)\ \mid\ 0 \to 0\,\}\ :\ \mathsf{Nat}},$$

where $\Gamma = f\ :\ \mathsf{Nat}^{\mathfrak{i}} \to \mathsf{Nat},\ \ x\ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}}$. The soundness of sized types for strong normalization allows one to conclude that M_{DBL} is indeed SN.

A Naïve Generalization to Probabilistic Terms. The aim of this article is to obtain a probabilistic, *quantitative* counterpart to this soundness result for sized types. Note that unlike the result for sized types, which was focusing on *all* reduction strategies of terms, we only consider a *call-by-value* calculus.[2] Terms can now contain a probabilistic choice operator $\oplus_p$, such that $M \oplus_p N$ reduces to the term M with probability $p \in \mathbb{R}_{[0,1]}$, and to N with probability $1-p$. The language and its operational semantics will be defined more extensively in Section 3. Suppose for the moment that we type the choice operator in a naïve way:

$$\text{Choice} \quad \frac{\Gamma\ \vdash\ M\ :\ \sigma \qquad \Gamma\ \vdash\ N\ :\ \sigma}{\Gamma\ \vdash\ M \oplus_p N\ :\ \sigma}.$$

[2] Choosing a reduction strategy is crucial in a probabilistic setting; otherwise, one risks getting nasty forms of nonconfluence [20].

Since the original system of sized types features subtyping, it allows some flexibility to “unify” the types of M and N to σ. However, it is easy to realize that this approach is too naïve: *all* probabilistic executions would have to be terminating, without any hope of capturing interesting AST programs. Indeed, nothing has been done to capture the *quantitative* nature of probabilistic termination. An instance of a term that is not strongly normalizing but is almost-surely terminating—meaning that it normalizes with probability 1—is

$$M_{BIAS} = \Big(\mathsf{letrec}\ f = \lambda x.\mathsf{case}\ x\ \mathsf{of}\ \Big\{\mathsf{S} \to \lambda y.f(y) \oplus_{\frac{2}{3}} (f(\mathsf{S}\,\mathsf{S}\,y))) \ \Big|\ 0 \to 0\Big\}\Big)\ \underline{n}, \tag{2}$$

simulating a *biased random walk*, which, on $x = m + 1$, goes to m with probability $\frac{2}{3}$ and to $m + 2$ with probability $\frac{1}{3}$. The naïve generalization of the sized type system only allows us to type the body of the recursive definition as follows:

$$f\ :\ \mathsf{Nat}^{\widehat{\widehat{\mathfrak{i}}}} \to \mathsf{Nat}^{\infty}\ \vdash\ \lambda y.f(y) \oplus_{\frac{2}{3}} (f(\mathsf{S}\,\mathsf{S}\,y)))\ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \mathsf{Nat}^{\infty} \tag{3}$$

and thus does *not* allow us to deduce any relevant information on the *quantitative* termination of this term: nothing tells us that the recursive call $f(\mathsf{S}\,\mathsf{S}\,y)$ is performed with a relatively low probability.

A Monadic Type System. Along the evaluation of M_{BIAS}, there is *indeed* a quantity that decreases during each recursive call to the function f: the *average* size of the input on which the call is performed. Indeed, on an input of size $\widehat{\mathfrak{i}}$, f calls itself on an input of smaller size $\mathfrak{i}$ with probability $\frac{2}{3}$ and on an input of greater size $\widehat{\widehat{\mathfrak{i}}}$ with probability only $\frac{1}{3}$. To capture such a relevant *quantitative* information on the recursive calls of f, and with the aim to capture almost-sure termination, we introduce a *monadic* type system, in which *distributions of types* can be used to type in a finer way the functions to be used recursively. In a sense, then, the distribution monad is not only applied to the operational semantics but also to types themselves. Contexts $\Gamma \mid \Theta$ will be generated by a context Γ attributing sized types to any number of variables, while Θ will attribute a *distribution* of sized types to at most *one* variable—typically the one we want to use to recursively define a function. Terms themselves will be typed by a distribution type, formed by combining the Dirac distributions of types introduced in the axiom rules using the following rule for probabilistic choice:

$$\text{Choice}\quad \frac{\Gamma \mid \Theta\ \vdash\ M\ :\ \mu \qquad \Gamma \mid \Psi\ \vdash\ N\ :\ \nu \qquad \langle \mu \rangle\ =\ \langle \nu \rangle}{\Gamma \mid \Theta \oplus_p \Psi\ \vdash\ M \oplus_p N\ :\ \mu \oplus_p \nu}.$$

The guard condition $\langle \mu \rangle = \langle \nu \rangle$ ensures that μ and ν are distributions of types decorating the *same* simple type. Without this condition, there is no hope to aim for a decidable type inference algorithm.

The Fix-Point Rule. Using these monadic types, instead of the insufficiently informative typing in Equation (3), we can derive the sequent

$$f\ :\ \left\{\left(\mathsf{Nat}^{\mathfrak{i}} \to \mathsf{Nat}^{\infty}\right)^{\frac{2}{3}},\ \left(\mathsf{Nat}^{\widehat{\widehat{\mathfrak{i}}}} \to \mathsf{Nat}^{\infty}\right)^{\frac{1}{3}}\right\}\ \vdash\ \lambda y.f(y) \oplus_{\frac{2}{3}} (f(\mathsf{S}\,\mathsf{S}\,y)))\ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \mathsf{Nat}^{\infty}, \tag{4}$$

in which the type of f contains finer information on the sizes of arguments over which it is called recursively, and with which probability. This information enables us to perform a first switch from a qualitative to a quantitative notion of termination: we will adapt the hypothesis

$$\Gamma, f\ :\ \mathsf{Nat}^{\mathfrak{i}} \to \sigma \vdash M\ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \sigma\left[\widehat{\mathfrak{i}}/\mathfrak{i}\right] \tag{5}$$

of the original fix rule (Equation (1)) of sized types, expressing that f is called on an argument of size one less than the one on which M is called, to a condition meaning that there is probability 1 to call f on arguments of a lesser size *after enough iterations of recursive calls*. We therefore define a random walk associated to the distribution type μ of f, the *sized walk* associated to μ, and which is as follows for the typing (Equation (4)):

- The random walk starts on 1, corresponding to the size $\widehat{\mathfrak{i}}$.
- On an integer $n+1$, the random walk jumps to n with probability $\frac{2}{3}$ and to $n+2$ with probability $\frac{1}{3}$.
- 0 is stationary: on it, the random walk loops.

This random walk—as all sized walks will be—is an instance of the *one-counter Markov decision process* [10], so that it is decidable in polynomial time whether the walk reaches 0 with probability 1. We will therefore replace the hypothesis in Equation (5) of the letrec rule by the quantitative counterpart we just sketched, obtaining

$$\mathsf{letrec}\ \dfrac{\begin{array}{c}\left\{ \left(\mathsf{Nat}^{\mathfrak{s}_j} \to \nu\left[\mathfrak{s}_j/\mathfrak{i}\right]\right)^{p_j} \;\middle|\; j \in \mathcal{J} \right\} \text{induces an AST sized walk} \\ \Gamma \mid f \;:\; \left\{ \left(\mathsf{Nat}^{\mathfrak{s}_j} \to \nu\left[\mathfrak{s}_j/\mathfrak{i}\right]\right)^{p_j} \;\middle|\; j \;\in\; \mathcal{J} \right\} \vdash V \;:\; \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \nu\left[\widehat{\mathfrak{i}}/\mathfrak{i}\right]\end{array}}{\Gamma,\, \Delta \mid \Theta \;\vdash\; \mathsf{letrec}\ f \;=\; V \;:\; \mathsf{Nat}^{\mathfrak{r}} \to \nu\left[\mathfrak{r}/\mathfrak{i}\right]},$$

where we omit two additional technical conditions to be found in Section 4 and which justify the weakening on contexts incorporated to this rule. The resulting type system allows one to type a variety of examples, among which is the following program computing the geometric distribution over the natural numbers:

$$M_{EXP} = \left(\mathsf{letrec}\ f = \lambda x.x \oplus_{\frac{1}{2}} \mathsf{S}\ (f\ x)\right)\ 0, \tag{6}$$

and for which the decreasing quantity is the size of the set of probabilistic branches of the term making recursive calls to f. Another example is the unbiased random walk:

$$M_{UNB} = \left(\mathsf{letrec}\ f = \lambda x.\mathsf{case}\ x\ \mathsf{of}\ \left\{ \mathsf{S} \to \lambda y.f(y) \oplus_{\frac{1}{2}} (f(\mathsf{S}\,\mathsf{S}\,y))) \;\middle|\; 0 \to 0 \right\}\right)\ \underline{n}, \tag{7}$$

for which there is no clear notion of decreasing measure during recursive calls but which yet terminates almost surely, as witnessed by the sized walk associated to an appropriate derivation in the sized type system. We therefore claim that the use of this external guard condition on associated sized walks, allowing us to give a general condition of termination, is satisfying as it both captures an interesting class of examples and is decidable in polynomial time.

In Section 6, we prove that this shift from a qualitative to a quantitative hypothesis in the type system results in a shift from the soundness for strong normalization of the original sized type system to a soundness for its quantitative counterpart: *almost-sure termination*. There is a price to pay, however: proving soundness turns out to be significantly more complicated than in the deterministic setting, as we will show in Section 6.

Why Affinity? To ensure the soundness of the letrec rule, we need one more structural restriction on the type system. For the sized walk argument to be adequate, we must ensure that the recursive calls of f are indeed precisely modeled by the sized walk. This is not the case when considering, for instance, the following term:

$$M_{NAFF} \;=\; \left(\mathsf{letrec}\ f = \lambda x.\mathsf{case}\ x\ \mathsf{of}\ \left\{ \mathsf{S} \to \lambda y.f(y) \oplus_{\frac{2}{3}} (f(\mathsf{S}\,\mathsf{S}\,y)\,;\, f(\mathsf{S}\,\mathsf{S}\,y)) \;\middle|\; 0 \to 0 \right\}\right)\ \underline{n}, \tag{8}$$

where the sequential composition ; is defined in this call-by-value calculus as

$$M\,;\,N = (\lambda x.\lambda y.0)\ M\,N.$$

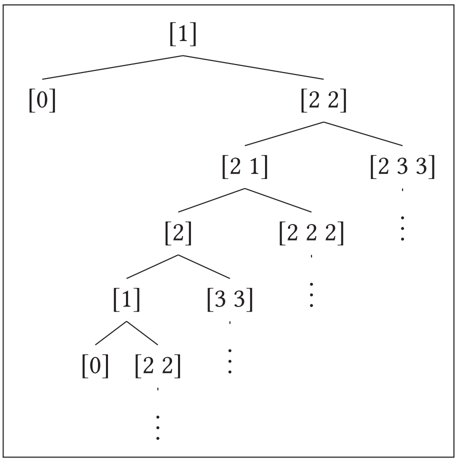

Fig. 1. A tree of recursive calls.

Note that M_{NAFF} calls recursively f *twice* in the right branch of its probabilistic choice and is not therefore modeled appropriately by the sized walk associated to its type. In fact, we would need a generalized notion of random walk to model the recursive calls of this process; it would be a random walk on *stacks* of integers. In the case where $n = 1$, the recursive calls to f can indeed be represented by a tree of stacks as depicted in Figure 1, where the leftmost edges have probability $\frac{2}{3}$ and the rightmost ones $\frac{1}{3}$.

The root indicates that the first call on f was on the integer 1. From it, there is either a call of f on 0 that terminates or *two* calls on 2 that are put into a stack of calls, and so on. We could prove that, without the *affine* restriction we are about to formulate, the term M_{NAFF} is typable with monadic sized types and the fixpoint rule we just designed. However, this term is not almost-surely terminating. Notice, indeed, that the *sum* of the integers appearing in a stack labeling a node of the tree in Figure 1 decreases by 1 when the left edge of probability $\frac{2}{3}$ is taken and increases by *at least* 3 when the right edge of probability $\frac{1}{3}$ is taken. It follows that the expected increase of the sum of the elements of the stack during one step is at least $-1 \times \frac{2}{3} + 3 \times \frac{1}{3} = \frac{1}{3} > 0$. This implies that the probability that f is called on an input of size 0 after enough iterations is strictly less than 1, so that the term M_{NAFF} cannot be almost surely terminating.

Such general random processes have stacks as states and are rather complex to analyze. To the best of the authors' knowledge, they do not seem to have been considered in the literature. We also believe that the complexity of determining whether 0 can be reached almost surely in such a process, if decidable, would be very high. This leads us to the design of an *affine* type system, in which the management of contexts ensures that a given probabilistic branch of a term may only use at most once a given higher-order symbol. We do not, however, formulate restrictions on variables of simple type Nat, as affinity is only used on the letrec rule and thus on higher-order symbols. This is in the spirit of certain systems from implicit computational complexity [15, 26].

Another restriction imposed by this reduction of almost-sure termination checking for higher-order programs to almost-sure termination checking for one-counter Markov decision processes is the fact that we do not allow a general form of nested recursion. This restriction is encoded in the system by allowing *at most* one variable to have a distribution of types in the context. It follows that programs making use of mutual recursion cannot be typed in this system.

3 A SIMPLE PROBABILISTIC FUNCTIONAL PROGRAMMING LANGUAGE

We consider the language $\lambda_\oplus$, which is an extension of the λ-calculus with recursion, constructors for the natural numbers, and a choice operator. In this section, we introduce this language and its operational semantics and use them to define the crucial notion of *almost-sure termination.*

Terms and Values. Given a set of variables $\mathcal{X}$, terms and values of the language $\lambda_\oplus$ are defined by mutual induction as follows:

$$\begin{array}{llcl} \text{Terms:} & M, N, \ldots & ::= & V \mid V\,W \mid \mathsf{let}\ x = M\ \mathsf{in}\ N \mid M \oplus_p N \\ & & & \mid \mathsf{case}\ V\ \mathsf{of}\ \{\, \mathsf{S} \to W \mid 0 \to Z \,\} \\ \\ \text{Values:} & V, W, Z, \ldots & ::= & x \mid 0 \mid \mathsf{S}\,V \mid \lambda x.M \mid \mathsf{letrec}\ f = V, \end{array}$$

where $x,\ f \in \mathcal{X}$, and $p \in]0,1[$ are rational. When $p = \frac{1}{2}$, we often write $\oplus$ as a shorthand for $\oplus_{\frac{1}{2}}$. The set of terms is denoted $\Lambda_\oplus$ and the set of values is denoted $\Lambda_\oplus^V$. Terms of the calculus are assumed to be in A-normal form [37]. This allows one to formulate crucial definitions in a simpler way, concentrating in the let construct the study of the probabilistic behavior of terms. We claim that all traditional constructions can be encoded in this formalism. For instance, the usual application $M\,N$ of two terms can be harmlessly recovered via the encoding $\mathsf{let}\ x = M\ \mathsf{in}\ (\mathsf{let}\ y = N\ \mathsf{in}\ x\,y)$. In the sequel, we write $c\,\overrightarrow{V}$ when a value may be either 0 or of the shape $\mathsf{S}\,V$.

Beyond Probabilistic Choice. The only operator in $\lambda_\oplus$ exhibiting a genuinely probabilistic behavior is $\oplus_p$, whose evaluation corresponds to flipping a biased coin. One may wonder whether this is a too limited form of probabilistic behavior. First of all, there is a large class of distributions D such that there is a term M_D in our language modeling sampling from D. For example, one could do so for the geometric distribution. In fact, we claim that all computable distributions on the natural numbers can be captured this way [20]. This of course does not mean that one could have continuous distributions in our language: $\lambda_\oplus$ is discrete. Extending the language with a set of parameterized discrete distribution symbols would not be too complicated and would not affect our type system nor our termination result.

Term Distributions. The introduction of a probabilistic choice operator in the syntax leads to a *probabilistic* reduction relation. It is therefore meaningful to consider the (operational) semantics of a term as a *distribution* of values modeling the outcome of *all* of the finite probabilistic reduction paths of the term. For instance, the term M_{EXP} defined in Equation (6) evaluates to the term distribution assigning probability $\frac{1}{2^{n+1}}$ to the value $\underline{n}$. Let us define this notion more formally:

Definition 3.1 (Distribution). A *distribution*[3] on a set X is a function $\mathscr{D} : X \to [0,1]$, which is strictly positive only on a countable subset of its domain, and that satisfies the constraint $\sum \mathscr{D} = \sum_{x \in X} \mathscr{D}(x) \leq 1$, where $\sum \mathscr{D}$ is called the *sum* of the distribution $\mathscr{D}$. We say that $\mathscr{D}$ is *proper* precisely when $\sum \mathscr{D} = 1$. We denote by $\mathcal{P}_X$ the set of all distributions over X whether they are proper or not. We often simply write $\mathcal{P}$ for $\mathcal{P}_X$ when X is clear from the context. We define the *support* $\mathsf{S}(\mathscr{D})$ of a distribution $\mathscr{D}$ as $\mathsf{S}(\mathscr{D}) = \{x \in X \mid \mathscr{D}(x) > 0\}$. When $\mathsf{S}(\mathscr{D})$ consists only of closed terms, we say that $\mathscr{D}$ is a *closed* distribution. When it is finite, we say that $\mathscr{D}$ is a *finite* distribution. We call *Dirac* a proper distribution $\mathscr{D}$ such that $\mathsf{S}(\mathscr{D})$ is a singleton. We denote by 0 the null distribution, mapping every term to the probability 0.

[3] What we are defining here is usually called a *sub*distribution in the literature. Since this concept is used in such a widespread way in this article, we prefer to stick to the less baroque terminology.

When $X = \Lambda_\oplus$, we say that $\mathscr{D}$ is a *term distribution*. In the sequel, we will use a more practical notion of *representation* of distributions, which enumerates the terms with their probabilities as a family of assignments. For technical reasons, notably related to the subject reduction property, we will also need *pseudo-representations*, which are essentially multiset-like decompositions of the representation of a distribution.

Definition 3.2 (Representations and Pseudo-Representations). Let $\mathscr{D} \in \mathcal{P}$ be of support $\{x_i \mid i \in \mathcal{I}\}$, where $x_i = x_j$ implies $i = j$ for every $i, j \in \mathcal{I}$. The *representation* of $\mathscr{D}$ is the set $\mathscr{D} = \{\, x_i^{\mathscr{D}(x_i)} \mid i \in \mathcal{I}\,\}$, where $x_i^{\mathscr{D}(x_i)}$ is just an intuitive way to write the pair $(x_i, \mathscr{D}(x_i))$. A *pseudo-representation* of $\mathscr{D}$ is any multiset $[\, y_j^{p_j} \mid j \in \mathcal{J}\,]$ such that

$$\forall j \in \mathcal{J},\ \ y_j \in \mathrm{S}(\mathscr{D}) \qquad\qquad \forall i \in \mathcal{I},\ \ \mathscr{D}(x_i) = \sum_{y_j = x_i} p_j.$$

By abuse of notation, we will simply write $\mathscr{D} = [\, y_j^{p_j} \ \mid\ j \in \mathcal{J}\,]$ to mean that $[\, y_j^{p_j} \ \mid\ j \in \mathcal{J}\,]$ is a pseudo-representation of $\mathscr{D}$. Any distribution has a canonical pseudo-representation obtained by simply replacing the set-theoretic notation with the multiset-theoretic one and keeping the underlying index set unchanged.

Definition 3.3 (ω-CPO of distributions). We define the pointwise order on distributions over X as

$$\mathscr{D} \preccurlyeq \mathscr{E} \qquad \text{if and only if} \qquad \forall x \in X, \quad \mathscr{D}(x) \leq \mathscr{E}(x).$$

This turns $(\mathcal{P}, \preccurlyeq)$ into a partial order. This partial order is an ω-CPO but not a lattice as the join of two distributions does not necessarily exist. The bottom element of this ω-CPO is the null distribution 0.

Definition 3.4 (Operations on Distributions). Given a distribution $\mathscr{D}$ and a real number $\alpha \leq 1$, we define the distribution $\alpha \cdot \mathscr{D}$ as $x \mapsto \alpha \cdot \mathscr{D}(x)$. We similarly define the sum $\mathscr{D} + \mathscr{E}$ of two distributions over a same set X as the function $x \mapsto \mathscr{D}(x) + \mathscr{E}(x)$. Note that this is a total operation on functions $X \to \mathbb{R}$ but a partial operation on distributions: it is defined if and only if $\sum \mathscr{D} + \sum \mathscr{E} \leq 1$. When $\mathscr{D} \preccurlyeq \mathscr{E}$, we define the partial operation of the difference of distributions $\mathscr{E} - \mathscr{D}$ as the function $V \mapsto \mathscr{E}(V) - \mathscr{D}(V)$. We naturally extend these operations to representations and pseudo-representations of distributions.

Definition 3.5 (Value Decomposition of a Term Distribution). Let $\mathscr{D}$ be a term distribution. We write its *value decomposition* as $\mathscr{D} \overset{VD}{=} \mathscr{D}_{|V} + \mathscr{D}_{|T}$, where $\mathscr{D}_{|V}$ is the subdistribution of $\mathscr{D}$ whose support consists of all the values of $\mathrm{S}(\mathscr{D})$, and $\mathscr{D}_{|T} = \mathscr{D} - \mathscr{D}_{|V}$ is the subdistribution whose support is the “nonvalues” contained in $\mathrm{S}(\mathscr{D})$. Both in $\mathscr{D}_{|V}$ and in $\mathscr{D}_{|T}$, every element appears with the same probability it has in $\mathscr{D}$.

Operational Semantics. The semantics of a term will be the value distribution to which it reduces via the probabilistic reduction relation, iterated up to the limit. As a first step, we define the call-by-value reduction relation $\rightarrow_v \subseteq \mathcal{P} \times \mathbb{R}^{\Lambda_\oplus}$ on Figure 2. The relation $\rightarrow_v$ is in fact a relation on distributions:

Lemma 3.6. *Let $\mathscr{D}$ be a distribution such that $\mathscr{D} \rightarrow_v \mathscr{E}$. Then $\mathscr{E}$ is a distribution.*

Note that we write Dirac distributions simply as terms on the left side of $\rightarrow_v$, to improve readability. As usual, we denote by $\rightarrow_v^n$ the nth iterate of the relation $\rightarrow_v$, with $\rightarrow_v^0$ being the identity relation. We then define the relation $\Rrightarrow_v^n$ as follows. Let $\mathscr{D} \rightarrow_v^n \mathscr{E} \overset{VD}{=} \mathscr{E}_{|V} + \mathscr{E}_{|T}$. Then $\mathscr{D} \Rrightarrow_v^n \mathscr{E}_{|V}$. Note that, for every $n \in \mathbb{N}$ and $\mathscr{D} \in \mathcal{P}$, there is a unique distribution $\mathscr{E}$ such that $\mathscr{D} \rightarrow_v^n \mathscr{E}$. Moreover, $\mathscr{E}_{|V}$ is the only distribution such that $\mathscr{D} \Rrightarrow_v^n \mathscr{E}_{|V}$.

$$\frac{}{\text{let } x = V \text{ in } M \ \to_v \ \left\{ (M[V/x])^1 \right\}} \qquad \frac{}{(\lambda x.M)\ V \ \to_v \ \left\{ (M[V/x])^1 \right\}}$$

$$\frac{}{M \oplus_p N \ \to_v \ \left\{ M^p, N^{1-p} \right\}}$$

$$\frac{M \ \to_v \ \left\{ L_i^{p_i} \ \middle|\ i \in \mathcal{I} \right\}}{\text{let } x = M \text{ in } N \ \to_v \ \left\{ (\text{let } x = L_i \text{ in } N)^{p_i} \ \middle|\ i \in \mathcal{I} \right\}}$$

$$\frac{}{\text{case S } V \text{ of } \{ \mathsf{S} \to W \mid 0 \to Z \} \ \to_v \ \left\{ (W\ V)^1 \right\}}$$

$$\frac{}{\text{case } 0 \text{ of } \{ \mathsf{S} \to W \mid 0 \to Z \} \ \to_v \ \left\{ (Z)^1 \right\}}$$

$$\frac{}{(\text{letrec } f = V) \left(c\, \overrightarrow{W}\right) \ \to_v \ \left\{ \left(V\,[(\text{letrec } f = V)/f] \left(c\, \overrightarrow{W}\right)\right)^1 \right\}}$$

$$\frac{\mathscr{D} \overset{VD}{=} \mathscr{D}_{|V} + \left\{ M_j^{p_j} \ \middle|\ j \in \mathcal{J} \right\} \qquad \forall j \in \mathcal{J},\ M_j \ \to_v \ \mathscr{E}_j}{\mathscr{D} \ \to_v \ \left(\sum_{j \in \mathcal{J}} p_j \cdot \mathscr{E}_j\right) + \mathscr{D}_{|V}}$$

Fig. 2. Call-by-value reduction relation $\to_v$ on distributions.

Lemma 3.7. *Let $n, m \in \mathbb{N}$ with $n < m$. Let $\mathscr{D}_n$ ($\mathscr{D}_m$, respectively) be the distribution such that $M \to_v^n \mathscr{D}_n$ ($M \to_v^m \mathscr{D}_m$, respectively). Then $\mathscr{D}_n \preccurlyeq \mathscr{D}_m$.*

Lemma 3.8. *Let $n, m \in \mathbb{N}$ with $n < m$. Let $\mathscr{D}_n$ ($\mathscr{D}_m$, respectively) be the distribution such that $M \Rrightarrow_v^n \mathscr{D}_n$ ($M \Rrightarrow_v^m \mathscr{D}_m$, respectively). Then $\mathscr{D}_n \preccurlyeq \mathscr{D}_m$.*

Definition 3.9 (Semantics of a Term, of a Distribution). The semantics of a distribution $\mathscr{D}$ is the distribution $[\![\mathscr{D}]\!] = \sup_{n \in \mathbb{N}}(\{\mathscr{D}_n \mid \mathscr{D} \Rrightarrow_v^n \mathscr{D}_n\})$. This supremum exists thanks to Lemma 3.8, combined with the fact that $(\mathcal{P}, \preccurlyeq)$ is an ω-CPO. We define the semantics of a term M as $[\![M]\!] = [\![\{ M^1 \}]\!]$.

Corollary 3.10. *Let $n \in \mathbb{N}$ and $\mathscr{D}_n$ be such that $M \Rrightarrow_v^n \mathscr{D}_n$. Then $\mathscr{D}_n \preccurlyeq [\![M]\!]$.*

We now have all the ingredients required to define the central concept of this article, the one of the almost-surely terminating term:

Definition 3.11 (Almost-Sure Termination). We say that a term M is *almost-surely terminating* precisely when $\sum [\![M]\!] = 1$.

Before introducing typing, let us formulate the following lemma on the operational semantics of the let construction, which will be used in the proof of typing soundness for monadic affine sized types:

Lemma 3.12. *Suppose that $M \Rrightarrow_v^n [\, V^{p_i} \mid i \in \mathcal{I} \,]$ and that, for every $i \in \mathcal{I}$, $N[V_i/x] \Rrightarrow_v^m \mathscr{E}_i$. Then* let $x = M$ in $N \Rrightarrow_v^{n+m+1} \sum_{i \in \mathcal{I}} p_i \cdot \mathscr{E}_i$.

What the lemma tells us is that the expected big-step evaluation rule for the let operator and the relation $\Rrightarrow_v^n$ is indeed derivable.

4 MONADIC AFFINE SIZED TYPING

Following the discussion from Section 2, we introduce in this section a nontrivial lifting of sized types to our probabilistic setting. As a first step, we design an *affine* simple type system for $\lambda_\oplus$. This means that no higher-order variable may be used more than once in the same probabilistic branch. However, variables of base type Nat may be used freely. In spite of this restriction, the resulting system allows one to type terms corresponding to any probabilistic Turing machine. In Section 4.2, we introduce a more sophisticated type system, which will be *monadic* and affine, and which will be sound for almost-sure termination as we prove in Section 6.

4.1 Affine Simple Types for $\lambda_\oplus$

The terms of the language $\lambda_\oplus$ can be typed using a variant of the simple types of the λ-calculus, extended to type letrec and $\oplus_p$, but also restricted to an *affine* management of contexts. Recall that the constraint of affinity ensures that a given higher-order symbol is used at most *once* in a probabilistic branch. We define simple types over the base type Nat in the usual way: $\kappa, \kappa', \ldots ::= \mathsf{Nat} \mid \kappa \rightarrow \kappa'$, where, by convention, the arrow associates to the right. Contexts $\Gamma, \Delta, \ldots$ are sequences of simply typed variables $x :: \kappa$. We write sequents as $\Gamma \vdash M :: \kappa$ to distinguish these sequents from the ones using distribution types appearing later in this section. Before giving the rules of the type system, we need to define an affine policy for contracting contexts.

Affine Context Contraction. The affine contraction Γ, Δ is partially defined as follows:

- $x :: \kappa \in \Gamma \setminus \Delta \;\Rightarrow\; x :: \kappa \in \Gamma, \Delta$;
- $x :: \kappa \in \Delta \setminus \Gamma \;\Rightarrow\; x :: \kappa \in \Gamma, \Delta$; and
- if $x :: \kappa \in \Gamma$ and $x :: \kappa' \in \Delta$,
  —if $\kappa = \kappa' = \mathsf{Nat}$, $x :: \kappa \in \Gamma, \Delta$;
  —in any other case, the operation is undefined.

As we explained earlier, only variables of base type Nat may be contracted.

The Affine Type System. The affine simple type system is then defined in Figure 3. All the rules are quite standard. Higher-order variables can occur at most once in any probabilistic branch because all binary typing rules—except probabilistic choice—treat contexts affinely. We set $\Lambda_\oplus^V(\Gamma, \kappa) = \{V \in \Lambda_\oplus^V \mid \Gamma \vdash V :: \kappa\}$ and $\Lambda_\oplus(\Gamma, \kappa) = \{M \in \Lambda_\oplus \mid \Gamma \vdash M :: \kappa\}$. We simply write $\Lambda_\oplus^V(\kappa) = \Lambda_\oplus^V(\emptyset, \kappa)$ and $\Lambda_\oplus(\kappa) = \Lambda_\oplus(\emptyset, \kappa)$ when the terms or values are closed. These closed, typable terms enjoy subject reduction and the progress property.

On the Expressive Power of Affine Typing. The reader may wonder whether affine types represent too much of a constraint themselves, i.e., whether the expressive power of affinely typable $\lambda_\oplus$ terms is too low. Actually, affinely typable terms in $\lambda_\oplus$ can be shown to capture probabilistic Turing machines, following the classic encoding of Kleene's function algebra into PCF. We do not include this result in this article, but the reader can refer to [31] for some results on generalizing basic recursion-theoretical constructions to probabilistic computation.

4.2 Monadic Affine Sized Types for $\lambda_\oplus$

This section is devoted to giving the basic definitions and results about monadic affine sized types (MASTs for short), which can be seen as decorations of the affine simple types with some *size information.*

$$\text{Var}\ \frac{}{\Gamma, x :: \kappa \vdash x :: \kappa} \qquad \frac{\Gamma \vdash V :: \mathsf{Nat}}{\Gamma \vdash \mathsf{S}\, V :: \mathsf{Nat}} \qquad \frac{}{\Gamma \vdash 0 :: \mathsf{Nat}}$$

$$\lambda\ \frac{\Gamma, x :: \kappa \vdash M :: \kappa'}{\Gamma \vdash \lambda x.M :: \kappa \to \kappa'} \qquad \frac{\Gamma \vdash V :: \kappa \to \kappa' \qquad \Delta \vdash W :: \kappa}{\Gamma, \Delta \vdash V\, W :: \kappa'}\ \text{App}$$

$$\text{Choice}\ \frac{\Gamma \vdash M :: \kappa \qquad \Gamma \vdash N :: \kappa}{\Gamma \vdash M \oplus_p N :: \kappa}$$

$$\text{Let}\ \frac{\Gamma \vdash M :: \kappa \qquad \Delta, x :: \kappa \vdash N :: \kappa'}{\Gamma, \Delta \vdash \mathsf{let}\ x = M\ \mathsf{in}\ N :: \kappa'}$$

$$\text{Case}\ \frac{\Gamma \vdash V :: \mathsf{Nat} \qquad \Delta \vdash W :: \mathsf{Nat} \to \kappa \qquad \Delta \vdash Z :: \kappa}{\Gamma, \Delta \vdash \mathsf{case}\ V\ \mathsf{of}\ \{\, \mathsf{S} \to W \mid 0 \to Z \,\} :: \kappa}$$

$$\text{letrec}\ \frac{\Gamma, f :: \mathsf{Nat} \to \kappa \vdash V :: \mathsf{Nat} \to \kappa \qquad \forall x \in \Gamma,\ x :: \mathsf{Nat}}{\Gamma \vdash \mathsf{letrec}\ f = V :: \mathsf{Nat} \to \kappa}$$

Fig. 3. Affine simple types for $\lambda_{\oplus}$.

Sized Types. We consider a set $\mathcal{S}$ of *size variables*, denoted $\mathfrak{i}, \mathfrak{j}, \ldots$, and define *sizes* (called *stages* in [3]) as

$$\mathfrak{s}, \mathfrak{r} \quad ::= \quad \mathfrak{i} \mid \infty \mid \widehat{\mathfrak{s}},$$

where $\widehat{\cdot}$ denotes the *successor* operation. We denote the iterations of $\widehat{\cdot}$ as follows: $\widehat{\widehat{\mathfrak{s}}}$ is denoted $\widehat{\mathfrak{s}}^{2}$, $\widehat{\widehat{\widehat{\mathfrak{s}}}}$ is denoted $\widehat{\mathfrak{s}}^{3}$, and so on. By definition, at most one variable $\mathfrak{i} \in \mathcal{S}$ appears in a given size $\mathfrak{s}$. We call it its *spine variable*, denoted as $\mathrm{spine}(\mathfrak{s})$. We write $\mathrm{spine}(\mathfrak{s}) = \emptyset$ when there is no variable in $\mathfrak{s}$. An order $\preccurlyeq$ on sizes can be defined as follows:

$$\frac{}{\mathfrak{s} \preccurlyeq \mathfrak{s}} \qquad \frac{\mathfrak{s} \preccurlyeq \mathfrak{r} \qquad \mathfrak{r} \preccurlyeq \mathfrak{t}}{\mathfrak{s} \preccurlyeq \mathfrak{t}} \qquad \frac{}{\mathfrak{s} \preccurlyeq \widehat{\mathfrak{s}}} \qquad \frac{}{\mathfrak{s} \preccurlyeq \infty}.$$

Notice that these rules imply notably that $\widehat{\infty}$ is equivalent to ∞, i.e., $\widehat{\infty} \preccurlyeq \infty$ and $\infty \preccurlyeq \widehat{\infty}$. We consider sizes modulo this equivalence. We can now define sized types and distribution types by mutual induction, calling distributions of (sized) types the distributions over the set of sized types:

Definition 4.1 (Sized Types, Distribution Types). Sized types and distribution types are defined by mutual induction contextually with the function $\langle \cdot \rangle$, which maps any sized or distribution type to its *underlying* affine type.

$$\begin{array}{llcl}
\text{Sized types:} & \sigma, \tau & ::= & \sigma \to \mu \mid \mathsf{Nat}^{\mathfrak{s}} \\
\text{Distribution types:} & \mu, \nu & ::= & \{\, \sigma_i^{p_i} \mid i \in \mathcal{I} \,\}, \\
\text{Underlying map:} & \langle \sigma \to \mu \rangle & = & \langle \sigma \rangle \to \langle \mu \rangle \\
 & \langle \mathsf{Nat}^{\mathfrak{s}} \rangle & = & \mathsf{Nat} \\
 & \langle \{\, \sigma_i^{p_i} \mid i \in \mathcal{I} \,\} \rangle & = & \langle \sigma_j \rangle
\end{array}$$

For distribution types we require additionally that $\sum_{i \in \mathcal{I}} p_i \leq 1$, that $\mathcal{I}$ is a finite nonempty set, and that $\langle \sigma_i \rangle = \langle \sigma_j \rangle$ for every $i, j \in \mathcal{I}$. In the last equation, j is any element of $\mathcal{I}$. We write $\sigma :: \kappa$ when $\kappa = \langle \sigma \rangle$.

$$\frac{}{\mathfrak{i}\ \mathrm{pos}\ \mathsf{Nat}^{\mathfrak{s}}} \qquad \frac{\mathfrak{i}\ \mathrm{neg}\ \sigma \qquad \mathfrak{i}\ \mathrm{pos}\ \mu}{\mathfrak{i}\ \mathrm{pos}\ \sigma \to \mu} \qquad \frac{\forall i \in \mathcal{I},\ \ \mathfrak{i}\ \mathrm{pos}\ \sigma_i}{\mathfrak{i}\ \mathrm{pos}\ \left\{ \sigma_i^{p_i} \ \middle|\ i \in \mathcal{I} \right\}}$$

$$\frac{\mathfrak{i} \notin \mathfrak{s}}{\mathfrak{i}\ \mathrm{neg}\ \mathsf{Nat}^{\mathfrak{s}}} \qquad \frac{\mathfrak{i}\ \mathrm{pos}\ \sigma \qquad \mathfrak{i}\ \mathrm{neg}\ \mu}{\mathfrak{i}\ \mathrm{neg}\ \sigma \to \mu} \qquad \frac{\forall i \in \mathcal{I},\ \ \mathfrak{i}\ \mathrm{neg}\ \sigma_i}{\mathfrak{i}\ \mathrm{neg}\ \left\{ \sigma_i^{p_i} \ \middle|\ i \in \mathcal{I} \right\}}$$

Fig. 4. Positive and negative occurrences of a size variable in a sized type and in a distribution type.

The definition of sized types is *monadic* in that a higher-order sized type is of the shape $\sigma \to \mu$, where σ is again a sized type and μ is a *distribution* of sized types. This is, by the way, reminiscent of (and inspired by) Moggi semantics for the computational lambda calculus, in which terms of type $A \to B$ are interpreted as $[\![A]\!] \Rightarrow \Box[\![B]\!]$, where $\Box$ is, indeed, a monad.

The definition of the fix point will refer to the notion of *positivity* of a size variable in a sized or distribution type. We define positive and negative occurrences of a size variable in such a type in Figure 4. The idea is that a size variable is positive when it annotates a base type that is itself in the positive position, and conversely for the negative position.

Contexts and Operations on Them. Contexts are sequences of variables together with a sized type and at most one distinguished variable with a distribution type:

Definition 4.2 (Contexts). Contexts are of the shape $\Gamma \mid \Theta$, with

$$\begin{array}{llcl|ll} \textit{Sized contexts:} & \Gamma, \Delta, \ldots & ::= & \emptyset & x : \sigma, \Gamma & (x \notin \mathit{dom}(\Gamma)) \\ \textit{Distribution contexts:} & \Theta, \Psi, \ldots & ::= & \emptyset & x : \mu. & \end{array}$$

As usual, we define the *domain* $\mathit{dom}(\Gamma)$ of a sized context Γ by induction: $\mathit{dom}(\emptyset) = \emptyset$ and $\mathit{dom}(x : \sigma, \Gamma) = \{x\} \uplus \mathit{dom}(\Gamma)$. We proceed similarly for the domain $\mathit{dom}(\Theta)$ of a distribution context Θ. When a sized context $\Gamma = x_1 : \sigma_1, \ldots, x_n : \sigma_n$ $(n \geq 1)$ is such that there is a simple type κ with $\forall i \in \{1, \ldots, n\},\ \langle \sigma_i \rangle = \kappa$, we say that Γ is *uniform* of simple type κ. We write this as $\langle \Gamma \rangle = \kappa$.

We write Γ, Δ for the *disjoint union* of these sized contexts: it is defined whenever $\mathit{dom}(\Gamma) \cap \mathit{dom}(\Delta) = \emptyset$. We proceed similarly for Θ, Ψ but note that due to the restriction on the cardinality of such contexts, there is the additional requirement that $\Theta = \emptyset$ or $\Psi = \emptyset$.

We finally define *contexts* as pairs $\Gamma \mid \Theta$ of a sized context and of a distribution context, with the constraint that $\mathit{dom}(\Gamma) \cap \mathit{dom}(\Theta) = \emptyset$.

Definition 4.3 (Probabilistic Sum of Distribution Types). Let μ and ν be two distribution types. We define their probabilistic sum $\mu \oplus_p \nu$ as the distribution type $p \cdot \mu + (1 - p) \cdot \nu$.

We extend this operation to a *partial* and n-ary operation on distribution contexts:

Definition 4.4 (Weighted Sum of Distribution Contexts). Let $(\Theta_i)_{i \in \mathcal{I}}$ be a nonempty family of distribution contexts and $(p_i)_{i \in \mathcal{I}}$ be a family of reals in $[0, 1]$. We define the weighted sum $\sum_{i \in \mathcal{I}} p_i \cdot \Theta_i$ as the distribution context $x : \sum_{i \in \mathcal{I}} p_i \cdot \mu_i$ when the following conditions are met:

(1) $\exists x,\ \ \forall i \in \mathcal{I},\ \Theta_i = x : \mu_i$;
(2) $\forall (i, j) \in \mathcal{I}^2,\ \ \langle \Theta_i \rangle = \langle \Theta_j \rangle$; and
(3) and $\sum_{i \in \mathcal{I}} p_i \leq 1$.

In any other case, the operation is undefined.

Definition 4.5 (Substitution of a Size Variable). We define the substitution $\mathfrak{s}[\mathfrak{r}/\mathfrak{i}]$ of a size variable in a size as follows:

$$\mathfrak{i}\,[\mathfrak{r}/\mathfrak{i}] = \mathfrak{r} \qquad \mathfrak{j}\,[\mathfrak{r}/\mathfrak{i}] = \mathfrak{j} \qquad \infty\,[\mathfrak{r}/\mathfrak{i}] = \infty \qquad \widehat{\mathfrak{s}}\,[\mathfrak{r}/\mathfrak{i}] = \widehat{\mathfrak{s}\,[\mathfrak{r}/\mathfrak{i}]},$$

where $\mathfrak{i} \neq \mathfrak{j}$. We then define the substitution $\sigma[\mathfrak{s}/\mathfrak{i}]$ ($\mu[\mathfrak{s}/\mathfrak{i}]$, respectively) of a size variable $\mathfrak{i}$ by a size $\mathfrak{s}$ in a sized or distribution type as

$$(\sigma \to \mu)\,[\mathfrak{s}/\mathfrak{i}] = \sigma\,[\mathfrak{s}/\mathfrak{i}] \to \mu\,[\mathfrak{s}/\mathfrak{i}] \qquad\qquad (\mathsf{Nat}^{\mathfrak{s}})\,[\mathfrak{r}/\mathfrak{i}] \;=\; \mathsf{Nat}^{\mathfrak{s}[\mathfrak{r}/\mathfrak{i}]}$$
$$\left(\left\{\, \sigma_i^{p_i} \;\middle|\; i \in \mathcal{I} \,\right\}\right)[\mathfrak{s}/\mathfrak{i}] = \left\{\, (\sigma_i\,[\mathfrak{s}/\mathfrak{i}])^{p_i} \;\middle|\; i \in \mathcal{I} \,\right\}.$$

We define the substitution of a size variable in a sized or distribution context in the obvious way:

$$\emptyset\,[\mathfrak{s}/\mathfrak{i}] = \emptyset \qquad\qquad (x : \sigma, \Gamma)\,[\mathfrak{s}/\mathfrak{i}] = x : \sigma\,[\mathfrak{s}/\mathfrak{i}]\,, \Gamma\,[\mathfrak{s}/\mathfrak{i}]$$
$$(x : \mu)\,[\mathfrak{s}/\mathfrak{i}] = x : \mu\,[\mathfrak{s}/\mathfrak{i}]\,.$$

The following lemma shows that index substitution properly commutes with weighted sums of types and contexts. It will be very useful in the sequel.

