arXiv:1611.07255v2 [cs.LO] 23 Nov 2016

STANDARDIZATION AND CONSERVATIVITY OF A REFINED CALL-BY-VALUE LAMBDA-CALCULUS

GIULIO GUERRIERI, LUCA PAOLINI, AND SIMONA RONCHI DELLA ROCCA

Aix Marseille Université, CNRS, Centrale Marseille, I2M UMR 7373, F-13453, Marseille, France
e-mail address: giulio.guerrieri@univ-amu.fr

Dipartimento di Informatica, Università degli Studi di Torino, C.so Svizzera 185, Torino, Italia
e-mail address: paolini@di.unito.it, ronchi@di.unito.it

Abstract. We study an extension of Plotkin's call-by-value lambda-calculus via two commutation rules (sigma-reductions). These commutation rules are sufficient to remove harmful call-by-value normal forms from the calculus, so that it enjoys elegant characterizations of many semantic properties. We prove that this extended calculus is a conservative refinement of Plotkin's one. In particular, the notions of solvability and potential valuability for this calculus coincide with those for Plotkin's call-by-value lambda-calculus. The proof rests on a standardization theorem proved by generalizing Takahashi's approach of parallel reductions to our set of reduction rules. The standardization is weak (i.e. redexes are not fully sequentialized) because of overlapping interferences between reductions.

1. Introduction

Call-by-value evaluation is the most common parameter passing mechanism for programming languages: parameters are evaluated before being passed. The λ_v-calculus (λ_v for short) has been introduced by Plotkin in [Plo75] in order to give a formal account of call-by-value evaluation in the context of λ-calculus. Plotkin's λ_v has the same term syntax as the ordinary, i.e. call-by-name, λ-calculus (λ for short), but its reduction rule, β_v, is a restriction of β-reduction for λ: β_v-reduction reduces a β-redex only in case the argument is a *value* (i.e. a variable or an abstraction). While β_v is enough for evaluation of closed terms not reducing under abstractions, it turned out to be too weak in order to study semantical and operational properties of terms in λ_v. This fact makes the theory of λ_v (see [EHR92]) more complex to be described than that of λ. For example, in λ, β-reduction is sufficient to characterize solvability and (in addition with η) separability (see [Bar84] for an extensive survey); but in order to characterize similar properties for λ_v, only reduction rules incorrect

---

1998 ACM Subject Classification: D.3.1 Formal Definitions and Theory, F.3.2 Semantics of Programming Language, F.4.1 Mathematical Logic.

Key words and phrases: call-by-value, standardization, sequentialization, observational equivalence, sigma-reduction, head reduction, parallel reduction, internal reduction, standard sequence, lambda-calculus, solvability, potential valuability.

This paper is a revised and extended version of [GPR15], invited for the special issue of TLCA 2015.

LOGICAL METHODS
IN COMPUTER SCIENCE
DOI:10.2168/LMCS-???

for call-by-value evaluation have been defined (see [PR99, Pao02, RP04]): for λ_v this is disappointing and requires complex analyses. The reason of this mismatching is that in λ_v there are *stuck* β*-redexes* such as $(\lambda y.M)(zz)$, i.e. β-redexes that β_v-reduction will never fire because their argument is normal but not a value (nor will it ever become one). The real problem with stuck β-redexes is that they may prevent the creation of other β_v-redexes, providing *"premature"* β_v*-normal forms*. The issue is serious, as it affects termination and thus can impact on the study of observational equivalence and other operational properties in λ_v. For instance, it is well-known that in λ all unsolvable terms are not β-normalizable (more precisely, solvable terms coincide with the head β-normalizable ones). But in λ_v (see [RP04, AP12, CG14]) there are *unsolvable* β_v*-normal* terms, e.g. M and N in Eq. 1.1:

$$M = (\lambda y.\Delta)(zz)\Delta \qquad N = \Delta((\lambda y.\Delta)(zz)) \qquad (\text{where } \Delta = \lambda x.xx) \tag{1.1}$$

Such M and N contain the stuck β-redex $(\lambda y.\Delta)(zz)$ forbidding evaluation to keep going. These β_v-normal forms can be considered "premature" because they are unsolvable and so one would expect them to diverge. The idea that M and N should behave like the famous divergent term $\Delta\Delta$ is corroborated by the fact that in λ_v they are observationally equivalent to $\Delta\Delta$ and have the same semantics as $\Delta\Delta$ in all non-trivial denotational models of λ_v.

In a call-by-value setting, the issue of stuck β-redexes and then of premature β_v-normal forms arises only when one considers *open terms* (in particular, when the reduction under abstractions is allowed, since it forces to deal with "locally open" terms). Even if to model functional programming languages with a call-by-value parameter passing, such as OCaml, it is usually enough to just consider closed terms and evaluation not reducing under abstractions (i.e. function bodies are evaluated only when all parameters are supplied), the importance to consider open terms in a call-by-value setting can be found, for example, in partial evaluation (which evaluates a function when not all parameters are supplied, see [JGS93]), in the theory of proof assistants such as Coq (in particular, for type checking in a system based on dependent types, see [GL02]), or to reason about (denotational or operational) equivalences of terms in λ_v that are congruences, or about other theoretical properties of λ_v such as separability, potential valuability and solvability, as already mentioned.

We study the *shuffling calculus* λ_v^σ, the extension of λ_v proposed in [CG14]. It keeps the same term syntax as λ_v (and λ) and it adds to β_v-reduction two commutation rules, called σ_1 and σ_3, which "shuffle" constructors in order to move stuck β-redexes and unblock β_v-redexes that are hidden by the "hyper-sequential structure" of terms. These commutation rules for λ_v (referred also as σ*-reduction rules*) are similar to Regnier's σ-rules for λ [Reg92, Reg94] and inspired by linear logic proof-nets [Gir87]. It is well-known that β_v-reduction can be simulated by linear logic cut-elimination via the call-by-value "boring" translation $(\cdot)^v$ of λ-terms into proof-nets [Gir87, pp. 81-82], which decomposes the intuitionistic implication as follows: $(A \Rightarrow B)^v = !(A \multimap B)^v$ (see also [Acc15]). It turns out that the images under $(\cdot)^v$ of a σ-redex and its contractum are equal modulo some non-structural cut-elimination steps. Note that Regnier's σ-rules are contained in β-equivalence, while in λ_v our σ-rules are more interesting, as they are not contained into (i.e. they enrich) β_v-equivalence.

One of the benefits of λ_v^σ is that its σ-rules make all normal forms solvable (indeed M and N in Eq. 1.1 are not normal in λ_v^σ). More generally, λ_v^σ allows one to characterize semantical and operational properties which are relevant in a call-by-value setting, such as solvability and potential valuability, in an internal and elegant way, as shown in [CG14].

The main result of this paper is the conservativity of λ_v^σ with respect to λ_v. Namely, λ_v^σ is sound with respect to the operational semantics of λ_v (Corollary 5.4), and the notions of potential valuability and solvability characterize, respectively, the same classes of terms in λ_v^σ and λ_v (Theorem 5.7). This fully justifies the project in [CG14] where λ_v^σ has been introduced as a tool for studying λ_v by means of reductions sound for λ_v. These conservativity results are a consequence of a *standardization* property for λ_v^σ (Theorem 4.6) that formalizes the good interaction arising between β_v-reduction and σ-reduction in λ_v^σ.

Let us recall the notion of standardization, which has been first studied in the ordinary λ-calculus (see [CF58, Hin78, Mit79, Bar84]). A reduction sequence is *standard* if redexes are fired in a given order, and the standardization theorem establishes that every reduction sequence can be transformed into a standard one in a constructive way. Standardization is a key tool to grasp the way in which reductions works and sheds some light on relationships and dependencies between redexes. It is useful especially to characterize semantic properties through reduction strategies, such as normalization and operational adequacy.

Standardization theorems for λ_v have been proved by Plotkin [Plo75], Paolini and Ronchi Della Rocca [RP04, PR04] and Crary [Cra09]. Plotkin and Crary define the same notion of standard reduction sequence, based on a partial order between redexes, while Paolini and Ronchi Della Rocca define a different notion, based on a total order between redexes. According to the terminology of [Klo80, Kri90], the former gives rise to a weak standardization, while the latter to a strong one. These standardization theorems for λ_v have been proved using a notion of parallel reduction adapted for β_v-reduction. Parallel reduction has been originally introduced for λ by Tait and Martin-Löf to prove confluence of β-reduction: intuitively, it reduces a number of β-redexes in a term simultaneously. Takahashi [Tak89, Tak95] has improved this approach and shown that it can be used also to prove standardization for λ without involving the tricky notion of residual of a redex, unlike the proofs in [CF58, Hin78, Mit79, Bar84]. Crary [Cra09] has adapted to λ_v Takahashi's method for standardization. In order to prove our standardization theorem for λ_v^σ, we extend the notion of parallel reduction to include all the reductions of λ_v^σ. So, we consider two groups of redexes, β_v-redexes and σ-redexes (putting together σ_1 and σ_3), and we induce a total order between redexes of the two groups, without imposing any order between σ_1- and σ_3-redexes. Whenever σ-redexes are missing, this notion of standardization coincides with that presented in [PR04, RP04]. We show it is impossible to strengthen our standardization by (locally) giving precedence to σ_1-reduction over σ_3-reduction or vice-versa.

As usual, our standardization proof is based on a *sequentialization* result: inner reductions can always be postponed to the head ones, according to a non-standard definition of head reduction. However, our proof is peculiar with respect to other ones in the literature. In particular, our parallel reduction does not enjoy the diamond property (we are unaware of interesting parallel reductions that do not enjoy it), thus it cannot be used to prove the confluence. This lack is crucially related to the second distinctive aspect of our study, viz. the presence of several kinds of redexes being mutually overlapping (in the sense of [Ter03]).

The aim of this paper is first of all theoretical: to supply a tool for reasoning about semantic and operational properties of Plotkin's λ_v, such as observational equivalence, solvability and potential valuability. The shuffling calculus λ_v^σ realizes this aim, as shown by the conservativity results with respect to λ_v. These results are achieved since λ_v^σ avoids the problem of premature β_v-normal forms by dealing uniformly with open and closed terms, so allowing one to use the classical reasoning by induction on the structure of terms, which is essential in proving semantic and operational properties.

In the light of its good behaviour, we believe that λ_v^σ is also an interesting calculus deserving to be studied in itself and in comparison with other call-by-value extensions of Plotkin's λ_v dealing with the problem of stuck β-redexes, as done for instance in [AG16].

The approach supplied by λ_v^σ to circumvent the issue of stuck β-redexes might be profitably used also in more practical settings based on a call-by-value evaluation dealing with open terms, such as the aforementioned partial evaluation and theory of proof assistants.

Related work. Several variants of λ_v, arising from different perspectives, have been introduced in the literature for modeling the call-by-value computation and dealing with stuck β-redexes. We would like here to mention at least the contributions of Moggi [Mog88, Mog89], Felleisen and Sabry [SF92, SF93], Maraist *et al.* [MOTW95, MOTW99], Sabry and Wadler [SW97], Curien and Herbelin [CH00], Dyckhoff and Lengrand [DL07], Herbelin and Zimmerman [HZ09], Accattoli and Paolini [AP12], Accattoli and Sacerdoti Coen [AS15]. All these proposals are based on the introduction of new constructs to the syntax of λ_v and/or new reduction rules extending β_v, so the comparison between them is not easy with respect to syntactical properties (some detailed comparison is given in [AP12, AG16]). We point out that the calculi introduced in [Mog88, Mog89, SF92, SF93, MOTW95, SW97, MOTW99, HZ09] present some variants of our σ_1 and/or σ_3 rules, often in a setting with explicit substitutions. The shuffling calculus λ_v^σ has been introduced by Carraro and Guerrieri in [CG14] and further studied in [GPR15, Gue15, AG16].

Regnier [Reg92, Reg94] introduced in λ the rule σ_1 (but not σ_3) and another similar shuffling rule called σ_2. The σ-rules for λ and λ_v are different because they are inspired by two different translations of λ-terms into linear logic proof-nets (see [Gir87]). A generalization of our and Regnier's σ-rules is used in [EG16] for a variant of the λ-calculus subsuming both call-by-name and call-by-value evaluations.

Our approach to prove standardization for λ_v^σ is inspired by Takahashi's one [Tak89, Tak95] for λ based on parallel reduction, adapted by Crary [Cra09] for Plotkin's λ_v.

A preliminary version of this paper, focused essentially on the standardization result for λ_v^σ, has been presented in [GPR15].

Outline. In Section 2 the syntax of λ_v^σ with its reduction rules is introduced; in Section 3 the sequentialization property is proved; Section 4 proves the standardization theorem for λ_v^σ; in Section 5 the main results are given, namely the conservativity of λ_v^σ with respect to Plotkin's λ_v-calculus. Section 6 provides some conclusions and hints for future work.