Lemma 4.6.

(1) $(\mu \oplus_p \nu)[\mathfrak{s}/\mathfrak{i}] = \mu[\mathfrak{s}/\mathfrak{i}] \oplus_p \nu[\mathfrak{s}/\mathfrak{i}]$.
(2) For distribution contexts, $(\Theta \oplus_p \Psi)[\mathfrak{s}/\mathfrak{i}] = \Theta[\mathfrak{s}/\mathfrak{i}] \oplus_p \Psi[\mathfrak{s}/\mathfrak{i}]$.
(3) For distribution contexts, $(\sum_{i\in\mathcal{I}}\ p_i \cdot \Gamma_i)[\mathfrak{s}/\mathfrak{i}] = \sum_{i\in\mathcal{I}}\ p_i \cdot \Gamma_i[\mathfrak{s}/\mathfrak{i}]$.

Proof.

(1) Let $\mu = \{\, \sigma_i^{p'_i} \mid i \in \mathcal{I} \,\}$ and $\nu = \{\, \tau_j^{p''_j} \mid j \in \mathcal{J} \,\}$. Then

$$\begin{aligned}
& \mu\,[\mathfrak{s}/\mathfrak{i}] \oplus_p \nu\,[\mathfrak{s}/\mathfrak{i}] \\
=\ & \left\{\, \sigma_i^{p'_i} \;\middle|\; i \in \mathcal{I} \,\right\}[\mathfrak{s}/\mathfrak{i}] \oplus_p \left\{\, \tau_j^{p''_j} \;\middle|\; j \in \mathcal{J} \,\right\}[\mathfrak{s}/\mathfrak{i}] \\
=\ & \left\{\, (\sigma_i\,[\mathfrak{s}/\mathfrak{i}])^{p'_i} \;\middle|\; i \in \mathcal{I} \,\right\} \oplus_p \left\{\, \left(\tau_j\,[\mathfrak{s}/\mathfrak{i}]\right)^{p''_j} \;\middle|\; j \in \mathcal{J} \,\right\} \\
=\ & \left[\, (\sigma_i\,[\mathfrak{s}/\mathfrak{i}])^{pp'_i} \;\middle|\; i \in \mathcal{I} \,\right] + \left[\, \left(\tau_j\,[\mathfrak{s}/\mathfrak{i}]\right)^{(1-p)p''_j} \;\middle|\; j \in \mathcal{J} \,\right] \\
=\ & \left(\left[\, (\sigma_i)^{pp'_i} \;\middle|\; i \in \mathcal{I} \,\right] + \left[\, \left(\tau_j\right)^{(1-p)p''_j} \;\middle|\; j \in \mathcal{J} \,\right]\right)[\mathfrak{s}/\mathfrak{i}] \\
=\ & (\mu \oplus_p \nu)\,[\mathfrak{s}/\mathfrak{i}]\,.
\end{aligned}$$

(2) Suppose that $\Theta = x : \mu$ and that $\Psi = x : \nu$. Then $\Theta \oplus_p \Psi = x : \mu \oplus_p \nu$. It follows from (1) that $\Theta[\mathfrak{s}/\mathfrak{i}] \oplus_p \Psi[\mathfrak{s}/\mathfrak{i}] = x : \mu[\mathfrak{s}/\mathfrak{i}] \oplus_p \nu[\mathfrak{s}/\mathfrak{i}] = x : (\mu \oplus_p \nu)[\mathfrak{s}/\mathfrak{i}] = (\Theta \oplus_p \Psi)[\mathfrak{s}/\mathfrak{i}]$.
(3) The proof is similar to the previous cases. □

A subtyping relation allows us to lift the order $\preccurlyeq$ on sizes to monadic sized types:

Definition 4.7 (Subtyping). We define the subtyping relation $\sqsubseteq$ on sized types and distribution types as follows:

$$\frac{}{\sigma \sqsubseteq \sigma} \qquad \frac{\mathfrak{s} \preccurlyeq \mathfrak{r}}{\mathsf{Nat}^{\mathfrak{s}} \sqsubseteq \mathsf{Nat}^{\mathfrak{r}}} \qquad \frac{\tau \sqsubseteq \sigma \qquad \mu \sqsubseteq \nu}{\sigma \to \mu \;\sqsubseteq\; \tau \to \nu}$$
$$\frac{\exists f : \mathcal{I} \to \mathcal{J},\ \left(\forall i \in \mathcal{I},\ \ \sigma_i \sqsubseteq \tau_{f(i)}\right) \text{ and } \left(\forall j \in \mathcal{J},\ \sum_{i\in f^{-1}(j)}\ p_i \leq q_j\right)}{\left\{\, \sigma_i^{p_i} \;\middle|\; i \in \mathcal{I} \,\right\} \;\sqsubseteq\; \left\{\, \tau_j^{q_j} \;\middle|\; j \in \mathcal{J} \,\right\}}.$$

Sized Walks and Distribution Types. As we explained in Section 2, the rule typing letrecs in the monadic affine type system relies on an external decision procedure, computable in polynomial time. This procedure ensures that the *sized walk*—a particular instance of the *one-counter Markov decision process* (OC-MDP, see [10]), but which does not make use of nondeterminism—associated to the type of the recursive function of interest indeed ensures almost-sure termination. Let us now define the sized walk associated to a distribution type μ. We then make precise the connection with OC-MDPs, from which the decidability (in polynomial time) of the almost-sure termination of the random walks follows.

Definition 4.8 (Sized Walk). Let $\mathcal{I} \subseteq_{fin} \mathbb{N}$ be a finite set of integers. Let $\{p_i\}_{i \in \mathcal{I}}$ be such that $\sum_{i \in \mathcal{I}} p_i \leq 1$. These parameters define a Markov chain whose set of states is $\mathbb{N}$ and whose transition relation is defined as follows:

- The state $0 \in \mathbb{N}$ is stationary (i.e., one goes from 0 to 0 with probability 1).
- From the state $s + 1 \in \mathbb{N}$ one moves:
  —to the state $s + i$ with probability p_i, for every $i \in \mathcal{I}$;
  —to 0 with probability $1 - (\sum_{i \in \mathcal{I}} p_i)$.

We call this Markov chain the *sized walk* on $\mathbb{N}$ associated to $(\mathcal{I}, (p_i)_{i \in \mathcal{I}})$. A sized walk is *almost surely terminating* when it reaches 0 with probability 1 from any initial state.

Notably, checking whether a sized walk is terminating is relatively easy:

Proposition 4.9 (Decidability of AST for Sized Walks). *It is decidable in polynomial time whether a sized walk is AST.*

Proof. See Section 4.3. □

The role of sized walks in our type system is intimately related to recursion, in that types allow one to *reflect* the recursive call structure of the typed term into a distribution *type*, which can then be seen as a sized walk and thus appropriately analyzed with dedicated decision procedures. This is the main idea behind the following definition:

Definition 4.10 (From Types to Sized Walks). Let $\mu = \{ (\mathsf{Nat}^{\mathfrak{s}_j} \to \nu_j)^{p_j} \mid j \in \mathcal{J} \}$ be a distribution type such that $\forall j \in \mathcal{J},\ \mathrm{spine}\,(\mathfrak{s}_j) = \mathfrak{i}$. Then μ induces a sized walk, defined as follows. First, by definition, $\mathfrak{s}_j$ must be of the shape $\widehat{\mathfrak{i}}^{k_j}$ with $k_j \geq 0$ for every $j \in \mathcal{J}$. We set $\mathcal{I} = \{k_j \mid j \in \mathcal{J}\}$ and $q_{k_j} = p_j$ for every $j \in \mathcal{J}$. The sized walk induced by the distribution type μ is then the sized walk associated to $(\mathcal{I}, (q_i)_{i \in \mathcal{I}})$.

Example 4.11. Let $\mu = \{ (\mathsf{Nat}^{\mathfrak{i}} \to \mathsf{Nat}^{\infty})^{\frac{1}{2}},\ (\mathsf{Nat}^{\widehat{\mathfrak{i}}^2} \to \mathsf{Nat}^{\infty})^{\frac{1}{3}} \}$. Then the induced sized walk is the one associated to $(\{0, 2\}, (p_0 = \frac{1}{2}, p_2 = \frac{1}{3}))$. In other words, it is the random walk on $\mathbb{N}$ that is stationary on 0, and that on nonnull integers $i + 1$ moves to i with probability $\frac{1}{2}$, moves to $i + 2$ with probability $\frac{1}{3}$, and jumps to 0 with probability $\frac{1}{6}$. Note that the type μ, and therefore the associated sized walk, models a recursive function that calls itself on a size lesser by one unit with probability $\frac{1}{2}$, calls itself on a size greater by one unit with probability $\frac{1}{3}$, and does not call itself with probability $\frac{1}{6}$.

Typing Rules. Judgments are of the shape $\Gamma \mid \Theta \vdash M : \mu$. When a distribution $\mu = \{ \sigma^1 \}$ is Dirac, we simply write it as σ. The type system is defined in Figure 5. As earlier, we define sets of typable terms and set $\Lambda_{\oplus}^{\mathfrak{s}, V}(\Gamma \mid \Theta, \sigma) = \{V \mid \Gamma \mid \Theta \vdash V : \sigma\}$ and $\Lambda_{\oplus}^{\mathfrak{s}}(\Gamma \mid \Theta, \mu) = \{M \mid \Gamma \mid \Theta \vdash M : \mu\}$. We abbreviate $\Lambda_{\oplus}^{\mathfrak{s}, V}(\emptyset \mid \emptyset, \sigma)$ as $\Lambda_{\oplus}^{\mathfrak{s}, V}(\sigma)$ and $\Lambda_{\oplus}^{\mathfrak{s}}(\emptyset \mid \emptyset, \sigma)$ as $\Lambda_{\oplus}^{\mathfrak{s}}(\sigma)$.

This sized type system is a refinement of the affine simple type system for $\lambda_{\oplus}$: if $x_1 : \sigma_1, \ldots, x_n : \sigma_n \mid f : \mu \vdash M : \nu$, then it is easily checked that $x_1 :: \langle \sigma_1 \rangle, \ldots, x_n :: \langle \sigma_n \rangle, f :: \langle \mu \rangle \vdash M :: \langle \nu \rangle$.

$$\text{Var}\ \frac{}{\Gamma, x : \sigma \,|\, \Theta \vdash x : \sigma} \qquad \frac{}{\Gamma \,|\, x : \sigma \vdash x : \sigma}\ \text{Var'}$$

$$\text{Succ}\ \frac{\Gamma \,|\, \Theta \vdash V : \mathsf{Nat}^{\mathfrak{s}}}{\Gamma \,|\, \Theta \vdash \mathsf{S}\, V : \mathsf{Nat}^{\widehat{\mathfrak{s}}}} \qquad \frac{}{\Gamma \,|\, \Theta \vdash 0 : \mathsf{Nat}^{\widehat{\mathfrak{s}}}}\ \text{Zero}$$

$$\lambda\ \frac{\Gamma, x : \sigma \,|\, \Theta \vdash M : \mu}{\Gamma \,|\, \Theta \vdash \lambda x.M : \sigma \to \mu} \qquad \frac{\Gamma \,|\, \Theta \vdash M : \mu \qquad \mu \sqsubseteq \nu}{\Gamma \,|\, \Theta \vdash M : \nu}\ \text{Sub}$$

$$\text{App}\ \frac{\Gamma, \Delta \,|\, \Theta \vdash V : \sigma \to \mu \qquad \Gamma, \Xi \,|\, \Psi \vdash W : \sigma \qquad \langle \Gamma \rangle = \mathsf{Nat}}{\Gamma, \Delta, \Xi \,|\, \Theta, \Psi \vdash V\, W : \mu}$$

$$\text{Choice}\ \frac{\Gamma \,|\, \Theta \vdash M : \mu \qquad \Gamma \,|\, \Psi \vdash N : \nu \qquad \langle \mu \rangle = \langle \nu \rangle}{\Gamma \,|\, \Theta \oplus_p \Psi \vdash M \oplus_p N : \mu \oplus_p \nu}$$

$$\text{Let}\ \frac{\begin{array}{c}\Gamma, \Delta \,|\, \Theta \vdash M : \left\{ \sigma_i^{p_i} \;\middle|\; i \in \mathcal{I} \right\} \qquad \langle \Gamma \rangle = \mathsf{Nat} \\ \Gamma, \Xi, x : \sigma_i \,|\, \Psi_i \vdash N : \mu_i \quad (\forall i \in \mathcal{I})\end{array}}{\Gamma, \Delta, \Xi \,|\, \Theta, \left(\sum_{i \in \mathcal{I}} p_i \cdot \Psi_i\right) \vdash \mathsf{let}\ x = M\ \mathsf{in}\ N : \sum_{i \in \mathcal{I}} p_i \cdot \mu_i}$$

$$\text{Case}\ \frac{\Gamma \,|\, \emptyset \vdash V : \mathsf{Nat}^{\widehat{\mathfrak{s}}} \qquad \Delta \,|\, \Theta \vdash W : \mathsf{Nat}^{\mathfrak{s}} \to \mu \qquad \Delta \,|\, \Theta \vdash Z : \mu}{\Gamma, \Delta \,|\, \Theta \vdash \mathsf{case}\ V\ \mathsf{of}\ \{ \mathsf{S} \to W \mid 0 \to Z \} : \mu}$$

$$\text{letrec}\ \frac{\begin{array}{c}\langle \Gamma \rangle = \mathsf{Nat} \\ \mathfrak{i} \notin \Gamma \text{ and } \mathfrak{i} \text{ positive in } \nu \text{ and } \forall j \in \mathcal{J},\ \mathrm{spine}\left(\mathfrak{s}_j\right) = \mathfrak{i} \\ \left\{ \left(\mathsf{Nat}^{\mathfrak{s}_j} \to \nu\left[\mathfrak{s}_j/\mathfrak{i}\right]\right)^{p_j} \;\middle|\; j \in \mathcal{J} \right\} \text{induces an AST sized walk} \\ \Gamma \,|\, f : \left\{ \left(\mathsf{Nat}^{\mathfrak{s}_j} \to \nu\left[\mathfrak{s}_j/\mathfrak{i}\right]\right)^{p_j} \;\middle|\; j \in \mathcal{J} \right\} \vdash V : \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \nu\left[\widehat{\mathfrak{i}}/\mathfrak{i}\right]\end{array}}{\Gamma, \Delta \,|\, \Theta \vdash \mathsf{letrec}\ f = V : \mathsf{Nat}^{\mathfrak{r}} \to \nu\left[\mathfrak{r}/\mathfrak{i}\right]}$$

Fig. 5. Affine distribution types for $\lambda_{\oplus}$.

Lemma 4.12 (Properties of Distribution Types).

- $\Gamma \,|\, \Theta \vdash V : \mu \implies \mu$ *is Dirac.*
- $\Gamma \,|\, \Theta \vdash M : \mu \implies \mu$ *is proper.*

Proof. Immediate inspection of the rules. □

4.3 On Sized Walks and Almost-Sure Termination

In this section, we prove Proposition 4.9 by showing how sized walks are a very special sort of one-counter Markov decision process (OC-MDP) and using then a result of [10] to conclude. Please note that in [10] the Markov decision processes are more general, as they allow nondeterminism. They are called OC-MDPs and contain in particular all the deterministic OC-MDPs. We omit this feature in our presentation.

Definition 4.13 (Markov Decision Process). A Markov decision process (MDP) is a tuple $(V, \mapsto, Pr)$ such that V is a finite or countable set of vertices, $\mapsto \subseteq V \times V$ is a total transition relation, and Pr

is a probability assignment mapping each $v \in V$ to a probability distribution associating a rational and nonnull probability to each edge outgoing of v. These distributions are moreover required to sum to 1.

Definition 4.14 (Deterministic One-Counter Markov Decision Process). A *deterministic one-counter Markov decision process* (DOC-MDP) is a tuple $(Q, \delta^{=0}, \delta^{>0}, P^{=0}, P^{>0})$ such that:

- Q is a finite set of states;
- $\delta^{=0} \subseteq Q \times \{0,1\} \times Q$ and $\delta^{>0} \subseteq Q \times \{-1,0,1\} \times Q$ are sets of *zero* and *positive* transitions, satisfying that every $q \in Q$ has at least a zero and a positive outgoing transition;
- $P^{=0}$ ($P^{>0}$, respectively) is a probability assignment mapping every $q \in Q$ to a probability distribution over the outgoing transitions of $\delta^{=0}$ ($\delta^{>0}$, respectively) from q. These distributions are required to attribute a nonnull, rational probability to every outgoing transition, and to sum to 1.

Definition 4.15 (Induced Markov Decision Process). A DOC-MDP $(Q, \delta^{=0}, \delta^{>0}, P^{=0}, P^{>0})$ induces an MDP $(Q \times \mathbb{N}, \mapsto, Pr)$ such that, for $q \in Q$ and $n \in \mathbb{N}$:

- for every state q' such that $(q, m, q') \in \delta^{=0}$, $(q, 0) \mapsto (q', m)$, and the probability of this transition is the one attributed by $P^{=0}(q)$ to the transition (q, m, q');
- for every state q' such that $(q, m, q') \in \delta^{>0}$, $(q, n) \mapsto (q', n + m)$, and the probability of this transition is the one attributed by $P^{>0}(q)$ to the transition (q, m, q').

This MDP is said to *terminate* when it reaches the value counter 0 in *any* state $q \in Q$.

Recall that, by definition, $|m| \leq 1$. This is the only restriction to overcome (using intermediate states) to encode sized walks in DOC-MDPs, so that the MDP they induce coincides with the original sized walk. We will then obtain the result of polynomial-time decidability of termination with probability 1 using the following proposition:

Proposition 4.16 ([10], Theorem 4.1). *It is decidable in polynomial time whether the MDP induced by an OC-MDP—and thus by a DOC-MDP—terminates with probability 1.*

We now encode sized walks as DOC-MDPs:

Definition 4.17 (DOC-MDP Corresponding to a Sized Walk). Consider the sized walk on $\mathbb{N}$ associated to $(\mathcal{I}, (p_i)_{i \in \mathcal{I}})$. We define the corresponding DOC-MDP $(Q, \delta^{=0}, \delta^{>0}, P^{=0}, P^{>0})$ as follows. Let us first consider the following set of states:

$$Q = \{q_\alpha,\ q_{zero}\} \ \cup\ \{q_1,\ \ldots, q_{j-2} \ \mid\ j = \max\{i \in \mathcal{I} \ \mid\ i \geq 2\}\},$$

where q_α is the "main" state of the DOC-MDP and the other ones will be used for encoding purposes. We define the transitions of $\delta^{>0}$ as follows:

- We add the transition $(q_{zero}, -1, q_{zero})$ with probability 1.
- For every $j \in \{2, \ldots, \max\{i \in \mathcal{I} \mid i \geq 2\} - 2\}$, we add the transition $(q_j, 1, q_{j-1})$ with probability 1.
- We add the transition $(q_1, 1, q_\alpha)$ with probability 1.
- For $i \in \mathcal{I} \cap \{0,1,2\}$, we add the transition $(q_\alpha, i - 1, q_\alpha)$ and attribute it with probability p_i.
- For $i \in \mathcal{I} \setminus \{0,1,2\}$, we add the transition $(q_\alpha, 1, q_{i-2})$ and attribute it with probability p_i.
- If $1 - (\sum_{i \in \mathcal{I}} \ p_i) > 0$, we add the transition $(q_\alpha, -1, q_{zero})$ with probability $1 - (\sum_{i \in \mathcal{I}} \ p_i)$.

Finally, we define $\delta^{=0}$ as follows: for every state $q \in Q$, we add the transition $(q, 0, q)$ and attribute it with probability 1.

It is easily checked that, by construction, these DOC-MDPs induce the same Markov decision processes as sized walks:

Proposition 4.18. *The sized walk on $\mathbb{N}$ associated to $(\mathcal{I}, (p_i)_{i \in \mathcal{I}})$ coincides with the induced MDP of the corresponding DOC-MDP.*

This allows us to deduce from the result of [10] the polynomial-time decidability of AST for sized walks:

Corollary 4.19 (Proposition 4.9). *It is decidable in polynomial time whether a sized walk is almost-surely terminating.*

Example 4.20. Recall the sized walk from Example 4.11. The algorithm allows one to decide that it is AST. It follows from the main result of this article, Theorem 6.36, that a program whose recursion is modeled by the sized walk of Example 4.11, such as

$$M_{BIAS} = \Big(\mathsf{letrec}\ f = \lambda x.\mathsf{case}\ x\ \mathsf{of}\ \Big\{ \mathsf{S} \to \lambda y. f(y) \oplus_{\frac{2}{3}} (f(\mathsf{S}\,\mathsf{S}\,y))) \ \Big|\ 0 \to 0 \Big\}\Big)\ \underline{n},$$

is almost-surely terminating.

5 SUBJECT REDUCTION FOR MONADIC AFFINE SIZED TYPES

The type system enjoys a form of subject reduction adapted to the probabilistic case and more specifically to the fact that terms reduce to *distributions* of terms. Let us sketch the idea of this adapted subject reduction property on an example. The type system allows us to derive the judgment

$$\emptyset \mid \emptyset \vdash 0 \oplus 0 \ :\ \Big\{ \Big(\mathsf{Nat}^{\widehat{\mathfrak{s}}}\Big)^{\frac{1}{2}}, \Big(\mathsf{Nat}^{\widehat{\widehat{\mathfrak{r}}}}\Big)^{\frac{1}{2}} \Big\}, \tag{9}$$

where this distribution type is formed by typing a copy of 0 with $\mathsf{Nat}^{\widehat{\mathfrak{s}}}$ and the other with $\mathsf{Nat}^{\widehat{\widehat{\mathfrak{r}}}}$. Then, the term $0 \oplus 0$ reduces to $\{\, 0^{\frac{1}{2}} \,\} + \{\, 0^{\frac{1}{2}} \,\} = \{\, 0^1 \,\} = [\![0 \oplus 0]\!]$: the operational semantics collapses the two copies of 0 appearing during the reduction. However, in the spirit of the usual subject reduction for deterministic languages, we would like to type the two copies of 0 appearing during the reduction with different types. We therefore use the notion of *pseudo-representation*: $[\, 0^{\frac{1}{2}}, 0^{\frac{1}{2}} \,]$ is a pseudo-representation of $[\![0 \oplus 0]\!]$, and we attribute the type $\mathsf{Nat}^{\widehat{\mathfrak{s}}}$ to the first element of this pseudo-representation and the type $\mathsf{Nat}^{\widehat{\widehat{\mathfrak{r}}}}$ to the other, obtaining the following *closed distribution of typed terms*:

$$\Big\{ \Big(0 \ :\ \mathsf{Nat}^{\widehat{\mathfrak{s}}}\Big)^{\frac{1}{2}}, \Big(0 \ :\ \mathsf{Nat}^{\widehat{\widehat{\mathfrak{r}}}}\Big)^{\frac{1}{2}} \Big\}. \tag{10}$$

We can then compute the *average* type of Equation (10), which we call the *expectation type* of this closed distribution of typed terms:

$$\frac{1}{2} \cdot \Big\{ \Big(\mathsf{Nat}^{\widehat{\mathfrak{s}}}\Big)^1 \Big\} \ +\ \frac{1}{2} \cdot \Big\{ \Big(\mathsf{Nat}^{\widehat{\widehat{\mathfrak{r}}}}\Big)^1 \Big\} = \Big\{ \Big(\mathsf{Nat}^{\widehat{\mathfrak{s}}}\Big)^{\frac{1}{2}}, \Big(\mathsf{Nat}^{\widehat{\widehat{\mathfrak{r}}}}\Big)^{\frac{1}{2}} \Big\}.$$

This type coincides with the one of the initial term (Equation (9)). This will be our result of subject reduction: when a closed term M of distribution type μ reduces to a distribution $\mathscr{D}$ of terms, we can type all the terms appearing in a pseudo-representation of $\mathscr{D}$ to obtain a closed distribution of typed terms whose expectation type is μ. Let us now introduce the definitions necessary to the formal statement of the subject reduction property.

Definition 5.1 (Distributions of Distribution Types, of Typed Terms).

- A *distribution of distribution types* is a distribution $\mathscr{D}$ over the set of distribution types and such that $\mu,\ \nu \in \mathrm{S}(\mathscr{D}) \ \Rightarrow\ \langle \mu \rangle = \langle \nu \rangle$.
- A *distribution of typed terms*, or *typed distribution*, is a distribution of typing sequents that are derivable in the monadic affine sized type system. The representation of such a distribution has thus the following form: $\{\, (\Gamma_i \mid \Theta_i \vdash M_i \ :\ \mu_i)^{p_i} \ \mid\ i \in \mathcal{I} \,\}$. In the sequel, we restrict to the *uniform* case in which all the terms appearing in the sequents are typed with distribution types of the same fixed underlying type. We denote this unique simple type κ as $\langle \vec{\mu} \rangle$.
- A *distribution of closed typed terms*, or *closed typed distribution*, is a typed distribution in which all contexts are $\emptyset \mid \emptyset$. In this case, we simply write the representation of the distribution as $\{\, (M_i \ :\ \mu_i)^{p_i} \ \mid\ i \in \mathcal{I} \,\}$, or even as $(M_i \ :\ \mu_i)^{p_i}$ when the indexing is clear from context. We write pseudo-representations in a similar way.
- The *underlying term distribution* of a closed typed distribution $\{\, (M_i \ :\ \mu_i)^{p_i} \ \mid\ i \in \mathcal{I} \,\}$ is the distribution $\{\, (M_i)^{p_i} \ \mid\ i \in \mathcal{I} \,\}$.

Definition 5.2 (Expectation Types). Let $(M_i \ :\ \mu_i)^{p_i}$ be a closed typed distribution. We define its *expectation type* as the distribution type $\mathbb{E}((M_i \ :\ \mu_i)^{p_i}) = \sum_{i \in \mathcal{I}} \ p_i \mu_i$.

Lemma 5.3. *Expectation is linear:*

- $\mathbb{E}((M_i \ :\ \mu_i)^{p_i} + (N_j \ :\ \nu_j)^{q_j}) = \mathbb{E}((M_i \ :\ \mu_i)^{p_i}) + \mathbb{E}((N_j \ :\ \nu_j)^{q_j})$.
- $\mathbb{E}((M_i \ :\ \mu_i)^{p q_i}) = p \cdot \mathbb{E}((M_i \ :\ \mu_i)^{q_i})$.

5.1 Subtyping Probabilistic Sums

Lemma 5.4 (Subtyping Probabilistic Sums). *Suppose that* $\sum (\nu \oplus_p \xi) = 1$ *and that* $\nu \oplus_p \xi \sqsubseteq \mu$. *Then there exists* ν' *and* ξ' *such that* $\mu = \nu' \oplus_p \xi'$, $\nu \sqsubseteq \nu'$, *and* $\xi \sqsubseteq \xi'$. *Note that this implies that* $\mathrm{S}(\nu') \cup \mathrm{S}(\xi') = \mathrm{S}(\mu)$.

Proof. Let $\nu = \{\, \sigma_i^{p'_i} \ \mid\ i \in \mathcal{I} \,\}$ and $\xi = \{\, \tau_j^{p''_j} \ \mid\ j \in \mathcal{J} \,\}$. We assume, without loss of generality, that $\mathcal{I}$ and $\mathcal{J}$ are chosen in such a way that, setting $\mathcal{K} = \mathcal{I} \cap \mathcal{J}$,

$$\exists (i,j) \in \mathcal{I} \times \mathcal{J},\ \ \sigma_i = \tau_j \iff i = j \in \mathcal{K}.$$

It follows that

$$\nu \oplus_p \xi = \left\{ \sigma_i^{pp'_i} \ \middle|\ i \in \mathcal{I} \setminus \mathcal{K} \right\} \ +\ \left\{ \tau_j^{(1-p)p''_j} \ \middle|\ j \in \mathcal{J} \setminus \mathcal{K} \right\} \ +\ \left\{ \sigma_i^{pp'_i + (1-p)p''_i} \ \middle|\ i \in \mathcal{K} \right\}.$$

Set $\mu = \{\, \theta_l^{p'''_l} \ \mid\ l \in \mathcal{L} \,\}$. Since $\nu \oplus_p \xi \sqsubseteq \mu$ and $\sum (\nu \oplus_p \xi) = 1$, there exists a decomposition:

$$\mu = \left[\, \theta_i^{pp'_i} \ \middle|\ i \in \mathcal{I} \setminus \mathcal{K} \,\right] \ +\ \left[\, \theta_j^{(1-p)p''_j} \ \middle|\ j \in \mathcal{J} \setminus \mathcal{K} \,\right] \ +\ \left[\, \theta_k^{pp'_i + (1-p)p''_i} \ \middle|\ k \in \mathcal{K} \,\right]$$

(note that the supports of these distributions may have a nonempty intersection), and this decomposition is such that $\forall i \in \mathcal{I},\ \ \sigma_i \sqsubseteq \theta_i$ and $\forall j \in \mathcal{J},\ \ \tau_j \sqsubseteq \theta_j$. We define $\nu' = \{\, \theta_i^{p'_i} \ \mid\ i \in \mathcal{I} \,\}$ and $\xi' = \{\, \theta_j^{p''_j} \ \mid\ j \in \mathcal{J} \,\}$, which satisfy $\nu \sqsubseteq \nu'$ and $\xi \sqsubseteq \xi'$ but also, by construction, $\mu = \nu' \oplus_p \xi'$. □

Corollary 5.5. *Suppose that* $\mu = \sum_{i \in \mathcal{I}} \ p_i \cdot \mu_i$ *is a distribution such that* $\mu \sqsubseteq \nu$ *and that* $\sum \mu = 1$. *Then there exists a family* $(\nu_i)_{i \in \mathcal{I}}$ *of distributions such that* $\nu = \sum_{i \in \mathcal{I}} p_i \cdot \nu_i$ *and that, for all* $i \in \mathcal{I}$, $\mu_i \sqsubseteq \nu_i$.

Note that the requirement that $\sum \mu = 1$ is not necessary to obtain this result, although it simplifies the reasoning.

5.2 Generation Lemma for Typing

A mandatory step in proofs of Subject Reduction is a generation lemma, which allow one to see any type derivation in an inductive way *even if* the underlying type system is not syntax directed.

Lemma 5.6 (Generation Lemma for Typing).

(1) $\emptyset \mid \emptyset \vdash \mathsf{let}\ x = V\ \mathsf{in}\ N : \mu \implies \exists(\nu,\sigma),\ \emptyset \mid \emptyset \vdash V : \sigma$ *and* $x : \sigma \mid \emptyset \vdash N : \nu$ *and* $\nu \sqsubseteq \mu$.

(2) $\emptyset \mid \emptyset \vdash V\,W : \mu \implies \exists(\nu,\sigma),\ \emptyset \mid \emptyset \vdash V : \sigma \to \nu$ *and* $\emptyset \mid \emptyset \vdash W : \sigma$ *and* $\nu \sqsubseteq \mu$.

(3) $\emptyset \mid \emptyset \vdash \lambda x.M : \sigma \to \mu \implies \exists(\nu,\tau),\ x : \tau \mid \emptyset \vdash M : \nu$ *and* $\sigma \sqsubseteq \tau$ *and* $\nu \sqsubseteq \mu$.

(4) $\emptyset \mid \emptyset \vdash M \oplus_p N : \mu \implies \exists(\nu,\xi),\ \emptyset \mid \emptyset \vdash M : \nu$ *and* $\emptyset \mid \emptyset \vdash N : \xi$ *with* $\sum(\nu \oplus_p \xi) = 1$ *and* $\nu \oplus_p \xi \sqsubseteq \mu$ *and* $\langle \mu \rangle = \langle \nu \rangle = \langle \xi \rangle$.

(5) $\emptyset \mid \emptyset \vdash \mathsf{let}\ x = M\ \mathsf{in}\ N : \nu \implies \exists(\mathcal{I}, (\sigma_i)_{i \in \mathcal{I}}, (p_i)_{i \in \mathcal{I}}, (\mu_i)_{i \in \mathcal{I}})$ *such that*
- $\sum_{i \in \mathcal{I}} p_i \cdot \mu_i \sqsubseteq \nu$,
- $\sum(\sum_{i \in \mathcal{I}} p_i \cdot \mu_i) = 1$,
- $\emptyset \mid \emptyset \vdash M : \{ \sigma_i^{p_i} \mid i \in \mathcal{I} \}$,
- $\forall i \in \mathcal{I},\ x : \sigma_i \mid \emptyset \vdash N : \mu_i$.

(6) $\emptyset \mid \emptyset \vdash \mathsf{case}\ V\ \mathsf{of}\ \{ \mathsf{S} \to W \mid 0 \to Z \} : \mu \implies \exists(\mathfrak{s},\nu)$ *such that* $\emptyset \mid \emptyset \vdash V : \mathsf{Nat}^{\widehat{\mathfrak{s}}}$ *and* $\emptyset \mid \emptyset \vdash W : \mathsf{Nat}^{\mathfrak{s}} \to \nu$ *and* $\emptyset \mid \emptyset \vdash Z : \nu$ *with* $\nu \sqsubseteq \mu$.

(7) $\emptyset \mid \emptyset \vdash \mathsf{letrec}\ f = V : \mu \implies \exists\,((p_j)_{j \in \mathcal{J}}, (\mathfrak{s}_j)_{j \in \mathcal{J}}, \mathfrak{i})$ *such that*
- $\mathsf{Nat}^{\mathfrak{r}} \to \nu[\mathfrak{r}/\mathfrak{i}] \sqsubseteq \mu$,
- $\forall j \in \mathcal{J},\ \mathrm{spine}(\mathfrak{s}_j) = \mathfrak{i}$,
- $\mathfrak{i} \notin \Gamma$ *and* $\mathfrak{i}$ *positive in* ν,
- $\{ (\mathsf{Nat}^{\mathfrak{s}_j} \to \nu[\mathfrak{s}_j/\mathfrak{i}])^{p_j} \mid j \in \mathcal{J} \}$ *induces an AST sized walk,*
- $\emptyset \mid f : \{ (\mathsf{Nat}^{\mathfrak{s}_j} \to \nu[\mathfrak{s}_j/\mathfrak{i}])^{p_j} \mid j \in \mathcal{J} \} \vdash V : \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \nu[\widehat{\mathfrak{i}}/\mathfrak{i}]$.

Proof. By inspection of the rules, the key point being that the subtyping rule is the only one that is not syntax directed, and that by transitivity of $\sqsubseteq$ we can compose several successive subtyping rules. In case (5), we have $\sum(\sum_{i \in \mathcal{I}} p_i \cdot \mu_i) = 1$ since it appears that $\emptyset \mid \emptyset \vdash \mathsf{let}\ x = M\ \mathsf{in}\ N : \sum_{i \in \mathcal{I}} p_i \cdot \mu_i$. Lemma 4.12 allows one then to conclude that this distribution of types has sum 1. □

5.3 Value Substitutions

Definition 5.7 (Context Extending Another). We say that a context $\Delta \mid \Psi$ extends a context $\Gamma \mid \Theta$ when (1) for every $x : \sigma \in \Gamma$ we have $x : \sigma \in \Delta$, and (2) either $\Theta = \emptyset$ or $\Theta = \Psi$. In other words, $\Delta \mid \Psi$ extends $\Gamma \mid \Theta$ when there exists Ξ and Φ such that $\Delta = \Gamma, \Xi$ and $\Psi = \Theta, \Phi$.

Lemma 5.8. *Let M be a closed term such that $\Gamma \mid \Theta \vdash M : \mu$. Then for every context $\Delta \mid \Psi$ extending $\Gamma \mid \Theta$, we have $\Delta \mid \Psi \vdash M : \mu$.*

Proof. We proceed by induction on the structure of M. We set $\Delta = \Gamma, \Xi$ and $\Psi = \Theta, \Phi$.

- If $M = x$ is a variable, the result is immediate.
- If $M = 0$, the result is immediate.
- If $M = \mathsf{S}\ V$, we have by typing rules that $\sigma = \mathsf{Nat}^{\widehat{\mathfrak{s}}}$ and that $\Gamma \mid \Theta \vdash V : \mathsf{Nat}^{\mathfrak{s}}$. By induction hypothesis, $\Delta \mid \Psi \vdash V : \mathsf{Nat}^{\mathfrak{s}}$, from which we conclude using the typing rule for S.
- If $M = \lambda x.N$, we have $\sigma = \tau \to \mu$ and $\Gamma, x : \tau \mid \Theta \vdash N : \mu$. By definition, $\Delta, x : \tau \mid \Psi$ extends $\Gamma, x : \tau \mid \Theta$ so that we have $\Delta, x : \tau \mid \Psi \vdash N : \mu$. The result follows using the Lambda rule.

- If $M = \mathsf{letrec}\ f = V$, the typing rule is of the shape

$$\mathsf{letrec}\ \frac{\begin{array}{c}\langle \Gamma_1 \rangle = \mathsf{Nat} \\ \mathfrak{i} \notin \Gamma_1 \text{ and } \mathfrak{i} \text{ positive in } \nu \text{ and } \forall j \in \mathcal{J},\ \mathrm{spine}(\mathfrak{s}_j) = \mathfrak{i} \\ \left\{ (\mathsf{Nat}^{\mathfrak{s}_j} \to \nu[\mathfrak{s}_j/\mathfrak{i}])^{p_j} \;\middle|\; j \in \mathcal{J} \right\} \text{induces an AST sized walk} \\ \Gamma_1 \,|\, f \;:\; \left\{ (\mathsf{Nat}^{\mathfrak{s}_j} \to \nu[\mathfrak{s}_j/\mathfrak{i}])^{p_j} \;\middle|\; j \in \mathcal{J} \right\} \;\vdash\; V \;:\; \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \nu\left[\widehat{\mathfrak{i}}/\mathfrak{i}\right]\end{array}}{\Gamma_1,\, \Gamma_2 \,|\, \Theta \;\vdash\; \mathsf{letrec}\ f = V \;:\; \mathsf{Nat}^{\mathfrak{r}} \to \nu\left[\mathfrak{r}/\mathfrak{i}\right]}.$$

  Let $\Delta = \Delta_1,\, \Delta_2$, with Δ_1 the maximal subcontext consisting only of variables of affine type Nat. Then

$$\Delta_1 \,|\, f \;:\; \left\{ (\mathsf{Nat}^{\mathfrak{s}_j} \to \nu[\mathfrak{s}_j/\mathfrak{i}])^{p_j} \;\middle|\; j \in \mathcal{J} \right\}$$

  extends

$$\Gamma_1 \,|\, f \;:\; \left\{ \left(\mathsf{Nat}^{\mathfrak{s}_j} \to \nu\left[\mathfrak{s}_j/\mathfrak{i}\right]\right)^{p_j} \;\middle|\; j \in \mathcal{J} \right\}$$

  so that by induction hypothesis $\Delta_1 \,|\, f \;:\; \{ (\mathsf{Nat}^{\mathfrak{s}_j} \to \nu[\mathfrak{s}_j/\mathfrak{i}])^{p_j} \mid j \in \mathcal{J} \} \vdash V \;:\; \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \nu[\widehat{\mathfrak{i}}/\mathfrak{i}]$ so that we can conclude using the letrec rule again that

$$\Delta_1,\, \Delta_2 \,|\, \Psi \vdash \mathsf{letrec}\ f = V \;:\; \mathsf{Nat}^{\mathfrak{r}} \to \nu\,[\mathfrak{r}/\mathfrak{i}]\,.$$

- If $M = V\ W$, the typing derivation provides contexts such that $\Gamma = \Gamma_1,\, \Gamma_2,\, \Gamma_3$ and that $\Theta = \Theta_1,\, \Theta_2$ with $\Gamma_1,\, \Gamma_2 \,|\, \Theta_1 \vdash V \;:\; \sigma \to \mu$ and $\Gamma_1,\, \Gamma_3 \,|\, \Theta_2 \vdash W \;:\; \sigma$. By induction hypothesis, $\Gamma_1,\, \Gamma_3,\, \Xi \,|\, \Theta_2,\, \Phi \vdash W \;:\; \sigma$, from which we conclude using the App rule.
- If $M = \mathsf{let}\ x = N\ \mathsf{in}\ L$, the typing derivation provides contexts such that $\Gamma = \Gamma_1,\, \Gamma_2,\, \Gamma_3$ and that $\Theta = \Theta_1,\, \sum_{i \in \mathcal{I}}\ p_i \cdot \Theta_{2,i}$ with $\Gamma_1,\, \Gamma_2 \,|\, \Theta_1 \vdash M \;:\; \{ \sigma_i^{p_i} \;\mid\; i \in \mathcal{I} \}$ and $\Gamma_1,\, \Gamma_3,\, x \;:\; \sigma_i \,|\, \Theta_{2,i} \vdash N \;:\; \mu_i$. By induction hypothesis, $\Gamma_1,\, \Gamma_2,\, \Xi \,|\, \Theta_1,\, \Phi \vdash M \;:\; \{ \sigma_i^{p_i} \;\mid\; i \in \mathcal{I} \}$, from which we conclude using the Let rule.
- If $M = N \oplus_p L$, then $\Theta = \Theta_1 \oplus_p \Theta_2$ with $\Gamma \,|\, \Theta_1 \;\vdash\; M \;:\; \mu$ and $\Gamma \,|\, \Theta_2 \;\vdash\; N \;:\; \nu$. By applying the induction hypothesis twice, we obtain $\Gamma,\, \Xi \,|\, \Theta_1,\, \Phi \;\vdash\; M \;:\; \mu$ and $\Gamma,\, \Xi \,|\, \Theta_2,\, \Phi \;\vdash\; N \;:\; \nu$. We apply the Choice rule; it remains to prove that $(\Theta_1,\, \Phi) \oplus_p (\Theta_2,\, \Phi) = \Theta_1 \oplus_p \Theta_2,\, \Phi$, which is easily done by definition of $\oplus_p$.
- If $M = \mathsf{case}\ V\ \mathsf{of}\ \{ \mathsf{S} \to W \;\mid\; 0 \to Z \}$, the typing derivation provides contexts such that $\Gamma = \Gamma_1,\, \Gamma_2$ with $\Gamma_1 \,|\, \emptyset \vdash V \;:\; \mathsf{Nat}^{\widehat{\mathfrak{s}}}$ and $\Gamma_2 \,|\, \Theta \vdash W \;:\; \mathsf{Nat}^{\mathfrak{s}} \to \mu$ and $\Gamma_2 \,|\, \Theta \vdash Z \;:\; \mu$. By induction hypothesis, $\Gamma_2,\, \Xi \,|\, \Theta,\, \Phi \vdash W \;:\; \mathsf{Nat}^{\mathfrak{s}} \to \mu$ and $\Gamma_2,\, \Xi \,|\, \Theta,\, \Phi \vdash Z \;:\; \mu$, from which we conclude using the Case rule. □

The key intermediate steps toward Subject Reduction are appropriate substitution lemmas. Here, we need two of them: one for values and one for distributions.