2. The shuffling calculus: a call-by-value λ-calculus with σ-rules

In this section we introduce the *shuffling calculus* λ_v^σ, namely the call-by-value λ-calculus defined in [CG14] that adds two σ-reduction rules to the pure (i.e. without constants) call-by-value λ-calculus λ_v proposed by Plotkin in [Plo75]. The syntax of terms of λ_v^σ is the same as Plotkin's λ_v and then the same as the ordinary (i.e. call-by-name) λ-calculus λ.

Definition 2.1 (Term, value)**.** Given a countable set $\mathcal{V}$ of *variables* (denoted by $x, y, z, \dots$), the sets Λ of *terms* and Λ_v of *values* are defined by mutual induction as follows:

$$\begin{array}{lrcll} (\Lambda_v) & V, U & ::= & x \mid \lambda x.M & \textit{values} \\ (\Lambda) & M, N, L & ::= & V \mid MN & \textit{terms} \end{array}$$

Clearly, $\Lambda_v \subsetneq \Lambda$. Terms of the form MN (resp. $\lambda x.M$) are called *applications* (resp. *abstractions*). All terms are considered up to α-conversion (i.e. renaming of bound variables).

As usual, λ's associate to the right and applications to the left, so $\lambda xy.N$ stands for $\lambda x.(\lambda y.N)$ and MNL for $(MN)L$. The set of free variables of a term N is denoted by $\mathsf{fv}(N)$: N is *open* if $\mathsf{fv}(N) = \emptyset$, *closed* otherwise. Given $V_1, \dots, V_n \in \Lambda_v$ and pairwise distinct variables $x_1, \dots, x_n$, $N\{V_1/x_1, \dots, V_n/x_n\}$ denotes the term obtained by the *capture-avoiding simultaneous substitution* of V_i for each free occurrence of x_i in the term N (for all $1 \le i \le n$). Note that if $N \in \Lambda_v$ then $N\{V_1/x_1, \dots, V_n/x_n\} \in \Lambda_v$ (values are closed under substitution).

Remark 2.2. Any term can be written in a unique way as $VN_1 \dots N_n$ (a value V recursively applied to n terms $N_1, \dots, N_n$) for some $n \in \mathbb{N}$; in particular, values are obtained for $n = 0$.

From now on, we set $I = \lambda x.x$ and $\Delta = \lambda x.xx$. One-hole contexts are defined as usual.

Definition 2.3 (Context)**.** *Contexts* (with exactly one *hole* $(\!|\cdot|\!)$), denoted by $\mathtt{C}$, are defined via the grammar:

$$\mathtt{C} ::= (\!|\cdot|\!) \mid \lambda x.\mathtt{C} \mid \mathtt{C}M \mid M\mathtt{C}.$$

Let $\mathtt{C}$ be a context. The set of free variables of $\mathtt{C}$ is denoted by $\mathsf{fv}(\mathtt{C})$. We use $\mathtt{C}(\!|M|\!)$ for the term obtained by the capture-allowing substitution of the term M for the hole $(\!|\cdot|\!)$ in $\mathtt{C}$.

The set of λ_v^σ-reduction rules contains Plotkin's β_v-reduction rule together with two simple commutation rules called σ_1 and σ_3, studied in [CG14].

Definition 2.4 (Reduction rules)**.** For any $M, N, L \in \Lambda$ and any $V \in \Lambda_v$, we define the following binary relations on Λ:

$$\begin{aligned} (\lambda x.M)V &\mapsto_{\beta_v} M\{V/x\} & \\ (\lambda x.M)NL &\mapsto_{\sigma_1} (\lambda x.ML)N & \text{with } x \notin \mathsf{fv}(L) \\ V((\lambda x.L)N) &\mapsto_{\sigma_3} (\lambda x.VL)N & \text{with } x \notin \mathsf{fv}(V). \end{aligned}$$

We set $\mapsto_\sigma = \mapsto_{\sigma_1} \cup \mapsto_{\sigma_3}$ and $\mapsto_{\mathsf{v}} = \mapsto_{\beta_v} \cup \mapsto_\sigma$.

For any $\mathsf{r} \in \{\beta_v, \sigma_1, \sigma_3, \sigma, \mathsf{v}\}$, if $M \mapsto_{\mathsf{r}} M'$ then M is a r-*redex* and M' is its r-*contractum*. In the same sense, a term of the shape $(\lambda x.M)N$ (for any $M, N \in \Lambda$) is a β-*redex*.

The side conditions for $\mapsto_{\sigma_1}$ and $\mapsto_{\sigma_3}$ in Definition 2.4 can be always fulfilled by α-renaming. Clearly, any β_v-redex is a β-redex but the converse does not hold: $(\lambda x.z)(yI)$ is a β-redex but not a β_v-redex. Redexes of different kind may *overlap* (in the sense of [Ter03]): e.g. the term $\Delta I \Delta$ is a σ_1-redex and contains the β_v-redex ΔI; the term $\Delta(I\Delta)(xI)$ is a σ_1-redex and contains the σ_3-redex $\Delta(I\Delta)$, which contains in turn the β_v-redex $I\Delta$.

Remark 2.5. The relation $\mapsto_\sigma$ can be defined as a unique reduction rule, namely

$$\mathtt{E}(\!|(\lambda x.M)N|\!) \mapsto_\sigma (\lambda x.\mathtt{E}(\!|M|\!))N$$

where $\mathtt{E}$ is a context of the form $(\!|\cdot|\!)L$ or $V(\!|\cdot|\!)$ (for any $L \in \Lambda$ and $V \in \Lambda_v$) such that $x \notin \mathsf{fv}(\mathtt{E})$.

Let R be a binary relation on Λ. We denote by R^* (resp. R^+; $\mathsf{R}^=$) its reflexive-transitive (resp. transitive; reflexive) closure.

Definition 2.6 (Rewriting notations and terminology)**.** Let $\mathsf{r} \in \{\beta_v, \sigma_1, \sigma_3, \sigma, \mathsf{v}\}$.