Lemma 5.9 (Closed Value Substitution). *Suppose that* $\Gamma,\, x \;:\; \sigma \,|\, \Theta \vdash M \;:\; \mu$ *and that* $\emptyset \,|\, \emptyset \vdash V \;:\; \sigma$. *Then* $\Gamma \,|\, \Theta \vdash M[V/x] \;:\; \mu$.

Proof. As usual, the proof is by induction on the structure of the typing derivation. We proceed by case analysis on the last rule:

- If it is Var, we have two cases:
  - If the conclusion is $\Gamma,\, x \;:\; \sigma \,|\, \Theta \;\vdash\; x \;:\; \sigma$, then $x[V/x] = V$. By Lemma 5.8, we obtain that $\Gamma \,|\, \Theta \vdash V \;:\; \sigma$.
  - If the conclusion is $\Gamma,\, x \;:\; \sigma,\, y \;:\; \tau \,|\, \Theta \;\vdash\; y \;:\; \tau$, then $y[V/x] = y$ and we obtain $\Gamma,\, y \;:\; \tau \,|\, \Theta \;\vdash\; y \;:\; \tau$ using the Var rule.

- If it is Var', the situation is similar to the latter case of the previous one. The conclusion is $\Gamma, x : \sigma \mid y : \tau \vdash y : \tau$ and $y[V/x] = y$ so that we obtain $\Gamma \mid y : \tau \vdash y : \tau$ using the Var' rule.
- If it is Succ, then $M = \mathsf{S}\ W$ and $\mu = \mathsf{Nat}^{\widehat{\mathfrak{s}}}$. We obtain by induction hypothesis that $\Gamma \mid \Theta \vdash W[V/x] : \mathsf{Nat}^{\mathfrak{s}}$ and we conclude using the Succ rule that $\Gamma \mid \Theta \vdash (\mathsf{S}\ W)[V/x] : \mathsf{Nat}^{\widehat{\mathfrak{s}}}$.
- If it is Zero, we obtain immediately the result.
- If it is λ, suppose that $\Gamma, x : \sigma \mid \Theta \vdash \lambda y.M : \tau \to \mu$. This comes from $\Gamma, x : \sigma, y : \tau \mid \Theta \vdash M : \mu$, to which we apply the induction hypothesis, obtaining that $\Gamma, y : \tau \mid \Theta \vdash M[V/x] : \mu$. Then applying the λ rule gives the expected result.
- For all the remaining cases, as for the λ rule, the result is obtained in a straightforward way from the induction hypothesis. □

Lemma 5.10 (Substitution for distributions). *Suppose that* $\Gamma \mid x : \{ \sigma_i^{p_i} \mid i \in \mathcal{I} \} \vdash M : \mu$ *and that, for every* $i \in \mathcal{I}$, *we have* $\emptyset \mid \emptyset \vdash V : \sigma_i$. *Then* $\Gamma \mid \emptyset \vdash M[V/x] : \mu$.

Proof. The proof is by induction on the structure of the typing derivation. We proceed by case analysis on the last rule:

- If it is Var, we have $M = y \neq x$ and $y \in \Gamma$. It follows that $y[V/x] = y$ and we obtain $\Gamma \mid \emptyset \vdash M[V/x] : \mu$ simply by the Var rule.
- If it is Var', we have $M = x$ so that $M[V/x] = V$. Moreover, the distribution $\{ \sigma_i^{p_i} \mid i \in \mathcal{I} \}$ must be Dirac; we denote by σ the unique element of its support. Note that we also obtain $\sigma = \mu$. As we supposed that $\emptyset \mid \emptyset \vdash V : \sigma$, Lemma 5.8 gives $\Gamma \mid \emptyset \vdash V : \sigma$, from which we conclude.
- If it is LetRec, then x does not occur free in M. It follows that $M[V/x] = M$, and we can derive $\Gamma \mid \emptyset \vdash M[V/x] : \mu$ using a letrec rule with the same hypothesis.
- All others cases are treated straightforwardly using the induction hypothesis. □

Lemma 5.11.

(1) $\Gamma \mid \Theta \vdash \mathsf{S}\ V : \mathsf{Nat}^{\widehat{\mathfrak{s}}} \quad \Longrightarrow \quad \Gamma \mid \Theta \vdash V : \mathsf{Nat}^{\mathfrak{s}}$.
(2) $\Gamma \mid \Theta \vdash 0 : \mathsf{Nat}^{\mathfrak{s}} \quad \Longrightarrow \quad \exists \mathfrak{r},\ \mathfrak{s} = \widehat{\mathfrak{r}}$.
(3) $\Gamma \mid \Theta \vdash \mathsf{S}\ V : \mathsf{Nat}^{\mathfrak{s}} \quad \Longrightarrow \quad \exists \mathfrak{r},\ \mathfrak{s} = \widehat{\mathfrak{r}}$.

Proof. All points are immediate due to the typing rules introducing 0 and S. Recall that by the subtyping rules, $\widehat{\infty} = \infty$. □

5.4 Size Substitutions

Another form of substitution that our type system implicitly implements is the one of sizes into types, e.g., in the letrec rule. Some easy intermediate lemmas are needed to make sure that this form of substitution is well behaved.

Lemma 5.12 (Successor and Size Order). *Suppose that* $\mathfrak{s} \preccurlyeq \mathfrak{r}$. *Then* $\widehat{\mathfrak{s}} \preccurlyeq \widehat{\mathfrak{r}}$.

Proof. By definition of $\preccurlyeq$, if $\mathfrak{s} \preccurlyeq \mathfrak{r}$, there are two cases: either $\mathfrak{r} = \infty$, or $\mathrm{spine}(\mathfrak{s}) = \mathrm{spine}(\mathfrak{r}) = \mathfrak{i}$ with $\mathfrak{s} = \widehat{\mathfrak{i}}^{k}$, $\mathfrak{r} = \widehat{\mathfrak{i}}^{k'}$, and $k \leq k'$. In both cases, the conclusion is immediate. □

Lemma 5.13 (Size Substitutions are Monotonic).

(1) *Suppose that* $\mathfrak{s} \preccurlyeq \mathfrak{r}$; *then for any size* $\mathfrak{t}$ *and size variable* $\mathfrak{i}$ *we have* $\mathfrak{s}[\mathfrak{t}/\mathfrak{i}] \preccurlyeq \mathfrak{r}[\mathfrak{t}/\mathfrak{i}]$.
(2) *Suppose that* $\mathfrak{s} \preccurlyeq \mathfrak{r}$; *then for any size* $\mathfrak{t}$ *and size variable* $\mathfrak{i}$ *we have* $\mathfrak{t}[\mathfrak{s}/\mathfrak{i}] \preccurlyeq \mathfrak{t}[\mathfrak{r}/\mathfrak{i}]$.

Proof.

(1) We proceed by induction on the derivation proving that $\mathfrak{s} \preccurlyeq \mathfrak{r}$ by case analysis on the last rule.
- If it is $\mathfrak{s} \preccurlyeq \mathfrak{s}$, then $\mathfrak{s} = \mathfrak{r}$ and the result is immediate.
- If it is
$$\frac{\mathfrak{s} \preccurlyeq \mathfrak{u} \qquad \mathfrak{u} \preccurlyeq \mathfrak{r}}{\mathfrak{s} \preccurlyeq \mathfrak{r}},$$
then by induction hypothesis $\mathfrak{s}[\mathfrak{t}/\mathfrak{i}] \preccurlyeq \mathfrak{u}[\mathfrak{t}/\mathfrak{i}]$ and $\mathfrak{u}[\mathfrak{t}/\mathfrak{i}] \preccurlyeq \mathfrak{r}[\mathfrak{t}/\mathfrak{i}]$ so that we conclude using this same deduction rule.
- If it is $\mathfrak{s} \preccurlyeq \widehat{\mathfrak{s}}$, we have $\mathfrak{r} = \widehat{\mathfrak{s}}$, and using the definition of size substitution, we obtain $\mathfrak{r}[\mathfrak{t}/\mathfrak{i}] = \widehat{\mathfrak{s}}[\mathfrak{t}/\mathfrak{i}] = \widehat{\mathfrak{s}[\mathfrak{t}/\mathfrak{i}]}$. We conclude using the same deduction rule.
- If it is $\mathfrak{s} \preccurlyeq \infty$, we have $\infty[\mathfrak{t}/\mathfrak{i}] = \infty$ and we obtain immediately $\mathfrak{s}[\mathfrak{t}/\mathfrak{i}] \preccurlyeq \infty$.

(2) We proceed by case analysis on $\mathfrak{t}$. There are four cases:
- If $\mathfrak{t} = \mathfrak{i}$, then $\mathfrak{t}[\mathfrak{s}/\mathfrak{i}] = \mathfrak{s} \preccurlyeq \mathfrak{r} = \mathfrak{t}[\mathfrak{r}/\mathfrak{i}]$.
- If $\mathfrak{t} = \mathfrak{j} \neq \mathfrak{i}$, then $\mathfrak{t}[\mathfrak{s}/\mathfrak{i}] = \mathfrak{j} \preccurlyeq \mathfrak{j} = \mathfrak{t}[\mathfrak{r}/\mathfrak{i}]$.
- If $\mathfrak{t} = \widehat{\mathfrak{u}}$, we have by induction hypothesis that $\mathfrak{u}[\mathfrak{s}/\mathfrak{i}] \preccurlyeq \mathfrak{u}[\mathfrak{r}/\mathfrak{i}]$. We conclude using Lemma 5.12.
- If $\mathfrak{t} = \infty$, $\mathfrak{t}[\mathfrak{s}/\mathfrak{i}] = \infty \preccurlyeq \infty = \mathfrak{t}[\mathfrak{r}/\mathfrak{i}]$. □

Lemma 5.14 (Size Substitutions and Subtyping).

(1) If $\sigma \sqsubseteq \tau$, then for any size $\mathfrak{s}$ and size variable $\mathfrak{i}$ we have $\sigma[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \tau[\mathfrak{s}/\mathfrak{i}]$.
If $\mu \sqsubseteq \nu$, then for any size $\mathfrak{s}$ and size variable $\mathfrak{i}$ we have $\mu[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \nu[\mathfrak{s}/\mathfrak{i}]$.

(2) If $\mathfrak{i}$ pos σ and $\mathfrak{s} \preccurlyeq \mathfrak{r}$, we have $\sigma[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \sigma[\mathfrak{r}/\mathfrak{i}]$.
If $\mathfrak{i}$ pos μ and $\mathfrak{s} \preccurlyeq \mathfrak{r}$, we have $\mu[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \mu[\mathfrak{r}/\mathfrak{i}]$.

(3) If $\mathfrak{i}$ neg σ and $\mathfrak{s} \preccurlyeq \mathfrak{r}$, we have $\sigma[\mathfrak{r}/\mathfrak{i}] \sqsubseteq \sigma[\mathfrak{s}/\mathfrak{i}]$.
If $\mathfrak{i}$ neg μ and $\mathfrak{s} \preccurlyeq \mathfrak{r}$, we have $\mu[\mathfrak{r}/\mathfrak{i}] \sqsubseteq \mu[\mathfrak{s}/\mathfrak{i}]$.

Proof.

(1) We prove both statements at the same time by induction on the derivation proving that $\mu \sqsubseteq \nu$ (or $\sigma \sqsubseteq \tau$).
- If the last rule is $\sigma \sqsubseteq \sigma$, then $\mu = \nu = \sigma$ and the result is immediate.
- If the last rule is
$$\frac{\mathfrak{t} \preccurlyeq \mathfrak{r}}{\mathsf{Nat}^{\mathfrak{t}} \sqsubseteq \mathsf{Nat}^{\mathfrak{r}}},$$
then by Lemma 5.13 we have $\mathfrak{t}[\mathfrak{s}/\mathfrak{i}] \preccurlyeq \mathfrak{r}[\mathfrak{s}/\mathfrak{i}]$ so that $(\mathsf{Nat}^{\mathfrak{t}})\,[\mathfrak{s}/\mathfrak{i}] = \mathsf{Nat}^{\mathfrak{t}[\mathfrak{s}/\mathfrak{i}]} \sqsubseteq \mathsf{Nat}^{\mathfrak{r}[\mathfrak{s}/\mathfrak{i}]} = (\mathsf{Nat}^{\mathfrak{r}})[\mathfrak{s}/\mathfrak{i}]$.
- If the last rule is
$$\frac{\tau \sqsubseteq \sigma \qquad \mu \sqsubseteq \nu}{\sigma \to \mu \ \sqsubseteq\ \tau \to \nu},$$
then by induction hypothesis $\tau[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \sigma[\mathfrak{s}/\mathfrak{i}]$ and $\mu[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \nu[\mathfrak{s}/\mathfrak{i}]$, from which we conclude using the same rule.
- If the last rule is
$$\frac{\exists f\ :\ \mathcal{I} \to \mathcal{J},\ \left(\forall i \in \mathcal{I},\ \ \sigma_i \ \sqsubseteq\ \tau_{f(i)}\right) \text{ and } \left(\forall j \in \mathcal{J},\ \ \sum_{i \in f^{-1}(j)}\ p_i \le p'_j\right)}{\left\{ \sigma_i^{p_i} \ \middle|\ i \in \mathcal{I} \right\} \ \sqsubseteq\ \left\{ \tau_j^{p'_j} \ \middle|\ j \in \mathcal{J} \right\}},$$
we obtain by induction hypothesis that for every $i \in \mathcal{I}\, \sigma_i[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \tau_{f(i)}[\mathfrak{s}/\mathfrak{i}]$, from which we conclude using the same rule.

(2) We prove (2) and (3) by mutual induction on μ (or σ). Let $\mathfrak{s} \preccurlyeq \mathfrak{r}$.

- If $\sigma = \mathsf{Nat}^{\mathfrak{t}}$,
  - —Suppose that $\mathfrak{i}$ pos $\mathsf{Nat}^{\mathfrak{t}}$. Note that this does not assume anything on $\mathfrak{t}$. Since $\mathfrak{s} \preccurlyeq \mathfrak{r}$, we have $(\mathsf{Nat}^{\mathfrak{t}})[\mathfrak{s}/\mathfrak{i}] = \mathsf{Nat}^{\mathfrak{t}[\mathfrak{s}/\mathfrak{i}]} \sqsubseteq \mathsf{Nat}^{\mathfrak{t}[\mathfrak{r}/\mathfrak{i}]} = (\mathsf{Nat}^{\mathfrak{t}})[\mathfrak{r}/\mathfrak{i}]$, where we used the monotonicity of size substitution (Lemma 5.13).
  - —Suppose that $\mathfrak{i}$ neg $\mathsf{Nat}^{\mathfrak{t}}$. Then $\mathfrak{i} \notin \mathfrak{t}$ and $(\mathsf{Nat}^{\mathfrak{t}})[\mathfrak{s}/\mathfrak{i}] = (\mathsf{Nat}^{\mathfrak{t}})[\mathfrak{r}/\mathfrak{i}]$ so that we can conclude.
- If $\sigma = \tau \to \mu$,
  - —Suppose that $\mathfrak{i}$ pos σ. Then $\mathfrak{i}$ neg τ and $\mathfrak{i}$ pos μ. By induction hypothesis, $\tau[\mathfrak{r}/\mathfrak{i}] \sqsubseteq \tau[\mathfrak{s}/\mathfrak{i}]$ and $\mu[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \mu[\mathfrak{r}/\mathfrak{i}]$. By the subtyping rules, $\sigma[\mathfrak{s}/\mathfrak{i}] = \tau[\mathfrak{s}/\mathfrak{i}] \to \mu[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \tau[\mathfrak{r}/\mathfrak{i}] \to \mu[\mathfrak{r}/\mathfrak{i}] = \sigma[\mathfrak{r}/\mathfrak{i}]$.
  - —Suppose that $\mathfrak{i}$ neg σ. The reasoning is symmetrical.
- If $\mu = \{\ \sigma_i^{p_i} \ \mid\ i \in \mathcal{I}\ \}$,
  - —Suppose that $\mathfrak{i}$ pos μ. Then for every $i \in \mathcal{I}$ we have $\mathfrak{i}$ pos σ_i and by induction hypothesis $\sigma_i[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \sigma_i[\mathfrak{r}/\mathfrak{i}]$. We obtain that $\mu[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \mu[\mathfrak{r}/\mathfrak{i}]$ using the identity as a reindexing function.
  - —Suppose that $\mathfrak{i}$ neg μ. The reasoning is symmetrical. □

Lemma 5.15 (Size substitution). *If $\Gamma \mid \Theta \vdash M\ :\ \mu$, then for any size variable $\mathfrak{i}$ and any size $\mathfrak{s}$ we have that $\Gamma[\mathfrak{s}/\mathfrak{i}] \mid \Theta[\mathfrak{s}/\mathfrak{i}] \vdash M\ :\ \mu[\mathfrak{s}/\mathfrak{i}]$.*

Proof. We assume that $\mathfrak{i} \notin \mathfrak{s}$, without loss of generality: otherwise, we introduce a fresh size variable $\mathfrak{j}$, substitute it with $\mathfrak{s}$, and then substitute $\mathfrak{i}$ with $\mathfrak{j}$. The proof is by induction on the typing derivation. We proceed by case analysis on the last rule.

- If it is Var: we have $\Gamma, x\ :\ \sigma \mid \Theta\ \vdash\ x\ :\ \sigma$ and deduce immediately using the Var rule again that $\Gamma[\mathfrak{s}/\mathfrak{i}], x\ :\ \sigma[\mathfrak{s}/\mathfrak{i}] \mid \Theta[\mathfrak{s}/\mathfrak{i}]\ \vdash\ x\ :\ \sigma[\mathfrak{s}/\mathfrak{i}]$.
- If it is Var': we have $\Gamma \mid x\ :\ \sigma\ \vdash\ x\ :\ \sigma$ and deduce immediately using the Var' rule again that $\Gamma[\mathfrak{s}/\mathfrak{i}] \mid x\ :\ \sigma[\mathfrak{s}/\mathfrak{i}]\ \vdash\ x\ :\ \sigma[\mathfrak{s}/\mathfrak{i}]$.
- If it is Succ: then $M = \mathsf{S}\ V$ and $\mu = \mathsf{Nat}^{\widehat{\mathfrak{r}}}$. By induction hypothesis, $\Gamma[\mathfrak{s}/\mathfrak{i}] \mid \Theta[\mathfrak{s}/\mathfrak{i}] \vdash V\ :\ (\mathsf{Nat}^{\mathfrak{r}})[\mathfrak{s}/\mathfrak{i}]$. But $(\mathsf{Nat}^{\mathfrak{r}})[\mathfrak{s}/\mathfrak{i}] = \mathsf{Nat}^{\mathfrak{r}[\mathfrak{s}/\mathfrak{i}]}$ so that by the Succ rule $\Gamma[\mathfrak{s}/\mathfrak{i}] \mid \Theta[\mathfrak{s}/\mathfrak{i}] \vdash \mathsf{S}\ V\ :\ \mathsf{Nat}^{\widehat{\mathfrak{r}[\mathfrak{s}/\mathfrak{i}]}}$. We use the equality $\mathsf{Nat}^{\widehat{\mathfrak{r}[\mathfrak{s}/\mathfrak{i}]}} = (\mathsf{Nat}^{\widehat{\mathfrak{r}}})[\mathfrak{s}/\mathfrak{i}]$ to conclude.
- If it is Zero: the result is immediate.
- If it is λ: we have $M = \lambda x.N$ and $\mu = \sigma \to \nu$. By induction hypothesis, $\Gamma[\mathfrak{s}/\mathfrak{i}], x\ :\ \sigma[\mathfrak{s}/\mathfrak{i}] \mid \Theta[\mathfrak{s}/\mathfrak{i}]\ \vdash\ N\ :\ \nu[\mathfrak{s}/\mathfrak{i}]$. By application of the λ rule, $\Gamma[\mathfrak{s}/\mathfrak{i}] \mid \Theta[\mathfrak{s}/\mathfrak{i}]\ \vdash\ \lambda x.N\ :\ \sigma[\mathfrak{s}/\mathfrak{i}] \to \nu[\mathfrak{s}/\mathfrak{i}]$. We conclude using $\sigma[\mathfrak{s}/\mathfrak{i}] \to \nu[\mathfrak{s}/\mathfrak{i}] = (\sigma \to \nu)[\mathfrak{s}/\mathfrak{i}]$.
- If it is Sub: the hypothesis of the rule is $\Gamma \mid \Theta \vdash M\ :\ \nu$ for $\nu \sqsubseteq \mu$. By induction hypothesis, $\Gamma[\mathfrak{s}/\mathfrak{i}] \mid \Theta[\mathfrak{s}/\mathfrak{i}] \vdash M\ :\ \nu[\mathfrak{s}/\mathfrak{i}]$. But by Lemma 5.14 we have $\nu[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \mu[\mathfrak{s}/\mathfrak{i}]$. We conclude using the Sub rule.
- If it is App, we have $M = V\ W$ and $\Gamma = \Gamma_1, \Gamma_2, \Gamma_3$ and $\Theta = \Theta_1, \Theta_2$ with $\langle \Gamma_1 \rangle = \mathsf{Nat}$, $\Gamma_1, \Gamma_2 \mid \Theta_1 \vdash V\ :\ \sigma \to \mu$ and $\Gamma_1, \Gamma_3 \mid \Theta_2 \vdash W\ :\ \sigma$. Applying the induction hypothesis twice gives $\Gamma_1[\mathfrak{s}/\mathfrak{i}], \Gamma_2[\mathfrak{s}/\mathfrak{i}] \mid \Theta_1[\mathfrak{s}/\mathfrak{i}] \vdash V\ :\ (\sigma \to \mu)[\mathfrak{s}/\mathfrak{i}]$ and $\Gamma_1[\mathfrak{s}/\mathfrak{i}], \Gamma_3[\mathfrak{s}/\mathfrak{i}] \mid \Theta_2[\mathfrak{s}/\mathfrak{i}] \vdash W\ :\ \sigma[\mathfrak{s}/\mathfrak{i}]$. Since $\sigma[\mathfrak{s}/\mathfrak{i}] \to \mu[\mathfrak{s}/\mathfrak{i}] = (\sigma \to \mu)[\mathfrak{s}/\mathfrak{i}]$, we can use the Application rule to conclude.
- If it is Choice, then $M = N \oplus_p L$ and $\mu = \nu \oplus_p \xi$ and $\Theta = \Theta_1 \oplus_p \Theta_2$ with $\Gamma \mid \Theta_1 \vdash N\ :\ \nu$ and $\Gamma \mid \Theta_2 \vdash L\ :\ \xi$ and $\langle \nu \rangle = \langle \xi \rangle$. The induction hypothesis, applied twice, gives $\Gamma[\mathfrak{s}/\mathfrak{i}] \mid \Theta_1[\mathfrak{s}/\mathfrak{i}] \vdash N\ :\ \nu[\mathfrak{s}/\mathfrak{i}]$ and $\Gamma[\mathfrak{s}/\mathfrak{i}] \mid \Theta_2[\mathfrak{s}/\mathfrak{i}] \vdash L\ :\ \xi[\mathfrak{s}/\mathfrak{i}]$, from which we conclude using the Choice rule again and the equality $\nu[\mathfrak{s}/\mathfrak{i}] \oplus_p \xi[\mathfrak{s}/\mathfrak{i}] = (\nu \oplus_p \xi)[\mathfrak{s}/\mathfrak{i}]$ from Lemma 4.6.
- If it is Let, then $M = (\mathsf{let}\ x = N\ \mathsf{in}\ L)$ and $\mu = \sum_{i \in \mathcal{I}}\ p_i \cdot \nu_i$ and $\Gamma = \Gamma_1, \Gamma_2, \Gamma_3$ and $\Theta = \Theta_1, \sum_{i \in \mathcal{I}}\ \Theta_{2,i}$ with $\Gamma_1, \Gamma_2 \mid \Theta_1 \vdash N\ :\ \{\ \sigma_i^{p_i}\ \mid\ i \in \mathcal{I}\ \}$ and, for every $i \in \mathcal{I}$, $\Gamma_1, \Gamma_3 \mid \Theta_{2,i} \vdash L\ :\ \nu_i$

and $\langle \Gamma_1 \rangle = \mathsf{Nat}$. By repeated applications of the induction hypothesis,

$$\Gamma_1\,[\mathfrak{s}/\mathfrak{i}]\,,\,\Gamma_2\,[\mathfrak{s}/\mathfrak{i}]\;|\,\Theta_1\,[\mathfrak{s}/\mathfrak{i}] \vdash N\;:\;\left\{\,\sigma_i^{p_i}\;\middle|\; i \in \mathcal{I}\,\right\}[\mathfrak{s}/\mathfrak{i}]\,,$$

and for every $i \in \mathcal{I}$,

$$\Gamma_1\,[\mathfrak{s}/\mathfrak{i}]\,,\,\Gamma_3\,[\mathfrak{s}/\mathfrak{i}]\;|\,\Theta_{2,i}\,[\mathfrak{s}/\mathfrak{i}] \vdash L\;:\;\nu_i\,[\mathfrak{s}/\mathfrak{i}]\,.$$

We use in the first time the equality $\{\,\sigma_i^{p_i}\;\mid\; i \in \mathcal{I}\,\}[\mathfrak{s}/\mathfrak{i}] = \{\,(\sigma_i[\mathfrak{s}/\mathfrak{i}])^{p_i}\;\mid\; i \in \mathcal{I}\,\}$ coming from the definition of size substitutions. We conclude using the Let rule again and the equality $(\sum_{i\in\mathcal{I}}\;p_i \cdot \nu_i)[\mathfrak{s}/\mathfrak{i}] = \sum_{i\in\mathcal{I}}\;p_i \cdot \nu_i[\mathfrak{s}/\mathfrak{i}]$ from Lemma 4.6.

- If it is Case, then $M = \mathsf{case}\ V\ \mathsf{of}\ \{\,\mathsf{S} \to W \;\mid\; 0 \to Z\,\}$ and $\Gamma = \Gamma_1,\,\Gamma_2$ with $\Gamma_1 \,|\, \emptyset \vdash V\;:\;\mathsf{Nat}^{\widehat{\mathfrak{r}}}$ and $\Gamma_2 \,|\, \Theta \vdash W\;:\;\mathsf{Nat}^{\mathfrak{r}} \to \mu$ and $\Gamma_2 \,|\, \Theta \vdash Z\;:\;\mu$. We apply the induction hypothesis three times and obtain $\Gamma_1[\mathfrak{s}/\mathfrak{i}] \,|\, \emptyset \vdash V\;:\;(\mathsf{Nat}^{\widehat{\mathfrak{r}}})[\mathfrak{s}/\mathfrak{i}]$ and $\Gamma_2[\mathfrak{s}/\mathfrak{i}] \,|\, \Theta[\mathfrak{s}/\mathfrak{i}] \vdash W\;:\;(\mathsf{Nat}^{\mathfrak{r}} \to \mu)[\mathfrak{s}/\mathfrak{i}]$ and $\Gamma_2[\mathfrak{s}/\mathfrak{i}] \,|\, \Theta[\mathfrak{s}/\mathfrak{i}] \vdash Z\;:\;\mu[\mathfrak{s}/\mathfrak{i}]$. We use the equalities $(\mathsf{Nat}^{\widehat{\mathfrak{r}}})[\mathfrak{s}/\mathfrak{i}] = \mathsf{Nat}^{\widehat{\mathfrak{r}[\mathfrak{s}/\mathfrak{i}]}}$ and $(\mathsf{Nat}^{\mathfrak{r}} \to \mu)[\mathfrak{s}/\mathfrak{i}] = \mathsf{Nat}^{\mathfrak{r}[\mathfrak{s}/\mathfrak{i}]} \to \mu[\mathfrak{s}/\mathfrak{i}]$ and then the Case rule to conclude.
- If it is letrec, we carefully adapt the proof scheme of [3, Lemma 3.8]. We have $M = \mathsf{letrec}\ f = V$ and $\mu = \mathsf{Nat}^{\mathfrak{r}} \to \nu[\mathfrak{r}/\mathfrak{j}]$ and $\Gamma = \Gamma_1,\,\Gamma_2$ with
  —$\langle \Gamma_1 \rangle = \mathsf{Nat}$,
  —$\mathfrak{j} \notin \Gamma_1$ and $\mathfrak{j}$ positive in ν and $\forall j \in \mathcal{J},\ \mathrm{spine}\,(\mathfrak{r}_j) = \mathfrak{j}$,
  —$\{\,(\mathsf{Nat}^{\mathfrak{r}_j} \to \nu[\mathfrak{r}_j/\mathfrak{j}])^{p_j}\;\mid\; j \in \mathcal{J}\,\}$ induces an AST sized walk,
  —and

$$\Gamma_1 \,|\, f\;:\;\left\{\,\left(\mathsf{Nat}^{\mathfrak{r}_j} \to \nu\left[\mathfrak{r}_j/\mathfrak{j}\right]\right)^{p_j}\;\middle|\; j \in \mathcal{J}\,\right\} \vdash V\;:\;\mathsf{Nat}^{\widehat{\mathfrak{j}}} \to \nu\left[\widehat{\mathfrak{j}}/\mathfrak{j}\right]. \tag{11}$$

We suppose, without loss of generality as this can be easily obtained by renaming $\mathfrak{j}$ to a fresh variable, that $\mathfrak{i} \neq \mathfrak{j}$ and that $\mathfrak{j} \notin \mathfrak{s}$. Let $\mathfrak{l}$ be a fresh size variable; it follows in particular that $\mathfrak{l} \notin \Gamma_1,\,\Gamma_2,\,\nu,\,\mathfrak{s}$. We apply the induction hypothesis to Equation (11) and obtain

$$\Gamma_1\,[\mathfrak{l}/\mathfrak{j}]\;|\,f\;:\;\left(\left\{\,\left(\mathsf{Nat}^{\mathfrak{r}_j} \to \nu\left[\mathfrak{r}_j/\mathfrak{j}\right]\right)^{p_j}\;\middle|\; j \in \mathcal{J}\,\right\}\right)[\mathfrak{l}/\mathfrak{j}] \vdash V\;:\;\left(\mathsf{Nat}^{\widehat{\mathfrak{j}}} \to \nu\left[\widehat{\mathfrak{j}}/\mathfrak{j}\right]\right)[\mathfrak{l}/\mathfrak{j}]\,,$$

which, after applying a series of equalities and using the fact that $\mathfrak{j} \notin \Gamma_1$, coincides with

$$\Gamma_1 \,|\, f\;:\;\left(\left\{\,\left(\mathsf{Nat}^{\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]} \to \nu\left[\mathfrak{r}_j\,[\mathfrak{l}/\mathfrak{j}]\,/\mathfrak{j}\right]\right)^{p_j}\;\middle|\; j \in \mathcal{J}\,\right\}\right) \vdash V\;:\;\mathsf{Nat}^{\widehat{\mathfrak{l}}} \to \nu\left[\widehat{\mathfrak{j}}/\mathfrak{j}\right][\mathfrak{l}/\mathfrak{j}]$$

but also with

$$\Gamma_1 \,|\, f\;:\;\left(\left\{\,\left(\mathsf{Nat}^{\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]} \to \nu\,[\mathfrak{l}/\mathfrak{j}]\left[\mathfrak{r}_j\,[\mathfrak{l}/\mathfrak{j}]\,/\mathfrak{l}\right]\right)^{p_j}\;\middle|\; j \in \mathcal{J}\,\right\}\right) \vdash V\;:\;\mathsf{Nat}^{\widehat{\mathfrak{l}}} \to \nu\,[\mathfrak{l}/\mathfrak{j}]\left[\widehat{\mathfrak{l}}/\mathfrak{l}\right].$$

We can apply the induction hypothesis again, and obtain after rewriting

$$\Gamma_1\,[\mathfrak{s}/\mathfrak{i}]\;|\,f : \left(\left\{\,\left(\mathsf{Nat}^{\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]} \to \nu\,[\mathfrak{l}/\mathfrak{j}]\left[\mathfrak{r}_j\,[\mathfrak{l}/\mathfrak{j}]\,/\mathfrak{l}\right][\mathfrak{s}/\mathfrak{i}]\right)^{p_j}\middle|\, j \in \mathcal{J}\,\right\}\right) \vdash V\;:\;\mathsf{Nat}^{\widehat{\mathfrak{l}}} \to \nu\,[\mathfrak{l}/\mathfrak{j}]\left[\widehat{\mathfrak{l}}/\mathfrak{l}\right][\mathfrak{s}/\mathfrak{i}]\,,$$

where we used the fact that $\forall j \in \mathcal{J},\ \mathrm{spine}\,(\mathfrak{r}_j) = \mathfrak{j} \neq \mathfrak{i}$ so that $(\mathsf{Nat}^{\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]})[\mathfrak{s}/\mathfrak{i}] = \mathsf{Nat}^{\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]}$. Since $\mathfrak{l} \notin \mathfrak{s}$, we can exchange $[\widehat{\mathfrak{l}}/\mathfrak{l}]$ and $[\mathfrak{s}/\mathfrak{i}]$. For every $j \in \mathcal{J}$, we can also exchange $[\mathfrak{s}/\mathfrak{i}]$ and $[\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]/\mathfrak{l}]$ since $\mathrm{spine}\,(\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]) = \mathfrak{l} \neq \mathfrak{i}$ and $\mathfrak{l} \notin \mathfrak{s}$. We obtain:

$$\Gamma_1\,[\mathfrak{s}/\mathfrak{i}]\;|\,f\;:\;\left\{\,\left(\mathsf{Nat}^{\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]} \to \nu\,[\mathfrak{l}/\mathfrak{j}]\,[\mathfrak{s}/\mathfrak{i}]\left[\mathfrak{r}_j\,[\mathfrak{l}/\mathfrak{j}]\,/\mathfrak{l}\right]\right)^{p_j}\;\middle|\; j \in \mathcal{J}\,\right\} \vdash V\;:\;\mathsf{Nat}^{\widehat{\mathfrak{l}}} \to \nu\,[\mathfrak{l}/\mathfrak{j}]\,[\mathfrak{s}/\mathfrak{i}]\left[\widehat{\mathfrak{l}}/\mathfrak{l}\right].$$

Additionally, we have:
—$\langle \Gamma_1[\mathfrak{s}/\mathfrak{i}] \rangle = \mathsf{Nat}$;
—$\mathfrak{l} \notin \Gamma_1[\mathfrak{s}/\mathfrak{i}]$;
—$\mathfrak{l}$ positive in $\nu[\mathfrak{l}/\mathfrak{j}][\mathfrak{s}/\mathfrak{i}]$ since $\mathfrak{j}$ was positive in ν;
—$\forall j \in \mathcal{J},\ \mathrm{spine}\,(\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]) = \mathfrak{l}$ since $\mathrm{spine}\,(\mathfrak{r}_j) = \mathfrak{j}$; and

- $\{\,(\mathsf{Nat}^{\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]} \to \nu[\mathfrak{l}/\mathfrak{j}][\mathfrak{s}/\mathfrak{i}][\mathfrak{r}_j[\mathfrak{l}/\mathfrak{j}]/\mathfrak{l}])^{p_j} \;\mid\; j \in \mathcal{J}\,\}$ induces the same sized walk, which is thus AST, as $\{\,(\mathsf{Nat}^{\mathfrak{r}_j} \to \nu[\mathfrak{r}_j/\mathfrak{j}])^{p_j} \;\mid\; j \in \mathcal{J}\,\}$. Indeed, only the spine variable changes under the substitution $[\mathfrak{l}/\mathfrak{j}]$.

Let $\mathfrak{t} = \mathfrak{r}[\mathfrak{s}/\mathfrak{i}]$. Since all these conditions are met, we can apply the letrec rule and obtain

$$\Gamma_1\,[\mathfrak{s}/\mathfrak{i}]\,,\,\Gamma_2\,[\mathfrak{s}/\mathfrak{i}]\;|\,\Theta\,[\mathfrak{s}/\mathfrak{i}] \vdash \mathsf{letrec}\ f = V \;:\; \mathsf{Nat}^{\mathfrak{t}} \to \nu\,[\mathfrak{l}/\mathfrak{j}]\,[\mathfrak{s}/\mathfrak{i}]\,[\mathfrak{t}/\mathfrak{l}]\,.$$

Since $\mathfrak{i}$, $\mathfrak{l} \notin \mathfrak{s}$ and $\mathfrak{l} \notin \nu$, we can commute $[\mathfrak{s}/\mathfrak{i}]$ and $[\mathfrak{t}/\mathfrak{l}]$ and compose substitutions to obtain

$$\Gamma\,[\mathfrak{s}/\mathfrak{i}]\;|\,\Theta\,[\mathfrak{s}/\mathfrak{i}] \vdash \mathsf{letrec}\ f = V \;:\; \mathsf{Nat}^{\mathfrak{t}} \to \nu\,[\mathfrak{t}/\mathfrak{j}]\,[\mathfrak{s}/\mathfrak{i}]\,,$$

which rewrites to

$$\Gamma\,[\mathfrak{s}/\mathfrak{i}]\;|\,\Theta\,[\mathfrak{s}/\mathfrak{i}] \vdash \mathsf{letrec}\ f = V \;:\; (\mathsf{Nat}^{\mathfrak{r}} \to \nu\,[\mathfrak{r}/\mathfrak{j}])\,[\mathfrak{s}/\mathfrak{i}]\,,$$

which allows us to conclude. □

5.5 Subject Reduction

We can now state the main lemma of subject reduction:

Lemma 5.16 (Subject Reduction, Fundamental Lemma). *Let $M \in \Lambda^{\mathsf{s}}_{\oplus}(\mu)$ and $\mathscr{D}$ be the unique closed term distribution such that $M \to_v \mathscr{D}$. Then there exists a closed typed distribution $\{\,(L_j : \nu_j)^{p_j} \;\mid\; j \in \mathcal{J}\,\}$ such that $\mathbb{E}((L_j : \nu_j)^{p_j}) = \mu$, and $[\,(L_j)^{p_j} \;\mid\; j \in \mathcal{J}\,]$ is a pseudo-representation of $\mathscr{D}$. Note that the condition on expectations implies that $\bigcup_{j\in\mathcal{J}}\ \mathsf{S}(\nu_j) = \mathsf{S}(\mu)$.*

Proof. We proceed by induction on M.

- Suppose that $M = \mathsf{let}\ x = V\ \mathsf{in}\ N$, that $\mathscr{D} = \{\,(N[V/x])^1\,\}$, and that $\emptyset\,|\,\emptyset \vdash \mathsf{let}\ x = V\ \mathsf{in}\ N \;:\; \mu$. By Lemma 5.6, there exists (ξ, σ) such that $\emptyset\,|\,\emptyset \vdash V \;:\; \sigma$ and $x \;:\; \sigma\,|\,\emptyset \vdash N \;:\; \xi$ with $\xi \sqsubseteq \mu$. By Lemma 5.9, $\emptyset\,|\,\emptyset \vdash N[V/x] \;:\; \xi$, and since $\xi \sqsubseteq \mu$, we obtain by subtyping that $\emptyset\,|\,\emptyset \vdash N[V/x] \;:\; \mu$. It follows that $\{\,(N[V/x] \;:\; \mu)^1\,\}$ is a closed typed distribution satisfying the requirements of the lemma.
- Suppose that $M = (\lambda x.N)\ V$, that $\mathscr{D} = \{\,(N[V/x])^1\,\}$, and that $\emptyset\,|\,\emptyset \vdash (\lambda x.N)\ V \;:\; \mu$. Applying Lemma 5.6 twice, we obtain that $x \;:\; \tau\,|\,\emptyset \vdash N \;:\; \xi$ and $\emptyset\,|\,\emptyset \vdash V \;:\; \sigma$ with $\sigma \sqsubseteq \tau$ and $\xi \sqsubseteq \mu$. Applying subtyping to the second judgment gives $\emptyset\,|\,\emptyset \vdash V \;:\; \tau$, and we can apply Lemma 5.9 to obtain $\emptyset\,|\,\emptyset \vdash N[V/x] \;:\; \xi$. Since $\xi \sqsubseteq \mu$, we obtain by weakening that $\emptyset\,|\,\emptyset \vdash N[V/x] \;:\; \mu$. It follows that $\{\,(N[V/x] \;:\; \mu)^1\,\}$ is a closed typed distribution satisfying the requirements of the lemma.
- Suppose that $M = N \;\;\oplus_p\;\; L$, that $\mathscr{D} = [\,N^p, L^{1-p}\,]$, and that $\emptyset\,|\,\emptyset \vdash N \;\;\oplus_p\;\; L \;:\; \mu$. By Lemma 5.6, there exists (ξ, ρ) such that $\emptyset\,|\,\emptyset \vdash N \;:\; \xi$ and $\emptyset\,|\,\emptyset \vdash L \;:\; \rho$ with $\xi \oplus_p \rho \sqsubseteq \mu$ and $\sum(\xi \oplus_p \rho) = 1$. By Lemma 5.4, there exists (ξ', ρ') such that $\mu = \xi' \oplus_p \rho'$, $\xi \sqsubseteq \xi'$ and $\rho \sqsubseteq \rho'$. By subtyping, $\emptyset\,|\,\emptyset \vdash N \;:\; \xi'$ and $\emptyset\,|\,\emptyset \vdash L \;:\; \rho'$. We consider the closed typed distribution of pseudo-representation $[\,(N \;:\; \xi')^p,\ (L \;:\; \rho')^{1-p}\,]$, which satisfies the requirements of the lemma since its expectation type is $p \cdot \xi' + (1-p) \cdot \rho' = \xi' \oplus_p \rho' = \mu$. Note that we use a pseudo-representation to cope with the very specific case in which $N = L$ and $\xi' = \rho'$, in which the representation of the closed typed distribution is $\{\,(N \;:\; \xi')^1\,\}$.
- Suppose that $M = \mathsf{let}\ x = N\ \mathsf{in}\ L$, that $\mathscr{D} = \{\,(\mathsf{let}\ x = P_j\ \mathsf{in}\ L)^{p'_j} \;\mid\; j \in \mathcal{J}\,\}$, and that $\emptyset\,|\,\emptyset \vdash \mathsf{let}\ x = N\ \mathsf{in}\ L \;:\; \mu$. By Lemma 5.6, there exists $(\mathcal{I}, (\sigma_i)_{i\in\mathcal{I}}, (p_i)_{i\in\mathcal{I}}, (\xi_i)_{i\in\mathcal{I}})$ such that
  - $\sum_{i\in\mathcal{I}}\ p_i \cdot \xi_i \sqsubseteq \mu$,
  - $\emptyset\,|\,\emptyset \vdash N \;:\; \{\,\sigma_i^{p_i} \;\mid\; i \in \mathcal{I}\,\}$, and
  - $\forall i \in \mathcal{I},\;\; x \;:\; \sigma_i\,|\,\emptyset \vdash L \;:\; \xi_i$.