- The r-*reduction* $\to_{\mathsf{r}}$ is the contextual closure of $\mapsto_{\mathsf{r}}$, i.e. $M \to_{\mathsf{r}} M'$ iff there is a context $\mathtt{C}$ and $N, N' \in \Lambda$ such that $M = \mathtt{C}(\!|N|\!)$, $M' = \mathtt{C}(\!|N'|\!)$ and $N \mapsto_{\mathsf{r}} N'$.
- The r-*equivalence* $=_{\mathsf{r}}$ is the congruence relation on Λ generated by $\mapsto_{\mathsf{r}}$, i.e. the reflexive-transitive and symmetric closure of $\to_{\mathsf{r}}$.

- Let M be a term: M is r-*normal* if there is no term N such that $M \to_{\mathsf{r}} N$; M is r-*normalizable* if there is a r-normal term N such that $M \to_{\mathsf{r}}^* N$, and we then say that N is a r-*normal form of* M; M is *strongly* r-*normalizing* if it does not exist an infinite sequence of r-reductions starting from M. Finally, $\to_{\mathsf{r}}$ is *strongly normalizing* if every $N \in \Lambda$ is strongly r-normalizing.

From Definitions 2.4 and 2.6, it follows immediately that $\to_{\mathsf{v}} = \to_{\beta_v} \cup \to_{\sigma}$ with $\to_{\sigma} \subsetneq \to_{\mathsf{v}}$ and $\to_{\beta_v} \subsetneq \to_{\mathsf{v}}$, and also that $\to_{\sigma} = \to_{\sigma_1} \cup \to_{\sigma_3}$ with $\to_{\sigma_1} \subsetneq \to_{\sigma}$ and $\to_{\sigma_3} \subsetneq \to_{\sigma}$.

Remark 2.7. Given $\mathsf{r} \in \{\beta_v, \sigma_1, \sigma_3, \sigma, \mathsf{v}\}$ (resp. $\mathsf{r} \in \{\sigma_1, \sigma_3, \sigma\}$), values are closed under r-reduction (resp. r-expansion): for any $V \in \Lambda_v$, if $V \to_{\mathsf{r}} M$ (resp. $M \to_{\mathsf{r}} V$) then $M \in \Lambda_v$ and more precisely $V = \lambda x.N$ and $M = \lambda x.L$ for some $N, L \in \Lambda$ with $N \to_{\mathsf{r}} L$ (resp. $L \to_{\mathsf{r}} N$).

Proposition 2.8 (Basic properties of reductions, [Plo75, CG14])**.** *The σ-reduction is confluent and strongly normalizing. The β_v- and v-reductions are confluent.*

Proof. Confluence of β_v-reduction has been proved in [Plo75]. The σ-reduction is strongly confluent in the sense of [Hue80], whence confluence of σ-reduction follows. The v-reduction is not strongly confluent and a more sophisticated proof is needed. All details (as well as the proof that σ-reduction is strongly normalizing) are in [CG14]. ☐

By confluence (Proposition 2.8), for any $\mathsf{r} \in \{\beta_v, \sigma, \mathsf{v}\}$ we have that: $M =_{\mathsf{r}} N$ iff $M \to_{\mathsf{r}}^* L \;{}_{\mathsf{r}}^{*}\!\leftarrow N$ for some term L; and any r-normalizable term has a *unique* r-normal form.

The *shuffling calculus* or λ_v^σ-*calculus* (λ_v^σ for short) is the set Λ of terms endowed with the reduction $\to_{\mathsf{v}}$. The set Λ endowed with the reduction $\to_{\beta_v}$ is the λ_v-calculus (λ_v for short), i.e. Plotkin's pure call-by-value λ-calculus [Plo75], a sub-calculus of λ_v^σ.

Example 2.9. Recalling the terms M and N in Eq. 1.1, one has that $M = (\lambda y.\Delta)(xI)\Delta \to_{\sigma_1} (\lambda y.\Delta\Delta)(xI) \to_{\beta_v} (\lambda y.\Delta\Delta)(xI) \to_{\beta_v} \dots$ and $N = \Delta((\lambda y.\Delta)(xI)) \to_{\sigma_3} (\lambda y.\Delta\Delta)(xI) \to_{\beta_v} (\lambda y.\Delta\Delta)(xI) \to_{\beta_v} \dots$ are the only possible v-reduction paths from M and N respectively: M and N are not v-normalizable and $M =_{\mathsf{v}} N$. But M and N are β_v-normal ($(\lambda y.\Delta)(xI)$ is a stuck β-redex) and different, hence $M \neq_{\beta_v} N$ by confluence of $\to_{\beta_v}$ (Proposition 2.8).

Example 2.9 shows how σ-reduction shuffles constructors and moves stuck β-redex in order to unblock β_v-redexes which are hidden by the "hyper-sequential structure" of terms, avoiding "premature" normal forms. An alternative approach to circumvent the issue of stuck β-redexes is given by λ_{vsub}, the call-by-value calculus with explicit substitutions introduced in [AP12], where hidden β_v-redexes are reduced using rules acting at a distance. In [AG16] it has been shown that λ_{vsub} and λ_v^σ can be embedded in each other preserving termination and divergence. Interestingly, both calculi are inspired by an analysis of Girard's "boring" call-by-value translation of λ-terms into linear logic proof-nets [Gir87, Acc15].

3. Sequentialization

Standardization is a consequence of a *sequentialization* property: every v-reduction sequence can always be rearranged in such a way that head v-reduction steps precede internal ones. To prove this sequentialization (Theorem 3.4), we adapt to λ_v^σ Takahashi's method [Tak95, Cra09] based on parallel reduction. This is the most technical part of the paper: for the sake of readability, this proof together with all needed lemmas are collected in Section 3.1.

First, we partition v-reduction into head v-reduction and internal v-reduction. In turn, head v-reduction divides up into head β_v-reduction and head σ-reduction. Their definitions are driven by the shape of terms, as given in Remark 2.2.