This reduction comes, by definition of $\to_v$, from $N \to_v \{ P_j^{p'_j} \mid j \in \mathcal{J} \}$, to which we can apply the induction hypothesis: there exists a closed typed distribution

$$\left\{ (R_k \,:\, \rho_k)^{p''_k} \;\middle|\; k \in \mathcal{K} \right\},$$

which is such that

$$\left\{ \sigma_i^{p_i} \;\middle|\; i \in \mathcal{I} \right\} = \sum_{k \in \mathcal{K}} p''_k \cdot \rho_k$$

and that $[(R_k)^{p''_k} \mid k \in \mathcal{K}]$ is a pseudo-representation of $\{ P_j^{p'_j} \mid j \in \mathcal{J} \}$. It follows that, for every $k \in \mathcal{K}$, we can write ρ_k as the pseudo-representation $[\sigma_i^{p'''_{ki}} \mid i \in \mathcal{I}]$ where some of the p'''_{ki} (but not all of them) may be worth 0. This implies that, for all $i \in \mathcal{I}$,

$$p_i = \sum_{k \in \mathcal{K}} p''_k p'''_{ki}.$$

Now, for every $k \in \mathcal{K}$, we can derive $\emptyset \mid \emptyset \vdash \mathsf{let}\ x = R_k\ \mathsf{in}\ L \,:\, \sum_{i \in \mathcal{I}} p'''_{ki} \cdot \xi_i$ from the rule

$$\frac{\emptyset \mid \emptyset \vdash R_k \,:\, \left\{ \sigma_i^{p'''_{ki}} \;\middle|\; i \in \mathcal{I} \right\} \qquad x \,:\, \sigma_i \mid \emptyset \vdash L \,:\, \xi_i \quad (\forall i \in \mathcal{I})}{\emptyset \mid \emptyset \vdash \mathsf{let}\ x = R_k\ \mathsf{in}\ L \,:\, \sum_{i \in \mathcal{I}} p'''_{ki} \cdot \xi_i}$$

so that $[(\mathsf{let}\ x = R_k\ \mathsf{in}\ L \,:\, \sum_{i \in \mathcal{I}} p'''_{ki} \cdot \xi_i)^{p''_k} \mid k \in \mathcal{K}]$ is a pseudo-representation of a closed typed distribution, whose expectation is

$$\sum_{k \in \mathcal{K}} p''_k \sum_{i \in \mathcal{I}} p'''_{ki} \cdot \xi_i = \sum_{i \in \mathcal{I}} \left(\sum_{k \in \mathcal{K}} p''_k p'''_{ki} \right) \cdot \xi_i = \sum_{i \in \mathcal{I}} p_i \cdot \xi_i.$$

By Lemma 5.6, the sum of $\sum_{i \in \mathcal{I}} p_i \cdot \xi_i$ is 1, and it follows that $\sum \mu = 1$ as well. Since $\sum_{i \in \mathcal{I}} p_i \cdot \xi_i \sqsubseteq \mu$, applying Corollary 5.5 gives us a family $(\nu_i)_{i \in \mathcal{I}}$ of distribution types such that, by subtyping, we can derive for every $k \in \mathcal{K}$ the judgment $\emptyset \mid \emptyset \vdash \mathsf{let}\ x = R_k\ \mathsf{in}\ L \,:\, \sum_{i \in \mathcal{I}} p'''_{ki} \cdot \nu_i$. This family $\overrightarrow{\nu}$ satisfies moreover $\sum_{i \in \mathcal{I}} p_i \cdot \nu_i = \mu$. We therefore consider the closed typed distribution of pseudo-representation

$$\left[\left(\mathsf{let}\ x = R_k\ \mathsf{in}\ L \,:\, \sum_{i \in \mathcal{I}} p'''_{ki} \cdot \nu_i \right)^{p''_k} \;\middle|\; k \in \mathcal{K} \right]$$

and of expectation type

$$\sum_{k \in \mathcal{K}} p''_k \cdot \left(\sum_{i \in \mathcal{I}} p'''_{ki} \cdot \nu_i \right) = \sum_{i \in \mathcal{I}} p_i \cdot \nu_i = \mu.$$

Since $[(R_k \,:\, \rho_k)^{p''_k} \mid k \in \mathcal{K}]$ is a pseudo-representation of $\{ P_j^{p'_j} \mid j \in \mathcal{J} \}$, we have that

$$\left[(\mathsf{let}\ x = R_k\ \mathsf{in}\ L \,:\, \rho_k)^{p''_k} \;\middle|\; k \in \mathcal{K} \right]$$

is a pseudo-representation of

$$\left\{ \left(\mathsf{let}\ x = P_j\ \mathsf{in}\ L \right)^{p'_j} \;\middle|\; j \in \mathcal{J} \right\},$$

which allows us to conclude.

- Suppose that $M = \mathsf{case}\ \mathsf{S}\ V\ \mathsf{of}\ \{\, \mathsf{S} \to W \ \mid\ 0 \to Z \,\}$, that $\mathscr{D} = \{\, (W\ V)^1 \,\}$, and that $\emptyset \mid \emptyset \vdash \mathsf{case}\ \mathsf{S}\ V\ \mathsf{of}\ \{\, \mathsf{S} \to W \mid 0 \to Z \,\} \ :\ \mu$. By Lemma 5.6, there exists $\mathfrak{s}$ and ξ such that $\emptyset \mid \emptyset \vdash \mathsf{S}\, V \ :\ \mathsf{Nat}^{\widehat{\mathfrak{s}}}$ and $\emptyset \mid \emptyset \vdash W \ :\ \mathsf{Nat}^{\mathfrak{s}} \to \xi$ with $\xi \sqsubseteq \mu$. Lemma 5.11 implies that $\emptyset \mid \emptyset \vdash V \ :\ \mathsf{Nat}^{\mathfrak{s}}$. Using an Application rule, we obtain that $\emptyset \mid \emptyset \vdash W\ V \ :\ \xi$, and subtyping gives $\emptyset \mid \emptyset \vdash W\ V \ :\ \mu$, allowing us to conclude for $\{\, (W\ V \ :\ \mu)^1 \,\}$.
- Suppose that $M = \mathsf{case}\ 0\ \mathsf{of}\ \{\, \mathsf{S} \to W \ \mid\ 0 \to Z \,\}$, that $\mathscr{D} = \{\, (Z)^1 \,\}$, and that

  $$\emptyset \mid \emptyset \vdash \mathsf{case}\ 0\ \mathsf{of}\ \{\, \mathsf{S} \to W \ \mid\ 0 \to Z \,\} \ :\ \mu.$$

  By Lemma 5.6, there exists ξ with $\xi \sqsubseteq \mu$ and such that $\emptyset \mid \emptyset \vdash Z \ :\ \xi$. By subtyping, $\emptyset \mid \emptyset \vdash Z \ :\ \mu$, which allows to conclude for $\{\, (Z \ :\ \mu)^1 \,\}$.
- Suppose that $M = (\mathsf{letrec}\ f = V)\ (c\ \overrightarrow{W})$, that $\mathscr{D} = \{\, (V[(\mathsf{letrec}\ f = V)/f]\ (c\ \overrightarrow{W}))^1 \,\}$, and that $\emptyset \mid \emptyset \vdash (\mathsf{letrec}\ f = V)\ (c\ \overrightarrow{W}) \ :\ \mu$. We apply again Lemma 5.6, but this time we rather depict the derivation typing M with μ it induces, for the sake of clarity. This derivation is of the form (modulo composition of subtyping rules):

  $$\dfrac{\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdots\\ \mathit{Hyp}\\ \emptyset \mid f \ :\ \{\, (\mathsf{Nat}^{\mathfrak{u}_j} \to \xi[\mathfrak{u}_j/\mathfrak{i}])^{p_j} \ \mid\ j \in \mathcal{J} \,\} \vdash V \ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \xi[\widehat{\mathfrak{i}}/\mathfrak{i}]\end{array}}{\emptyset \mid \emptyset \vdash \mathsf{letrec}\ f \ =\ V \ :\ \mathsf{Nat}^{\mathfrak{t}} \to \xi[\mathfrak{t}/\mathfrak{i}]}}{\emptyset \mid \emptyset \vdash \mathsf{letrec}\ f \ =\ V \ :\ \mathsf{Nat}^{\widehat{\mathfrak{s}}} \to \mu} \qquad \dfrac{\begin{array}{c}\pi_2\\ \vdots\\ \emptyset \mid \emptyset \vdash c\ \vec{W} \ :\ \mathsf{Nat}^{\widehat{\mathfrak{r}}}\end{array}}{\emptyset \mid \emptyset \vdash c\ \vec{W} \ :\ \mathsf{Nat}^{\widehat{\mathfrak{s}}}}}{\emptyset \mid \emptyset \vdash (\mathsf{letrec}\ f \ =\ V)\ \left(c\ \vec{W}\right) \ :\ \mu},$$

  where the two sizes appearing in the types for $c\ \overrightarrow{W}$ are successors due to Lemma 5.11, and where
  $-$*Hyp* denotes the additional premises of the letrec rule, and contains notably $\mathfrak{i}$ pos ξ,
  $-\mathfrak{r} \preccurlyeq \widehat{\mathfrak{r}} \preccurlyeq \widehat{\mathfrak{s}} \preccurlyeq \mathfrak{t}$, and
  $-\xi[\mathfrak{t}/\mathfrak{i}] \sqsubseteq \mu$.
  It follows that, for every $j \in \mathcal{J}$, we can deduce that the closed value $\mathsf{letrec}\ f = V$ has type $\mathsf{Nat}^{\mathfrak{u}_j} \to \xi[\mathfrak{u}_j/\mathfrak{i}]$, as proved by the derivation

  $$\dfrac{\begin{array}{c}\pi_1\\ \vdots\\ \mathit{Hyp}\\ \emptyset \mid f \ :\ \left\{ (\mathsf{Nat}^{\mathfrak{u}_j} \to \xi[\mathfrak{u}_j/\mathfrak{i}])^{p_j} \ \middle|\ j \in \mathcal{J} \right\} \vdash V \ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \xi\left[\widehat{\mathfrak{i}}/\mathfrak{i}\right]\end{array}}{\emptyset \mid \emptyset \ \vdash\ \mathsf{letrec}\ f \ =\ V \ :\ \mathsf{Nat}^{\mathfrak{u}_j} \to \xi\left[\mathfrak{u}_j/\mathfrak{i}\right]}.$$

  Since

  $$\emptyset \mid f \ :\ \left\{ (\mathsf{Nat}^{\mathfrak{u}_j} \to \xi[\mathfrak{u}_j/\mathfrak{i}])^{p_j} \ \middle|\ j \in \mathcal{J} \right\} \vdash V \ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \xi\left[\widehat{\mathfrak{i}}/\mathfrak{i}\right],$$

  we obtain by Lemma 5.10 that

  $$\emptyset \mid \emptyset \vdash V\left[(\mathsf{letrec}\ f = V)\,/f\right] \ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \xi\left[\widehat{\mathfrak{i}}/\mathfrak{i}\right].$$

  We now apply Lemma 5.15 to $\emptyset \mid \emptyset \vdash V[(\mathsf{letrec}\ f = V)/f] \ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \xi[\widehat{\mathfrak{i}}/\mathfrak{i}]$ with the substitution $[\mathfrak{r}/\mathfrak{i}]$ and we obtain that $\emptyset \mid \emptyset \vdash V[(\mathsf{letrec}\ f =)/V]f \ :\ \mathsf{Nat}^{\widehat{\mathfrak{r}}} \to \xi[\widehat{\mathfrak{r}}/\mathfrak{i}]$. Using the

Application rule, we derive $\emptyset \,|\, \emptyset \vdash V[(\mathsf{letrec}\ f = V)/f]\ (c\ \overrightarrow{W}) : \xi[\widehat{\mathfrak{r}}/\mathfrak{i}]$. Since $\mathfrak{i}$ pos ξ and $\widehat{\mathfrak{r}} \preccurlyeq \mathfrak{t}$, by Lemma 5.14, we get that $\xi[\widehat{\mathfrak{r}}/\mathfrak{i}] \sqsubseteq \xi[\mathfrak{t}/\mathfrak{i}]$. By transitivity of $\sqsubseteq$, $\xi[\widehat{\mathfrak{r}}/\mathfrak{i}] \sqsubseteq \mu$, which allows us to conclude by subtyping that $\emptyset \,|\, \emptyset \vdash V[(\mathsf{letrec}\ f = V)/f]\ (c\ \overrightarrow{W}) : \mu$. The result follows, for $\{\, (V[(\mathsf{letrec}\ f = V)/f]\ (c\ \overrightarrow{W}) : \mu)^1 \,\}$. □

Theorem 5.17 (Subject Reduction for $\to_v^n$). *Let $n \in \mathbb{N}$, and $\{\, (M_i : \mu_i)^{p_i} \mid i \in \mathcal{I} \,\}$ be a closed typed distribution. Suppose that $\{\, (M_i)^{p_i} \mid i \in \mathcal{I} \,\} \to_v^n \{\, (N_j)^{p'_j} \mid j \in \mathcal{J} \,\}$; then there exists a closed typed distribution $\{\, (L_k : \nu_k)^{p''_k} \mid k \in \mathcal{K} \,\}$ such that*

- $\mathbb{E}((M_i : \mu_i)^{p_i}) = \mathbb{E}((L_k : \nu_k)^{p''_k})$, *and*
- $[\, (L_k)^{p''_k} \mid k \in \mathcal{K} \,]$ *is a pseudo-representation of* $\{\, (N_j)^{p'_j} \mid j \in \mathcal{J} \,\}$.

Proof. The proof is by induction on n. For $n = 0$, $\to_v^0$ is the identity relation and the result is immediate. For $n + 1$, we have

$$\left\{ (M_i)^{p_i} \;\middle|\; i \in \mathcal{I} \right\} \;\to_v^n\; \left\{ (P_l)^{p''_l} \;\middle|\; l \in \mathcal{L} \right\} \;\to_v\; \left\{ (N_j)^{p'_j} \;\middle|\; j \in \mathcal{J} \right\}.$$

We apply the induction hypothesis and obtain a closed typed distribution $\{\, (R_g : \xi_g)^{p_g^{(3)}} \mid g \in \mathcal{G} \,\}$ satisfying $\mathbb{E}((M_i : \mu_i)^{p_i}) = \mathbb{E}((R_g : \xi_g)^{p_g^{(3)}})$ and such that $[\, (R_g)^{p_g^{(3)}} \mid g \in \mathcal{G} \,]$ is a pseudo-representation of $\{\, (P_l)^{p''_l} \mid l \in \mathcal{L} \,\}$. For every $g \in \mathcal{G}$:

- if R_g is a value, we set $\mathscr{D}_g = \{\, R^1 \,\}$ and $\mathscr{T}_g$ to be the closed typed distribution $\mathscr{T}_g = \{\, (T_h : \rho_h)^{p_h^{(4)}} \mid h \in \mathcal{H}_g \,\} = (R_g : \xi_g)^1$,
- else $R_g \to_v \mathscr{D}_g$. We apply Lemma 5.16 and obtain a closed typed distribution:

$$\mathscr{T}_g = \left\{ (T_h : \rho_h)^{p_h^{(4)}} \;\middle|\; h \in \mathcal{H}_g \right\}$$

  such that $\mathbb{E}((T_h : \rho_h)^{p_h^{(4)}}) = \xi_g$ and that $[\, (T_h)^{p_h^{(4)}} \mid h \in \mathcal{H}_g \,]$ is a pseudo-representation of $\mathscr{D}_g$.

We claim that the closed typed distribution defined as

$$\left\{ (L_k : \nu_k)^{p''_k} \;\middle|\; k \in \mathcal{K} \right\} = \sum_{g \in \mathcal{G}} p_g^{(3)} \cdot \mathscr{T}_g$$

satisfies the required conditions. Indeed, the expectation type is preserved:

$$\begin{aligned}
\mathbb{E}\Big((M_i : \mu_i)^{p_i}\Big) &= \mathbb{E}\left((R_g : \xi_g)^{p_g^{(3)}}\right) \\
&= \textstyle\sum_{g \in \mathcal{G}} p_g^{(3)} \cdot \xi_g \\
&= \textstyle\sum_{g \in \mathcal{G}} p_g^{(3)} \cdot \mathbb{E}\Big((T_h : \rho_h)^{p_h^{(4)}} \Big) \\
&= \mathbb{E}\left(\textstyle\sum_{g \in \mathcal{G}} p_g^{(3)} \cdot \mathscr{T}_g \right) \\
&= \mathbb{E}\Big(\Big\{ (L_k : \nu_k)^{p''_k} \;\Big|\; k \in \mathcal{K} \Big\}\Big).
\end{aligned}$$

Moreover, by definition of the family $(\mathscr{D}_g)_{g \in \mathcal{G}}$,

$$\{\, (P_l)^{p''_l} \mid l \in \mathcal{L} \,\} = \sum_{g \in \mathcal{G}} p_g^{(3)} \cdot \{\, (R_g)^1 \,\} \;\to_v\; \{\, (N_j)^{p'_j} \mid j \in \mathcal{J} \,\} = \sum_{g \in \mathcal{G}} p_g^{(3)} \cdot \mathscr{D}_g.$$

The result follows from the fact that $[\, (T_h)^{p_h^{(4)}} \mid h \in \mathcal{H}_g \,]$ is a pseudo-representation of $\mathscr{D}_g$ for every $g \in \mathcal{G}$. □

5.6 Subject Reduction for $\Rrightarrow_v$

Recall that there is an order $\preccurlyeq$ on distributions, defined pointwise.

Lemma 5.18. *Suppose that* $M \Rrightarrow_v \{ V_i^{p_i} \mid i \in \mathcal{I} \}$ *and that* $M \in \Lambda_\oplus^{\mathfrak{s}}(\mu)$. *Then there exists a closed typed distribution* $\{ (W_j : \sigma_j)^{p'_j} \mid j \in \mathcal{J} \}$ *such that*

- $\mathbb{E}((W_j : \sigma_j)^{p'_j}) \preccurlyeq \mu$, *and*
- $[(W_j)^{p'_j} \mid j \in \mathcal{J}]$ *is a pseudo-representation of* $\{ (V_i)^{p_i} \mid i \in \mathcal{I} \}$.

Proof. We have $M \rightarrow_v \mathscr{D} \overset{VD}{=} \mathscr{D}_{|T} + \{ V_i^{p_i} \mid i \in \mathcal{I} \}$. By Lemma 5.16, there exists a closed typed distribution $\{ (L_k : \nu_k)^{p''_k} \mid k \in \mathcal{K} \}$ such that $\mathbb{E}((L_k : \nu_k)^{p''_k}) = \mu$ and that $[(L_k)^{p''_k} \mid k \in \mathcal{K}]$ is a pseudo-representation of $\mathscr{D}$. We consider the pseudo-representation $[(W_j)^{p'_j} \mid j \in \mathcal{J}]$ obtained from $[(L_k)^{p''_k} \mid k \in \mathcal{K}]$ by removing all the terms that are not values. Note that $\mathcal{J} \subseteq \mathcal{K}$. We obtain in this way a pseudo-representation of $\{ V_i^{p_i} \mid i \in \mathcal{I} \}$, which induces a closed typed representation $\{ (W_j : \nu_j)^{p'_j} \mid j \in \mathcal{J} \}$ such that $\mathbb{E}((W_j : \nu_j)^{p'_j}) \preccurlyeq \mu$. □

Theorem 5.19 (Subject Reduction). *Let* $M \in \Lambda_\oplus^{\mathfrak{s}}(\mu)$. *Then there exists a closed typed distribution* $\{ (W_j : \sigma_j)^{p_j} \mid j \in \mathcal{J} \}$ *such that*

- $\mathbb{E}((W_j : \sigma_j)^{p_j}) \preccurlyeq \mu$, *and*
- $[(W_j)^{p_j} \mid j \in \mathcal{J}]$ *is a pseudo-representation of* $[\![M]\!]$.

Note that $\mathbb{E}((W_j : \sigma_j)^{p_j}) \preccurlyeq \mu$ since the semantics of a term may not be a proper distribution at this stage. In fact, it will follow from the soundness theorem of Section 6 that the typability of M implies that $\sum [\![M]\!] = 1$ and thus that the previous statement is an equality.

6 TYPABILITY IMPLIES TERMINATION: REDUCIBILITY STRIKES AGAIN

This section is the most technically advanced of the article and proves that the typing discipline we have introduced indeed enforces almost-sure termination. As already mentioned, the technique we will employ is a substantial generalization of Girard-Tait's reducibility. In particular, reducibility must be made quantitative, in that terms can be said to be reducible *with a certain probability*. This means that reducibility sets will be defined as sets parameterized by a real number p, called the *degree of reducibility* of the set. As Lemma 6.4 will emphasize, this degree of reducibility ensures that terms contained in a reducibility set parameterized by p terminate with probability at least p. These "intermediate" degrees of reducibility are required to handle the fix-point construction and show that recursively defined terms that are typable are indeed AST—that is, that they belong to the appropriate reducibility set, parameterized by 1.

6.1 Reducibility Sets for Closed Terms

The first preliminary notion we need is that of a size environment:

Definition 6.1 (Size Environment). A *size environment* is any function ρ from $\mathcal{S}$ to $\mathbb{N} \cup \{\infty\}$. Given a size environment ρ and a size expression $\mathfrak{s}$, there is a naturally defined element of $\mathbb{N} \cup \{\infty\}$, which we indicate as $[\![\mathfrak{s}]\!]_\rho$:

- $[\![\widehat{\mathfrak{i}}^{\,n}]\!]_\rho = \rho(\mathfrak{i}) + n$,
- $[\![\infty]\!]_\rho = \infty$.

In other words, the purpose of size environments is to give a semantic meaning to size expressions. Our reducibility sets will be parameterized not only on a probability but also on a size environment.

Definition 6.2 (Reducibility Sets).

- For values of simple type Nat, we define the reducibility sets
$$\mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{s}},\rho} = \left\{ \mathsf{S}^{n}\ 0 \;\middle|\; p > 0 \Longrightarrow n < [\![\mathfrak{s}]\!]_{\rho} \right\}.$$
- Values of higher-order type are in a reducibility set when their applications to appropriate values are reducible terms, with an adequate degree of reducibility:
$$\mathsf{VRed}^{p}_{\sigma\to\mu,\rho} = \left\{ V \in \Lambda^{V}_{\oplus}\left(\langle \sigma \to \mu \rangle\right) \;\middle|\; \forall q \in (0,1],\ \ \forall W \in \mathsf{VRed}^{q}_{\sigma,\rho},\ \ V\ W \in \mathsf{TRed}^{pq}_{\mu,\rho} \right\}.$$
- Distributions of values are reducible with degree p when they consist of values that are themselves globally reducible "enough." Formally, $\mathsf{DRed}^{p}_{\mu,\rho}$ is the set of finite distributions of values—in the sense that they have a finite support—admitting a pseudo-representation $\mathscr{D} = [\,(V_i)^{p_i} \;\mid\; i \in \mathcal{I}\,]$ such that, setting $\mu = \{\,(\sigma_j)^{p'_j} \;\mid\; j \in \mathcal{J}\,\}$, there exists a family $(p_{ij})_{i\in\mathcal{I},j\in\mathcal{J}} \in [0,1]^{|\mathcal{I}|\times|\mathcal{J}|}$ of probabilities and a family $(q_{ij})_{i\in\mathcal{I},j\in\mathcal{J}} \in [0,1]^{|\mathcal{I}|\times|\mathcal{J}|}$ of degrees of reducibility, satisfying:

  (1) $\forall i \in \mathcal{I},\ \ \forall j \in \mathcal{J},\ \ V_i \in \mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho}$;
  (2) $\forall i \in \mathcal{I},\ \ \sum_{j\in\mathcal{J}}\ p_{ij} = p_i$;
  (3) $\forall j \in \mathcal{J},\ \ \sum_{i\in\mathcal{I}}\ p_{ij} = \mu(\sigma_j)$; and
  (4) $p \leq \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}}\ q_{ij}p_{ij}$.

  Note that (2) and (3) imply that $\sum \mathscr{D} = \sum \mu$. We say that $[\,(V_i)^{p_i} \mid i \in \mathcal{I}\,]$ *witnesses* that $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu,\rho}$.
- A term is reducible with degree p when its finite approximations compute distributions of values of the degree of reducibility arbitrarily close to p:
$$\mathsf{TRed}^{p}_{\mu,\rho} = \Big\{ M \in \Lambda_{\oplus}\left(\langle \mu \rangle\right) \;\Big|\; \forall 0 \leq r < p,\quad \exists \nu_r \preccurlyeq \mu,\ \ \exists n_r \in \mathbb{N},\quad M \Rrightarrow^{n_r}_{v} \mathscr{D}_r \text{ and } \mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho} \Big\}.$$

  Note that here, unlike the case of DRed, the fact that $M \in \Lambda_{\oplus}(\langle\mu\rangle)$ implies that μ is proper.

The first thing to observe about reducibility sets as given in Definition 6.2 is that they only deal with closed terms, and not with arbitrary terms. As such, we cannot rely *directly* on them when proving AST for typable terms, at least if we want to prove it by induction on the structure of type derivations. We will therefore define in Section 6.9 an extension of these sets to open terms, which will be based on these sets of closed terms, and therefore enjoy similar properties. Before embarking in the proof that typability implies reducibility, it is convenient to prove some fundamental properties of reducibility sets, which inform us about how these sets are structured, and which will be crucial in the sequel. This is the purpose of the following subsections.

As a preliminary, the following easy lemma relates the reducibility of natural numbers and will be used to treat the case of the rules Succ and Zero in the proof of typing soundness:

Lemma 6.3.

- $V \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{s}},\rho} \quad\Longrightarrow\quad \mathsf{S}\ V \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\widehat{\mathfrak{s}}},\rho}$.
- *For every size* $\mathfrak{s}$, $0 \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\widehat{\mathfrak{s}}},\rho}$.

Proof. First point:

- Suppose that $V \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\widehat{\mathfrak{i}}^{k}},\rho}$ and that $p > 0$. Then $V = \mathsf{S}^{n}\ 0$ for some $n < [\![\mathfrak{s}]\!]_{\rho}$. Then $\mathsf{S}\ V = \mathsf{S}^{n+1}\ 0$ satisfies $n+1 < [\![\widehat{\mathfrak{s}}]\!]_{\rho} = [\![\mathfrak{s}]\!]_{\rho} + 1$, so that $\mathsf{S}\ V \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\widehat{\mathfrak{i}}^{k+1}},\rho}$.

- Suppose that $V \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\infty},\rho}$ or that $p = 0$. By definition, $V = \mathsf{S}^n\ 0$ for $n \in \mathbb{N}$. It follows that $\mathsf{S}\ V \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\widehat{\mathfrak{s}}},\rho}$.

Second point:

- Suppose that $p = 0$. Then $0 \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\widehat{\mathfrak{s}}},\rho}$, by definition.
- Else we need to prove that $[\![\mathfrak{s}]\!]_\rho > 0$. But $\widehat{\mathfrak{s}}$ either is ∞, in which case $[\![\mathfrak{s}]\!]_\rho = \infty$, or is of the shape $\widehat{\mathfrak{i}}^{k}$ for $k > 0$, and $[\![\widehat{\mathfrak{i}}^{k}]\!]_\rho = \rho(\mathfrak{i}) + k > 0$. □

6.2 Reducibility Sets and Termination

The following lemma, relatively easy to prove, is crucial for the understanding of the reducibility sets, for it shows that the degree of reducibility of a term gives information on the sum of its operational semantics:

Lemma 6.4 (Reducibility and Termination).

- *Let $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu,\rho}$. Then $\sum \mathscr{D} \geq p$.*
- *Let $M \in \mathsf{TRed}^{p}_{\mu,\rho}$. Then $\sum [\![M]\!] \geq p$.*

Proof.

- Let $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu,\rho}$. Then there exists a pseudo-representation $\mathscr{D} = [\,(V_i)^{p_i} \mid i \in \mathcal{I}\,]$ and families $(p_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ and $(q_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ of reals of $[0,1]$ such that $\forall i \in \mathcal{I},\ \ \sum_{j\in\mathcal{J}}\ p_{ij} = p_i$, and that $p \leq \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}}\ q_{ij}p_{ij}$. We therefore have
$$\sum \mathscr{D} = \sum_{i\in\mathcal{I}} p_i = \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} p_{ij} \quad \geq \quad \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} q_{ij}p_{ij} \quad \geq \quad p.$$
- Since $M \in \mathsf{TRed}^{p}_{\mu,\rho}$, for every $0 \leq r < p$, there exists n_r with $M \Rrightarrow_v^{n_r} \mathscr{D}_r$ and $\mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$. From the previous point, we get that $\sum \mathscr{D}_r \geq r$ for every $0 \leq r < p$. It follows from Corollary 3.10 that $\sum [\![M]\!] \geq r$ for every $0 \leq r < p$ and, by taking the supremum, $\sum [\![M]\!] \geq p$. □

It follows from this lemma that terms with degree of reducibility 1 are AST:

Corollary 6.5 (Reducibility and AST). *Let $M \in \mathsf{TRed}^{1}_{\mu,\rho}$. Then M is AST.*

6.3 Reducibility Sets and Reducibility Degrees

We now prove two results related to the reducibility degrees of reducibility sets. First of all, if the degree of reducibility p is 0, then no assumption is made on the probability of termination of terms, distributions, or values. It follows that the three kinds of reducibility sets collapse to the set of all affinely simply typable terms, distributions, or values:

Lemma 6.6 (Candidates of Null Reducibility).

- *If $V \in \Lambda^{V}_{\oplus}(\kappa)$, then $V \in \mathsf{VRed}^{0}_{\sigma,\rho}$ for every σ such that $\langle\sigma\rangle = \kappa$ and every size environment ρ.*
- *Let $\mathscr{D} = \{\,(V_i)^{p_i} \mid i \in \mathcal{I}\,\}$ be a finite distribution of values. If $\forall i \in \mathcal{I},\ \ V_i \in \Lambda^{V}_{\oplus}(\kappa)$, then $\mathscr{D} \in \mathsf{DRed}^{0}_{\mu,\rho}$ for every μ such that $\langle\mu\rangle = \kappa$ and $\sum \mu = \sum \mathscr{D}$ and every ρ.*
- *If $M \in \Lambda_{\oplus}(\kappa)$, then $M \in \mathsf{TRed}^{0}_{\mu,\rho}$ for μ such that $\langle\mu\rangle = \kappa$ and every ρ.*

Structure of the proof. In this lemma, as for most lemmas proving properties about VRed, DRed, and TRed, we use a proof by induction on types. As the property is defined in a mutual way over VRed, DRed, and TRed, we typically prove it for $\mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{s}},\rho}$ for any size $\mathfrak{s}$ refining Nat, and then for $\mathsf{VRed}^{p}_{\sigma\to\mu,\rho}$ by using the associated hypothesis on $\mathsf{TRed}^{p}_{\mu,\rho}$. Then we prove the property for any distribution type for $\mathsf{DRed}^{p}_{\mu,\rho}$ using the induction hypothesis on the $\mathsf{VRed}^{p}_{\sigma,\rho}$ for $\sigma \in \mathsf{S}(\mu)$, and we prove it for $\mathsf{TRed}^{p}_{\mu,\rho}$ using the induction hypothesis on $\mathsf{VRed}^{p}_{\sigma,\rho}$. The point is that these ingredients allow one to give a proof by induction on the simple type underlying the sized type of interest. In the base case, the sized type is necessarily of the form $\mathsf{Nat}^{\mathfrak{s}}$ for some size $\mathfrak{s}$: we prove the statement on $\mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{s}},\rho}$ for all these sized types without using any induction-like hypothesis. Then we prove the statement for distribution types $\mu = \{\, (\mathsf{Nat}^{\mathfrak{s}_i})^{p_i} \mid i \in \mathcal{I} \,\}$ first on $\mathsf{DRed}^{p}_{\mu,\rho}$ by using the results for the sets $\mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{s}_i},\rho}$. Then we prove the result for $\mathsf{TRed}^{p}_{\mu,\rho}$ typically using the one for $\mathsf{DRed}^{p}_{\mu,\rho}$.

We then switch to higher-order types and give the proof for $\mathsf{VRed}^{p}_{\sigma\to\mu,\rho}$, which may use the results for the other sets on types σ and μ. Typically, only results on $\mathsf{TRed}^{p}_{\mu,\rho}$ are used. Then the proofs for $\mathsf{DRed}^{p}_{\sigma\to\mu,\rho}$ and $\mathsf{TRed}^{p}_{\sigma\to\mu,\rho}$ are typically the same as in the case of distributions over sized types refining Nat: therefore, we do not write them again.

This proof scheme will become more clear with the proof of this lemma on candidates of null reducibility:

Proof.

- Let $V \in \Lambda^{V}_{\oplus}(\mathsf{Nat})$. Every $\sigma :: \mathsf{Nat}$ is of the shape $\sigma = \mathsf{Nat}^{\mathfrak{s}}$ for a size $\mathfrak{s}$. Let ρ be a size environment. By inspection of the grammar of values and of the simple type system, we see that V must be of the shape $\mathsf{S}^{n}\ 0$ for $n \in \mathbb{N}$. Note that V is closed: it cannot be a variable. By definition, $V \in \mathsf{VRed}^{0}_{\sigma,\rho}$.
- Let $\kappa = \kappa' \to \kappa''$ be a higher-order type, with $\sigma :: \kappa'$ and $\mu :: \kappa''$. Let ρ be a size environment, and $V \in \Lambda^{V}_{\oplus}(\kappa)$. Let $q \in (0,1]$ and $W \in \mathsf{VRed}^{q}_{\sigma,\rho}$; we need to prove that $V\ W \in \mathsf{TRed}^{0}_{\mu,\rho}$. But, by definition of $\mathsf{VRed}^{q}_{\sigma,\rho}$, $W \in \Lambda^{V}_{\oplus}(\kappa')$. It follows that $V\ W \in \Lambda_{\oplus}(\kappa'')$, and we can apply the induction hypothesis to deduce that $V\ W \in \mathsf{TRed}^{0}_{\mu,\rho}$, so that by definition $V \in \mathsf{VRed}^{0}_{\sigma,\rho}$.
- Let $\mathscr{D} = \{\, (V_i)^{p_i} \mid i \in \mathcal{I} \,\}$ be a distribution of values and $\mu = \{\, (\sigma_j)^{p'_j} \mid j \in \mathcal{J} \,\} :: \kappa$ be a distribution type. Suppose that $\forall i \in \mathcal{I},\ V_i \in \Lambda^{V}_{\oplus}(\kappa)$. Let ρ be a size environment. For every $(i,j) \in \mathcal{I} \times \mathcal{J}$, we set $p_{ij} = \frac{p_i p'_j}{\sum \mu}$ and $q_{ij} = 0$. We consider the canonical pseudo-representation $\mathscr{D} = [\, (V_i)^{p_i} \mid i \in \mathcal{I} \,]$ and check the four conditions to be in $\mathsf{DRed}^{0}_{\mu,\rho}$:
  (1) $\forall i \in \mathcal{I},\ \forall j \in \mathcal{J},\ V_i \in \mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho}$: this is obtained by induction hypothesis.
  (2) $\forall i \in \mathcal{I},\ \sum_{j\in\mathcal{J}} p_{ij} = p_i$: let $i \in \mathcal{I}$; we have $\sum_{j\in\mathcal{J}} p_{ij} = \frac{p_i}{\sum \mu} \sum_{j\in\mathcal{J}} p'_j = \frac{p_i}{\sum \mu} \times \sum \mu = p_i$.
  (3) $\forall j \in \mathcal{J},\ \sum_{i\in\mathcal{I}} p_{ij} = \mu(\sigma_j)$: let $j \in \mathcal{J}$; we have $\sum_{i\in\mathcal{I}} p_{ij} = \frac{p'_j}{\sum \mu} \sum_{i\in\mathcal{I}} p_i = \frac{p'_j}{\sum \mu} \times \sum \mathscr{D}$. But $\sum \mu = \sum \mathscr{D}$ so that the sum equals p'_j as requested.
  (4) $p \leq \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} q_{ij} p_{ij}$: this amounts to $0 \leq 0$, which holds.
- Let $M \in \Lambda_{\oplus}(\kappa)$ and $\mu :: \kappa$. Let ρ be a size environment. Then $M \in \mathsf{TRed}^{0}_{\mu,\rho}$: the condition on M in the definition of $\mathsf{TRed}^{0}_{\mu,\rho}$ is for any $0 \leq r < 0$ so that it's an empty condition in this case. □

As p gives us a lower bound on the sum of the semantics of terms, it is easily guessed that a term having degree of reducibility p must also have degree of reducibility $q < p$. The following lemma makes this statement precise:

Lemma 6.7 (Downward Closure). *Let σ be a sized type, μ be a distribution type, and ρ be a size environment. Let $0 \leq q < p \leq 1$. Then:*

- *For any value V, $V \in \mathsf{VRed}^{p}_{\sigma,\rho} \implies V \in \mathsf{VRed}^{q}_{\sigma,\rho}$.*
- *For any finite distribution of values $\mathscr{D}$, $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu,\rho} \implies \mathscr{D} \in \mathsf{DRed}^{q}_{\mu,\rho}$.*
- *For any term M, $M \in \mathsf{TRed}^{p}_{\mu,\rho} \implies M \in \mathsf{TRed}^{q}_{\mu,\rho}$.*

Proof. Let σ be a sized type, μ be a distribution type, and ρ be a size environment. If $q = 0$, the result is immediate as a consequence of Lemma 6.6. Let $0 < q < p \leq 1$.

- Suppose that $V \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{s}},\rho}$. Since by definition $p,\ q > 0 \implies \mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{s}},\rho} = \mathsf{VRed}^{q}_{\mathsf{Nat}^{\mathfrak{s}},\rho}$, and the result holds.
- Suppose that $V \in \mathsf{VRed}^{p}_{\sigma\to\mu,\rho}$. Then:

$$\begin{array}{ll} & V \in \mathsf{VRed}^{p}_{\sigma\to\mu,\rho} \\ \iff & \forall q \in (0,1],\ \ \forall W \in \mathsf{VRed}^{q}_{\sigma,\rho},\ \ VW \in \mathsf{TRed}^{pq}_{\mu,\rho} \\ \implies & \forall q' \in (0,1],\ \ \forall W \in \mathsf{VRed}^{q'}_{\sigma,\rho},\ \ VW \in \mathsf{TRed}^{qq'}_{\mu,\rho} \qquad \text{(by IH, since } 0 < qq' < pq \leq 1) \\ \iff & V \in \mathsf{VRed}^{q}_{\sigma\to\mu,\rho}. \end{array}$$

- Suppose that $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu,\rho}$. Then there are a pseudo-representation $\mathscr{D} = [\,(V_i)^{p_i} \mid i \in \mathcal{I}\,]$ and families of reals $(p_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ and $(q_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ satisfying conditions (1) through (4). We have $\mathscr{D} \in \mathsf{DRed}^{q}_{\mu,\rho}$ for the same pseudo-representation, since conditions (1) through (3) are the same, and (4) holds as well since $q < p$.
- Suppose that $M \in \mathsf{TRed}^{p}_{\mu,\rho}$. Then for every $0 \leq r < p$, there exists $\nu_r \preccurlyeq \mu$ and $n_r \in \mathbb{N}$ with $M \Rrightarrow^{n_r}_{v} \mathscr{D}_r$ and $\mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$. So this statement also holds for every $0 \leq r < q$ and $M \in \mathsf{TRed}^{q}_{\mu,\rho}$. □

6.4 Continuity of the Reducibility Sets

To prove the lemma of continuity on the reducibility sets, which says that if an element is in all the reducibility sets for degrees $q < p$ then it is also in the set parameterized by the degree p, we use the following companion lemma computing a family of probabilities maximizing the degree of reducibility of a distribution:

Lemma 6.8 (Maximizing the Degree of Reducibility of a Distribution). *Let $\mathscr{D} = [\,(V_i)^{p_i} \mid i \in \mathcal{I}\,]$ be a finite distribution of values and $\mu = \{\,(\sigma_j)^{p'_j} \mid j \in \mathcal{J}\,\}$ be a distribution type. Set $q_{ij} = \max\{q \mid V_i \in \mathsf{VRed}^{q}_{\sigma_j,\rho}\}$ for every $(i,j) \in \mathcal{I} \times \mathcal{J}$. Then there exists a family $(p_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ of reals of $[0,1]$ satisfying:*

(1) $\forall i \in \mathcal{I},\ \ \sum_{j\in\mathcal{J}}\ p_{ij} = p_i$, and
(2) $\forall j \in \mathcal{J},\ \ \sum_{i\in\mathcal{I}}\ p_{ij} = \mu(\sigma_j)$,

and which maximizes $\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}}\ q_{ij}p_{ij}$.

Proof. We use the theory of linear programming in the finite real vector space $\mathbb{R}^n$, taking [38] as a reference. We stick to the notations of this book. The problem then amounts to showing the existence of

$$\max\left\{cx \;\middle|\; x \geq \overrightarrow{0},\ Ax = b\right\}, \tag{12}$$

where, supposing that we can index vectors and matrices by $i \times j$ thanks to a bijection $i \times j \to \{1, \ldots, n\}$, where $n = \#(\mathcal{I} \times \mathcal{J})$:

- x is the column vector indexed by the finite set $\mathcal{I} \times \mathcal{J}$, where x_{ij} plays the role of p_{ij};
- c is the row vector indexed by $\mathcal{I} \times \mathcal{J}$, with $c_{ij} = \max\{q \mid V_i \in \mathsf{VRed}^{q}_{\sigma_j,\rho}\}$;
- $\overrightarrow{0}$ is the null column vector of size $\#(\mathcal{I} \times \mathcal{J})$;
- A is the matrix with columns indexed by $\mathcal{I} \times \mathcal{J}$ and rows indexed by $\mathcal{I} + \mathcal{J}$, and such that:
  - $a_{i',(i,j)} = 1$ if and only if $i = i'$, and 0 else, and
  - $a_{j',(i,j)} = 1$ if and only if $j = j'$, and 0 else;
- b is the column vector indexed by $\mathcal{I} + \mathcal{J}$ and such that $b_i = p_i$ and $b_j = \mu(\sigma_j)$.

Following [38, Section 7.4], the maximum (Equation (12)) exists if and only if:

- the problem is *feasible*: its constraints admit a solution, and
- if it is *bounded*: there should be an upper bound over Equation (12),

and also, its existence is equivalent to the one of the maximum of the following problem:

$$\max \left\{ cx \;\middle|\; x \geq \overrightarrow{0},\; Ax \leq b \right\}. \tag{13}$$

This reformulation makes the feasibility immediate for the null vector $x = \overrightarrow{0}$. It is also immediate to see that the problem is bounded: by construction, all the $q_{ij} \in [0,1]$, and $\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \; p_{ij} = 1$ so that $\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \; c_{ij} p_{ij} \leq 1$. The existence of the maximum in Equation (12) follows, and the lemma therefore holds. □

It follows that a distribution has a maximal degree of reducibility: the supremum of the degrees of reducibility is again a degree of reducibility:

Corollary 6.9 (Maximizing the Degree of Reducibility of a Distribution II). *Let $\mathscr{D}$ be a finite distribution of values, μ be a distribution type, and ρ be a size environment. Suppose that $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu,\rho}$ for some real $p \in [0,1]$. Then there exists a maximal real $p_{max} \in [p,1]$ such that $\mathscr{D} \in \mathsf{DRed}^{p_{max}}_{\mu,\rho}$ and $p' > p_{max} \Rightarrow \mathscr{D} \notin \mathsf{DRed}^{p'}_{\mu,\rho}$.*

Proof. Let $\mathscr{D} = [\, (V_i)^{p_i} \mid i \in \mathcal{I} \,]$ be a finite distribution of values and $\mu = \{\, (\sigma_j)^{p'_j} \mid j \in \mathcal{J} \,\}$ be a distribution type. By Lemma 6.8, setting $q_{ij} = \max\{q \mid V_i \in \mathsf{VRed}^{q}_{\sigma_j,\rho}\}$ for every $(i,j) \in \mathcal{I} \times \mathcal{J}$, there exists a family $(p_{ij})_{i \in \mathcal{I}, j \in \mathcal{J}}$ of reals of $[0,1]$ that maximizes $w = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \; q_{ij} p_{ij}$. It is immediate to see that any increase of a q_{ij} to q' is contradictory with $V_i \in \mathsf{VRed}^{q'}_{\sigma_j,\rho}$, and that any decrease of a q_{ij} cannot increase w. It follows that $p_{max} = w$. □

To analyze the letrec construction, we will prove that, for every $\varepsilon \in (0,1]$, performing enough unfoldings of the fix point allows to prove that the recursively defined term is in a reducibility set parameterized by $1 - \varepsilon$. We will be able to conclude on the AST nature of recursive constructions using the following continuity lemma, proved using Corollary 6.9 and thus using the theory of linear programming:

Lemma 6.10 (Continuity). *Let σ be a sized type, μ be a distribution type, and ρ be a size environment. Let $p \in (0,1]$. Then:*

- $\mathsf{VRed}^{p}_{\sigma,\rho} = \bigcap_{0<q<p} \; \mathsf{VRed}^{q}_{\sigma,\rho}$,
- $\mathsf{DRed}^{p}_{\mu,\rho} = \bigcap_{0<q<p} \; \mathsf{DRed}^{q}_{\mu,\rho}$, *and*
- $\mathsf{TRed}^{p}_{\mu,\rho} = \bigcap_{0<q<p} \; \mathsf{TRed}^{q}_{\mu,\rho}$.