Definition 3.1 (Head β_v-reduction)**.** We define inductively the *head β_v-reduction* $\xrightarrow{h}_{\beta_v}$ by the following rules ($m \in \mathbb{N}$ in both rules):

$$\frac{}{(\lambda x.M)VM_1 \dots M_m \xrightarrow{h}_{\beta_v} M\{V/x\}M_1 \dots M_m}\,\beta_v \qquad \frac{N \xrightarrow{h}_{\beta_v} N'}{VNM_1 \dots M_m \xrightarrow{h}_{\beta_v} VN'M_1 \dots M_m}\,\mathit{right}\ .$$

Head β_v-reduction is the reduction strategy choosing at every step the (unique, if any) leftmost-outermost β_v-redex not in the scope of a λ: thus, it is a *deterministic* reduction (i.e. a partial function from Λ to Λ) and does not reduce values. It coincides with the "left reduction" defined in [Plo75, p. 136] for λ_v, called "evaluation" in [SF93, Las05, Cra09], and it models call-by-value evaluation as implemented in functional programming languages such as OCaml. Head β_v-reduction is often equivalently defined either by using the rules

$$\frac{}{(\lambda x.M)V \xrightarrow{h}_{\beta_v} M\{V/x\}} \qquad \frac{N \xrightarrow{h}_{\beta_v} N'}{VN \xrightarrow{h}_{\beta_v} VN'} \qquad \frac{M \xrightarrow{h}_{\beta_v} M'}{MN \xrightarrow{h}_{\beta_v} M'N}$$

or as the closure of the relation $\mapsto_{\beta_v}$ under evaluation contexts $\mathtt{E} ::= \llparenthesis\cdot\rrparenthesis \mid \mathtt{E}M \mid V\mathtt{E}$. We prefer our presentation since it allows more concise proofs and stresses in a more explicit way how head β_v-reduction acts on the general shape of terms, as given in Remark 2.2.

Definition 3.2 (Head σ-reduction)**.** We define inductively the *head σ-reduction* $\xrightarrow{h}_{\sigma}$ by the following rules ($m \in \mathbb{N}$ in all the rules, $x \notin \mathsf{fv}(L)$ in the rule σ_1, $x \notin \mathsf{fv}(V)$ in the rule σ_3):

$$\frac{}{(\lambda x.M)NLM_1 \dots M_m \xrightarrow{h}_{\sigma} (\lambda x.ML)NM_1 \dots M_m}\,\sigma_1 \qquad \frac{N \xrightarrow{h}_{\sigma} N'}{VNM_1 \dots M_m \xrightarrow{h}_{\sigma} VN'M_1 \dots M_m}\,\mathit{right}$$

$$\frac{}{V((\lambda x.L)N)M_1 \dots M_m \xrightarrow{h}_{\sigma} (\lambda x.VL)NM_1 \dots M_m}\,\sigma_3$$

The *head σ_1-*(resp. *head σ_3-*)*reduction* is $\xrightarrow{h}_{\sigma_1} = \to_{\sigma_1} \cap \xrightarrow{h}_{\sigma}$ (resp. $\xrightarrow{h}_{\sigma_3} = \to_{\sigma_3} \cap \xrightarrow{h}_{\sigma}$).

Head σ-reduction is a non-deterministic reduction, since it reduces at every step "one of the leftmost-outermost" σ_1- or σ_3-redexes not in the scope of a λ: such head σ-redexes may be not unique and overlap, e.g. the term N in Figure 1 is a head σ_1-redex containing the head σ_3-redex $(\lambda y.y')(\Delta(xI))$, or the term $I(\Delta(I(xI)))$ is a head σ_3-redex containing another head σ_3-redex $\Delta(I(xI))$ (see Definition 3.3 below for the formal definition of head redex).

Definition 3.3 (Head v-reduction, internal v-reduction, head redex)**.** The *head v-reduction* is $\xrightarrow{h}_{\mathsf{v}} = \xrightarrow{h}_{\beta_v} \cup \xrightarrow{h}_{\sigma}$. The *internal v-reduction* is $\xrightarrow{int}_{\mathsf{v}} = \to_{\mathsf{v}} \smallsetminus \xrightarrow{h}_{\mathsf{v}}$.

Given $\mathsf{r} \in \{\beta_v, \sigma_1, \sigma_3, \sigma, \mathsf{v}\}$, a *head r-redex of a term M* is a r-redex R occurring in M such that $M \xrightarrow{h}_{\mathsf{r}} N$ for some term N obtained from M by replacing R with its r-contractum.

Note that $\mapsto_{\beta_v} \subsetneq \xrightarrow{h}_{\beta_v} \subsetneq \to_{\beta_v}$ and $\mapsto_{\sigma} \subsetneq \xrightarrow{h}_{\sigma} \subsetneq \to_{\sigma}$ and $\mapsto_{\mathsf{v}} \subsetneq \xrightarrow{h}_{\mathsf{v}} \subsetneq \to_{\mathsf{v}}$. It is immediate to check that, for any $M, M', N, L \in \Lambda$, $M \to_{\mathsf{v}} M'$ implies $NLM \xrightarrow{int}_{\mathsf{v}} NLM'$.

Head v-reduction is non-deterministic since head σ-reduction is so, and since head β_v- and head σ_1-(resp. σ_3-)redexes may overlap, as in the term $I\Delta I$ (resp. $I(\Delta I)$). Also, Figure 1 shows that head σ- and head v-reductions are not confluent and a term may have several head σ/v-normal forms, indeed N_0' and N_1' are head σ/v-normal forms of N but $N_0' \neq N_1'$. However, this does not contradict the confluence of σ- and v-reductions because $N_1' \to_{\sigma} N_0'$ by performing an internal v-reduction step. Corollary 5.1.2 in Section 5 implies that if a term N has a head v-normal form $V \in \Lambda_v$, then V is the unique head v-normal form of N.

Now we can state the first main result of this paper, namely the sequentialization theorem (Theorem 3.4), saying that any v-reduction sequence can be sequentialized into a head

$$\begin{tikzcd}
 & N = (\lambda y.y')(\Delta(xI))I \arrow[dl, bend right, "h"', "\sigma_1"] \arrow[dr, bend left, "h", "\sigma_3"'] & \\
N_0 = (\lambda y.y'I)(\Delta(xI)) \arrow[d, "h"', "\sigma_3"] & & (\lambda z.(\lambda y.y')(zz))(xI)I = N_1 \arrow[d, "\sigma_1"', "h"] \\
N_0' = (\lambda z.(\lambda y.y'I)(zz))(xI) & & (\lambda z.(\lambda y.y')(zz)I)(xI) = N_1' \arrow[ll, dotted, "\sigma_1"]
\end{tikzcd}$$

Figure 1: Overlapping of (head) σ-redexes.

β_v-reduction sequence followed by a head σ-reduction sequence, followed by an internal v-reduction sequence. In ordinary λ-calculus, the well-known result corresponding to Theorem 3.4 says that a β-reduction sequence can be factorized in a head β-reduction sequence followed by an internal β-reduction sequence (see for example [Tak95, Corollary 2.6]).

Theorem 3.4 (Sequentialization; its proof is in Section 3.1)**.** *If* $M \to_{\mathsf{v}}^{*} M'$ *then there exist* $L, N \in \Lambda$ *such that* $M \xrightarrow{h}{}^{*}_{\beta_v} L \xrightarrow{h}{}^{*}_{\sigma} N \xrightarrow{int}{}^{*}_{\mathsf{v}} M'$.

Sequentialization (Theorem 3.4) imposes no order on head σ-reduction steps, in accordance with the notion of head σ-reduction (Definition 3.2) which puts together head σ_1/σ_3-reduction steps. So, a natural question arises: is it possible to sequentialize them? More precisely, we wonder if it is possible to anticipate *a priori* all the head σ_1- or all the head σ_3-reduction steps. The answer is negative, as proved by the next two counterexamples.

- $M = x((\lambda y.z')(zI))\Delta \xrightarrow{h}_{\sigma_3} (\lambda y.xz')(zI)\Delta \xrightarrow{h}_{\sigma_1} (\lambda y.xz'\Delta)(zI) = N$, but there exists no L such that $M \xrightarrow{h}{}^{*}_{\sigma_1} L \xrightarrow{h}{}^{*}_{\sigma_3} N$. In fact, M contains only a head σ_3-redex and $(\lambda y.xz')(zI)\Delta$ has only a head σ_1-redex, created by firing the head σ_3-redex in M.
- $M = x((\lambda y.z')(zI)\Delta) \xrightarrow{h}_{\sigma_1} x((\lambda y.z'\Delta)(zI)) \xrightarrow{h}_{\sigma_3} (\lambda y.x(z'\Delta))(zI) = N$, but there is no L such that $M \xrightarrow{h}{}^{*}_{\sigma_3} L \xrightarrow{h}{}^{*}_{\sigma_1} N$. In fact, M contains only a head σ_1-redex and $x((\lambda y.z'\Delta)(zI))$ has only a head σ_3-redex, created by firing the head σ_1-redex in M.

The impossibility of prioritizing a kind of head σ-reduction over the other is due to the fact that a head σ_1-reduction step can create a new head σ_3-redex, and vice-versa. Thus, sequentialization (and then standardization) does not force a total order on head σ-redexes. This is not a serious issue, since head σ-reduction is strongly normalizing (by Proposition 2.8, as $\xrightarrow{h}_{\sigma} \subseteq \to_{\sigma}$) and hence the order in which head σ-reduction steps are performed is irrelevant. Moreover, following Remark 2.5, it seems natural to treat head σ_1- and head σ_3-reductions as a same reduction also because the two axiom schemes σ_1 and σ_3 in the definition of head σ-reduction (Definition 3.2) can be equivalently replaced by the unique axiom scheme

$$\frac{}{\mathtt{E}(\!|(\lambda x.M)N|\!)M_1 \ldots M_m \xrightarrow{h}_{\sigma} (\lambda x.\mathtt{E}(\!|M|\!))NM_1 \ldots M_m}\ \sigma$$

where $\mathtt{E}$ is a context of the form $(\!|\cdot|\!)L$ or $V(\!|\cdot|\!)$ such that $x \notin \mathsf{fv}(\mathtt{E})$.

Sequentialization (Theorem 3.4) says that any v-reduction sequence from a term M to a term M' can be rearranged into an initial head β_v-reduction sequence (whose steps reduce, in a deterministic way, the unique leftmost-outermost β_v-redex not under the scope of a λ) from M to some term L, followed by a head σ-reduction sequence (whose steps reduce, non-deterministically, one of the leftmost-outermost σ-redexes not in the scope of a λ) from L to some term N, followed by an internal v-reduction sequence from N to M'. For this internal v-reduction sequence, the same kind of decomposition can be iterated on the subterms of N.

3.1. Proof of the Sequentialization Theorem.

In this subsection we present a detailed proof, with all auxiliary lemmas, of Theorem 3.4. First, we define parallel reduction.

Definition 3.5 (Parallel reduction)**.** We define inductively the *parallel reduction* $\Rightarrow$ by the following rules ($x \notin \mathsf{fv}(L)$ in the rule σ_1, $x \notin \mathsf{fv}(V)$ in the rule σ_3):

$$\frac{V \Rightarrow V' \quad M_i \Rightarrow M_i' \quad (m \in \mathbb{N},\ 0 \leq i \leq m)}{(\lambda x.M_0)VM_1 \dots M_m \Rightarrow M_0'\{V'/x\}M_1' \dots M_m'}\,\beta_v \qquad \frac{N \Rightarrow N' \quad L \Rightarrow L' \quad M_i \Rightarrow M_i' \quad (m \in \mathbb{N},\ 0 \leq i \leq m)}{(\lambda x.M_0)NLM_1 \dots M_m \Rightarrow (\lambda x.M_0'L')N'M_1' \dots M_m'}\,\sigma_1$$

$$\frac{V \Rightarrow V' \qquad N \Rightarrow N' \qquad L \Rightarrow L' \qquad M_i \Rightarrow M_i' \quad (m \in \mathbb{N},\ 1 \leq i \leq m)}{V((\lambda x.L)N)M_1 \dots M_m \Rightarrow (\lambda x.V'L')N'M_1' \dots M_m'}\,\sigma_3$$

$$\frac{M_i \Rightarrow M_i' \quad (m \in \mathbb{N},\ 0 \leq i \leq m)}{(\lambda x.M_0)M_1 \dots M_m \Rightarrow (\lambda x.M_0')M_1' \dots M_m'}\,\lambda \qquad\qquad \frac{M_i \Rightarrow M_i' \quad (m \in \mathbb{N},\ 1 \leq i \leq m)}{x\,M_1 \dots M_m \Rightarrow x\,M_1' \dots M_m'}\,\mathit{var}$$

The rule *var*, in Definition 3.5, has no premises when $m = 0$: this is the base case of the inductive definition of $\Rightarrow$. The rules σ_1 and σ_3 have exactly three premises when $m = 0$. Intuitively, $M \Rightarrow M'$ means that M' is obtained from M by reducing a number of β_v-, σ_1- and σ_3-redexes (existing in M) simultaneously.