Proof. Let σ be a sized type, μ be a distribution type, and ρ be a size environment. Let $p \in (0,1]$.

- If $\sigma = \mathsf{Nat}^{\mathfrak{s}}$ for some size $\mathfrak{s}$, then for every $0 < q < p$ we have $\mathsf{VRed}^{q}_{\sigma,\rho} = \mathsf{VRed}^{p}_{\sigma,\rho}$ so that $\mathsf{VRed}^{p}_{\sigma,\rho} = \bigcap_{0<q<p} \mathsf{VRed}^{q}_{\sigma,\rho}$.
- If $\sigma = \tau \to \mu$, we proceed by mutual inclusions.
  —$\mathsf{VRed}^{p}_{\sigma,\rho} \subseteq \bigcap_{0<q<p} \mathsf{VRed}^{q}_{\sigma,\rho}$ is an immediate consequence of Lemma 6.7.
  —Let us prove now that $\bigcap_{0<q<p} \mathsf{VRed}^{q}_{\sigma,\rho} \subseteq \mathsf{VRed}^{p}_{\sigma,\rho}$. Let $V \in \bigcap_{0<q<p} \mathsf{VRed}^{q}_{\sigma,\rho}$; it follows that

$$\begin{array}{rl} & \forall q \in (0,p),\ \ \forall q' \in [0,1],\ \ \forall W \in \mathsf{VRed}^{q'}_{\sigma,\rho},\ \ V\ W \in \mathsf{TRed}^{qq'}_{\mu,\rho} \\ \Longrightarrow & \forall q' \in [0,1],\ \ \forall W \in \mathsf{VRed}^{q'}_{\sigma,\rho},\ \ \forall q \in (0,p),\ \ V\ W \in \mathsf{TRed}^{qq'}_{\mu,\rho} \\ \Longrightarrow & \forall q' \in [0,1],\ \ \forall W \in \mathsf{VRed}^{q'}_{\sigma,\rho},\ \ V\ W \in \bigcap_{0<q<p} \mathsf{TRed}^{qq'}_{\mu,\rho}. \end{array}$$

  But
$$\bigcap_{0<q<p} \mathsf{TRed}^{qq'}_{\mu,\rho} = \bigcap_{0<r<pq'} \mathsf{TRed}^{r}_{\mu,\rho} = \mathsf{TRed}^{pq'}_{\mu,\rho} \qquad \text{(by IH)}$$
  so that
$$\forall q' \in [0,1],\ \ \forall W \in \mathsf{VRed}^{q'}_{\sigma,\rho},\ \ V\ W \in \mathsf{TRed}^{pq'}_{\mu,\rho}.$$
  By definition, $V \in \mathsf{VRed}^{p}_{\sigma,\rho}$.
- The inclusion $\mathsf{DRed}^{p}_{\mu,\rho} \subseteq \bigcap_{0<q<p} \mathsf{DRed}^{q}_{\mu,\rho}$ is an immediate consequence of Lemma 6.7. Let $\mathscr{D} \in \bigcap_{0<q<p} \mathsf{DRed}^{q}_{\mu,\rho}$. Let $(q_n)_{n\in\mathbb{N}}$ be an increasing sequence of reals of $[0,p)$ converging to p. For every $n \in \mathbb{N}$, $\mathscr{D} \in \mathsf{DRed}^{q_n}_{\mu,\rho}$ so that by Corollary 6.9 there exists a real $p_{max,n} \in [q_n, 1]$ such that $\mathscr{D} \in \mathsf{DRed}^{p_{max,n}}_{\mu,\rho}$ and $p' > p_{max,n} \Rightarrow \mathscr{D} \notin \mathsf{DRed}^{p'}_{\mu,\rho}$. It follows that all the $p_{max,n}$ coincide, and that they are greater than $\sup_{n\in\mathbb{N}} q_n = p$. So $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu,\rho}$.
- The inclusion $\mathsf{TRed}^{p}_{\mu,\rho} \subseteq \bigcap_{0<q<p} \mathsf{TRed}^{q}_{\mu,\rho}$ is an immediate consequence of Lemma 6.7. Let $M \in \bigcap_{0<q<p} \mathsf{TRed}^{q}_{\mu,\rho}$. We need to prove that $M \in \mathsf{TRed}^{p}_{\mu,\rho}$, that is, that for every $0 \leq r < p$ there exists $\nu_r \preccurlyeq \mu$, $n_r \in \mathbb{N}$, $\mathscr{D}_r$ such that $M \Rrightarrow^{n_r}_{v} \mathscr{D}_r$ and that $\mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$. Let $r \in [0,p)$. Since $M \in \mathsf{TRed}^{\frac{p+r}{2}}_{\mu,\rho}$ and $\frac{p+r}{2} > r$, we obtain the desired $\nu_r \preccurlyeq \mu$, $n_r \in \mathbb{N}$, $\mathscr{D}_r$ having the properties of interest. The result follows. □

6.5 Reducibility Sets and Sizes

In this subsection, we show how the sizes appearing in the (sized or distribution) type parameterizing a reducibility set relate with the interpretation of size variables contained in the size environment that also parameterizes it. We prove first the following lemma, which will be used as a companion for this result:

Lemma 6.11 (Commuting Sizes with Environments). *Let $\mathfrak{i}$ be a size variable; $\mathfrak{s}$, $\mathfrak{r}$ be two sizes; and ρ be a size environment. Suppose that $\mathfrak{s} = \infty$ or that* $\mathsf{spine}(\mathfrak{s}) \neq \mathfrak{i}$. *Then $\llbracket \mathfrak{r}[\mathfrak{s}/\mathfrak{i}] \rrbracket_{\rho} = \llbracket \mathfrak{r} \rrbracket_{\rho[\mathfrak{i} \mapsto \llbracket \mathfrak{s} \rrbracket_{\rho}]}$.*

Proof. By case analysis.

- If $\mathfrak{r} = \widehat{\mathfrak{j}}^{n}$ for $\mathfrak{j} \neq \mathfrak{i}$, then $\mathfrak{r}[\mathfrak{s}/\mathfrak{i}] = \mathfrak{r}$ and $\llbracket \mathfrak{r} \rrbracket_{\rho} = \rho(\mathfrak{j}) + n = \llbracket \mathfrak{r} \rrbracket_{\rho[\mathfrak{i} \mapsto \llbracket \mathfrak{s} \rrbracket_{\rho}]}$.
- If $\mathfrak{r} = \widehat{\mathfrak{i}}^{n}$, then
  —if $\mathfrak{s} = \widehat{\mathfrak{j}}^{m}$ for $\mathfrak{j} \neq \mathfrak{i}$, then $\mathfrak{r}[\mathfrak{s}/\mathfrak{i}] = \widehat{\mathfrak{j}}^{n+m}$ and
$$\llbracket \mathfrak{r}\,[\mathfrak{s}/\mathfrak{i}] \rrbracket_{\rho} = \rho(\mathfrak{j}) + n + m = \llbracket \widehat{\mathfrak{j}}^{m} \rrbracket_{\rho} + n = \llbracket \mathfrak{s} \rrbracket_{\rho} + n = \llbracket \widehat{\mathfrak{i}}^{n} \rrbracket_{\rho[\mathfrak{i} \mapsto \llbracket \mathfrak{s} \rrbracket_{\rho}]} = \llbracket \mathfrak{r} \rrbracket_{\rho[\mathfrak{i} \mapsto \llbracket \mathfrak{s} \rrbracket_{\rho}]};$$
  —if $\mathfrak{s} = \infty$, then $\mathfrak{r}[\mathfrak{s}/\mathfrak{i}] = \infty$ and $\llbracket \mathfrak{r}[\mathfrak{s}/\mathfrak{i}] \rrbracket_{\rho} = \infty = \llbracket \mathfrak{r} \rrbracket_{\rho[\mathfrak{i} \mapsto \llbracket \mathfrak{s} \rrbracket_{\rho}]}$.
- If $\mathfrak{r} = \infty$, then $\mathfrak{r}[\mathfrak{s}/\mathfrak{i}] = \mathfrak{r}$ and $\llbracket \mathfrak{r} \rrbracket_{\rho} = \infty = \llbracket \mathfrak{r} \rrbracket_{\rho[\mathfrak{i} \mapsto \llbracket \mathfrak{s} \rrbracket_{\rho}]}$. □

The last fundamental property about reducibility sets that will be crucial to treat the recursive case is the following, stating that the sizes appearing in a sized type may be recovered in the reducibility set by using an appropriate semantics of the size variables, and conversely:

Lemma 6.12 (Size Commutation). *Let $\mathfrak{i}$ be a size variable, $\mathfrak{s}$ be a size such that $\mathfrak{s} = \infty$ or that* $\mathsf{spine}(\mathfrak{s}) = \mathfrak{j} \neq \mathfrak{i}$, *and ρ be a size environment. Then:*

- $\mathsf{VRed}^{p}_{\sigma[\mathfrak{s}/\mathfrak{i}],\rho} = \mathsf{VRed}^{p}_{\sigma,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]}$,
- $\mathsf{DRed}^{p}_{\mu[\mathfrak{s}/\mathfrak{i}],\rho} = \mathsf{DRed}^{p}_{\mu,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]}$, *and*
- $\mathsf{TRed}^{p}_{\mu[\mathfrak{s}/\mathfrak{i}],\rho} = \mathsf{TRed}^{p}_{\mu,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]}$.

Proof.

- The first case to consider is $\sigma = \mathsf{Nat}^{\mathfrak{r}}$ for some size $\mathfrak{r}$. Using Lemma 6.11, we have that
$$\begin{array}{rcl}\mathsf{VRed}^{p}_{(\mathsf{Nat}^{\mathfrak{r}})[\mathfrak{s}/\mathfrak{i}],\rho} & = & \mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{r}[\mathfrak{s}/\mathfrak{i}]},\rho} \\ & = & \left\{\mathsf{S}^n 0 \;\middle|\; p > 0 \Longrightarrow n < [\![\mathfrak{r}\,[\mathfrak{s}/\mathfrak{i}]]\!]_\rho\right\} \\ & = & \left\{\mathsf{S}^n 0 \;\middle|\; p > 0 \Longrightarrow n < [\![\mathfrak{r}]\!]_{\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]}\right\} \\ & = & \mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{r}},\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]}.\end{array}$$
- We then consider the case of the sized type $\sigma \to \mu :: \kappa' \to \kappa''$. We have
$$\begin{array}{l}\mathsf{VRed}^{p}_{(\sigma\to\mu)[\mathfrak{s}/\mathfrak{i}],\rho} \\ \quad = \mathsf{VRed}^{p}_{\sigma[\mathfrak{s}/\mathfrak{i}]\to\mu[\mathfrak{s}/\mathfrak{i}],\rho} \\ \quad = \left\{V \in \Lambda^{V}_{\oplus}(\langle\sigma\,[\mathfrak{s}/\mathfrak{i}] \to \mu\,[\mathfrak{s}/\mathfrak{i}]\rangle) \;\middle|\; \forall q \in (0,1],\;\; \forall W \in \mathsf{VRed}^{q}_{\sigma[\mathfrak{s}/\mathfrak{i}],\rho},\;\; V\,W \in \mathsf{TRed}^{pq}_{\mu[\mathfrak{s}/\mathfrak{i}],\rho}\right\} \\ \quad = \left\{V \in \Lambda^{V}_{\oplus}(\langle\sigma \to \mu\rangle) \;\middle|\; \forall q \in (0,1],\;\; \forall W \in \mathsf{VRed}^{q}_{\sigma,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]},\;\; V\,W \in \mathsf{TRed}^{pq}_{\mu,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]}\right\} \\ \quad = \mathsf{VRed}^{p}_{\sigma\to\mu,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]},\end{array}$$
where we used the induction hypothesis twice, once on κ' and the other time on κ''.
- Let $\mathscr{D}$ be a finite distribution of values and $\mu = \{\, (\sigma_j)^{p'_j} \mid j \in \mathcal{J} \,\}$ be a distribution type. We have that $\mu[\mathfrak{s}/\mathfrak{i}] = \{\, (\sigma_j[\mathfrak{s}/\mathfrak{i}])^{p'_j} \mid j \in \mathcal{J} \,\}$. Suppose that $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu[\mathfrak{s}/\mathfrak{i}],\rho}$. Then there exist a pseudo-representation $\mathscr{D} = [\, (V_i)^{p_i} \mid i \in \mathcal{I} \,]$ and families $(p_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ and $(q_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ of reals of $[0,1]$ satisfying:

  (1) $\forall i \in \mathcal{I},\;\; \forall j \in \mathcal{J},\;\; V_i \in \mathsf{VRed}^{q_{ij}}_{\sigma_j[\mathfrak{s}/\mathfrak{i}],\rho}$;
  (2) $\forall i \in \mathcal{I},\;\; \sum_{j\in\mathcal{J}}\; p_{ij} = p_i$;
  (3) $\forall j \in \mathcal{J},\;\; \sum_{i\in\mathcal{I}}\; p_{ij} = \mu(\sigma_j)$; and
  (4) $p \leq \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}}\; q_{ij} p_{ij}$.

  But (1) is equivalent to $\forall i \in \mathcal{I},\;\; \forall j \in \mathcal{J},\;\; V_i \in \mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]}$ by induction hypothesis. It follows that $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]}$. The converse direction proceeds in the exact same way.
- Then, $M \in \mathsf{TRed}^{p}_{\mu[\mathfrak{s}/\mathfrak{i}],\rho}$ if and only if
$$M \in \Lambda_{\oplus}(\langle\mu\rangle) \text{ and } \forall 0 \leq r < p,\;\; \exists \nu_r \preccurlyeq \mu,\;\; \exists n_r \in \mathbb{N},\;\; M \Rrightarrow^{n_r}_{v} \mathscr{D}_r \text{ and } \mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r[\mathfrak{s}/\mathfrak{i}],\rho}$$
if and only if, by induction hypothesis,
$$M \in \Lambda_{\oplus}(\langle\mu\rangle) \text{ and } \forall 0 \leq r < p,\;\; \exists \nu_r \preccurlyeq \mu,\;\; \exists n_r \in \mathbb{N},\;\; M \Rrightarrow^{n_r}_{v} \mathscr{D}_r \text{ and } \mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]},$$
that is, if and only if $M \in \mathsf{TRed}^{p}_{\mu,\rho[\mathfrak{i}\mapsto[\![\mathfrak{s}]\!]_\rho]}$. □

6.6 Reducibility Sets Are Stable under Unfoldings

The most difficult step in proving all typable terms to be reducible is, unexpectedly, proving that terms involving *recursion* are reducible whenever their respective *unfoldings* are. This very natural concept expresses simply that any term in the form $\mathsf{letrec}\ f = W$ is assumed to compute the fix point of the function defined by W.

Definition 6.13 (n-Unfolding). Suppose that $V = (\mathsf{letrec}\ f = W)$ is closed; then the *n-unfolding* of V is:

- V if $n = 0$;
- $W[Z/f]$ if $n = m + 1$ and Z is the m-unfolding of V.

We write the set of unfoldings of V as $\mathit{Unfold}(V)$. Note that if V admits a simple type, then all its unfoldings have this same simple type as well. In the sequel, we implicitly consider that V is simply typed.

Any unfolding of $V = (\mathsf{letrec}\ f = W)$ should behave like V itself: all unfoldings of V should be equivalent. This, however, cannot be proved using simply the operational semantics. It requires some work, and techniques akin to logical relations, to prove this behavioral equivalence between a recursive definition and its unfoldings. The first lemma is technical and lists the unfoldings of terms defined recursively as equal to themselves or to a variable:

Lemma 6.14.

- *Let* $V = f$ *and* $W \in \mathit{Unfold}(\mathsf{letrec}\ f = V)$. *Then* $W = \mathsf{letrec}\ f = V$.
- *Let* $V = x \neq f$ *and* $W \in \mathit{Unfold}(\mathsf{letrec}\ f = V)$. *Then* $W = \mathsf{letrec}\ f = V$ *or* $W = x$. *More precisely, the n-unfoldings for* $n \geq 1$ *are all* x.

The next lemma is the technical core of this section. Think of two terms as related when they are of the shape $M[\overrightarrow{Z}/\overrightarrow{x}]$ and $M[\overrightarrow{Z'}/\overrightarrow{x}]$, where $\overrightarrow{x}$ is a sequence of "holes" in M, filled with unfoldings from a *same* recursively defined term. Then their rewritings by $\rightarrow_v$ form distributions of pairwise related terms.

Lemma 6.15. *Let* $V = (\mathsf{letrec}\ f = W)$ *be a closed value. Let* $\overrightarrow{x}$, $\overrightarrow{Z}$, $\overrightarrow{Z'}$ *be a vector of variables and two vectors of terms of Unfold*(V)*, all of the same length. Let* M *be a simply typed term with free variables contained in* $\overrightarrow{x}$*, all typed with the simple type of* V*. Suppose that* $M[\overrightarrow{Z}/\overrightarrow{x}] \rightarrow_v \mathscr{D}$*. Then there exists* $N_1, \ldots, N_n$*, a vector of variables* $\overrightarrow{y}$ *and* $\overrightarrow{Z_1}, \ldots, \overrightarrow{Z_n}, \overrightarrow{Z'_1}, \ldots, \overrightarrow{Z'_n} \in \mathit{Unfold}(V)$ *of the same length as* $\overrightarrow{y}$ *and such that* $\mathscr{D} = \{\, (N_i[\overrightarrow{Z_i}/\overrightarrow{y}])^{p_i} \,\}$ *and moreover* $M[\overrightarrow{Z'}/\overrightarrow{x}] \rightarrow_v \mathscr{E} = \{\, (N_i[\overrightarrow{Z'_i}/\overrightarrow{y}])^{p_i} \,\}$.

Proof. We prove the result by induction on the structure of M.

- The case where M is a variable cannot fit in this setting: either $M = y \notin \overrightarrow{x}$ and there is no reduction from $M[\overrightarrow{Z}/\overrightarrow{x}]$ or $M = x_i \in \overrightarrow{x}$ and there is no reduction either from $M[\overrightarrow{Z}/\overrightarrow{x}] = Z_i$ since it is a value. We can similarly rule out all the cases where M is a value.
- Suppose that $M = V_1\ V_2$. We proceed by case exhaustion on V_1. Three possibilities exist, the other ones contradicting the fact that there should be a reduction step from M:

— If $V_1 = x_i \in \overrightarrow{x}$, we distinguish four cases:

- Suppose that $Z_i = Z'_i$ are both the 0-unfolding of V. Then $M[\overrightarrow{Z}/\overrightarrow{x}] = M[\overrightarrow{Z'}/\overrightarrow{x}]$ and the result follows immediately from

$$\begin{array}{rcl} M\left[\overrightarrow{Z}/\overrightarrow{x}\right] & = & (\mathsf{letrec}\ f = W)\ V_2\left[\overrightarrow{Z}/\overrightarrow{x}\right] \\ & = & (\mathsf{letrec}\ f = W)\ (\mathsf{S}^m\ 0) \\ & \to_v & \{\, ((W\,[\mathsf{letrec}\ f = W/f])\ (\mathsf{S}^m\ 0))^1 \,\}, \end{array}$$

  where in the second line the shape of V_2 needs to be $\mathsf{S}^m\ 0$ by typing constraints. Note that $\overrightarrow{y}$ is the empty vector here.

- Suppose that Z_i is the n-unfolding of V for $n > 0$, and that Z'_i is the 0-unfolding. We have that

$$M\left[\overrightarrow{Z}/\overrightarrow{x}\right] = W\left[Z''/f\right]\ \ V_2\left[\overrightarrow{Z}/\overrightarrow{x}\right],$$

  where Z'' is the $(n-1)$-unfolding of V, and that

$$\begin{array}{rcl} M\left[\overrightarrow{Z'}/\overrightarrow{x}\right] & = & (\mathsf{letrec}\ f = W)\ \ V_2\left[\overrightarrow{Z'}/\overrightarrow{x}\right] \\ & \to_v & \left\{ \left((W\,[\mathsf{letrec}\ f = W/f])\ \ V_2\left[\overrightarrow{Z'}/\overrightarrow{x}\right]\right)^1 \right\}. \end{array}$$

  Notice that this reduction is possible since the constraint of simple typing implies that V_2 is of the shape $\mathsf{S}^m\ 0$ for some $m \geq 0$. We can therefore rewrite the two terms as

$$M\left[\overrightarrow{Z}/\overrightarrow{x}\right] = W\ [Z''/f]\ \ \ (\mathsf{S}^m\ 0)$$

  and

$$M\left[\overrightarrow{Z'}/\overrightarrow{x}\right]\ \ \to_v\ \left\{ ((W\,[\mathsf{letrec}\ f = W/f])\ \ (\mathsf{S}^m\ 0))^1 \right\}.$$

  We need to distinguish four cases, depending on the structure of W.

  - Suppose that W is a variable different from f. Then by Lemma 6.14, there cannot be a step of reduction from $M[\overrightarrow{Z}/\overrightarrow{x}]$.
  - Suppose that $W = f$. Then by Lemma 6.14, we have $Z_i = Z'_i = Z''$ so that $M[\overrightarrow{Z}/\overrightarrow{x}] = M[\overrightarrow{Z'}/\overrightarrow{x}]$ and the result follows just as for the case where both Z and Z' were 0-unfoldings.
  - Suppose that $W = \lambda y.L$. Then

$$\begin{array}{rcl} M\left[\overrightarrow{Z}/\overrightarrow{x}\right] & = & (\lambda y.L\,[Z''/f])\ \ (\mathsf{S}^m\ 0) \\ & \to_v & \{\, (L\,[Z''/f]\,[\mathsf{S}^m\ 0/y])^1 \,\} \\ & = & \{\, ((L\,[\mathsf{S}^m\ 0/y])\,[Z''/f])^1 \,\}. \end{array}$$

    Moreover,

$$\begin{array}{rcl} M\left[\overrightarrow{Z'}/\overrightarrow{x}\right] & \to_v & \{\, ((W\,[\mathsf{letrec}\ f = W/f])\ \ (\mathsf{S}^m\ 0))^1 \,\} \\ & = & \{\, (((\lambda y.L)\ \ (\mathsf{S}^m\ 0))\,[\mathsf{letrec}\ f = W/f])^1 \,\} \\ & \to_v & \{\, ((L\,[(\mathsf{S}^m\ 0)/y])\,[\mathsf{letrec}\ f = W/f])^1 \,\} \end{array}$$

    so that we can conclude with $\overrightarrow{y} = f$ and $N_1 = L[(\mathsf{S}^m\ 0)/y]$.

- Suppose that $W = \text{letrec } g = W'$. Then

$$\begin{array}{rcl} M\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] & = & ((\text{letrec } g = W')\,[Z''/f])\ \ (\mathsf{S}^m\ 0) \\ & \to_v & \{\,(W'\,[\text{letrec } g = W'/g])\,[Z''/f]\ \ (\mathsf{S}^m\ 0))^1\,\} \\ & = & \{\,(W'\,[\text{letrec } g = W'/g]\ \ (\mathsf{S}^m\ 0))\,[Z''/f])^1\,\}. \end{array}$$

  Moreover,

$$\begin{array}{rcl} M\Big[\overrightarrow{Z'}/\overrightarrow{x}\Big] & \to_v & ((\text{letrec } g = W')\,[\text{letrec } f = W/f])\ \ (\mathsf{S}^m\ 0) \\ & \to_v & \{\,(W'\,[\text{letrec } g = W'/g])\,[\text{letrec } f = W/f]\ \ (\mathsf{S}^m\ 0))^1\,\} \\ & = & \{\,(W'\,[\text{letrec } g = W'/g]\ \ (\mathsf{S}^m\ 0))\,[\text{letrec } f = W/f])^1\,\}, \end{array}$$

  and we conclude with $\overrightarrow{y} = f$ and $N_1 = W'[\text{letrec } g = W'/g]\ (\mathsf{S}^m\ 0))$.

* Suppose that Z_i is the 0-unfolding of V and that Z_i' is the n-unfolding for $n > 0$. Again, the constraint of simple typing implies that V_2 is of the shape $\mathsf{S}^m\ 0$ for some $m \geq 0$. We have that

$$\begin{array}{rcl} M\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] & \to_v & \{\,((W\,[\text{letrec } f = W/f])\ \ (\mathsf{S}^m\ 0))^1\,\} \\ & = & \{\,((W\ \ (\mathsf{S}^m\ 0))\,[\text{letrec } f = W/f]))^1\,\} \end{array}$$

and that

$$M\Big[\overrightarrow{Z'}/\overrightarrow{x}\Big] = W\ [Z''/f]\ \ \ (\mathsf{S}^m\ 0) = (W\ \ (\mathsf{S}^m\ 0))\ [Z''/f]\,,$$

where Z'' is the $(n-1)$-unfolding of V, so that we can conclude with $\overrightarrow{y} = f$ and $N_1 = W\ \ (\mathsf{S}^m\ 0)$.

* Suppose that Z_i is the n-unfolding of V for $n > 0$, and that Z_i' is the n'-unfolding for $n' > 0$. We have

$$M\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] = W\ [Z''/f]\ \ V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big],$$

where Z'' is the $(n-1)$-unfolding of V, and

$$M\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] = W\ [Z'''/f]\ \ V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big],$$

where Z'' is the $(n'-1)$-unfolding of V. We proceed by case analysis on W. As we discussed in the case where Z_i was a $(n''+1)$-unfolding and Z_i' a 0-unfolding, the case where W is a variable does not lead to a rewriting step. It remains to treat two cases:

- Suppose that $W = \lambda y.L$. Then

$$\begin{array}{rcl} M\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] & = & \lambda y.L\,[Z''/f]\ \ V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] \\ & \to_v & \Big\{\Big(L\Big[Z''/f\Big]\Big[V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]/y\Big]\Big)^1\Big\} \\ & = & \Big\{\Big(\Big(L\Big[V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]/y\Big]\Big)\Big[Z''/f\Big]\Big)^1\Big\} \end{array}$$

and

$$\begin{array}{rcl} M\Big[\overrightarrow{Z'}/\overrightarrow{x}\Big] & = & \lambda y.L\,[Z'''/f]\;\; V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] \\ & \to_v & \Big\{\Big(L\Big[Z'''/f\Big]\Big[V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]/y\Big]\Big)^1\Big\} \\ & = & \Big\{\Big(\Big(L\Big[V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]/y\Big]\Big)\Big[Z'''/f\Big]\Big)^1\Big\} \end{array}$$

so that we can conclude with $\overrightarrow{y} = f$ and $N_1 = L[V_2[\overrightarrow{Z}/\overrightarrow{x}]/y]$.

- Suppose that $W = \text{letrec}\ g = W'$. Then

$$\begin{array}{rcl} M\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] & = & ((\text{letrec}\ g = W')\,[Z''/f])\;\; V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] \\ & \to_v & \Big\{\Big(W'\,[\text{letrec}\ g = W'/g])\,[Z''/f]\;\; V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\Big)^1\Big\} \\ & = & \Big\{\Big(W'\,[\text{letrec}\ g = W'/g]\;\; V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big])\,[Z''/f]\Big)^1\Big\}, \end{array}$$

where the reduction is possible because the simple typing constraints imply that $V_2[\overrightarrow{Z}/\overrightarrow{x}]$ is of the shape $\mathrm{S}^m\ 0$ for some $m \in \mathbb{N}$. Moreover,

$$\begin{array}{rcl} M\Big[\overrightarrow{Z'}/\overrightarrow{x}\Big] & = & ((\text{letrec}\ g = W')\,[Z'''/f])\;\; V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] \\ & \to_v & \Big\{\Big((W'\,[\text{letrec}\ g = W'/g])\,[Z'''/f]\;\; V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\Big)^1\Big\} \\ & = & \Big\{\Big(\Big(W'\,[\text{letrec}\ g = W'/g]\;\; V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\Big)\,[Z'''/f]\Big)^1\Big\}, \end{array}$$

and we conclude with $\overrightarrow{y} = f$ and $N_1 = W'[\text{letrec}\ g = W'/g]\ V_2[\overrightarrow{Z}/\overrightarrow{x}]$.

—If $V_1 = \lambda y.L$,

$$\begin{array}{rcl} M\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] & = & \Big(\lambda y.L\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\Big)\; V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] \\ & \to_v & \Big\{\Big(L\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\Big[V_2\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]/y\Big]\Big)^1\Big\} \\ & = & \Big\{\Big(L\Big[V_2/y\Big]\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\Big)^1\Big\} \end{array}$$

and in the same way

$$M\Big[\overrightarrow{Z'}/\overrightarrow{x}\Big] \;\; \to_v \;\; \Big\{\Big(L\,[V_2/y]\,\Big[\overrightarrow{Z'}/\overrightarrow{x}\Big]\Big)^1\Big\},$$

which allows us to conclude with $N_1 = L[V_2/y]$.

—If $V_1 = \text{letrec}\ g = W'$, by typing constraints $V_2 = \mathrm{S}^m\ 0$ for some $m \geq 0$. It follows that we can reduce $M[\overrightarrow{Z}/\overrightarrow{x}]$ and $M[\overrightarrow{Z'}/\overrightarrow{x}]$ as follows:

$$\begin{array}{rcl} M\Big[\overrightarrow{Z}/\overrightarrow{x}\Big] & = & \Big(\text{letrec}\ g = W'\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\Big)\ \mathrm{S}^m\ 0 \\ & \to_v & \Big\{\Big(\Big(W'\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\Big[\text{letrec}\ g = W'\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]/g\Big]\Big)\ (\mathrm{S}^m\ 0)\Big)^1\Big\} \\ & = & \Big\{\Big(\Big(W'\Big[\text{letrec}\ g = W'/g\Big]\Big)\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\ (\mathrm{S}^m\ 0)\Big)^1\Big\} \\ & = & \Big\{\Big(\Big(W'\Big[\text{letrec}\ g = W'/g\Big]\ (\mathrm{S}^m\ 0)\Big)\Big[\overrightarrow{Z}/\overrightarrow{x}\Big]\Big)^1\Big\}, \end{array}$$

and similarly,

$$M\Big[\overrightarrow{Z'}/\overrightarrow{x}\Big] \;\; \to_v \;\; \Big\{\Big((W'\,[\text{letrec}\ g = W'/g]\;\; (\mathrm{S}^m\ 0))\,\Big[\overrightarrow{Z'}/\overrightarrow{x}\Big]\Big)^1\Big\}$$

so that we can conclude with $N_1 = W'[\text{letrec}\ g = W'/g]\ (\mathrm{S}^m\ 0)$.

- Suppose that $M = \mathsf{let}\ y = X\ \mathsf{in}\ P$. Then

$$\begin{array}{rcl} M\left[\overrightarrow{Z}/\overrightarrow{x}\right] & = & \mathsf{let}\ y = X\left[\overrightarrow{Z}/\overrightarrow{x}\right]\ \mathsf{in}\ P\left[\overrightarrow{Z}/\overrightarrow{x}\right] \\ & \rightarrow_v & \left\{ \left(P\left[\overrightarrow{Z}/\overrightarrow{x}\right]\left[X\left[\overrightarrow{Z}/\overrightarrow{x}\right]/y\right]\right)^1 \right\} \\ & = & \left\{ \left(P\left[X/y\right]\left[\overrightarrow{Z}/\overrightarrow{x}\right]\right)^1 \right\}, \end{array}$$

  and similarly, $M[\overrightarrow{Z'}/\overrightarrow{x}] \rightarrow_v \{ (P[X/y][\overrightarrow{Z'}/\overrightarrow{x}])^1 \}$, from which we can conclude.
- Suppose that $M = \mathsf{let}\ y = L\ \mathsf{in}\ P$ and that

$$\begin{array}{rcl} M\left[\overrightarrow{Z}/\overrightarrow{x}\right] & = & \mathsf{let}\ y = L\left[\overrightarrow{Z}/\overrightarrow{x}\right]\ \mathsf{in}\ P\left[\overrightarrow{Z}/\overrightarrow{x}\right] \\ & \rightarrow_v & \left\{ \left(\mathsf{let}\ y = L'_i\left[\overrightarrow{Z}/\overrightarrow{x}\right]\ \mathsf{in}\ P\left[\overrightarrow{Z}/\overrightarrow{x}\right]\right)^{p_i} \right\} \\ & = & \left\{ \left(\mathsf{let}\ y = L''_i\left[\overrightarrow{Z}/\overrightarrow{z}\right]\ \mathsf{in}\ P\left[\overrightarrow{Z}/\overrightarrow{x}\right]\right)^{p_i} \right\} \\ & = & \left\{ \left(\left(\mathsf{let}\ y = L''_i\ \mathsf{in}\ P\right)\left[\overrightarrow{Z},\overrightarrow{Z}/\overrightarrow{z},\overrightarrow{x}\right]\right)^{p_i} \right\}, \end{array}$$

  where the third step is obtained by α-renaming, and where by definition of $\rightarrow_v$ we have

$$L\left[\overrightarrow{Z}/\overrightarrow{x}\right] \quad \rightarrow_v \quad \left\{ \left(L'_i\left[\overrightarrow{Z}/\overrightarrow{x}\right]\right)^{p_i} \ \middle|\ i \in \mathcal{I} \right\}.$$

  By induction hypothesis, there exists $\overrightarrow{Z'_1}, \ldots, \overrightarrow{Z'_n} \in \mathit{Unfold}(V)$ such that

$$L\left[\overrightarrow{Z'}/\overrightarrow{x}\right] \quad \rightarrow_v \quad \left\{ \left(L'_i\left[\overrightarrow{Z'_i}/\overrightarrow{x}\right]\right)^{p_i} \ \middle|\ i \in \mathcal{I} \right\}.$$

  Now we see that

$$\begin{array}{rcl} M\left[\overrightarrow{Z'}/\overrightarrow{x}\right] & \rightarrow_v & \left\{ \left(\mathsf{let}\ y = L'_i\left[\overrightarrow{Z'_i}/\overrightarrow{x}\right]\ \mathsf{in}\ P\left[\overrightarrow{Z}/\overrightarrow{x}\right]\right)^{p_i} \right\} \\ & = & \left\{ \left(\mathsf{let}\ y = L''_i\left[\overrightarrow{Z'_i}/\overrightarrow{z}\right]\ \mathsf{in}\ P\left[\overrightarrow{Z}/\overrightarrow{x}\right]\right)^{p_i} \right\} \\ & = & \left\{ \left(\left(\mathsf{let}\ y = L''_i\ \mathsf{in}\ P\right)\left[\overrightarrow{Z},\overrightarrow{Z'_i}/\overrightarrow{x},\overrightarrow{z}\right]\right)^{p_i} \right\}. \end{array}$$

  The result follows for $\overrightarrow{y} = \overrightarrow{x},\overrightarrow{z}$ and $N_i = \mathsf{let}\ y = L''_i\ \mathsf{in}\ P$.
- Suppose that $M = L \oplus_p P$. Suppose that $L \neq P$. Then

$$M\left[\overrightarrow{Z}/\overrightarrow{x}\right] \quad \rightarrow_v \quad \left\{ L\left[\overrightarrow{Z}/\overrightarrow{x}\right]^p,\ P\left[\overrightarrow{Z}/\overrightarrow{x}\right]^{1-p} \right\}$$

  and

$$M\left[\overrightarrow{Z'}/\overrightarrow{x}\right] \quad \rightarrow_v \quad \left\{ L\left[\overrightarrow{Z'}/\overrightarrow{x}\right]^p,\ P\left[\overrightarrow{Z'}/\overrightarrow{x}\right]^{1-p} \right\}$$

  so that the result holds for $N_1 = L$ and $N_2 = P$.
  If $L = P$,

$$M\left[\overrightarrow{Z}/\overrightarrow{x}\right] \quad \rightarrow_v \quad \left\{ L\left[\overrightarrow{Z}/\overrightarrow{x}\right]^1 \right\}$$

  and

$$M\left[\overrightarrow{Z'}/\overrightarrow{x}\right] \quad \rightarrow_v \quad \left\{ L\left[\overrightarrow{Z'}/\overrightarrow{x}\right]^1 \right\},$$

  and the result holds as well. Note that the distinction is necessary so as to avoid the use of pseudo-representations in the statement of the lemma.
- Suppose that $M = \mathsf{case}\ V'\ \mathsf{of}\ \{\, \mathsf{S} \rightarrow X \mid 0 \rightarrow Y \,\}$. By typing constraints, $V' = \mathsf{S}^m\ 0$ or $V' = y$ is a variable.
  —If $V' = 0$, $M[\overrightarrow{Z}/\overrightarrow{x}] \quad \rightarrow_v \quad \{ (R[\overrightarrow{Z}/\overrightarrow{x}])^1 \}$ and $M[\overrightarrow{Z'}/\overrightarrow{x}] \quad \rightarrow_v \quad \{ (R[\overrightarrow{Z'}/\overrightarrow{x}])^1 \}$ so that we can conclude.
  —If $V' = \mathsf{S}^m\ 0$ with $m > 0$, we can conclude in the same way.

— In the latter case, there is no reduction from $M[\overrightarrow{Z}/\overrightarrow{x}]$ unless $V'[\overrightarrow{Z}/\overrightarrow{x}]$ is of the shape $V' = \mathsf{S}^m\ 0$. But this is of type Nat and cannot therefore be an unfolding of V, so that this case is impossible. □

This result can be extended to an n-step rewriting process; however, pseudo-representations are required to keep the statement true, as we explain in the proof.

Lemma 6.16. *Let $V = (\text{letrec } f = W)$ be a closed value. Let M be a simply typed term of free variables contained in $\overrightarrow{x}$, all typed with the simple type of V. Let $\overrightarrow{Z}, \overrightarrow{Z'} \in \mathit{Unfold}(V)$ and $n \in \mathbb{N}$. Then there exists a distribution of values of pseudo-representation $[\, X_i^{p_i} \mid i \in \mathcal{I}\,]$, a vector of variables $\overrightarrow{y}$, and families of vectors $(\overrightarrow{Z_i})_{i \in \mathcal{I}}, (\overrightarrow{Z_i'})_{i \in \mathcal{I}}$ of the same length as $\overrightarrow{y}$, all such that $M[\overrightarrow{Z}/\overrightarrow{x}] \Rrightarrow_v^n [\, (X_i[\overrightarrow{Z_i}/\overrightarrow{y}])^{p_i} \mid i \in \mathcal{I}\,]$ and that $M[\overrightarrow{Z'}/\overrightarrow{x}] \Rrightarrow_v^n [\, (X_i[\overrightarrow{Z_i'}/\overrightarrow{y}])^{p_i} \mid i \in \mathcal{I}\,]$.*

Proof. By iteration of Lemma 6.15. The pseudo-representations come from the fact that some terms in different reduction branches may converge to the same value, say, in the reduction from $M[\overrightarrow{Z}/\overrightarrow{x}]$ but not in the one from $M[\overrightarrow{Z'}/\overrightarrow{x}]$. □

The following lemma is of technical interest. It states that, given two pseudo-representations of a distribution—one of the shape exhibited in the previous lemmas and used for relating terms with unfoldings, the other one being a pseudo-representation witnessing the belonging to a set DRed—there exists a third one that "combines" both:

Lemma 6.17. *Suppose that $\mathscr{D}_r = [\, (X_i[\overrightarrow{Z_i}/\overrightarrow{y}])^{p_i} \mid i \in \mathcal{I}\,] = [\, (X_j')^{p_j'} \mid j \in \mathcal{J}\,]$. Then there exists a set $\mathcal{K}$, two applications $\pi_1 : \mathcal{K} \to \mathcal{I}$ and $\pi_2 : \mathcal{K} \to \mathcal{J}$, and a pseudo-representation $\mathscr{D}_r = [\, (X_k''[\overrightarrow{Z_{\pi_1(k)}}/\overrightarrow{y}])^{p_k''} \mid k \in \mathcal{K}\,]$ such that*

- $\forall k \in \mathcal{K},\ \ X_k'' = X_{\pi_1(k)}$;
- $\forall i \in \mathcal{I},\ \ \sum_{k \in \pi_1^{-1}(i)} p_k'' = p_i$;
- $\forall k \in \mathcal{K},\ \ X_k''[\overrightarrow{Z_k''}/\overrightarrow{y}] = X'_{\pi_2(k)}$; *and*
- $\forall j \in \mathcal{J},\ \ \sum_{k \in \pi_2^{-1}(j)} p_k'' = p_j'$.

Proof. Let $\mathscr{D} = \{\, (Y_l)^{p_l''} \mid l \in \mathcal{L}\,\}$ be the *representation* of $\mathscr{D}$. We build $\mathcal{K}$, π_1, and π_2 as follows. The construction starts from the empty set and the empty maps and is iterated on every $l \in \mathcal{L}$. First, we set $\mathcal{I}_l = \{i \in \mathcal{I} \mid Y_l = X_i[\overrightarrow{Z_i}/\overrightarrow{y}]\}$ and $\mathcal{J}_l = \{j \in \mathcal{J} \mid Y_l = X_j'\}$. We suppose that both these sets are enumerated and will write them $\mathcal{I}_l = \{i_0, \ldots, i_{n_l}\}$ and $\mathcal{J}_l = \{j_0, \ldots, j_{m_l}\}$. We consider the set of reals

$$R = \left\{0,\, p_{i_1},\, p_{i_1} + p_{i_2},\, \ldots,\, \sum_{r=0}^{n_l} p_{i_r}\right\} \cup \left\{0,\, p_{j_1}',\, p_{j_1}' + p_{j_2}',\, \ldots,\, \sum_{r'=0}^{m_l} p_{i_{r'}}'\right\} \subset [0, p_l''].$$

This set is ordered, as a set of reals, so that we have a maximal enumeration

$$0 = \alpha_0 < \alpha_1 < \cdots < \alpha_s = p,$$

where maximality means that $\beta \in R \Rightarrow \exists t,\quad \beta = \alpha_t$. We add s elements to the set $\mathcal{K}$ produced during the examination of previous elements of $\mathcal{L}$: $\mathcal{K} := \mathcal{K} \uplus \{0, \ldots, s-1\}$. For every $t \in \{0, \ldots, s-1\}$, we define:

- $p_t'' = \alpha_{t+1} - \alpha_t$;
- $\pi_1(t) = i_k \in \mathcal{I}_l$, where $\sum_{r=0}^{k-1} p_{i_r} \leq \alpha_t$ and $\sum_{r=0}^{k} p_{i_r} \geq \alpha_t$; and
- $\pi_2(t) = j_k \in \mathcal{J}_l$, where $\sum_{r=0}^{k-1} p_{j_r} \leq \alpha_t$ and $\sum_{r=0}^{k} p_{j_r} \geq \alpha_t$.

We claim that the set $\mathcal{K}$ resulting from this constructive process satisfies the equalities of the lemma. □

The series of previous lemmas allows one to deduce that a term is reducible if and only if the terms to which it is related are:

Lemma 6.18. *Let $V = (\mathsf{letrec}\ f = W)$ be a closed value. Let M be a simply typed term of free variables contained in $\overrightarrow{x}$, all typed with the simple type of V. Let $\overrightarrow{Z}, \overrightarrow{Z'} \in Unfold(V)$. Then $M[\overrightarrow{Z}/\overrightarrow{x}] \in \mathsf{TRed}^{p}_{\mu,\rho}$ if and only if $M[\overrightarrow{Z'}/\overrightarrow{x}] \in \mathsf{TRed}^{p}_{\mu,\rho}$.*

Proof. We prove that $M[\overrightarrow{Z}/\overrightarrow{x}] \in \mathsf{TRed}^{p}_{\mu,\rho}$ implies that $M[\overrightarrow{Z'}/\overrightarrow{x}] \in \mathsf{TRed}^{p}_{\mu,\rho}$, the converse direction being exactly symmetrical. The proof proceeds by induction on the simple type refined by μ.

Suppose that $\mu :: \mathsf{Nat}$. Let $r \in [0, p)$. Since $M[\overrightarrow{Z}/\overrightarrow{x}] \in \mathsf{TRed}^{p}_{\mu,\rho}$, there exists n_r and $\nu_r \preccurlyeq \mu$ such that $M[\overrightarrow{Z}/\overrightarrow{x}] \Rrightarrow_{v}^{n_r} \mathscr{D}_r$ and that $\mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$. Lemma 6.16 implies that there exists a distribution of values of pseudo-representation $[\, X_i^{p_i} \mid i \in \mathcal{I}\,]$, a vector of variables $\overrightarrow{y}$, and families of vectors $(\overrightarrow{Z_i})_{i\in\mathcal{I}}, (\overrightarrow{Z'_i})_{i\in\mathcal{I}}$ of the same length as $\overrightarrow{y}$ all such that $\mathscr{D}_r = [\, (X_i[\overrightarrow{Z_i}/\overrightarrow{y}])^{p_i} \mid i \in \mathcal{I}\,]$ and that $M[\overrightarrow{Z'}/\overrightarrow{x}] \Rrightarrow_{v}^{n} \mathscr{E}_r = [\, (X_i[\overrightarrow{Z'_i}/\overrightarrow{y}])^{p_i} \mid i \in \mathcal{I}\,]$. By typing constraints coming from the subject reduction property, all the $X_i[\overrightarrow{Z_i}/\overrightarrow{y}]$ and $X_i[\overrightarrow{Z'_i}/\overrightarrow{y}]$ have the simple type Nat. This implies that all these terms are of the shape $\mathsf{S}^m\ 0$ for $m \geq 0$, and thus that the X_i cannot contain a variable from $\overrightarrow{y}$, as their simple type is of the shape $\mathsf{Nat} \rightarrow \kappa$. It follows that, for every index $i \in \mathcal{I}$, $X_i[\overrightarrow{Z_i}/\overrightarrow{y}] = X_i[\overrightarrow{Z'_i}/\overrightarrow{y}]$. This implies that $\mathscr{E}_r = \mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$, and thus that $M[\overrightarrow{Z'}/\overrightarrow{x}] \in \mathsf{TRed}^{p}_{\mu,\rho}$.

Suppose that $\mu :: \kappa \rightarrow \kappa'$. Let $r \in [0, p)$. Since $M[\overrightarrow{Z}/\overrightarrow{x}] \in \mathsf{TRed}^{p}_{\mu,\rho}$, there exists n_r and $\nu_r \preccurlyeq \mu$ such that $M[\overrightarrow{Z}/\overrightarrow{x}] \Rrightarrow_{v}^{n_r} \mathscr{D}_r$ and that $\mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$. Lemma 6.16 implies that there exists a distribution of values of pseudo-representation $[\, X_i^{p_i} \mid i \in \mathcal{I}\,]$, a vector of variables $\overrightarrow{y}$, and families of vectors $(\overrightarrow{Z_i})_{i\in\mathcal{I}}, (\overrightarrow{Z'_i})_{i\in\mathcal{I}}$ of the same length as $\overrightarrow{y}$ all such that $\mathscr{D}_r = [\, (X_i[\overrightarrow{Z_i}/\overrightarrow{y}])^{p_i} \mid i \in \mathcal{I}\,]$ and that $M[\overrightarrow{Z'}/\overrightarrow{x}] \Rrightarrow_{v}^{n} \mathscr{E}_r = [\, (X_i[\overrightarrow{Z'_i}/\overrightarrow{y}])^{p_i} \mid i \in \mathcal{I}\,]$. Since $\mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$, there is a pseudo-representation $\mathscr{D}_r = [\, (Z'_j)^{p'_j} \mid j \in \mathcal{J}\,]$ witnessing this fact. By Lemma 6.17, there exists a pseudo-representation $\mathscr{D}_r = [\, (X''_k[\overrightarrow{Z_{\pi_1(k)}}/\overrightarrow{y}])^{p''_k} \mid k \in \mathcal{K}\,]$ satisfying a series of additional properties. These properties ensure two crucial facts for our purpose:

- $M[\overrightarrow{Z}/\overrightarrow{x}] \Rrightarrow_{v}^{n} [\, (X''_k[\overrightarrow{Z_{\pi_1(k)}}/\overrightarrow{y}])^{p''_k} \mid k \in \mathcal{K}\,]$,
- $M[\overrightarrow{Z'}/\overrightarrow{x}] \Rrightarrow_{v}^{n} \mathscr{E}_r = [\, (X''_k[\overrightarrow{Z'_{\pi_1(k)}}/\overrightarrow{y}])^{p''_k} \mid k \in \mathcal{K}\,]$, and
- $[\, (X''_k[\overrightarrow{Z_{\pi_1(k)}}/\overrightarrow{y}])^{p''_k} \mid k \in \mathcal{K}\,]$ is a pseudo-distribution witnessing the fact that $\mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$. Setting $\mu = \{\, (\sigma_l)^{p'''_l} \mid l \in \mathcal{L}\,\}$, there thus exists families $(p''_{kl})_{k\in\mathcal{K}, l\in\mathcal{L}}$ and $(q_{kl})_{k\in\mathcal{K}, l\in\mathcal{L}}$ of reals of $[0, 1]$ satisfying:

  (1) $\forall k \in \mathcal{K},\ \ \forall l \in \mathcal{L},\ \ X''_k[\overrightarrow{Z_{\pi_1(k)}}/\overrightarrow{y}] \in \mathsf{VRed}^{q_{kl}}_{\sigma_l,\rho}$;
  (2) $\forall k \in \mathcal{K},\ \ \sum_{l\in\mathcal{L}}\ p''_{kl} = p''_k$;
  (3) $\forall l \in \mathcal{L},\ \ \sum_{k\in\mathcal{K}}\ p''_{kl} = \mu(\sigma_l)$; and
  (4) $p \leq \sum_{k\in\mathcal{K}} \sum_{l\in\mathcal{L}}\ q_{kl} p''_{kl}$.

We now prove that $\forall k \in \mathcal{K},\ \forall l \in \mathcal{L},\ X''_k[\overrightarrow{Z'_{\pi_1(k)}/\vec{y}}] \in \mathsf{VRed}^{q_{kl}}_{\sigma_l,\rho}$. Let $k \in \mathcal{K}$ and $l \in \mathcal{L}$. Let $\sigma_l = \theta \rightarrow \nu$:

$$\begin{array}{lll}
& X''_k\left[\overrightarrow{Z_{\pi_1(k)}/\vec{y}}\right] \in \mathsf{VRed}^{q_{kl}}_{\sigma_l,\rho} & \\
\iff & \forall q \in (0,1],\ \ \forall Y \in \mathsf{VRed}^{q}_{\theta,\rho},\ \ X''_k\left[\overrightarrow{Z_{\pi_1(k)}/\vec{y}}\right]\ Y \in \mathsf{TRed}^{qq_{kl}}_{\nu,\rho} & \\
\iff & \forall q \in (0,1],\ \ \forall Y \in \mathsf{VRed}^{q}_{\theta,\rho},\ \ \left(X''_k\ Y\right)\left[\overrightarrow{Z_{\pi_1(k)}/\vec{y}}\right]\ \in \mathsf{TRed}^{qq_{kl}}_{\nu,\rho} & \\
\iff & \forall q \in (0,1],\ \ \forall Y \in \mathsf{VRed}^{q}_{\theta,\rho},\ \ \left(X''_k\ Y\right)\left[\overrightarrow{Z'_{\pi_1(k)}/\vec{y}}\right]\ \in \mathsf{TRed}^{qq_{kl}}_{\nu,\rho} & \text{by IH} \\
\iff & \forall q \in (0,1],\ \ \forall Y \in \mathsf{VRed}^{q}_{\theta,\rho},\ \ X''_k\left[\overrightarrow{Z'_{\pi_1(k)}/\vec{y}}\right]\ Y \in \mathsf{TRed}^{qq_{kl}}_{\nu,\rho} & \\
\iff & X''_k\left[\overrightarrow{Z'_{\pi_1(k)}/\vec{y}}\right] \in \mathsf{VRed}^{q_{kl}}_{\sigma_l,\rho}. &
\end{array}$$

This implies that $[\ (X''_k[\overrightarrow{Z'_{\pi_1(k)}/\vec{y}}])^{p''_k}\ \mid\ k \in \mathcal{K}\ ]$ witnesses that $\mathscr{E}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$, for the same families of reals p''_{kl} and q_{kl}. Now for every $r \in [0,p)$, there exists n_r and $\nu_r \preccurlyeq \mu$ such that $M[\overrightarrow{Z'}/\vec{x}] \Rrightarrow^{n_r}_{v} \mathscr{E}_r$ and that $\mathscr{E}_r \in \mathsf{DRed}^{r}_{\nu_r,\rho}$: we have that $M[\overrightarrow{Z'}/\vec{x}] \in \mathsf{TRed}^{p}_{\mu,\rho}$. □

The following lemma shows that reducible values are reducible terms:

Lemma 6.19 (Reducible Values are Reducible Terms). *Let V be a value. Then $V \in \mathsf{TRed}^{p}_{\{\ \sigma^1\ \},\rho}$ if and only if $V \in \mathsf{VRed}^{p}_{\sigma,\rho}$.*

Note that, conversely, we may have $V \in \mathsf{TRed}^{p}_{\mu,\rho}$, where μ is not Dirac. For instance, $0 \in \mathsf{TRed}^{1}_{\mu,\rho}$ for $\mu = \{\ (\mathsf{Nat}^{\mathfrak{i}})^{\frac{1}{2}}, (\mathsf{Nat}^{\widehat{\mathfrak{i}}})^{\frac{1}{2}}\ \}$.

Proof.

- Suppose that $V \in \mathsf{VRed}^{p}_{\sigma,\rho}$. Let $r \in [0,p)$. We must prove that there exists n_r and ν_r such that $V \rightarrow^{n_r}_{v} \{\ V^1\ \}$ and that $\{\ V^1\ \} \in \mathsf{DRed}^{r}_{\nu_r,\rho}$. Necessarily $n_r = 0$ and $\nu_r = \{\ \sigma^1\ \}$. Since $V \in \mathsf{VRed}^{p}_{\sigma,\rho}$, $\{\ V^1\ \} \in \mathsf{DRed}^{r}_{\nu_r,\rho}$: take the canonical pseudo-representation $[\ V^1\ ]$ and $p_{11} = 1$, $q_{11} = r$.
- Suppose that $V \in \mathsf{TRed}^{p}_{\{\ \sigma^1\ \},\rho}$. It follows that, for every $r \in [0,p)$, there exists n_r and ν_r such that $V \rightarrow^{n_r}_{v} \{\ V^1\ \}$ and that $\{\ V^1\ \} \in \mathsf{DRed}^{r}_{\nu_r,\rho}$. Again, since V is a value, we necessarily have $n_r = 0$ and $\nu_r = \{\ \sigma^1\ \}$. Since $\{\ V^1\ \} \in \mathsf{DRed}^{r}_{\nu_r,\rho}$, there is a pseudo-representation $[\ V^{p_1},\ldots,V^{p_n}\ ]$ such that $\sum_{i=1}^{n}\ p_i = 1$, and a family $(q_{i1})_{i\in\mathcal{I}}$ that is such that $r \leq \sum_{i\in\mathcal{I}}\ p_{i1}q_{i1}$, where $p_{i1} = p_i$.
  Suppose that there is no q_{i1} greater than or equal to r. Then $\forall i \in \mathcal{I},\ \ q_{i1} < r$, and
  $$\sum_{i\in\mathcal{I}} p_{i1}q_{i1}\ <\ \sum_{i\in\mathcal{I}} p_{i1}r = r\sum_{i\in\mathcal{I}} p_{i1} = r,$$
  which is a contradiction. So there exists $q_{i1} \geq r$ and therefore $V \in \mathsf{VRed}^{q_{i1}}_{\sigma,\rho}$. By Lemma 6.7, $V \in \mathsf{VRed}^{r}_{\sigma,\rho}$. Since the result is true for all $r \in [0,p)$, we obtain by Lemma 6.10 that $V \in \mathsf{VRed}^{p}_{\sigma,\rho}$. □

We finally deduce from the two previous lemmas the proposition of interest, relating the reducibility of a recursively defined term with the one of its unfoldings:

Proposition 6.20 (Reducibility is Stable by Unfolding). *Let $n \in \mathbb{N}$ and $V = (\mathsf{letrec}\ f = W)$ be a closed value. Suppose that Z is the n-unfolding of V. Then $V \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{s}}\rightarrow\mu,\rho}$ if and only if $Z \in \mathsf{VRed}^{p}_{\mathsf{Nat}^{\mathfrak{s}}\rightarrow\mu,\rho}$.*

Proof. A direct consequence of Lemma 6.18 and Lemma 6.19. □

6.7 Reducibility Sets versus Reductions and Probabilistic Combinations

If a distribution obtained as a partial approximation of the semantics $[\![M]\!]$ of a term M is reducible for a type μ_n, then all the partial approximations of $[\![M]\!]$ obtained by iterating at least as many times the reduction relation $\Rrightarrow_v$ have the same degree of reducibility, for a greater type:

Lemma 6.21. *Suppose that $M \Rrightarrow_v^n \mathscr{D}_n \in \mathsf{DRed}^p_{\mu_n,\rho}$ for $\mu_n \preccurlyeq \mu$, with $\sum \mu = 1$. Suppose that, for $m \geq n$, $M \Rrightarrow_v^m \mathscr{D}_m$. Then there exists $\mu_n \preccurlyeq \mu_m \preccurlyeq \mu$ such that $\mathscr{D}_m \in \mathsf{DRed}^p_{\mu_m,\rho}$.*

Proof. Let $\mu_n = \{ (\sigma_j)^{p'_j} \mid j \in \mathcal{J} \}$. As $\mathscr{D}_n \in \mathsf{DRed}^p_{\mu_n,\rho}$, there exists a pseudo-representation $\mathscr{D}_n = [\, V_i^{p_i} \mid i \in \mathcal{I} \,]$ and two families of reals $(p_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ and $(q_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ such that

(1) $\forall i \in \mathcal{I},\ \forall j \in \mathcal{J},\ V_i \in \mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho}$;
(2) $\forall i \in \mathcal{I},\ \sum_{j\in\mathcal{J}} p_{ij} = p_i$;
(3) $\forall j \in \mathcal{J},\ \sum_{i\in\mathcal{I}} p_{ij} = p'_j$; and
(4) $p \leq \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} q_{ij} p_{ij}$.

By Lemma 3.8, we have $\mathscr{D}_n \preccurlyeq \mathscr{D}_m$ so that the distribution $\mathscr{D}_m$ admits a pseudo-representation $\mathscr{D}_m = [\, V_i^{p_i} \mid i \in \mathcal{I} \uplus \mathcal{K} \,]$ extending the one of $\mathscr{D}_n$. We now need to define appropriate families of reals $(p'_{ij})_{i\in\mathcal{I}\uplus\mathcal{K},j\in\mathcal{J}}$ and $(q'_{ij})_{i\in\mathcal{I}\uplus\mathcal{K},j\in\mathcal{J}}$. We set:

- $\forall i \in \mathcal{I},\ \forall j \in \mathcal{J},\ p'_{ij} = p_{ij}$;
- $\forall i \in \mathcal{I},\ \forall j \in \mathcal{J},\ q'_{ij} = q_{ij}$; and
- $\forall i \in \mathcal{K},\ \forall j \in \mathcal{J},\ q'_{ij} = 0$,

and we choose the $(p'_{ij})_{i\in\mathcal{K},j\in\mathcal{J}}$ arbitrarily in $[0,1]$ under the constraints that $\forall i \in \mathcal{K},\ \sum_{j\in\mathcal{J}} p'_{ij} = p_i$ and that $\forall j \in \mathcal{J},\ \sum_{i\in\mathcal{I}\uplus\mathcal{K}} p'_{ij} \leq \mu(\sigma_j)$. These constraints are feasible since $\sum_{i\in\mathcal{I}\uplus\mathcal{K}} \sum_{j\in\mathcal{J}} p'_{ij} = \sum_{i\in\mathcal{I}\uplus\mathcal{K}} p_i \leq 1 = \sum \mu$. We then set $\mu_m = \{ (\sigma_j)^{\sum_{i\in\mathcal{I}\uplus\mathcal{K}} p'_{ij}} \mid j \in \mathcal{J} \} \preccurlyeq \mu$. Let us check that $\mathscr{D}_m \in \mathsf{DRed}^p_{\mu_m,\rho}$.

(1) $\forall i \in \mathcal{I},\ \forall j \in \mathcal{J},\ V_i \in \mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho}$, and $\forall i \in \mathcal{K},\ \forall j \in \mathcal{J},\ V_i \in \mathsf{VRed}^{0}_{\sigma_j,\rho}$ as this set contains all terms of simple type $\langle\sigma_j\rangle$ by Lemma 6.6;
(2) $\forall i \in \mathcal{I},\ \sum_{j\in\mathcal{J}} p'_{ij} = p_i$ by definition and $\forall i \in \mathcal{K},\ \sum_{j\in\mathcal{J}} p'_{ij} = p_i$ by construction;
(3) $\forall j \in \mathcal{J},\ \sum_{i\in\mathcal{I}\uplus\mathcal{K}} p'_{ij} = \mu_m(\sigma_j)$ by definition of μ_m; and
(4)

$$\begin{array}{rcl} p & \leq & \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} q_{ij} p_{ij} \\ & = & \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} q'_{ij} p'_{ij} + 0 \\ & = & \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} q'_{ij} p'_{ij} + \sum_{i\in\mathcal{K}} \sum_{j\in\mathcal{J}} q'_{ij} p'_{ij} \\ & = & \sum_{i\in\mathcal{I}\uplus\mathcal{K}} \sum_{j\in\mathcal{J}} q'_{ij} p'_{ij}. \end{array}$$

So $\mathscr{D}_m \in \mathsf{DRed}^p_{\mu_m,\rho}$. □

When two distributions $\mathscr{D}$ and $\mathscr{E}$ are reducible, with respective degrees of reducibility p' and p'', their probabilistic combination $\mathscr{D} \oplus_p \mathscr{E}$ is reducible as well with degree of reducibility $pp' + (1-p)p''$, for the distribution type computed by $\oplus_p$:

Lemma 6.22. *Suppose that $\langle\mu\rangle = \langle\nu\rangle$, that $\mathscr{D} \in \mathsf{DRed}^{p'}_{\mu,\rho}$, and that $\mathscr{E} \in \mathsf{DRed}^{p''}_{\nu,\rho}$. Then $p\mathscr{D} + (1-p)\mathscr{E} \in \mathsf{DRed}^{pp'+(1-p)p''}_{\mu\oplus_p\nu,\rho}$.*

Proof. Let $\mu = \{\, (\sigma_j)^{p'_j} \;\mid\; j \in \mathcal{J} \,\}$. Since $\mathscr{D} \in \mathrm{DRed}^{p'}_{\mu,\rho}$, there exists a pseudo-representation $\mathscr{D} = [\, V_i^{p_i} \;\mid\; i \in \mathcal{I} \,]$ and two families of reals $(p_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ and $(q_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ such that

(1) $\forall i \in \mathcal{I},\ \forall j \in \mathcal{J},\ V_i \in \mathrm{VRed}^{q_{ij}}_{\sigma_j,\rho}$;
(2) $\forall i \in \mathcal{I},\ \sum_{j\in\mathcal{J}} p_{ij} = p_i$;
(3) $\forall j \in \mathcal{J},\ \sum_{i\in\mathcal{I}} p_{ij} = p'_j$; and
(4) $p' \leq \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} q_{ij} p_{ij}$.

Let $\nu = \{\, (\tau_l)^{p'''_l} \;\mid\; l \in \mathcal{L} \,\}$. Since $\mathscr{E} \in \mathrm{DRed}^{p''}_{\nu,\rho}$, there exists a pseudo-representation $\mathscr{E} = [\, W_k^{p''_k} \;\mid\; k \in \mathcal{K} \,]$ and two families of reals $(p'_{kl})_{k\in\mathcal{K},l\in\mathcal{L}}$ and $(q'_{kl})_{k\in\mathcal{K},l\in\mathcal{L}}$ such that

(1) $\forall k \in \mathcal{K},\ \forall l \in \mathcal{L},\ W_k \in \mathrm{VRed}^{q'_{kl}}_{\tau_l,\rho}$;
(2) $\forall k \in \mathcal{K},\ \sum_{l\in\mathcal{L}} p'_{kl} = p''_k$;
(3) $\forall l \in \mathcal{L},\ \sum_{k\in\mathcal{K}} p'_{kl} = p'''_l$; and
(4) $p'' \leq \sum_{k\in\mathcal{K}} \sum_{l\in\mathcal{L}} q'_{kl} p'_{kl}$.

We suppose that $\mathcal{I}$ and $\mathcal{K}$ are disjoint, and that $j \in \mathcal{J} \cap \mathcal{L} \Leftrightarrow \sigma_j = \tau_j$. To prove that $p\mathscr{D} + (1-p)\mathscr{E} \in \mathrm{DRed}^{pp'+(1-p)p''}_{\mu\oplus_p\nu,\rho}$, we consider the pseudo-representation

$$p\mathscr{D} + (1-p)\mathscr{E} = \left[\, V_i^{pp_i} \;\middle|\; i \in \mathcal{I} \,\right] + \left[\, W_k^{(1-p)p''_k} \;\middle|\; k \in \mathcal{K} \,\right], \tag{14}$$

and we write the distribution type $\mu \oplus_p \nu$ as

$$\left\{ (\sigma_j)^{pp'_j} \middle| j \in \mathcal{J} \setminus (\mathcal{J}\cap\mathcal{L}) \right\} + \left\{ (\sigma_j)^{pp'_j+(1-p)p'''_j} \middle| j \in \mathcal{J}\cap\mathcal{L} \right\} + \left\{ (\tau_l)^{(1-p)p'''_l} \middle| l \in \mathcal{L} \setminus (\mathcal{J}\cap\mathcal{L}) \right\}.$$

We set $\mathcal{G} = \mathcal{I} + \mathcal{K}$ and $\mathcal{H} = \mathcal{J} + \mathcal{L}$. We now need to define appropriate families of reals $(p''_{gh})_{g\in\mathcal{G},h\in\mathcal{H}}$ and $(q''_{gh})_{g\in\mathcal{G},h\in\mathcal{H}}$. We proceed as follows:

- If $g \in \mathcal{I}$ and $h \in \mathcal{J}$, $p''_{gh} = pp_{gh}$ and $q''_{gh} = q_{gh}$.
- If $g \in \mathcal{I}$ and $h \in \mathcal{L}$, $p''_{gh} = 0$ and $q''_{gh} = 0$.
- If $g \in \mathcal{K}$ and $h \in \mathcal{J}$, $p''_{gh} = 0$ and $q''_{gh} = 0$.
- If $g \in \mathcal{K}$ and $h \in \mathcal{L}$, $p''_{gh} = (1-p)p'_{gh}$ and $q''_{gh} = q'_{gh}$.

Let us prove that Equation (14) together with these two families provides a witness that $p\mathscr{D} + (1-p)\mathscr{E} \in \mathrm{DRed}^{pp'+(1-p)p''}_{\mu\oplus_p\nu,\rho}$ by checking the four usual conditions. We write Z_g either for V_i or W_k, depending on the context. We write similarly θ_h for σ_j or τ_l.

(1) $\forall g \in \mathcal{G},\ \forall h \in \mathcal{H},\ Z_g \in \mathrm{VRed}^{q''_{gh}}_{\theta_h,\rho}$ is proved by case exhaustion:
   - $\forall g \in \mathcal{I},\ \forall h \in \mathcal{J},\ V_g \in \mathrm{VRed}^{q_{gh}}_{\sigma_h,\rho}$ since $\mathscr{D} \in \mathrm{DRed}^{p'}_{\mu,\rho}$;
   - $\forall g \in \mathcal{K},\ \forall h \in \mathcal{L},\ W_g \in \mathrm{VRed}^{q'_{gh}}_{\tau_h,\rho}$ since $\mathscr{E} \in \mathrm{DRed}^{p''}_{\nu,\rho}$; and
   - in the two remaining cases, $q''_{gh} = 0$ and by Lemma 6.6 the result holds.

(2) We proceed again by case exhaustion:
   - If $g \in \mathcal{I}$, $\sum_{h\in\mathcal{H}} p''_{gh} = \sum_{h\in\mathcal{J}} p''_{gh} + \sum_{h\in\mathcal{L}} p''_{gh} = \sum_{h\in\mathcal{J}} pp_{gh} = pp_g$.
   - If $g \in \mathcal{K}$, $\sum_{h\in\mathcal{H}} p''_{gh} = \sum_{h\in\mathcal{L}} (1-p)p'_{gh} = (1-p)p''_g$.

(3) We proceed again by case exhaustion:
   - Suppose that $h \in \mathcal{J} \setminus (\mathcal{J}\cap\mathcal{L})$. Then $\sum_{g\in\mathcal{G}} p''_{gh} = \sum_{g\in\mathcal{I}} p''_{gh} = \sum_{g\in\mathcal{I}} pp_{gh} = pp'_g$.

- Suppose that $h \in \mathcal{L} \setminus (\mathcal{J} \cap \mathcal{L})$. Then $\sum_{g \in \mathcal{G}} \ p''_{gh} = \sum_{g \in \mathcal{K}} \ p''_{gh} = \sum_{g \in \mathcal{K}} \ (1-p)p'_{gh} = (1-p)p'''_{g}$.
- Suppose that $h \in \mathcal{J} \cap \mathcal{L}$. Then $\sum_{g \in \mathcal{G}} \ p''_{gh} = \sum_{g \in \mathcal{I}} \ p''_{gh} \ + \ \sum_{g \in \mathcal{K}} \ p''_{gh} = pp'_g + (1-p)p'''_{g}$.

(4)

$$\begin{array}{rl} & \sum_{g \in \mathcal{G}} \sum_{h \in \mathcal{H}} \ q''_{gh} p''_{gh} \\ = & \sum_{g \in \mathcal{I}} \sum_{h \in \mathcal{J}} \ q''_{gh} p''_{gh} + \sum_{g \in \mathcal{K}} \sum_{h \in \mathcal{L}} \ q''_{gh} p''_{gh} \\ = & \sum_{g \in \mathcal{I}} \sum_{h \in \mathcal{J}} \ q_{gh} p p_{gh} + \sum_{g \in \mathcal{K}} \sum_{h \in \mathcal{L}} \ q'_{gh} (1-p) p'_{gh} \\ = & p \sum_{g \in \mathcal{I}} \sum_{h \in \mathcal{J}} \ q_{gh} p_{gh} + (1-p) \sum_{g \in \mathcal{K}} \sum_{h \in \mathcal{L}} \ q'_{gh} p'_{gh} \\ \geq & pp' + (1-p)p'' \end{array}$$

It follows that $p\mathscr{D} + (1-p)\mathscr{E} \in \mathsf{DRed}^{pp'+(1-p)p''}_{\mu \oplus_p \nu, \rho}$. □

This lemma generalizes to the n-ary case of a weighted sum of distributions:

Lemma 6.23. *Let $(\mu_i)_{i \in \mathcal{I}}$ be a family of distribution types of the same underlying type. For every $i \in \mathcal{I}$, let $\mathscr{D}_i \in \mathsf{DRed}^{q_i}_{\mu_i, \rho}$. Let $(p_i)_{i \in \mathcal{I}}$ be a family of reals of $[0,1]$ such that $\sum_{i \in \mathcal{I}} \ p_i \leq 1$. Then $\sum_{i \in \mathcal{I}} \ p_i \mathscr{D}_i \in \mathsf{DRed}^{\sum_{i \in \mathcal{I}} \ p_i q_i}_{\sum_{i \in \mathcal{I}} \ p_i \mu_i, \rho}$.*

Proof. Similar to the proof of Lemma 6.22. □

TRed is closed by antireduction for Dirac distributions but also in the case corresponding to the reduction of a choice operator:

Lemma 6.24 (Reductions and Sets of Candidates).

- *Suppose that $M \to_v \{ N^1 \}$ and that $N \in \mathsf{TRed}^{p}_{\mu, \rho}$. Then $M \in \mathsf{TRed}^{p}_{\mu, \rho}$.*
- *Suppose that $M \to_v \{ N^p, L^{1-p} \}$, that $N \in \mathsf{TRed}^{p'}_{\mu, \rho}$, and that $L \in \mathsf{TRed}^{p''}_{\nu, \rho}$. Then $M \in \mathsf{TRed}^{pp'+(1-p)p''}_{\mu \oplus_p \nu, \rho}$.*

Proof.

- Since $N \in \mathsf{TRed}^{p}_{\mu, \rho}$, for every $0 \leq r < p$ there exists $\nu_r \preccurlyeq \mu$ and $n_r \in \mathbb{N}$ such that $N \Rrightarrow_v^{n_r} \mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu_r, \rho}$. Recall that $\Rrightarrow_v^{n_r+1} = \to_v \circ \Rrightarrow_v^{n_r}$. It follows that $M \Rrightarrow_v^{n_r+1} \mathscr{D}_r$, which has the required properties, so that $M \in \mathsf{TRed}^{p}_{\mu, \rho}$.
- Let $0 \leq r < pp' + (1-p)p''$. Let (r', r'') be such that $r = pr' + (1-p)r''$, $0 \leq r' < p'$, and $0 \leq r'' < p''$. Since $N \in \mathsf{TRed}^{p'}_{\mu, \rho}$, there exists $n_{r'}$ and $\mu_{r'} \preccurlyeq \mu$ such that $N \Rrightarrow_v^{n_{r'}} \mathscr{D}_{r'} \in \mathsf{DRed}^{r'}_{\mu_{r'}, \rho}$. Since $L \in \mathsf{TRed}^{p''}_{\nu, \rho}$, there exists $m_{r''}$ and $\nu_{r''} \preccurlyeq \nu$ such that $L \Rrightarrow_v^{m_{r''}} \mathscr{E}_{r''} \in \mathsf{DRed}^{r''}_{\nu_{r''}, \rho}$. Suppose that $n_{r'} \leq m_{r''}$, the dual case being exactly symmetrical. By Lemma 6.21, by denoting $\mathscr{D}_{r''}$ the distribution such that $N \Rrightarrow_v^{m_{r''}} \mathscr{D}_{r''}$, there exists $\mu_{r'} \preccurlyeq \mu_{r''} \preccurlyeq \mu$ such that $\mathscr{D}_{r''} \in \mathsf{DRed}^{r'}_{\mu_{r''}, \rho}$. Now $M \Rrightarrow_v^{m_{r''}+1} p\mathscr{D}_{r''} + (1-p)\mathscr{E}_{r''}$, and by Lemma 6.22 we have $p\mathscr{D}_{r''} + (1-p)\mathscr{E}_{r''} \in \mathsf{DRed}^{pr'+(1-p)r''}_{\mu_{r''} \oplus_p \nu_{r''}, \rho}$. Since by construction $\mu_{r''} \oplus_p \nu_{r''} \preccurlyeq \mu \oplus_p \nu$, we can conclude that $M \in \mathsf{TRed}^{pp'+(1-p)p''}_{\mu \oplus_p \nu, \rho}$. □

6.8 Subtyping Soundness

Reducibility sets are monotonic with respect to the subtyping order $\sqsubseteq$:

Lemma 6.25 (Subtyping Soundness).

- *Suppose that $\sigma \sqsubseteq \tau$. Then, for every $p \in [0,1]$ and ρ, $\mathsf{VRed}^{p}_{\sigma,\rho} \subseteq \mathsf{VRed}^{p}_{\tau,\rho}$.*
- *Suppose that $\mu \sqsubseteq \nu$ and that $\sum \mu = \sum \nu$. Then, for every $p \in [0,1]$ and ρ, $\mathsf{DRed}^{p}_{\mu,\rho} \subseteq \mathsf{DRed}^{p}_{\nu,\rho}$.*
- *Suppose that $\mu \sqsubseteq \nu$. Then, for every $p \in [0,1]$ and ρ, $\mathsf{TRed}^{p}_{\mu,\rho} \subseteq \mathsf{TRed}^{p}_{\nu,\rho}$.*

Proof. The proof is by mutual induction on the statements following the shape of the simple type refined by σ and μ, as earlier.

- Suppose that σ :: Nat. Then $\sigma = \mathsf{Nat}^{\mathfrak{s}}$ and $\tau = \mathsf{Nat}^{\mathfrak{r}}$ with $\mathfrak{s} \preccurlyeq \mathfrak{r}$. Let $V \in \mathsf{VRed}^{p}_{\sigma,\rho}$. There are three possibilities:
  —Either $\mathfrak{s} = \widehat{\mathfrak{i}}^{k}$ and $\mathfrak{r} = \widehat{\mathfrak{i}}^{k'}$ with $k \leq k'$. Then V is of the shape $\mathsf{S}^n\ 0$. If $p = 0$, the result is immediate. Else we have $n < [\![\mathfrak{s}]\!]_{\rho} = \rho(\mathfrak{i}) + k \leq \rho(\mathfrak{i}) + k' = [\![\mathfrak{r}]\!]_{\rho}$ so that $V \in \mathsf{VRed}^{p}_{\tau,\rho}$.
  —Or $\mathfrak{s} = \widehat{\mathfrak{i}}^{k}$ and $\mathfrak{r} = \infty$. In this case V is of the shape $\mathsf{S}^n\ 0$ and therefore $V \in \mathsf{VRed}^{p}_{\tau,\rho}$.
  —Or $\mathfrak{s} = \mathfrak{r} = \infty$. In this case $\sigma = \tau$ and the result is immediate.
- Suppose that $\sigma = \sigma' \to \mu$ and that $\tau = \tau' \to \nu$. Let $p \in [0,1]$, ρ be a size environment, and $V \in \mathsf{VRed}^{p}_{\sigma,\rho}$. We have that $\tau' \sqsubseteq \sigma'$ and $\mu \sqsubseteq \nu$. It follows, by induction hypothesis, that $\mathsf{VRed}^{p'}_{\tau',\rho} \subseteq \mathsf{VRed}^{p'}_{\sigma',\rho}$ and that $\mathsf{TRed}^{p'}_{\mu,\rho} \subseteq \mathsf{TRed}^{p'}_{\nu,\rho}$ for every $p' \in [0,1]$. Since $V \in \mathsf{VRed}^{p}_{\sigma,\rho}$, for every $q \in (0,1]$ and $W \in \mathsf{VRed}^{q}_{\sigma',\rho}$, $V\,W \in \mathsf{TRed}^{pq}_{\mu,\rho} \subseteq \mathsf{TRed}^{pq}_{\nu,\rho}$. As $\mathsf{VRed}^{q}_{\tau',\rho} \subseteq \mathsf{VRed}^{q}_{\sigma',\rho}$, $V \in \mathsf{VRed}^{p}_{\tau,\rho}$.
- Suppose that $\mu = \{\, \sigma_j^{p'_j} \mid j \in \mathcal{J} \,\}$ and that $\nu = \{\, \tau_k^{p''_k} \mid k \in \mathcal{K} \,\}$. By definition of subtyping, there exists $f : \mathcal{J} \to \mathcal{K}$ such that for all $j \in \mathcal{J}$, $\sigma_j \sqsubseteq \tau_{f(j)}$ and that for all $k \in \mathcal{K}$, $\sum_{j \in f^{-1}(k)} p'_j \leq p''_k$. Note that since $\sum \mu = \sum \nu$, this is in fact an equality. Let $\mathscr{D} \in \mathsf{DRed}^{p}_{\mu,\rho}$; then there exists a pseudo-representation $\mathscr{D} = [\, (V_i)^{p_i} \mid i \in \mathcal{I} \,]$ and families $(p_{ij})_{i \in \mathcal{I}, j \in \mathcal{J}}$ and $(q_{ij})_{i \in \mathcal{I}, j \in \mathcal{J}}$ of reals of $[0,1]$ satisfying:

  (1) $\forall i \in \mathcal{I},\ \forall j \in \mathcal{J},\ V_i \in \mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho}$;
  (2) $\forall i \in \mathcal{I},\ \sum_{j \in \mathcal{J}} p_{ij} = p_i$;
  (3) $\forall j \in \mathcal{J},\ \sum_{i \in \mathcal{I}} p_{ij} = p'_j$; and
  (4) $p \leq \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} q_{ij} p_{ij}$.

  By induction hypothesis, for every $j \in \mathcal{J}$, $\mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho} \subseteq \mathsf{VRed}^{q_{ij}}_{\tau_{f(j)},\rho}$. We now prove that $[\, (V_i)^{p_i} | i \in \mathcal{I} \,]$ witnesses that $\mathscr{D} \in \mathsf{DRed}^{p}_{\nu,\rho}$. We need to define families of reals $(p'_{ik})_{i \in \mathcal{I}, k \in \mathcal{K}}$ and $(q'_{ik})_{i \in \mathcal{I}, k \in \mathcal{K}}$ satisfying the four usual conditions. To this end, for every $i \in \mathcal{I},\ k \in \mathcal{K}$, we set

  $$p'_{ik} = \sum_{j \in f^{-1}(k)} p_{ij}$$

  and

  $$q'_{ik} = \max_{j \in f^{-1}(k)} q_{ij}.$$

  Let us check that the four conditions hold:

  (1) $\forall i \in \mathcal{I},\ \forall k \in \mathcal{K},\ V_i \in \mathsf{VRed}^{q'_{ik}}_{\tau_{f(j)},\rho}$ by induction hypothesis and by definition of q'_{ik};
  (2) $\forall i \in \mathcal{I},\ \sum_{k \in \mathcal{K}} p'_{ik} = \sum_{k \in \mathcal{K}} \sum_{j \in f^{-1}(k)} p_{ij} = \sum_{j \in \mathcal{J}} p_{ij} = p_i$;
  (3) $\forall k \in \mathcal{K},\ \sum_{i \in \mathcal{I}} p'_{ik} = \sum_{i \in \mathcal{I}} \sum_{j \in f^{-1}(k)} p_{ij} = \sum_{j \in f^{-1}(k)} \sum_{i \in \mathcal{I}} p_{ij} = \sum_{j \in f^{-1}(k)} p'_j = p''_k$; and

(4)

$$\begin{array}{rcl} p & \leq & \sum_{i \in \mathcal{I}} \ \sum_{j \in \mathcal{J}} \ q_{ij} p_{ij} \\ & = & \sum_{i \in \mathcal{I}} \ \sum_{k \in \mathcal{K}} \ \sum_{j \in f^{-1}(k)} \ q_{ij} p_{ij} \\ & \leq & \sum_{i \in \mathcal{I}} \ \sum_{k \in \mathcal{K}} \ \sum_{j \in f^{-1}(k)} \ q'_{if(j)} p_{ij} \\ & = & \sum_{i \in \mathcal{I}} \ \sum_{k \in \mathcal{K}} \ q'_{ik} \sum_{j \in f^{-1}(k)} \ p_{ij} \\ & = & \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \ q'_{ik} p'_{ik}. \end{array}$$

It follows that $\mathscr{D} \in \mathsf{DRed}^{p}_{\nu,\rho}$.

- Suppose that $\mu = \{ \sigma_j^{p'_j} \mid j \in \mathcal{J} \}$ and that $\nu = \{ \tau_k^{p''_k} \mid k \in \mathcal{K} \}$. By definition of subtyping, there exists $f : \mathcal{J} \to \mathcal{K}$ such that for all $j \in \mathcal{J}$, $\sigma_j \sqsubseteq \tau_{f(j)}$ and that for all $k \in \mathcal{K}$, $\sum_{j \in f^{-1}(k)} \ p'_j \leq p''_k$. Let $M \in \mathsf{TRed}^{p}_{\mu,\rho}$. Then, for every $0 \leq r < p$, there exists $\mu'_r \preccurlyeq \mu$ and n_r such that $M \Rrightarrow_v^{n_r} \mathscr{D}_r \in \mathsf{DRed}^{r}_{\mu'_r,\rho}$. By definition of $\mu'_r \preccurlyeq \mu$, $\mu'_r = [\sigma_j^{q'_j} \mid j \in \mathcal{J}]$ with $q'_j \leq p'_j$ for every $j \in \mathcal{J}$. We set $\nu'_r = [\tau_{f(j)}^{q'_j} \mid j \in \mathcal{J}]$, which is such that $\sum \mu'_r = \sum \nu'_r$ and, by construction, $\mu'_r \sqsubseteq \nu'_r$ so that we can apply the induction hypothesis and obtain that $M \Rrightarrow_v^{n_r} \mathscr{D}_r \in \mathsf{DRed}^{r}_{\nu'_r,\rho}$. The result follows, since by construction $\nu'_r \preccurlyeq \nu$. □

6.9 Reducibility Sets for Open Terms

We are now ready to extend the notion of the reducibility set from the realm of *closed* terms to the one of *open* terms. This turns out to be subtle. The guiding intuition is that one would like to define a term M with free variables in $\overrightarrow{x}$ to be reducible if and only if any closure $M[\overrightarrow{V}/\overrightarrow{x}]$ is itself reducible in the sense of Definition 6.2. What happens, however, to the underlying degree of reducibility p? How do we relate the degrees of reducibility of $\overrightarrow{V}$ with the one of $M[\overrightarrow{V}/\overrightarrow{x}]$? Informally, the behavior of the reducibility degrees is multiplicative, except for the distribution context, which requires some more care to match the behavior of the sized walk:

- In the case of a context $\Gamma \mid \emptyset$ for $\Gamma = x_1 : \sigma_1, \ldots, x_n : \sigma_n$, which is the simplest one, if we substitute each x_i with $V_i \in \mathsf{VRed}^{q_i}_{\sigma_i,\rho}$, we require that $M[\overrightarrow{V}/\overrightarrow{x}] \in \mathsf{TRed}^{\prod_{i=1}^{n} q_i}_{\mu,\rho}$. In other words, the reducibility degree that we require is the product of the ones of the values substituted for the variables. If we think about termination, we may see this as the fact that, when each of the terms replacing the x_i terminates with probability at least q_i, then $M[\overrightarrow{V}/\overrightarrow{x}]$ should terminate with probability at least $\prod_{i=1}^{n} q_i$. Without this guarantee, the open variables of M could be substituted with values and nevertheless not be AST. Note that here we considered the case of an open term M, but it is exactly the same for a value W with open variables.
- If the context is $\emptyset \mid y : \{\tau_i^{p_i}\}_{i \in I}$, then we substitute the only free variable y in M with a value $V \in \bigcap_{i \in \mathcal{I}} \ \mathsf{VRed}^{q_i}_{\sigma_i,\rho}$. We then require that $M[V/y] \in \mathsf{TRed}^{\sum_{i \in I} p_i q_i + 1 - (\sum_{j \in \mathcal{J}} p_j)}_{\mu,\rho}$. Note that the degree of reducibility we find here is precisely connected to Equation (15) in the proof of Lemma 6.31, if one takes q_i to be $Pr_{n,m'+k_i}$. Here the informal explanation is that if the replacement V of y of type τ_i is in $\mathsf{VRed}^{q_i}_{\tau_i,\rho}$, then when used with type τ_i, the value V terminates with probability at least q_i. Since it is used with this type with probability p_i, the substitution terminates with probability at least $\sum_{i \in I} p_i q_i + 1 - (\sum_{j \in \mathcal{J}} p_j)$ because:
  - each i contributes to termination with probability $p_i q_i$, and
  - with probability $1 - (\sum_{j \in \mathcal{J}} p_j)$ the variable y is not used, so that no recursive call will be performed with this probability and termination is therefore ensured.