Definition 3.6 (Internal parallel reduction, strong parallel reduction)**.** We define inductively the *internal parallel reduction* $\overset{int}{\Rightarrow}$ by the following rules:

$$\frac{N \Rightarrow N'}{\lambda x.N \overset{int}{\Rightarrow} \lambda x.N'}\,\lambda \qquad \frac{}{x \overset{int}{\Rightarrow} x}\,\mathit{var} \qquad \frac{V \Rightarrow V' \quad N \overset{int}{\Rightarrow} N' \quad M_i \Rightarrow M_i' \quad (m \in \mathbb{N},\ 1 \leq i \leq m)}{VNM_1 \dots M_m \overset{int}{\Rightarrow} V'N'M_1' \dots M_m'}\,\mathit{right}$$

The *strong parallel reduction* $\Rrightarrow$ is defined by: $M \Rrightarrow N$ iff $M \Rightarrow N$ and there exist $M', M'' \in \Lambda$ such that $M \overset{h}{\to}_{\beta_v}^{*} M' \overset{h}{\to}_{\sigma}^{*} M'' \overset{int}{\Rightarrow} N$.

Notice that the rule *right* in Definition 3.6 has exactly two premises when $m = 0$.

Lemma 3.7 (Reflexivity)**.** *The relations $\Rightarrow$, $\Rrightarrow$ and $\overset{int}{\Rightarrow}$ are reflexive.*

Proof. The reflexivity of $\Rrightarrow$ follows immediately from the reflexivity of $\Rightarrow$ and $\overset{int}{\Rightarrow}$. The proofs of reflexivity of $\Rightarrow$ and $\overset{int}{\Rightarrow}$ are both by structural induction on a term: in the case of $\Rightarrow$, recall that any term is of the form $(\lambda x.N)M_1 \dots M_m$ or $x\,M_1 \dots M_m$ for some $m \in \mathbb{N}$ (Remark 2.2), and then apply the rule λ or *var* respectively, together with the inductive hypothesis; in the case of $\overset{int}{\Rightarrow}$, recall that every term is of the form $\lambda x.M$ or x or $VNM_1 \dots M_m$ for some $m \in \mathbb{N}$, and then apply the rule λ (together with the reflexivity of $\Rightarrow$) or *var* or *right* (together with the reflexivity of $\Rightarrow$ and the inductive hypothesis) respectively. $\square$

We have $\overset{int}{\Rightarrow} \subsetneq \Rrightarrow \subseteq \Rightarrow$ (first, prove that $\overset{int}{\Rightarrow} \subseteq \Rightarrow$ by induction on the derivation of $N \overset{int}{\Rightarrow} N'$, the other inclusions follow from the definition of $\Rrightarrow$; note that $II \Rrightarrow I$ but $II \overset{int}{\not\Rightarrow} I$) and, by reflexivity of $\Rightarrow$ (Lemma 3.7), $\overset{h}{\to}_{\beta_v} \subsetneq \Rightarrow$ and $\overset{h}{\to}_{\sigma} \subsetneq \Rightarrow$. Observe that $\Delta\Delta \mathrel{\mathsf{R}} \Delta\Delta$ for any $\mathsf{R} \in \{\mapsto_{\beta_v}, \overset{h}{\to}_{\beta_v}, \to_{\beta_v}, \Rightarrow, \overset{int}{\Rightarrow}, \Rrightarrow\}$, even if for different reasons: for example, $\Delta\Delta \overset{int}{\Rightarrow} \Delta\Delta$ by reflexivity of $\overset{int}{\Rightarrow}$ (Lemma 3.7), whereas $\Delta\Delta \overset{h}{\to}_{\beta_v} \Delta\Delta$ by reducing the only β_v-redex.

Some useful properties relating values and reductions follow. Note that Lemmas 3.8.1-2 imply that *all values are head* v-*normal*; the converse fails, since xI is head v-normal but not a value. It can be shown that closed head v-normal forms are values (actually, abstractions).

Lemma 3.8 (Values vs. reductions)**.**

(1) *The head β_v-reduction $\xrightarrow{h}_{\beta_v}$ does not reduce a value (i.e. values are head β_v-normal).*

(2) *The head σ-reduction $\xrightarrow{h}_{\sigma}$ does neither reduce a value nor reduce to a value.*

(3) *Variables and abstractions are preserved by $\overset{int}{\Leftarrow}$, more precisely: if $M \overset{int}{\Rightarrow} x$ (resp. $M \overset{int}{\Rightarrow} \lambda x.N'$) then $M = x$ (resp. $M = \lambda x.N$ for some $N \in \Lambda$ such that $N \Rightarrow N'$).*

(4) *If $M \Rightarrow M'$ then $\lambda x.M \ \mathsf{R}\ \lambda x.M'$ for any $\mathsf{R} \in \{\Rightarrow, \overset{int}{\Rightarrow}, \Rrightarrow\}$.*

(5) *For any $V, V' \in \Lambda_v$, one has $V \overset{int}{\Rightarrow} V'$ iff $V \Rightarrow V'$ iff $V \Rrightarrow V'$.*

Proof. (1) For every $M \in \Lambda$ and every $V \in \Lambda_v$, we have $V \overset{h}{\not\to}_{\beta_v} M$, because the head β_v-reduction does not reduce under λ's.

(2) For any $N \in \Lambda$ and $V \in \Lambda_v$, $V \overset{h}{\not\to}_{\sigma} N$ and $N \overset{h}{\not\to}_{\sigma} V$ since in the conclusion of any rule of Definition 3.2 the terms on the right and on the left of $\xrightarrow{h}_{\sigma}$ are applications.

(3) By simple inspection of the rules of $\overset{int}{\Rightarrow}$ (Definition 3.6), if $M \overset{int}{\Rightarrow} x$ (resp. $M \overset{int}{\Rightarrow} \lambda x.N'$) then the last rule in the derivation is necessarily *var* (resp. λ).

(4) For $\mathsf{R} \in \{\Rightarrow, \overset{int}{\Rightarrow}\}$ we apply the rule λ to conclude that $\lambda x.M \ \mathsf{R}\ \lambda x.M'$, therefore $\lambda x.M \Rrightarrow \lambda x.M'$ according to the definition of $\Rrightarrow$, since $\xrightarrow{h}{}^{*}_{\beta_v}$ and $\xrightarrow{h}{}^{*}_{\sigma}$ are reflexive.

(5) First we show that $V \overset{int}{\Rightarrow} V'$ iff $V \Rightarrow V'$. The left-to-right direction holds since $\overset{int}{\Rightarrow} \subseteq \Rightarrow$. Conversely, assume $V \Rightarrow V'$: if V is a variable then $V = V'$ and hence $V \overset{int}{\Rightarrow} V'$ by applying the rule *var* for $\overset{int}{\Rightarrow}$; otherwise $V = \lambda x.N$ for some $N \in \Lambda$, and then necessarily $V' = \lambda x.N'$ with $N \Rightarrow N'$, so $V \overset{int}{\Rightarrow} V'$ by applying the rule λ for $\overset{int}{\Rightarrow}$.