  Again, the same holds when the open term is a value.

- When the context is of the form $\Gamma \mid \Theta$, the resulting degree of reducibility is obtained as the result of the multiplication of the degrees obtained in both previous cases.

We may now introduce the reducibility sets for open terms in a formal way:

Definition 6.26 (Reducibility Sets for Open Terms). Suppose that Γ is a sized context in the form $x_1 : \sigma_1, \ldots, x_n : \sigma_n$, and that y is a variable distinct from $x_1, \ldots, x_n$. Then we define the following sets of terms and values:

$$\begin{aligned}
\mathsf{OTRed}^{\Gamma \mid \emptyset}_{\mu,\rho} &= \Big\{ M \ \Big|\ \forall (q_i)_i \in [0,1]^n,\ \forall\,(V_1,\ \ldots,\ V_n) \in \textstyle\prod_{i=1}^n\ \mathsf{VRed}^{q_i}_{\sigma_i,\rho}, \\
&\qquad\qquad M\Big[\overrightarrow{V}/\overrightarrow{x}\Big] \in \mathsf{TRed}^{\prod_{i=1}^n q_i}_{\mu,\rho} \Big\} \\
\mathsf{OVRed}^{\Gamma \mid \emptyset}_{\mu,\rho} &= \Big\{ W \ \Big|\ \forall (q_i)_i \in [0,1]^n,\ \forall\,(V_1,\ \ldots,\ V_n) \in \textstyle\prod_{i=1}^n\ \mathsf{VRed}^{q_i}_{\sigma_i,\rho}, \\
&\qquad\qquad W\Big[\overrightarrow{V}/\overrightarrow{x}\Big] \in \mathsf{VRed}^{\prod_{i=1}^n q_i}_{\mu,\rho} \Big\}
\end{aligned}$$

$$\begin{aligned}
\mathsf{OTRed}^{\Gamma \mid y : \{\tau_j^{p_j}\}_{j\in\mathcal{J}}}_{\mu,\rho} &= \Big\{ M \ \Big|\ \forall (q_i)_i \in [0,1]^n,\ \forall \overrightarrow{V} \in \textstyle\prod_{i=1}^n\ \mathsf{VRed}^{q_i}_{\sigma_i,\rho}, \\
&\qquad \forall (q'_j)_j \in [0,1]^{\mathcal{J}},\ \forall W \in \textstyle\bigcap_{j\in\mathcal{J}}\ \mathsf{VRed}^{q'_j}_{\tau_j,\rho}, \\
&\qquad\qquad M\Big[\overrightarrow{V}, W/\overrightarrow{x}, y\Big] \in \mathsf{TRed}^{\alpha}_{\mu,\rho} \Big\} \\
\mathsf{OVRed}^{\Gamma \mid y : \{\tau_j^{p_j}\}_{j\in\mathcal{J}}}_{\mu,\rho} &= \Big\{ Z \ \Big|\ \forall (q_i)_i \in [0,1]^n,\ \forall \overrightarrow{V} \in \textstyle\prod_{i=1}^n\ \mathsf{VRed}^{q_i}_{\sigma_i,\rho}, \\
&\qquad \forall (q'_j)_j \in [0,1]^{\mathcal{J}},\ \forall W \in \textstyle\bigcap_{j\in\mathcal{J}}\ \mathsf{VRed}^{q'_j}_{\tau_j,\rho}, \\
&\qquad\qquad Z\Big[\overrightarrow{V}, W/\overrightarrow{x}, y\Big] \in \mathsf{VRed}^{\alpha}_{\mu,\rho} \Big\},
\end{aligned}$$

where the degree of reducibility α is defined as

$$\alpha = \left(\prod_{i=1}^n q_i\right)\left(\left(\sum_{j\in\mathcal{J}} p_j q'_j\right) + 1 - \left(\sum_{j\in\mathcal{J}} p_j\right)\right).$$

Note that this contains

$$\begin{aligned}
\mathsf{OTRed}^{\emptyset \mid \emptyset}_{\mu,\rho} &= \mathsf{TRed}^{1}_{\mu,\rho} \\
\mathsf{OVRed}^{\emptyset \mid \emptyset}_{\sigma,\rho} &= \mathsf{VRed}^{1}_{\sigma,\rho} \\
\mathsf{OTRed}^{\emptyset \mid y : \{\tau_i^{p_i}\}_{i\in I}}_{\mu,\rho} &= \Big\{ M \ \Big|\ \forall\,(q_i)_i \in [0,1]^{\mathcal{I}},\ \forall V \in \textstyle\bigcap_{i\in\mathcal{I}}\ \mathsf{VRed}^{q_i}_{\tau_i,\rho}, \\
&\qquad\qquad M\,[V/y] \in \mathsf{TRed}^{\sum_{i\in I} p_i q_i + 1 - (\sum_{j\in\mathcal{J}} p_j)}_{\mu,\rho} \Big\} \\
\mathsf{OVRed}^{\emptyset \mid y : \{\tau_i^{p_i}\}_{i\in I}}_{\mu,\rho} &= \Big\{ W \ \Big|\ \forall\,(q_i)_i \in [0,1]^{\mathcal{I}},\ \forall V \in \textstyle\bigcap_{i\in\mathcal{I}}\ \mathsf{VRed}^{q_i}_{\tau_i,\rho}, \\
&\qquad\qquad W\,[V/y] \in \mathsf{VRed}^{\sum_{i\in I} p_i q_i + 1 - (\sum_{j\in\mathcal{J}} p_j)}_{\mu,\rho} \Big\}.
\end{aligned}$$

Note also that these sets extend the ones for closed terms: in particular, $\mathsf{OTRed}^{\emptyset \mid \emptyset}_{\mu,\rho} = \mathsf{TRed}^{1}_{\mu,\rho}$.

As for closed terms (Lemma 6.19), reducible values are reducible terms:

Lemma 6.27 (Reducible Values are Reducible Terms). *For every Γ, Θ, σ, and ρ, $V \in \mathsf{OVRed}^{\Gamma \mid \Theta}_{\sigma,\rho}$ if and only if $V \in \mathsf{OTRed}^{\Gamma \mid \Theta}_{\{\,\sigma^1\,\},\rho}$. An immediate consequence is that $\mathsf{OVRed}^{\Gamma \mid \Theta}_{\sigma,\rho} \subseteq \mathsf{OTRed}^{\Gamma \mid \Theta}_{\{\,\sigma^1\,\},\rho}$.*

Proof. Corollary of Lemma 6.19 and of the definitions of the candidates for open sets. □

6.10 Reducibility and Sized Walks

To handle the fix-point rule, we need to relate the notion of sized walk, which guards it with the reducibility sets, and in particular with the degrees of reducibility we can attribute to recursively defined terms.

Definition 6.28 (Probabilities of Convergence in Finite Time). Let us consider a sized walk. We define the associated *probabilities of convergence in finite time*$(Pr_{n,m})_{n\in\mathbb{N},m\in\mathbb{N}}$ as follows: $\forall n \in \mathbb{N}$, $\forall m \in \mathbb{N}$, and the real number $Pr_{n,m}$ is defined as the probability that, starting from m, the sized walk reaches 0 in *at most* n steps.

The point is that, for an AST sized walk, the more we iterate, the closer we get to reaching 0 in finite time n with probability 1.

Lemma 6.29 (Finite Approximations of AST). *Let $m \in \mathbb{N}$ and $\varepsilon \in (0,1]$. Consider a sized walk and its associated probabilities of convergence in finite time $(Pr_{n,m})_{n\in\mathbb{N},m\in\mathbb{N}}$. If the sized walk is AST, there exists $n \in \mathbb{N}$ such that $Pr_{n,m} \geq 1-\varepsilon$.*

Proof. Suppose, by contradiction, that there exists $\varepsilon \in (0,1]$ such that there is no $n \in \mathbb{N}$ with $Pr_{n,m} \geq 1-\varepsilon$. Then $\lim_{n\in\mathbb{N}} Pr_{n,m} \leq 1-\varepsilon$. But this limit should be worth 1 as we supposed the sized walk to be AST. □

The following lemma allows to treat the base case of Lemma 6.31:

Lemma 6.30. *Suppose that V is a closed value of simple type $\mathsf{Nat} \to \kappa$. Then, for every $\mathsf{Nat}^{\mathfrak{i}} \to \mu :: \mathsf{Nat} \to \kappa$, and for every size environment ρ such that $\rho(\mathfrak{i}) = 0$, we have $V \in \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{i}}\to\mu,\rho}$.*

Proof. To prove that $V \in \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{i}}\to\mu,\rho}$, we need to show that for every $q \in (0,1]$ and every $W \in \mathsf{VRed}^{q}_{\mathsf{Nat}^{\mathfrak{i}},\rho}$ we have that $V\ W \in \mathsf{TRed}^{q}_{\mu,\rho}$. This is always the case, as $\mathsf{VRed}^{q}_{\mathsf{Nat}^{\mathfrak{i}},\rho}$ is the empty set by definition: there is no term of the shape $\mathsf{S}^{n}\ 0$ with $n < \rho(\mathfrak{i}) = 0$. □

The following lemma is the crucial result relating sized walks with the reducibility sets. It proves that, when the sized walk is AST, and after substitution of the variables of the context by reducible values in the recursively defined term, we can prove the degree of reducibility to be any probability $Pr_{n,m}$ of convergence in finite time.

Lemma 6.31 (Convergence in Finite Time and letrec). *Consider the distribution type $\mu = \{ (\mathsf{Nat}^{\mathfrak{s}_j} \to \nu[\mathfrak{s}_j/\mathfrak{i}])^{p_j} \mid j \in \mathcal{J} \}$. Let Γ be the sized context $x_1 : \mathsf{Nat}^{\mathfrak{r}_1}, \ldots, x_l : \mathsf{Nat}^{\mathfrak{r}_l}$. Suppose that $\Gamma \,|\, f : \mu \vdash V : \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \nu[\widehat{\mathfrak{i}}/\mathfrak{i}]$ and that μ induces an AST sized walk. Denote by $(Pr_{n,m})_{n\in\mathbb{N},m\in\mathbb{N}}$ its associated probabilities of convergence in finite time. Suppose that $V \in \mathsf{OVRed}^{\Gamma\,|\,f\,:\,\mu}_{\mathsf{Nat}^{\widehat{\mathfrak{i}}}\to\nu[\widehat{\mathfrak{i}}/\mathfrak{i}],\rho}$ for every ρ. Let $\overrightarrow{W} \in \prod_{i=1}^{l} \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{r}_i},\rho}$; then for every $(n,m) \in \mathbb{N}^2$, we have that*

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right] \;\in\; \mathsf{VRed}^{Pr_{n,m}}_{\mathsf{Nat}^{\mathfrak{i}}\to\nu,\,\rho[\mathfrak{i}\mapsto m]}.$$

Proof. We prove the statement by induction on n.

- If $n = 0$, we have two cases:
  - If $m = 0$, then Lemma 6.30 implies that $\mathsf{letrec}\ f = V[\overrightarrow{W}/\overrightarrow{x}] \;\in\; \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{i}}\to\nu,\,\rho[\mathfrak{i}\mapsto 0]}$ so that by downward closure (Lemma 6.7), we obtain $\mathsf{letrec}\ f = V[\overrightarrow{W}/\overrightarrow{x}] \;\in\; \mathsf{VRed}^{Pr_{n,0}}_{\mathsf{Nat}^{\mathfrak{i}}\to\nu,\,\rho[\mathfrak{i}\mapsto 0]}$.
  - If $m \neq 0$, then $Pr_{n,m} = 0$. The hypothesis of the lemma imply that $\mathsf{letrec}\ f = V[\overrightarrow{W}/\overrightarrow{x}] :: \mathsf{Nat} \to \langle \nu \rangle$, and we conclude using Lemma 6.6.

- Suppose that $n \geq 1$:
  - If $m = 0$, the result is immediate as in the previous case.
  - Suppose that $m > 0$. Then $m = m' + 1$. By definition, $\mathfrak{s}_j$ must be of the shape $\widehat{\mathrm{i}}^{k_j}$ with $k_j \geq 0$ for every $j \in \mathcal{J}$. We set $\mathcal{I} = \{k_j \mid j \in \mathcal{J}\}$ and $q_{k_j} = p_j$ for every $j \in \mathcal{J}$. The sized walk induced by the distribution type μ is then the sized walk associated to $(\mathcal{I}, (q_i)_{i \in \mathcal{I}})$, which from the state $m' + 1 \in \mathbb{N} \setminus \{0\}$ moves:
    - to the state $m' + k_j$ with probability p_j, for every $j \in \mathcal{J}$, and
    - to 0 with probability $1 - (\sum_{j \in \mathcal{J}} \ p_j)$.

    It follows that

$$Pr_{n+1,\, m'+1} = \sum_{j \in \mathcal{J}} p_j Pr_{n,\, m'+k_j} \ + \ 1 - \left(\sum_{j \in \mathcal{J}} p_j \right). \tag{15}$$

For every $j \in \mathcal{J}$, let us apply the induction hypothesis and obtain

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right] \ \in \ \mathsf{VRed}^{Pr_{n,\, m'+k_j}}_{\mathsf{Nat}^{\mathrm{i}} \to \nu,\, \rho[\mathrm{i} \mapsto m'+k_j]}.$$

By Lemma 6.12,

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right] \ \in \ \mathsf{VRed}^{Pr_{n,\, m'+k_j}}_{\mathsf{Nat}^{\widehat{\mathrm{i}}^{k_j}} \to \nu[\widehat{\mathrm{i}}^{k_j}/\mathrm{i}],\, \rho[\mathrm{i} \mapsto m']} = \mathsf{VRed}^{Pr_{n,\, m'+k_j}}_{\mathsf{Nat}^{\mathfrak{s}_j} \to \nu\left[\mathfrak{s}_j/\mathrm{i}\right],\, \rho[\mathrm{i} \mapsto m']}.$$

Since this is valid for every $j \in \mathcal{J}$, we have that

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right] \ \in \ \bigcap_{j \in \mathcal{J}} \mathsf{VRed}^{Pr_{n,\, m'+k_j}}_{\mathsf{Nat}^{\mathfrak{s}_j} \to \nu[\mathfrak{s}_j/\mathrm{i}],\, \rho[\mathrm{i} \mapsto m']},$$

and since $V \in \mathsf{OVRed}^{\Gamma \mid f : \mu}_{\mathsf{Nat}^{\widehat{\mathrm{i}}} \to \nu[\widehat{\mathrm{i}}/\mathrm{i}],\, \rho[\mathrm{i} \mapsto m']}$, we obtain

$$V\left[\overrightarrow{W}, \mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right]/\overrightarrow{x}, f\right] \ \in \ \mathsf{VRed}^{\sum_{j \in \mathcal{J}} \ p_j Pr_{n,\, m+k_j} \ + \ 1 - (\sum_{j \in \mathcal{J}} \ p_j)}_{\mathsf{Nat}^{\widehat{\mathrm{i}}} \to \nu\left[\widehat{\mathrm{i}}/\mathrm{i}\right],\, \rho[\mathrm{i} \mapsto m']},$$

which, by Equation (15) and by Lemma 6.12, gives

$$V\left[\overrightarrow{W}, \mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right]/\overrightarrow{x}, f\right] \ \in \ \mathsf{VRed}^{Pr_{n+1,\, m'+1}}_{\mathsf{Nat}^{\mathrm{i}} \to \nu,\, \rho[\mathrm{i} \mapsto m'+1]}.$$

But this term is an unfolding of $\mathsf{letrec}\ f = V[\overrightarrow{W}/\overrightarrow{x}]$, so that by Corollary 6.20 we obtain

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right] \ \in \ \mathsf{VRed}^{Pr_{n+1,\, m}}_{\mathsf{Nat}^{\mathrm{i}} \to \nu,\, \rho[\mathrm{i} \mapsto m]}.$$

Now by definition, $Pr_{n+1,\, m} \geq Pr_{n,\, m}$, and by downward closure (Lemma 6.7):

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right] \ \in \ \mathsf{VRed}^{Pr_{n,\, m}}_{\mathsf{Nat}^{\mathrm{i}} \to \nu,\, \rho[\mathrm{i} \mapsto m]}.$$ □

6.11 Size Environments Mapping Sizes to Infinity

When $m = \infty$, the previous lemma does not allow one to conclude, and an additional argument is required. Indeed, it does not make sense to consider a sized walk beginning from ∞: the meaning of this size is in fact *any integer*, not the ordinal ω. Before we justify this understanding, we need the following companion lemma.

Lemma 6.32. *If* i pos σ, *then*

- $\mathsf{VRed}^{p}_{\sigma,\, \rho[\mathrm{i} \mapsto n]} \subseteq \mathsf{VRed}^{p}_{\sigma,\, \rho[\mathrm{i} \mapsto \infty]}$,
- $\mathsf{DRed}^{p}_{\mu,\, \rho[\mathrm{i} \mapsto n]} \subseteq \mathsf{DRed}^{p}_{\mu,\, \rho[\mathrm{i} \mapsto \infty]}$, *and*
- $\mathsf{TRed}^{p}_{\mu,\, \rho[\mathrm{i} \mapsto n]} \subseteq \mathsf{TRed}^{p}_{\mu,\, \rho[\mathrm{i} \mapsto \infty]}$.

Proof.

- Let $\mathfrak{s} = \widehat{\mathfrak{i}}^n$. We have $[\![\mathfrak{s}]\!]_{\rho[\mathfrak{i}\mapsto 0]} = n$. Using Lemma 6.12, we obtain
$$\mathsf{VRed}^p_{\sigma[\mathfrak{s}/\mathfrak{i}],\rho[\mathfrak{i}\mapsto 0]} = \mathsf{VRed}^p_{\sigma,\rho[\mathfrak{i}\mapsto [\![\mathfrak{s}]\!]_{\rho[\mathfrak{i}\mapsto 0]}]} = \mathsf{VRed}^p_{\sigma,\rho[\mathfrak{i}\mapsto n]}.$$
By the same lemma,
$$\mathsf{VRed}^p_{\sigma[\infty/\mathfrak{i}],\rho[\mathfrak{i}\mapsto 0]} = \mathsf{VRed}^p_{\sigma,\rho[\mathfrak{i}\mapsto [\![\infty]\!]_{\rho[\mathfrak{i}\mapsto 0]}]} = \mathsf{VRed}^p_{\sigma,\rho[\mathfrak{i}\mapsto \infty]}.$$
Since $\mathfrak{i}$ pos σ and $\mathfrak{s} \preccurlyeq \infty$, Lemma 5.14 implies that $\sigma[\mathfrak{s}/\mathfrak{i}] \sqsubseteq \sigma[\infty/\mathfrak{i}]$. By Lemma 6.25, we obtain $\mathsf{VRed}^p_{\sigma[\mathfrak{s}/\mathfrak{i}],\rho[\mathfrak{i}\mapsto 0]} \subseteq \mathsf{VRed}^p_{\sigma[\infty/\mathfrak{i}],\rho[\mathfrak{i}\mapsto 0]}$ and thus $\mathsf{VRed}^p_{\sigma,\rho[\mathfrak{i}\mapsto n]} \subseteq \mathsf{VRed}^p_{\sigma,\rho[\mathfrak{i}\mapsto \infty]}$.
- Let $\mathscr{D} \in \mathsf{DRed}^p_{\mu,\rho[\mathfrak{i}\mapsto n]}$. It follows that $\mathscr{D} = [\,(V_i)^{p_i} \mid i \in \mathcal{I}\,]$ and that, setting $\mu = \{\,(\sigma_j)^{p'_j} \mid j \in \mathcal{J}\,\}$, there exists families $(p_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ and $(q_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$ of reals of $[0,1]$ satisfying:
  (1) $\forall i \in \mathcal{I},\ \ \forall j \in \mathcal{J},\ \ V_i \in \mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho[\mathfrak{i}\mapsto n]}$;
  (2) $\forall i \in \mathcal{I},\ \ \sum_{j\in\mathcal{J}}\ p_{ij} = p_i$;
  (3) $\forall j \in \mathcal{J},\ \ \sum_{i\in\mathcal{I}}\ p_{ij} = \mu(\sigma_j)$; and
  (4) $p \leq \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\ q_{ij}p_{ij}$.
  Since $\forall i \in \mathcal{I},\quad \forall j \in \mathcal{J},\quad V_i \in \mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho[\mathfrak{i}\mapsto n]} \subseteq \mathsf{VRed}^{q_{ij}}_{\sigma_j,\rho[\mathfrak{i}\mapsto \infty]}$, we obtain that $\mathscr{D} \in \mathsf{DRed}^p_{\mu,\rho[\mathfrak{i}\mapsto\infty]}$ using the same witnesses.
- Let $M \in \mathsf{TRed}^p_{\mu,\rho[\mathfrak{i}\mapsto n]}$. It follows that for every $0 \leq r < p$, there exists $\nu_r \preccurlyeq \mu$ and $n_r \in \mathbb{N}$ such that $M \Rrightarrow^{n_r}_{v} \mathscr{D}_r$ and $\mathscr{D}_r \in \mathsf{DRed}^r_{\nu_r,\rho[\mathfrak{i}\mapsto n]}$. But $\mathsf{DRed}^r_{\nu_r,\rho[\mathfrak{i}\mapsto n]} \subseteq \mathsf{DRed}^r_{\nu_r,\rho[\mathfrak{i}\mapsto\infty]}$, so that $M \in \mathsf{TRed}^p_{\mu,\rho[\mathfrak{i}\mapsto\infty]}$. □

The following lemma proves that ∞ stands for "every integer." It proves indeed that, if a term is in a reducibility set for any finite interpretation of a size, then it is also in the set where the size is interpreted as ∞:

Lemma 6.33 (Reducibility for Infinite Sizes). *Suppose that* $\mathfrak{i}$ pos ν *and that* W *is the value* $\mathsf{letrec}\ f = V$. *If* $W \in \mathsf{VRed}^p_{\mathsf{Nat}^{\mathfrak{i}}\to\nu,\rho[\mathfrak{i}\mapsto n]}$ *for every* $n \in \mathbb{N}$, *then* $W \in \mathsf{VRed}^p_{\mathsf{Nat}^{\mathfrak{i}}\to\nu,\rho[\mathfrak{i}\mapsto\infty]}$.

Proof. Suppose that $\mathfrak{i}$ pos ν and that, for every $n \in \mathbb{N}$, $\mathsf{letrec}\ f = V \ \in\ \mathsf{VRed}^p_{\mathsf{Nat}^{\mathfrak{i}}\to\nu,\rho[\mathfrak{i}\mapsto n]}$. Let $W \in \mathsf{VRed}^p_{\mathsf{Nat}^{\mathfrak{i}},\rho[\mathfrak{i}\mapsto\infty]}$. Then $W = \mathsf{S}^m\ 0$ for some $m \in \mathbb{N}$. It follows that $W \in \mathsf{VRed}^p_{\mathsf{Nat}^{\mathfrak{i}},\rho[\mathfrak{i}\mapsto m+1]}$. But $\mathsf{letrec}\ f = V \ \in\ \mathsf{VRed}^p_{\mathsf{Nat}^{\mathfrak{i}}\to\nu,\rho[\mathfrak{i}\mapsto m+1]}$, so that
$$(\mathsf{letrec}\ f = V)\ \ W \in \mathsf{TRed}^p_{\nu,\rho[\mathfrak{i}\mapsto m+1]}.$$
By Lemma 6.32, since $\mathfrak{i}$ pos ν, we obtain that
$$(\mathsf{letrec}\ f = V)\ \ W \in \mathsf{TRed}^p_{\nu,\rho[\mathfrak{i}\mapsto\infty]}.$$
It follows that
$$\mathsf{letrec}\ f = V \ \in\ \mathsf{VRed}^p_{\mathsf{Nat}^{\mathfrak{i}}\to\nu,\rho[\mathfrak{i}\mapsto\infty]}.$$
□

6.12 A Last Technical Lemma

The following technical lemma will allow us to deal with the Let rule in the proof of typing soundness.

LEMMA 6.34. *Let $(q_i)_i \in (0,1]^n$, $(q'_j)_j \in (0,1]^m$, and $(q''_k)_k \in (0,1]^l$. Let $\mathcal{L}$ and $\mathcal{G}$ be two sets of indexes. Let $0 \leq r'' < (\prod_{i=1}^n q_i)(\prod_{j=1}^m q'_j)(\prod_{k=1}^l q''_k)$. Suppose that, for every $0 \leq r < \prod_{j=1}^m q'_j$, there exists two families $(p^r_{lg})_{l \in \mathcal{L}, g \in \mathcal{G}}$ and $(q^r_{lg})_{l \in \mathcal{L}, g \in \mathcal{G}}$ of reals of $[0,1]$ satisfying*

$$r \leq \sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} p^r_{lg} q^r_{lg} \tag{16}$$

and

$$\sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} p^r_{lg} \leq 1. \tag{17}$$

Then there exists $0 \leq r < \prod_{j=1}^m q'_j$ and a family $(r'_{lg})_{l \in \mathcal{L}, g \in \mathcal{G}}$ satisfying

$$\forall l \in \mathcal{L},\ \ \forall g \in \mathcal{G},\ \ 0 \leq r'_{lg} < \left(\prod_{i=1}^n q_i\right)\left(\prod_{k=1}^l q''_k\right) q^r_{lg} \tag{18}$$

and

$$r'' \leq \sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} p^r_{lg} r'_{lg}. \tag{19}$$

PROOF. Since $r'' < (\prod_{i=1}^n q_i)(\prod_{j=1}^m q'_j)(\prod_{k=1}^l q''_k)$, there exists $\varepsilon > 0$ and $\forall l \in \mathcal{L},\ \forall g \in \mathcal{G}$, $\varepsilon_{lg} > 0$ satisfying

$$0 < \varepsilon + \sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} \varepsilon_{lg} < \left(\prod_{i=1}^n q_i\right)\left(\prod_{j=1}^m q'_j\right)\left(\prod_{k=1}^l q''_k\right) - r''. \tag{20}$$

We pick r such that

$$\prod_{j=1}^m q'_j - \varepsilon < r < \prod_{j=1}^m q'_j, \tag{21}$$

and this induces families $(p^r_{lg})_{l \in \mathcal{L}, g \in \mathcal{G}}$ and $(q^r_{lg})_{l \in \mathcal{L}, g \in \mathcal{G}}$ of reals of $[0,1]$ satisfying Equations (16) and (17). We choose a family $(r'_{lg})_{l \in \mathcal{L}, g \in \mathcal{G}}$ such that

$$\forall l \in \mathcal{L},\ \ \forall g \in \mathcal{G},\ \ \left(\prod_{i=1}^n q_i\right)\left(\prod_{k=1}^l q''_k\right) q^r_{lg} - \varepsilon_{lg} < r'_{lg} < \left(\prod_{i=1}^n q_i\right)\left(\prod_{k=1}^l q''_k\right) q^r_{lg}.$$

By Equation (17) and since $(\prod_{i=1}^n q_i)(\prod_{k=1}^l q''_k)$, we obtain from Equation (20) that

$$\left(\prod_{i=1}^n q_i\right)\left(\prod_{k=1}^l q''_k\right)\varepsilon + \sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} p^r_{lg}\varepsilon_{lg} < \left(\prod_{i=1}^n q_i\right)\left(\prod_{j=1}^m q'_j\right)\left(\prod_{k=1}^l q''_k\right) - r''.$$

Thus,

$$r'' < \left(\prod_{i=1}^n q_i\right)\left(\prod_{k=1}^l q''_k\right)\left(\left(\prod_{j=1}^m q'_j\right) - \varepsilon\right) - \sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} p^r_{lg}\varepsilon_{lg}.$$

By Equations (21) and then (16):

$$r'' < \left(\prod_{i=1}^n q_i\right)\left(\prod_{k=1}^l q''_k\right)\left(\sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} p^r_{lg} q^r_{lg}\right) - \sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} p^r_{lg}\varepsilon_{lg},$$

which rewrites to

$$r'' < \sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} p^r_{lg} \left(\left(\prod_{i=1}^{n} q_i \right) \left(\prod_{k=1}^{l} q''_k \right) q^r_{lg} - \varepsilon_{lg} \right),$$

and by definition of $(r'_{lg})_{l \in \mathcal{L}, g \in \mathcal{G}}$ we obtain

$$r'' < \sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} p^r_{lg} r'_{lg}$$

as requested. □

6.13 Typing Soundness

All these fundamental lemmas allow us to prove the following proposition, which expresses that all typable terms are reducible and is the key step toward the fact that typability implies AST:

Proposition 6.35 (Typing Soundness). *If $\Gamma \,|\, \Theta \vdash M : \mu$, then $M \in \mathsf{OTRed}^{\Gamma \,|\, \Theta}_{\mu,\rho}$ for every ρ. Similarly, if $\Gamma \,|\, \Theta \vdash V : \sigma$, then $V \in \mathsf{OVRed}^{\Gamma \,|\, \Theta}_{\sigma,\rho}$ for every ρ.*

Proof. We proceed by induction on the derivation of the sequent $\Gamma \,|\, \Theta \vdash M : \mu$. When $M = V$ is a value, we know by Lemma 4.12 that $\mu = \{ \sigma^1 \}$; and we prove that $V \in \mathsf{OVRed}^{\Gamma \,|\, \Theta}_{\sigma,\rho}$ for every ρ. By Lemma 6.27, we obtain that $V \in \mathsf{OTRed}^{\Gamma \,|\, \Theta}_{\mu,\rho}$ for every ρ. We proceed by case analysis on the last rule of the derivation:

We suppose in the following that Γ is a sized context that can be enumerated in the form $x_1 : \sigma_1, \ldots, x_n : \sigma_n$, and that y is a variable distinct from $x_1, \ldots, x_n$. We proceed according to the last rule of the derivation:

- **Var:** Suppose that $\Gamma, y : \tau \,|\, \Theta \vdash y : \tau$. Let $(q_i)_i \in [0,1]^{n+1}$ and $(V_1, \ldots, V_n, W) \in (\prod_{i=1}^{n} \mathsf{VRed}^{q_i}_{\sigma_i,\rho}) \times \mathsf{VRed}^{q_{n+1}}_{\tau,\rho}$.
  - —If $\Theta = \emptyset$, we need to prove that $y[\overrightarrow{V}, W / \overrightarrow{x}, y] = W \in \mathsf{VRed}^{\prod_{i=1}^{n+1} q_i}_{\tau,\rho}$. This is immediate since $\prod_{i=1}^{n+1} q_i \leq q_{n+1}$, using Lemma 6.7.
  - —If $\Theta = z : \{ \theta_j^{p_j} \,|\, j \in \mathcal{J} \}$, let $(q'_j)_{j \in \mathcal{J}} \in [0,1]^{\mathcal{J}}$ and $Z \in \bigcap_{j \in \mathcal{J}} \mathsf{VRed}^{q'_j}_{\sigma_j,\rho}$. We need to prove that $y[\overrightarrow{V}, W, Z / \overrightarrow{x}, y, z] = W \in \mathsf{VRed}^{(\prod_{i=1}^{n+1} q_i)(\sum_{j \in \mathcal{J}} p_j q'_j)}_{\tau,\rho}$. But again, $(\prod_{i=1}^{n+1} q_i)(\sum_{j \in \mathcal{J}} p_j q'_j) \leq q_{n+1}$ since $q_i \leq 1$ for every i, $q'_j \leq 1$ for every j, and $\sum_{j \in \mathcal{J}} p_j = 1$. We conclude using Lemma 6.7.
- **Var':** Suppose that $\Gamma \,|\, y : \{ \tau^1 \} \vdash y : \tau$. Let $(q_i)_i \in [0,1]^{n+1}$ and $(V_1, \ldots, V_n, W) \in (\prod_{i=1}^{n} \mathsf{VRed}^{q_i}_{\sigma_i,\rho}) \times \mathsf{VRed}^{q_{n+1}}_{\tau,\rho}$. We need to prove that $y[\overrightarrow{V}, W / \overrightarrow{x}, y] = W \in \mathsf{VRed}^{\prod_{i=1}^{n+1} q_i}_{\tau,\rho}$. This is immediate since $\prod_{i=1}^{n+1} q_i \leq q_{n+1}$, using Lemma 6.7.
- **Succ:** Suppose that $\Gamma \,|\, \Theta \vdash \mathsf{S}\ V : \mathsf{Nat}^{\widehat{\mathfrak{s}}}$. Suppose moreover that $\Theta = \emptyset$. Let $(q_i)_i \in [0,1]^n$ and $(W_1, \ldots, W_n) \in \prod_{i=1}^{n} \mathsf{VRed}^{q_i}_{\sigma_i,\rho}$. We need to prove that $(\mathsf{S}\ V)[\overrightarrow{W} / \overrightarrow{x}] \in \mathsf{VRed}^{\prod_{i=1}^{n} q_i}_{\mathsf{Nat}^{\widehat{\mathfrak{s}}},\rho}$. But $(\mathsf{S}\ V)[\overrightarrow{W} / \overrightarrow{x}] = \mathsf{S}\ (V[\overrightarrow{W} / \overrightarrow{x}])$ and, by induction hypothesis, $V[\overrightarrow{W} / \overrightarrow{x}] \in \mathsf{VRed}^{\prod_{i=1}^{n} q_i}_{\mathsf{Nat}^{\mathfrak{s}},\rho}$. By Lemma 6.3, $(\mathsf{S}\ V)[\overrightarrow{W} / \overrightarrow{x}] \in \mathsf{VRed}^{\prod_{i=1}^{n} q_i}_{\mathsf{Nat}^{\widehat{\mathfrak{s}}},\rho}$ and we can conclude. The case where $\Theta \neq \emptyset$ is similar.
- **Zero:** Suppose that $\Gamma \,|\, \Theta \vdash 0 : \mathsf{Nat}^{\widehat{\mathfrak{s}}}$. Suppose moreover that $\Theta = \emptyset$. Let $(q_i)_i \in [0,1]^n$ and $(V_1, \ldots, V_n) \in \prod_{i=1}^{n} \mathsf{VRed}^{q_i}_{\sigma_i,\rho}$. By Lemma 6.3, $0[\overrightarrow{V} / \overrightarrow{x}] = 0 \in \mathsf{VRed}^{\prod_{i=1}^{n} q_i}_{\mathsf{Nat}^{\widehat{\mathfrak{s}}},\rho}$. The case where $\Theta \neq \emptyset$ is similar.

- **$\boldsymbol{\lambda}$:** Suppose that $\Gamma \mid \Theta \vdash \lambda y.M : \sigma \to \mu$, with $\Theta = z : \{ (\tau_j)^{p_j} \mid j \in \mathcal{J} \}$. Let $(q_i)_i \in [0,1]^n$ and $(V_1, \dots, V_n) \in \prod_{i=1}^n \mathsf{VRed}^{q_i}_{\sigma_i,\rho}$. Let $(q'_j)_{j\in\mathcal{J}} \in [0,1]^{\mathcal{J}}$ and $W \in \bigcap_{j\in\mathcal{J}} \mathsf{VRed}^{q'_j}_{\sigma_j,\rho}$. We need to prove that

$$(\lambda y.M)\left[\overrightarrow{V}, W/\overrightarrow{x}, z\right] = \lambda y.M\left[\overrightarrow{V}, W/\overrightarrow{x}, z\right] \in \mathsf{VRed}^{(\prod_{i=1}^n q_i)(\sum_{j\in\mathcal{J}} p_j q'_j).}_{\sigma\to\mu,\rho}$$

  Therefore, let $q'' \in (0,1]$ and $Z \in \mathsf{VRed}^{q''}_{\sigma,\rho}$. We now have to prove that

$$\left(\lambda y.M\left[\overrightarrow{V}, W/\overrightarrow{x}, z\right]\right) Z \ \in \ \mathsf{TRed}^{q''(\prod_{i=1}^n q_i)(\sum_{j\in\mathcal{J}} p_j q'_j).}_{\mu,\rho} \tag{22}$$

  But

$$\left(\lambda y.M\left[\overrightarrow{V}, W/\overrightarrow{x}, z\right]\right) Z \ \to_v \ M\left[\overrightarrow{V}, W, Z/\overrightarrow{x}, z, y\right].$$

  Since $\Gamma,\ x : \sigma \mid \Theta \vdash M : \mu$ by typing, the induction hypothesis ensures that

$$M\left[\overrightarrow{V}, W, Z/\overrightarrow{x}, z, y\right] \in \mathsf{TRed}^{q''(\prod_{i=1}^n q_i)(\sum_{j\in\mathcal{J}} p_j q'_j)}_{\mu,\rho}.$$

  and by Lemma 6.24 we obtain that Equation (22) holds, which allows us to conclude.

  The case where $\Theta = \emptyset$ is similar.

- **Sub:** Suppose that $\Gamma \mid \Theta \vdash M : \nu$ is derived from $\Gamma \mid \Theta \vdash M : \mu$, where $\mu \sqsubseteq \nu$. Suppose that $\Theta = \emptyset$. Let $(q_i)_i \in [0,1]^n$ and $(V_1, \dots, V_n) \in \prod_{i=1}^n \mathsf{VRed}^{q_i}_{\sigma_i,\rho}$. By induction hypothesis, $M[V/\overrightarrow{x}] \in \mathsf{TRed}^{\prod_{i=1}^n q_i}_{\mu,\rho}$ so that by Lemma 6.25, we have $M[V/\overrightarrow{x}] \in \mathsf{TRed}^{\prod_{i=1}^n q_i}_{\nu,\rho}$, which allows us to conclude.

  The case where $\Theta \neq \emptyset$ is similar.

- **App:** Suppose that $\Gamma,\ \Delta,\ \Xi \mid \Theta,\ \Psi \vdash V\ W : \mu$. Suppose that $\Theta,\ \Psi = \emptyset$. We set $\Gamma = x_1 : \sigma_1, \dots, x_n : \sigma_n$, $\Delta = y_1 : \tau_1, \dots, y_m : \tau_m$, and $\Xi = z_1 : \theta_1, \dots, z_l : \theta_l$. Let $(q_i)_i \in [0,1]^n$, $(q'_j)_j \in [0,1]^m$, $(q''_k)_k \in [0,1]^l$, $(V_1, \dots, V_n) \in \prod_{i=1}^n \mathsf{VRed}^{q_i}_{\sigma_i,\rho}$, $(W_1, \dots, W_m) \in \prod_{j=1}^m \mathsf{VRed}^{q'_j}_{\tau_j,\rho}$, and $(Z_1, \dots, Z_l) \in \prod_{k=1}^l \mathsf{VRed}^{q_k}_{\theta_k,\rho}$. We need to prove that

$$(V\ W)\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}/\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}\right] = V\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}/\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}\right]\ W\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}/\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}\ \right] \tag{23}$$

  is in $\mathsf{TRed}^{(\prod_{i=1}^n q_i)(\prod_{j=1}^m q'_j)(\prod_{k=1}^l q''_k)}_{\mu,\rho}$.

  - Suppose that $\prod_{i=1}^n q_i = 0$. Then we need to prove that Equation (23) is in $\mathsf{TRed}^0_{\mu,\rho}$, which is immediate by Lemma 6.6 as it is of simple type $\langle\mu\rangle$.
  - Suppose that $\prod_{i=1}^n q_i \neq 0$. It follows that $\forall i \in \mathcal{I},\ q_i \neq 0$. We have that $\Gamma,\ \Delta \mid \emptyset \vdash V : \sigma \to \mu$, which, by induction hypothesis, gives that $V \in \mathsf{OVRed}^{\Gamma,\Delta\mid\emptyset}_{\sigma\to\mu,\rho}$. Note that for every $i \in \mathcal{I}$ we have $\sigma_i :: \mathsf{Nat}$; since $q_i \neq 0$, we have by definition of the sets of candidates that $\mathsf{VRed}^{q_i}_{\sigma_i,\rho} = \mathsf{VRed}^1_{\sigma_i,\rho}$. It follows that $V[\overrightarrow{V}, \overrightarrow{W}/\overrightarrow{x}, \overrightarrow{y}] = V[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}/\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}] \in \mathsf{VRed}^{(\prod_{i=1}^n 1)(\prod_{j=1}^m q'_j)}_{\sigma\to\mu,\rho} = \mathsf{VRed}^{\prod_{j=1}^m q'_j}_{\sigma\to\mu,\rho}$. Since $\Gamma,\ \Xi \mid \Psi \vdash W : \sigma$, we obtain similarly from the induction hypothesis that $W[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}/\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}] \in \mathsf{VRed}^{\prod_{k=1}^l q''_k}_{\sigma,\rho}$. By definition of $\mathsf{VRed}^{\prod_{j=1}^m q'_j}_{\sigma\to\mu,\rho}$, we obtain that

$$V\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}/\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}\right]\ W\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}/\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}\right] \ \in \ \mathsf{TRed}^{(\prod_{j=1}^m q'_j)(\prod_{k=1}^l q''_k)}_{\mu,\rho},$$

    and by downwards closure (Lemma 6.7) we get that Equation (23) is in $\mathsf{TRed}^{(\prod_{i=1}^n q_i)(\prod_{j=1}^m q'_j)(\prod_{k=1}^l q''_k)}_{\mu,\rho}$ so that we can conclude.

  The case where $\Theta,\ \Psi \neq \emptyset$ is similar.

- **Choice:** Suppose that $\Gamma \mid \Theta \oplus_p \Psi \vdash M \oplus_p N : \mu \oplus_p \nu$. Suppose that $\Theta \neq \emptyset$ and that $\Psi \neq \emptyset$. We set $\Theta = y : \{ \tau_j^{p_j} \mid j \in \mathcal{J} \}$ and $\Psi = y : \{ (\tau'_k)^{p'_k} \mid k \in \mathcal{K} \}$, where we suppose that $j \in \mathcal{J} \cap \mathcal{L} \Leftrightarrow \sigma_j = \tau_j$. We obtain that

$$\Theta \oplus_p \Psi = y : \left\{ \tau_j^{pp_j} \;\middle|\; j \in \mathcal{J} \setminus (\mathcal{J} \cap \mathcal{K}) \right\} + \left\{ (\tau_l)^{pp_l + (1-p)p'_l} \;\middle|\; l \in \mathcal{J} \cap \mathcal{K} \right\}$$
$$+ \left\{ \left(\tau'_k\right)^{(1-p)p'_k} \;\middle|\; k \in \mathcal{K} \setminus (\mathcal{J} \cap \mathcal{K}) \right\}.$$

Let $(q_i)_i \in [0,1]^n$, $(q'_j)_j \in [0,1]^{|\mathcal{J} \setminus (\mathcal{J} \cap \mathcal{K})|}$, $(q''_l)_l \in [0,1]^{|\mathcal{J} \cap \mathcal{K}|}$, $(q'''_k)_k \in [0,1]^{|\mathcal{K} \setminus (\mathcal{J} \cap \mathcal{K})|}$, $(V_1, \ldots, V_n) \in \prod_{i=1}^n \mathsf{VRed}_{\sigma_i,\rho}^{q_i}$, and

$$W \in \bigcap_{j \in \mathcal{J} \setminus (\mathcal{J} \cap \mathcal{K})} \mathsf{VRed}_{\tau_j,\rho}^{q'_j} \cap \bigcap_{l \in \mathcal{J} \cap \mathcal{K}} \mathsf{VRed}_{\tau_l,\rho}^{q''_l} \cap \bigcap_{k \in \mathcal{K} \setminus (\mathcal{J} \cap \mathcal{K})} \mathsf{VRed}_{\tau'_k,\rho}^{q'''_k}.$$

We need to prove that $(M \oplus_p N)[\overrightarrow{V}, W/\overrightarrow{x}, y]$ is in

$$\mathsf{TRed}_{\mu \oplus_p \nu,\rho}^{(\prod_{i=1}^n q_i)(\sum_{j \in \mathcal{J} \setminus (\mathcal{J} \cap \mathcal{K})} pp_j q'_j + \sum_{l \in \mathcal{J} \cap \mathcal{K}} (pp_l + (1-p)p'_l) q''_l + \sum_{k \in \mathcal{K} \setminus (\mathcal{J} \cap \mathcal{K})} (1-p) p'_k q'''_k)}$$
$$= \mathsf{TRed}_{\mu \oplus_p \nu,\rho}^{p(\prod_{i=1}^n q_i)(\sum_{j \in \mathcal{J} \setminus (\mathcal{J} \cap \mathcal{K})} p_j q'_j + \sum_{l \in \mathcal{J} \cap \mathcal{K}} p_l q''_l) + (1-p)(\prod_{i=1}^n q_i)(\sum_{l \in \mathcal{J} \cap \mathcal{K}} p'_l q''_l + \sum_{k \in \mathcal{K} \setminus (\mathcal{J} \cap \mathcal{K})} p'_k q'''_k)}.$$

Typing gives us that $\Gamma \mid \Theta \vdash M : \mu$, which by the induction hypothesis implies that

$$M\left[\overrightarrow{V}, W/\overrightarrow{x}, y\right] \in \mathsf{TRed}_{\mu,\rho}^{(\prod_{i=1}^n q_i)(\sum_{j \in \mathcal{J} \setminus (\mathcal{J} \cap \mathcal{K})} p_j q'_j + \sum_{l \in \mathcal{J} \cap \mathcal{K}} p_l q''_l)}.$$

Typing also implies that $\Gamma \mid \Psi \vdash N : \nu$ and provides by the induction hypothesis

$$N\left[\overrightarrow{V}, W/\overrightarrow{x}, y\right] \in \mathsf{TRed}_{\mu \oplus_p \nu,\rho}^{(\prod_{i=1}^n q_i)(\sum_{l \in \mathcal{J} \cap \mathcal{K}} p'_l q''_l + \sum_{k \in \mathcal{K} \setminus (\mathcal{J} \cap \mathcal{K})} p'_k q'''_k)}.$$

Since

$$(M \oplus_p N)\left[\overrightarrow{V}, W/\overrightarrow{x}, y\right] \to_v \left\{ \left(M\left[\overrightarrow{V}, W/\overrightarrow{x}, y\right]\right)^p, \left(N\left[\overrightarrow{V}, W/\overrightarrow{x}, y\right]\right)^{1-p} \right\},$$

Lemma 6.24 allows us to conclude.
The cases where $\Theta = \emptyset$ or $\Psi = \emptyset$ are treated similarly.

- **Let:** Suppose that $\Gamma, \Delta, \Xi \mid \Theta, (\sum_{i \in \mathcal{I}} p_i \cdot \Psi_i) \vdash \mathsf{let}\ x = M\ \mathsf{in}\ N : \sum_{i \in \mathcal{I}} p_i \cdot \mu_i$. Let $\Gamma = x_1 : \sigma_1, \ldots, x_n : \sigma_n$, $\Delta = y_1 : \tau_1, \ldots, y_m : \tau_m$, and $\Xi = z_1 : \theta_1, \ldots, z_m : \theta_l$. Let $(q_i)_i \in [0,1]^n$, $(q'_j)_j \in [0,1]^m$ and $(q''_k)_k \in [0,1]^l$. Let $(V_1, \ldots, V_n) \in \prod_{i=1}^n \mathsf{VRed}_{\sigma_i,\rho}^{q_i}$, $(W_1, \ldots, W_m) \in \prod_{j=1}^m \mathsf{VRed}_{\tau_j,\rho}^{q'_j}$, and $(Z_1, \ldots, Z_l) \in \prod_{k=1}^l \mathsf{VRed}_{\theta_k,\rho}^{q''_k}$. There are two subcases here.
—Suppose that M is a value. Then the last typing rule is

$$\frac{\Gamma, \Delta \mid \Theta \vdash M : \sigma \qquad \Gamma, \Xi, x : \sigma \mid \Psi \vdash N : \mu \qquad \langle \Gamma \rangle = \mathsf{Nat}}{\Gamma, \Delta, \Xi \mid \Theta, \Psi \vdash \mathsf{let}\ x = M\ \mathsf{in}\ N : \mu}.$$

We treat the case where $\Theta = \Psi = \emptyset$, the two other ones being similar. We need to prove that

$$(\mathsf{let}\ x = M\ \mathsf{in}\ N)\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}/\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}\right] \in \mathsf{TRed}_{\mu,\rho}^{(\prod_{i \in \mathcal{I}} q_i)(\prod_{j \in \mathcal{J}} q'_j)(\prod_{k \in \mathcal{K}} q''_k)}. \qquad (24)$$

We now distinguish two cases.
* Suppose that $\prod_{i \in \mathcal{I}} q_i = 0$. Then Equation (24) holds immediately since by Lemma 6.6 all the terms of simple type $\langle \mu \rangle$ are in $\mathsf{TRed}_{\mu,\rho}^0$.

  * Else for every $i \in \mathcal{I}$ we have $\mathsf{VRed}^{q_i}_{\sigma_i,\rho} = \mathsf{VRed}^{1}_{\sigma_i,\rho}$. Since $\Gamma,\ \Delta \,|\, \Theta \vdash M \ :\ \sigma$, we obtain by induction hypothesis that $M[\overrightarrow{V},\overrightarrow{W}/\overrightarrow{x},\overrightarrow{y}] \in \mathsf{TRed}^{(\prod_{i\in\mathcal{I}} 1)(\prod_{j\in\mathcal{J}} q'_j)}_{\sigma,\rho}$. None of the $\overrightarrow{z}$ occur in M, so $M[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}] \in \mathsf{TRed}^{\prod_{j\in\mathcal{J}} q'_j}_{\sigma,\rho}$. Since $\Gamma,\ \Xi,\ x\ :\ \sigma \,|\, \Psi \vdash N\ :\ \mu$, we obtain by induction hypothesis that

$$N\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z},M\left[\overrightarrow{V},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right]/\overrightarrow{x},\overrightarrow{z},x\right] \ \in\ \mathsf{TRed}^{(\prod_{i\in\mathcal{I}} q_i)(\prod_{j\in\mathcal{J}} q'_j)(\prod_{k\in\mathcal{K}} q''_k)}_{\mu,\rho}.$$

  Since none of the variables of $\overrightarrow{y}$ occur in this term, we obtain

$$N\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z},M\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right]/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z},x\right] \ \in\ \mathsf{TRed}^{(\prod_{i\in\mathcal{I}} q_i)(\prod_{j\in\mathcal{J}} q'_j)(\prod_{k\in\mathcal{K}} q''_k)}_{\mu,\rho}.$$

  Now

$$\begin{array}{rl}
& (\mathsf{let}\ x = M\ \mathsf{in}\ N)\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right] \\
= & \mathsf{let}\ x = M\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right]\ \mathsf{in}\ N\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right] \\
\to_v & \left\{\left(N\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right]\left[M\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right]/x\right]\right)^1\right\} \\
= & \left\{\left(N\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z},M\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right]/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z},x\right]\right)^1\right\},
\end{array}$$

  and it follows from Lemma 6.24 that Equation (24) holds, allowing us to conclude.

— Suppose that M is not a value. We treat first the case where $\Theta = \Psi = \emptyset$. The case where $\Theta \neq \emptyset$ is exactly similar, while the case where $\Psi \neq \emptyset$ reveals the reason that a sum $\sum_{j\in\mathcal{J}} p_j q'_j$ appears in the definitions of OTRed and OVRed. The last typing rule is

$$\frac{\Gamma,\ \Delta \,|\, \emptyset \vdash M\ :\ \left\{\sigma_h^{p_h}\ \middle|\ h \in \mathcal{H}\right\} \qquad \Gamma,\ \Xi,\ x\ :\ \sigma_h \,|\, \emptyset \vdash N\ :\ \mu_h \qquad \langle\Gamma\rangle = \mathsf{Nat}}{\Gamma,\ \Delta,\ \Xi \,|\, \emptyset \vdash \mathsf{let}\ x = M\ \mathsf{in}\ N\ :\ \sum_{h\in\mathcal{H}}\ p_h \cdot \mu_h}.$$

We need to prove that

$$\begin{array}{rl}
& (\mathsf{let}\ x = M\ \mathsf{in}\ N)\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right] \\
= & \mathsf{let}\ x = M\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right]\ \mathsf{in}\ N\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right] \\
& \quad \in \mathsf{TRed}^{(\prod_{i\in\mathcal{I}} q_i)(\prod_{j\in\mathcal{J}} q'_j)(\prod_{k\in\mathcal{K}} q''_k)}_{\sum_{h\in\mathcal{H}}\ p_h\cdot\mu_h,\rho}.
\end{array} \tag{25}$$

We now distinguish two cases:

  * Suppose that $(\prod_{i\in\mathcal{I}} q_i)(\prod_{j\in\mathcal{J}} q'_j)(\prod_{k\in\mathcal{K}} q''_k) = 0$. Then Equation (25) holds immediately as by Lemma 6.6 all the terms of simple type $\langle\sum_{h\in\mathcal{H}}\ p_h \cdot \mu_h\rangle$ are in $\mathsf{TRed}^0_{\sum_{h\in\mathcal{H}}\ p_h\cdot\mu_h,\rho}$.
  * Else, we use the induction hypothesis on $\Gamma,\ \Delta \,|\, \emptyset \vdash M\ :\ \{\ \sigma_h^{p_h}\ \ |\ \ h \in \mathcal{H}\ \}$. Since $\langle\sigma_i\rangle = \mathsf{Nat}$, for every $i \in \mathcal{I}$ we have $\mathsf{VRed}^{q_i}_{\sigma_i,\rho} = \mathsf{VRed}^{1}_{\sigma_i,\rho}$. Together with the fact that $\overrightarrow{z}$ does not appear in M, we obtain that

$$M\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right] \in \mathsf{TRed}^{\prod_{j\in\mathcal{J}} q'_j}_{\{\ \sigma_h^{p_h}\ \ |\ \ h\in\mathcal{H}\ \},\rho}.$$

  By definition, for every $0 \leq r < \prod_{j=1}^{m}\ q'_j$, there exists n_r and $\nu_r = \{\ \sigma_g^{p_{r,g}}\ \ |\ \ g \in \mathcal{G}_r\ \} \preccurlyeq \{\ \sigma_h^{p_h}\ \ |\ \ h \in \mathcal{H}\ \}$ with $\mathcal{G}_r \subseteq \mathcal{H}$ such that

$$M\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right] \ \Rrightarrow_v^{n_r}\ \mathscr{D}_r = \left[\ X_l^{p''_{r,l}}\ \middle|\ l \in \mathcal{L}_r\ \right] \ \in\ \mathsf{DRed}^r_{\nu_r,\rho}.$$

This implies the existence of two families $(p^r_{lg})_{l\in\mathcal{L}_r, g_r\in\mathcal{G}}$ and $(q^r_{lg})_{l\in\mathcal{L}_r, g_r\in\mathcal{G}}$ of reals of $[0,1]$ satisfying in particular

$$r \leq \sum_{l\in\mathcal{L}_r}\sum_{g\in\mathcal{G}_r} p^r_{lg} q^r_{lg} \tag{26}$$

$$\sum_{l\in\mathcal{L}_r}\sum_{g\in\mathcal{G}_r} p^r_{lg} \leq 1 \tag{27}$$

$$\forall l \in \mathcal{L},\ \sum_{g\in\mathcal{G}_r} p^r_{lg} = p''_{r,l} \tag{28}$$

$$\forall g \in \mathcal{G},\ \sum_{l\in\mathcal{L}_r} p^r_{lg} = p_{r,g} \tag{29}$$

and

$$\forall l \in \mathcal{L}_r,\ \ \forall g \in \mathcal{G}_r,\ \ X_l \in \mathsf{VRed}^{q^r_{lg}}_{\sigma_g,\rho}. \tag{30}$$

By Equations (26) and (27), we can apply Lemma 6.34 and we obtain $0 \leq r < \prod_{j=1}^m\ q'_j$ and a family $(r'_{lg})_{l\in\mathcal{L}_r, g\in\mathcal{G}_r}$ satisfying

$$\forall l \in \mathcal{L}_r,\ \ \forall g \in \mathcal{G}_r,\ \ 0 \leq r'_{lg} < \left(\prod_{i=1}^{n}\ q_i\right)\left(\prod_{k=1}^{l}\ q''_k\right) q^r_{lg} \tag{31}$$

and

$$r'' \leq \sum_{l\in\mathcal{L}}\sum_{g\in\mathcal{G}} p^r_{lg} r'_{lg}. \tag{32}$$

We now consider r to be fixed to this value given by the lemma, this providing $\mathscr{D}_r$, ν_r, and so on.

Since Γ, Ξ, $x\ :\ \sigma_h \,|\, \emptyset \vdash N\ :\ \mu_h$, we obtain by induction hypothesis using Equation (30) that for every $l \in \mathcal{L}$ and $g \in \mathcal{G}$ we have

$$N\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}, X_l / \overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}, x\right]\ \in\ \mathsf{TRed}^{(\prod_{i=1}^n\ q_i)(\prod_{k=1}^l\ q''_k) q^r_{lg}}_{\mu_g,\rho}. \tag{33}$$

By Equation (31), there exists for every $l \in \mathcal{L}$ and $g \in \mathcal{G}$ an index m_{lg} and a type $\mu'_{lg} \sqsubseteq \mu_g$ such that

$$N\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}, X_l / \overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}, x\right]\ \Rrightarrow^{m_{lg}}_{v}\ \mathscr{E}_{lg}\ \in\ \mathsf{DRed}^{r'_{lg}}_{\mu'_{lg},\rho}. \tag{34}$$

Now set

$$m = \max_{l\in\mathcal{L}, g\in\mathcal{G}}\ m_{lg}.$$

By Lemma 6.21, we obtain types $\mu'_{lg} \preccurlyeq \mu''_{lg} \preccurlyeq \mu_g$ and distributions $\mathscr{E}'_{lg}$ such that all the reduction lengths are the same:

$$N\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z}, X_l / \overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}, x\right]\ \Rrightarrow^{m}_{v}\ \mathscr{E}'_{lg}\ \in\ \mathsf{DRed}^{r'_{lg}}_{\mu''_{lg},\rho}. \tag{35}$$

Now it follows of Equation (28) that

$$\mathscr{D}_r = \left[\ X_l^{p^r_{l,g}}\ \middle|\ l \in \mathcal{L}_r,\ g \in \mathcal{G}_r\ \right],$$

which allows us to use Lemma 3.12, obtaining that

$$(\mathsf{let}\ x = M\ \mathsf{in}\ N)\left[\overrightarrow{V}, \overrightarrow{W}, \overrightarrow{Z} / \overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}\right]\ \Rrightarrow^{n_r+m+1}_{v}\ \sum_{l\in\mathcal{L}}\sum_{g\in\mathcal{G}} p^r_{l,g}\cdot\mathscr{E}'_{lg}.$$

By Equation (35) and Lemma 6.23, we obtain that

$$\sum_{l\in\mathcal{L}}\sum_{g\in\mathcal{G}} p^r_{l,g}\cdot\mathscr{E}'_{lg} \;\in\; \mathsf{DRed}^{\sum_{l\in\mathcal{L}}\ \sum_{g\in\mathcal{G}}\ p^r_{l,g}r'_{l,g}}_{\sum_{l\in\mathcal{L}}\ \sum_{g\in\mathcal{G}}\ p^r_{l,g}\mu''_{l,g},\rho}.$$

By Equation (32) and downward closure (Lemma 6.7) we obtain

$$\sum_{l\in\mathcal{L}}\sum_{g\in\mathcal{G}} p^r_{l,g}\cdot\mathscr{E}'_{lg} \;\in\; \mathsf{DRed}^{r''}_{\sum_{l\in\mathcal{L}}\ \sum_{g\in\mathcal{G}}\ p^r_{l,g}\mu''_{l,g},\rho},$$

and since by Equation (29) we have $\sum_{l\in\mathcal{L}}\ \sum_{g\in\mathcal{G}}\ p^r_{l,g}\mu''_{l,g} \preccurlyeq \sum_{h\in\mathcal{H}}\ p_h\mu_h$, we can conclude that

$$(\mathsf{let}\ x = M\ \mathsf{in}\ N)\left[\overrightarrow{V},\overrightarrow{W},\overrightarrow{Z}/\overrightarrow{x},\overrightarrow{y},\overrightarrow{z}\right] \;\in\; \mathsf{TRed}^{(\prod_{i\in\mathcal{I}} q_i)(\prod_{j\in\mathcal{J}} q'_j)(\prod_{k\in\mathcal{K}} q''_k)}_{\sum_{h\in\mathcal{H}}\ p_h\cdot\mu_h,\rho}.$$

- **Case:** Suppose that $\Gamma,\,\Delta\,|\,\Theta \vdash \mathsf{case}\ V\ \mathsf{of}\ \{\,\mathsf{S}\to W \ \mid\ 0\to Z\,\}\ :\ \mu$. Suppose that $\Theta=\emptyset$. We set $\Gamma = x_1\ :\ \sigma_1,\ldots,x_n\ :\ \sigma_n$ and $\Delta = y_1\ :\ \tau_1,\ldots,y_m\ :\ \tau_m$.

  Let $(q_i)_i \in [0,1]^n$, $(q'_j)_j \in [0,1]^m$, $(V_1,\ \ldots,\ V_n) \in \prod_{i=1}^n\ \mathsf{VRed}^{q_i}_{\sigma_i,\rho}$ and $(V'_1,\ \ldots,\ V'_m) \in \prod_{j=1}^m\ \mathsf{VRed}^{q'_j}_{\tau_j,\rho}$. We need to prove that

$$(\mathsf{case}\ V\ \mathsf{of}\ \{\,\mathsf{S}\to W\ \mid\ 0\to Z\,\})\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right] \;\in\; \mathsf{TRed}^{(\prod_{i=1}^n q_i)(\prod_{j=1}^m q'_j)}_{\mu,\rho},$$

  i.e., that

$$\mathsf{case}\ V\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mathsf{of}\ \left\{\mathsf{S}\to W\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mid\ 0\to Z\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\right\}$$

  is in $\mathsf{TRed}^{(\prod_{i=1}^n q_i)(\prod_{j=1}^m q'_j)}_{\mu,\rho}$. Since $\Gamma\,|\,\emptyset\vdash V\ :\ \mathsf{Nat}^{\widehat{\mathfrak{s}}}$, we have by induction hypothesis that $V[\overrightarrow{V}/\overrightarrow{x}] \in \mathsf{TRed}^{\prod_{i=1}^n q_i}_{\{\,(\mathsf{Nat}^{\widehat{\mathfrak{s}}})^1\,\},\rho}$. Since it is a value, we have by Lemma 6.19 the stronger statement that $V[\overrightarrow{V}/\overrightarrow{x}] \in \mathsf{VRed}^{\prod_{i=1}^n q_i}_{\mathsf{Nat}^{\widehat{\mathfrak{s}}},\rho}$, which implies that $V[\overrightarrow{V}/\overrightarrow{x}]$ is of the shape $\mathsf{S}^k\ 0$ for $k\in\mathbb{N}$ satisfying $\prod_{i=1}^n q_i\ \neq\ 0 \Rightarrow k < [\![\widehat{\mathfrak{s}}]\!]_\rho$. The typing also ensures that none of the variables of $\overrightarrow{y}$ occurs in V, so that $V[\overrightarrow{V}/\overrightarrow{x}] = V[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}]$.

  —If $k=0$, then

$$\begin{aligned}&\mathsf{case}\ 0\ \mathsf{of}\ \left\{\mathsf{S}\to W\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mid\ 0\to Z\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\right\}\\ &\to_v\ \ \left\{\left(Z\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\right)^1\right\}.\end{aligned}$$

  Since $\Delta\,|\,\emptyset\vdash Z\ :\ \mu$, by induction hypothesis, we have that

$$Z\left[\overrightarrow{V'}/\overrightarrow{y}\right] \in \mathsf{TRed}^{\prod_{j=1}^m q'_j}_{\mu,\rho}$$

  and also, by the typing hypothesis, that none of the variables of $\overrightarrow{x}$ is free in $Z[\overrightarrow{V'}/\overrightarrow{y}]$ so that $Z[\overrightarrow{V'}/\overrightarrow{y}] = Z[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}]$. But $\prod_{i=1}^n q_i\ \leq 1$, so that the downward-closure property of Lemma 6.7 induces that

$$Z\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right] \in \mathsf{TRed}^{(\prod_{i=1}^n q_i)(\prod_{j=1}^m q'_j)}_{\mu,\rho}.$$

  Now the closure by antireduction of Lemma 6.24 ensures that

$$\mathsf{case}\ V\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mathsf{of}\ \left\{\mathsf{S}\to W\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mid\ 0\to Z\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\right\}$$

  is in $\mathsf{TRed}^{(\prod_{i=1}^n q_i)(\prod_{j=1}^m q'_j)}_{\mu,\rho}$.

—If $k > 0$, then

$$\begin{array}{l}\mathsf{case}\ \mathsf{S}^k\ 0\ \mathsf{of}\ \left\{\mathsf{S} \to W\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mid\ 0 \to Z\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\right\}\\ \to_v\ \ \left\{\left(\left(W\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\right)\ (\mathsf{S}^{k-1}\ 0)\right)^1\right\}.\end{array}$$

By typing hypothesis, we have $\Delta \mid \emptyset \vdash W\ :\ \mathsf{Nat}^{\mathfrak{s}} \to \mu$ and the induction hypothesis provides $W[\overrightarrow{V'}/\overrightarrow{y}] \in \mathsf{TRed}^{\prod_{j=1}^{m} q'_j}_{\{\ (\mathsf{Nat}^{\mathfrak{s}} \to \mu)^1\ \},\rho}$, which, since none of the $\overrightarrow{x}$ appears freely in W, and by Lemma 6.19, implies that $W[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}] \in \mathsf{VRed}^{\prod_{j=1}^{m} q'_j}_{\mathsf{Nat}^{\mathfrak{s}} \to \mu,\rho}$.

* Suppose that $\prod_{i=1}^{n} q_i\ \neq\ 0$. It follows that $k < [\![\widehat{\mathfrak{s}}]\!]_\rho$ and therefore $k - 1 < [\![\mathfrak{s}]\!]_\rho$, which implies that $\mathsf{S}^{k-1}\ 0 \in \mathsf{VRed}^1_{\mathsf{Nat}^{\mathfrak{s}},\rho}$. Since $W[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}] \in \mathsf{VRed}^{\prod_{j=1}^{m} q'_j}_{\mu,\rho}$, we obtain that $(W[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}])\ (\mathsf{S}^{k-1}\ 0)$ is in $\mathsf{TRed}^{\prod_{j=1}^{m} q'_j}_{\mu,\rho}$. By closure by antireduction (Lemma 6.24), we have that

$$\mathsf{case}\ V\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mathsf{of}\ \left\{\mathsf{S} \to W\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mid\ 0 \to Z\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\right\}$$

is in $\mathsf{TRed}^{\prod_{j=1}^{m} q'_j}_{\mu,\rho}$, and by downward closure (Lemma 6.7), we obtain that it is also in $\mathsf{TRed}^{(\prod_{i=1}^{n} q_i)(\prod_{j=1}^{m} q'_j)}_{\mu,\rho}$, from which we conclude.

* Suppose that $\prod_{i=1}^{n} q_i\ =\ 0$. Then all we need to prove is that

$$\mathsf{case}\ V\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mathsf{of}\ \left\{\mathsf{S} \to W\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\ \mid\ 0 \to Z\left[\overrightarrow{V},\overrightarrow{V'}/\overrightarrow{x},\overrightarrow{y}\right]\right\}$$

is in $\mathsf{TRed}^0_{\mu,\rho}$. But this term has simple type $\langle\mu\rangle$ and by Lemma 6.6 the result holds.

The case where $\Theta \neq \emptyset$ is similar.

- **letrec:** Suppose that $\Gamma, \Delta \mid \Theta \vdash \mathsf{letrec}\ f = V\ :\ \mathsf{Nat}^{\mathfrak{r}} \to \nu[\mathfrak{r}/\mathfrak{i}]$. We treat the case where $\Delta = \Theta = \emptyset$. The general case is easily deduced using the downward closure of the reducibility sets (Lemma 6.7). Let $\Gamma = x_1\ :\ \mathsf{Nat}^{\mathfrak{r}_1}, \ldots,\ x_n\ :\ \mathsf{Nat}^{\mathfrak{r}_n}$. We need to prove that, for every family $(q_i)_i \in [0,1]^n$ and every $(W_1,\ldots,W_n) \in \prod_{i=1}^{n} \mathsf{VRed}^{q_i}_{\mathsf{Nat}^{\mathfrak{r}_i},\rho}$, we have

$$(\mathsf{letrec}\ f = V)\left[\overrightarrow{W}/\overrightarrow{x}\right] = \left(\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right]\right)\ \in\ \mathsf{VRed}^{\prod_{i=1}^{n} q_i}_{\mathsf{Nat}^{\mathfrak{r}} \to \nu[\mathfrak{r}/\mathfrak{i}],\rho}.$$

If there exists $i \in \mathcal{I}$ such that $q_i = 0$, the result is immediate as the term is simply typed and Lemma 6.6 applies. Else, for every $i \in \mathcal{I}$, we have by definition that $\mathsf{VRed}^{q_i}_{\mathsf{Nat}^{\mathfrak{r}_i},\rho} = \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{r}_i},\rho}$. Since the sets VRed are downward closed (Lemma 6.7), it is in fact enough to prove that for every $(W_1,\ldots,W_n) \in \prod_{i=1}^{n} \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{r}_i},\rho}$, we have

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right]\ \in\ \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{r}} \to \nu[\mathfrak{r}/\mathfrak{i}],\rho}.$$

Moreover, by size commutation (Lemma 6.12),

$$\mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{r}} \to \nu[\mathfrak{r}/\mathfrak{i}],\rho} = \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{i}} \to \nu,\rho[\mathfrak{i} \mapsto [\![\mathfrak{r}]\!]_\rho]}.$$

Let us therefore prove the stronger fact that, for *every* integer $m \in \mathbb{N} \cup \{\infty\}$,

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right]\ \in\ \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{i}} \to \nu,\rho[\mathfrak{i} \mapsto m]}.$$

Now, the typing derivation gives us that $\Gamma \mid f\ :\ \mu \vdash V\ :\ \mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \nu[\widehat{\mathfrak{i}}/\mathfrak{i}]$ and that μ induces an AST sized walk. Denote by $(Pr_{n,m})_{n \in \mathbb{N}, m \in \mathbb{N}}$ its associated probabilities of convergence in finite time. By induction hypothesis, $V \in \mathsf{OVRed}^{\Gamma \mid f : \mu}_{\mathsf{Nat}^{\widehat{\mathfrak{i}}} \to \nu[\widehat{\mathfrak{i}}/\mathfrak{i}], \rho}$ for every ρ and we can apply Lemma 6.31. It follows that, for every $(n, m) \in \mathbb{N}$,

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right] \quad \in \quad \mathsf{VRed}^{Pr_{n,m}}_{\mathsf{Nat}^{\mathfrak{i}} \to \nu, \rho[\mathfrak{i} \mapsto m]}.$$

Let $\varepsilon \in (0, 1)$. By Lemma 6.29, there exists $n \in \mathbb{N}$ such that $Pr_{n,m} \geq 1 - \varepsilon$. Using downward closure (Lemma 6.7) and quantifying over all the ε, we obtain

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right] \quad \in \quad \bigcap_{0<\varepsilon<1} \mathsf{VRed}^{1-\varepsilon}_{\mathsf{Nat}^{\mathfrak{i}} \to \nu, \rho[\mathfrak{i} \mapsto m]}$$

so that, by continuity of VRed (Lemma 6.10), we obtain

$$\mathsf{letrec}\ f = V\left[\overrightarrow{W}/\overrightarrow{x}\right] \quad \in \quad \mathsf{VRed}^{1}_{\mathsf{Nat}^{\mathfrak{i}} \to \nu, \rho[\mathfrak{i} \mapsto m]} \tag{36}$$

for every $m \in \mathbb{N}$, allowing us to conclude. It remains, however, to treat the case where $m = \infty$. Since $\mathfrak{i}$ pos ν and Equation (36) holds for every $m \in \mathbb{N}$, Lemma 6.33 applies and we obtain the result. □

This proposition, together with the definition of OTRed, implies the main result of the article, namely, that typability implies almost-sure termination:

Theorem 6.36. *Suppose that $M \in \Lambda^{\mathfrak{s}}_{\oplus}(\mu)$. Then M is AST.*

Proof. Suppose that $M \in \Lambda^{\mathfrak{s}}_{\oplus}(\mu)$; then by Proposition 6.35, we have $M \in \mathsf{OTRed}^{\emptyset \mid \emptyset}_{\mu, \rho}$ for every ρ. By definition, $\mathsf{OTRed}^{\emptyset \mid \emptyset}_{\mu, \rho} = \mathsf{TRed}^{1}_{\mu, \rho}$. Corollary 6.5 then implies that M is AST. □

7 CONCLUSIONS AND PERSPECTIVES

We have presented a type system for an affine, simply typed λ-calculus enriched with a probabilistic choice operator, constructors for the natural numbers, and recursion. This affinity constraint implies that a given higher-order variable may occur (freely) at most once in any probabilistic branch of a program. The type system we designed decorates the affine simple types with *size information*, allowing one to incorporate in the types relevant information about the recursive behavior of the functions contained in the program. A guard condition on the typing rule for letrec, formulated with reference to an appropriate Markov chain, ensures that typable terms are AST. The proof of soundness of this type system for AST relies on a quantitative extension of the reducibility method, to accommodate sets of candidates to the infinitary and probabilistic nature of the computations we consider.

A first natural question is the one of the decidability of type inference for our system. In the deterministic case, this question was only addressed by Barthe and colleagues in an unpublished tutorial [6], and their solution is technically involved, especially when it comes to dealing with the fixpoint rule. We believe that their approach could be extended to our system of monadic sized types and hope that it could provide a decidable type inference procedure for it. However, this extension will certainly be challenging, as we need to appropriately infer distribution types associated with AST sized walks in the letrec rule.

Another perspective would be to study the general, *nonaffine* case. This is challenging, for two reasons. First, the system of size annotations needs to be more expressive in order to distinguish between various occurrences of a same function symbol in a same probabilistic branch. A solution would be to use the combined power of dependent types—which already allowed Xi to formulate an interesting type system for termination in the deterministic case [41]—and of linearity: we could use *linear dependent types* [17] to formulate an extension of the monadic sized type system keeping track of *how many* recursive calls are performed and of the size of *each* recursive argument. The second challenge would then be to associate, in the typing rule for letrec, this information contained in linear dependent types with an appropriate random process. This random process should be kept decidable to guarantee that at least *derivation* checking can be automated, and there will probably be a tradeoff between the duplication power we allow in programs and the complexity of deciding AST for the guard in the letrec rule.

The extension of our type system to deal with general inductive data types is essentially straightforward. Other perspectives would be to enrich the type system so as to be able to treat coinductive data, polymorphic types, or ordinal sizes, three features present in most systems of sized types dealing with the traditional deterministic case but which we chose not to address in this article to focus on the already complex task of accommodating sized types in a probabilistic and higher-order framework.

REFERENCES

[1] Andreas Abel. 2004. Termination checking with types. *ITA* 38, 4 (2004), 277–319. DOI:https://doi.org/10.1051/ita:2004015

[2] Roberto M. Amadio and Solange Coupet-Grimal. 1998. Analysis of a guard condition in type theory. In *FoSSaCS'98 (LNCS)*, Maurice Nivat (Ed.), Vol. 1378. Springer, 48–62. DOI:https://doi.org/10.1007/BFb0053541

[3] Gilles Barthe, Maria João Frade, Eduardo Giménez, Luis Pinto, and Tarmo Uustalu. 2004. Type-based termination of recursive definitions. *MSCS* 14, 1 (2004), 97–141. DOI:https://doi.org/10.1017/S0960129503004122

[4] Gilles Barthe, Benjamin Grégoire, and Santiago Zanella Béguelin. 2009. Formal certification of code-based cryptographic proofs. In *POPL'09*, Zhong Shao and Benjamin C. Pierce (Eds.). ACM, 90–101.

[5] Gilles Barthe, Benjamin Grégoire, Sylvain Heraud, and Santiago Zanella Béguelin. 2011. Computer-aided security proofs for the working cryptographer. In *CRYPTO 2011 (LNCS)*, Phillip Rogaway (Ed.), Vol. 6841. Springer, 71–90.

[6] Gilles Barthe, Benjamin Grégoire, and Colin Riba. 2008. A tutorial on type-based termination. In *ALFA Summer School 2008 (LNCS)*, Ana Bove, Luís Soares Barbosa, Alberto Pardo, and Jorge Sousa Pinto (Eds.), Vol. 5520. Springer, 100–152. DOI:https://doi.org/10.1007/978-3-642-03153-3_3

[7] Gilles Barthe, Benjamin Grégoire, and Colin Riba. 2008. Type-based termination with sized products. In *CSL'08 (LNCS)*, Michael Kaminski and Simone Martini (Eds.), Vol. 5213. Springer, 493–507. DOI:https://doi.org/10.1007/978-3-540-87531-4_35

[8] Olivier Bournez and Florent Garnier. 2005. Proving positive almost-sure termination. In *RTA'05 (LNCS)*, Jürgen Giesl (Ed.), Vol. 3467. Springer, 323–337. DOI:https://doi.org/10.1007/978-3-540-32033-3_24

[9] Olivier Bournez and Claude Kirchner. 2002. Probabilistic rewrite strategies. Applications to ELAN. In *RTA'02 (LNCS)*, Sophie Tison (Ed.), Vol. 2378. Springer, 252–266. Retrieved from DOI:https://doi.org/10.1007/3-540-45610-4_18

[10] T. Brázdil, V. Brožek, K. Etessami, A. Kučera, and D. Wojtczak. 2010. One-counter Markov decision processes. In *21st ACM-SIAM Symposium on Discrete Algorithms*. DOI:https://doi.org/10.1137/1.9781611973075.70

[11] Alberto Cappai and Ugo Dal Lago. 2015. On equivalences, metrics, and polynomial time. In *FCT'15 (LNCS)*, Adrian Kosowski and Igor Walukiewicz (Eds.), Vol. 9210. Springer, 311–323. DOI:https://doi.org/10.1007/978-3-319-22177-9_24

[12] Aleksandar Chakarov and Sriram Sankaranarayanan. 2013. Probabilistic program analysis with martingales. In *CAV'13 (LNCS)*, Natasha Sharygina and Helmut Veith (Eds.), Vol. 8044. Springer, 511–526. DOI:https://doi.org/10.1007/978-3-642-39799-8_34

[13] Krishnendu Chatterjee, Hongfei Fu, and Amir Kafshdar Goharshady. 2016. Termination analysis of probabilistic programs through Positivstellensatz's. In *CAV'16 (LNCS)*, Swarat Chaudhuri and Azadeh Farzan (Eds.), Vol. 9779. Springer, 3–22. DOI:https://doi.org/10.1007/978-3-319-41528-4_1

[14] Krishnendu Chatterjee, Hongfei Fu, Petr Novotný, and Rouzbeh Hasheminezhad. 2016. Algorithmic analysis of qualitative and quantitative termination problems for affine probabilistic programs. In *POPL'16*, Rastislav Bodík and Rupak Majumdar (Eds.). ACM, 327–342. DOI:https://doi.org/10.1145/2837614.2837639

[15] Ugo Dal Lago. 2005. The geometry of linear higher-order recursion. In *LICS'05*. IEEE Computer Society, 366–375. DOI:https://doi.org/10.1109/LICS.2005.52

[16] Ugo Dal Lago. 2011. A short introduction to implicit computational complexity. In *ESSLLI'10*. 89–109.

[17] Ugo Dal Lago and Marco Gaboardi. 2011. Linear dependent types and relative completeness. In *LICS'11*. IEEE Computer Society, 133–142. DOI:https://doi.org/10.1109/LICS.2011.22

[18] Ugo Dal Lago and Martin Hofmann. 2011. Realizability models and implicit complexity. *Theor. Comput. Sci.* 412, 20 (2011), 2029–2047. DOI:https://doi.org/10.1016/j.tcs.2010.12.025

[19] Ugo Dal Lago and Paolo Parisen Toldin. 2015. A higher-order characterization of probabilistic polynomial time. *Inf. Comput.* 241 (2015), 114–141. DOI: https://doi.org/10.1016/j.ic.2014.10.009

[20] Ugo Dal Lago and Margherita Zorzi. 2012. Probabilistic operational semantics for the lambda calculus. *RAIRO - Theor. Inf. Appl.* 46, 3 (2012), 413–450. DOI: https://doi.org/10.1051/ita/2012012

[21] Karel de Leeuw, Edward Forrest Moore, Claude Elwood Shannon, and Norman Shapiro. 1956. Computability by probabilistic machines. *Automata Stud.* 34 (1956), 183–212.

[22] Javier Esparza, Andreas Gaiser, and Stefan Kiefer. 2012. Proving termination of probabilistic programs using patterns. In *CAV'12 (LNCS)*, P. Madhusudan and Sanjit A. Seshia (Eds.), Vol. 7358. Springer, 123–138. DOI: https://doi.org/10.1007/978-3-642-31424-7_14

[23] Luis María Ferrer Fioriti and Holger Hermanns. 2015. Probabilistic termination: Soundness, completeness, and compositionality. In *POPL'15*, Sriram K. Rajamani and David Walker (Eds.). ACM, 489–501. DOI: https://doi.org/10.1145/2676726.2677001

[24] Jean-Yves Girard, Paul Taylor, and Yves Lafont. 1989. *Proofs and Types*. Cambridge University Press, New York.

[25] Noah D. Goodman, Vikash K. Mansinghka, Daniel M. Roy, Keith Bonawitz, and Joshua B. Tenenbaum. 2008. Church: A language for generative models. In *UAI'08*, David A. McAllester and Petri Myllymäki (Eds.). AUAI Press, 220–229.

[26] Martin Hofmann. 1997. A mixed modal/linear lambda calculus with applications to Bellantoni-Cook safe recursion. In *CSL'97 (LNCS)*, Mogens Nielsen and Wolfgang Thomas (Eds.), Vol. 1414. Springer, 275–294.

[27] John Hughes, Lars Pareto, and Amr Sabry. 1996. Proving the correctness of reactive systems using sized types. In *POPL'96*, Hans-Juergen Boehm and Guy L. Steele Jr. (Eds.). ACM Press, 410–423. DOI: https://doi.org/10.1145/237721.240882

[28] Benjamin Lucien Kaminski and Joost-Pieter Katoen. 2015. On the hardness of almost-sure termination. In *MFCS (1) (Lecture Notes in Computer Science)*, Vol. 9234. Springer, 307–318.

[29] Benjamin Lucien Kaminski, Joost-Pieter Katoen, Christoph Matheja, and Federico Olmedo. 2018. Weakest precondition reasoning for expected runtimes of randomized algorithms. *J. ACM* 65, 5 (2018), 30:1–30:68.

[30] Ugo Dal Lago and Charles Grellois. 2017. Probabilistic termination by monadic affine sized typing. In *ESOP'17 (LNCS)*, Hongseok Yang (Ed.), Vol. 10201. Springer, 393–419. DOI: https://doi.org/10.1007/978-3-662-54434-1_15

[31] Ugo Dal Lago, Sara Zuppiroli, and Maurizio Gabbrielli. 2014. Probabilistic recursion theory and implicit computational complexity. *Sci. Ann. Comp. Sci.* 24, 2 (2014), 177–216. DOI: https://doi.org/10.7561/SACS.2014.2.177

[32] Christopher Manning and Hinrich Schütze. 2001. *Foundations of Statistical Natural Language Processing*. MIT Press.

[33] Annabelle McIver and Carroll Morgan. 2005. *Abstraction, Refinement and Proof for Probabilistic Systems*. Springer. DOI: https://doi.org/10.1007/b138392

[34] Annabelle McIver, Carroll Morgan, Benjamin Lucien Kaminski, and Joost-Pieter Katoen. 2018. A new proof rule for almost-sure termination. *PACMPL* 2, *(POPL'18)*, 33:1–33:28.

[35] Rajeev Motwani and Prabhakar Raghavan. 1995. *Randomized Algorithms*. Cambridge University Press.

[36] Judea Pearl. 1989. *Probabilistic Reasoning in Intelligent Systems - Networks of Plausible Inference*. Morgan Kaufmann.

[37] Amr Sabry and Matthias Felleisen. 1993. Reasoning about programs in continuation-passing style. *Lisp Symbol. Comput.* 6, 3–4 (1993), 289–360.

[38] Alexander Schrijver. 1986. *Theory of Linear and Integer Programming*. John Wiley & Sons, New York, NY.

[39] Kazushige Terui. 2007. Light affine lambda calculus and polynomial time strong normalization. *Arch. Math. Log.* 46, 3–4 (2007), 253–280. DOI: https://doi.org/10.1007/s00153-007-0042-6

[40] Sebastian Thrun. 2002. Robotic mapping: A survey. In *Exploring Artificial Intelligence in the New Millenium*. Morgan Kaufmann.

[41] Hongwei Xi. 2002. Dependent types for program termination verification. *Higher-Order Symbol. Comput.* 15, 1 (2002), 91–131. DOI: https://doi.org/10.1023/A:1019916231463

Received May 2017; revised October 2018; accepted October 2018