Now we prove that $V \Rightarrow V'$ iff $V \Rrightarrow V'$. The right-to-left direction follows immediately from the definition of $\Rrightarrow$ (Definition 3.6). Conversely, if $V \Rightarrow V'$ then we have just shown that $V \overset{int}{\Rightarrow} V'$, so $V \Rrightarrow V'$ since $\xrightarrow{h}{}^{*}_{\beta_v}$ and $\xrightarrow{h}{}^{*}_{\sigma}$ are reflexive. ☐

We collect some basic closure properties and relations that hold for reductions.

Lemma 3.9.

(1) *If $M \Rightarrow M'$ and $N \Rightarrow N'$ then $MN \Rightarrow M'N'$.*

(2) *If $\mathsf{R} \in \{\xrightarrow{h}_{\beta_v}, \xrightarrow{h}_{\sigma}\}$ and $M \ \mathsf{R}\ M'$, then $MN \ \mathsf{R}\ M'N$ for any $N \in \Lambda$.*

(3) *If $M \overset{int}{\Rightarrow} M'$ and $N \Rightarrow N'$ where $M' \notin \Lambda_v$, then $MN \overset{int}{\Rightarrow} M'N'$.*

(4) *$\to_{\mathsf{v}} \subseteq \Rightarrow \subseteq \to_{\mathsf{v}}^{*}$ and hence $\Rightarrow^{*} = \to_{\mathsf{v}}^{*}$.*

(5) *$\xrightarrow{int}_{\mathsf{v}} \subseteq \overset{int}{\Rightarrow} \subseteq \xrightarrow{int}{}^{*}_{\mathsf{v}}$ and hence $\overset{int}{\Rightarrow}{}^{*} = \xrightarrow{int}{}^{*}_{\mathsf{v}}$.*

(6) *$\Rightarrow$ is confluent.*

(7) *If $M \xrightarrow{int}{}^{*}_{\mathsf{v}} x$ (resp. $M \xrightarrow{int}{}^{*}_{\mathsf{v}} \lambda x.N'$) then $M = x$ (resp. $M = \lambda x.N$ with $N \to_{\mathsf{v}}^{*} N'$).*

(8) *For any $\mathsf{R} \in \{\xrightarrow{h}_{\beta_v}, \xrightarrow{h}_{\sigma}\}$, if $M \ \mathsf{R}\ M'$ then $M\{V/x\} \ \mathsf{R}\ M'\{V/x\}$ for any $V \in \Lambda_v$.*

Proof. (1) Just add the derivation of $N \Rightarrow N'$ as the "rightmost" premise of the last rule of the derivation of $M \Rightarrow M'$.

(2) In the conclusion of the derivation of $M \ \mathsf{R}\ M'$, replace $M \ \mathsf{R}\ M'$ with $MN \ \mathsf{R}\ M'N$.

(3) The last rule in the derivation of $M \overset{int}{\Rightarrow} M'$ can be neither λ nor *var* because $M' \notin \Lambda_v$, so it is *right* and hence we can add the derivation of $N \Rightarrow N$ (which exists since $\Rightarrow$ is reflexive, Lemma 3.7) as its rightmost premise. Note that the hypothesis $M' \notin \Lambda_v$ is crucial: for example, $x \overset{int}{\Rightarrow} x$ and $I\Delta \Rightarrow \Delta$ but $I\Delta \overset{int}{\not\Rightarrow} \Delta$ and thus $x(I\Delta) \overset{int}{\not\Rightarrow} x\Delta$.

(4) The proof that $M \to_{\mathsf{v}} M'$ implies $M \Rightarrow M'$ is by induction on $M \in \Lambda$, using the reflexivity of $\Rightarrow$ (Lemma 3.7) and Lemma 3.9.1. The proof that $M \Rightarrow M'$ implies $M \to_{\mathsf{v}}^{*} M'$ is by straightforward induction on the derivation of $M \Rightarrow M'$.

(5) We prove that $M \xrightarrow{int}_{\mathsf{v}} M'$ implies $M \overset{int}{\Rightarrow} M'$ by induction on $M \in \Lambda$. According to Remark 2.2, $M = VN_1 \dots N_n$ for some $n \in \mathbb{N}$, $V \in \Lambda_v$ and $N_1, \dots, N_n \in \Lambda$. Since $M \to_{\mathsf{v}} M'$ and $M \overset{h}{\not\to}_{\mathsf{v}} M'$, there are only three cases:
- either $M' = V'N_1 \dots N_n$ with $V \to_{\mathsf{v}} V'$, then $V = \lambda x.N$ and $V' = \lambda x.N'$ with $N \to_{\mathsf{v}} N'$ by Remark 2.7, so $N \Rightarrow N'$ according to Lemma 3.9.4, and thus $V = \lambda x.N \overset{int}{\Rightarrow} \lambda x.N' = V'$ by applying the rule λ for $\overset{int}{\Rightarrow}$; if $n = 0$ then $M = \lambda x.N$ and $M' = \lambda x.N'$ and we are done; otherwise $n > 0$ and hence $V \Rightarrow V'$ (Lemma 3.8.5), so $M \overset{int}{\Rightarrow} M'$ by applying the rule *right* for $\overset{int}{\Rightarrow}$, since $N_1 \overset{int}{\Rightarrow} N_1$ and $N_i \Rightarrow N_i$ for any $2 \leq i \leq n$ by reflexivity of $\overset{int}{\Rightarrow}$ and $\Rightarrow$ (Lemma 3.7);
- or $n > 0$ and $M' = VN_1'N_2 \dots N_n$ with $N_1 \xrightarrow{int}_{\mathsf{v}} N_1'$, then $N_1 \overset{int}{\Rightarrow} N_1'$ by induction hypothesis, and $V \Rightarrow V$ and $N_i \Rightarrow N_i$ for any $2 \leq i \leq n$ by reflexivity of $\Rightarrow$ (Lemma 3.7); hence $M \overset{int}{\Rightarrow} M'$ by applying the rule *right* for $\overset{int}{\Rightarrow}$;
- or $n > 1$ and $M' = VN_1 \dots N_i' \dots N_n$ with $N_i \to_{\mathsf{v}} N_i'$ for some $2 \leq i \leq n$, then $N_i \Rightarrow N_i'$ by Lemma 3.9.4, $V \Rightarrow V$ and $N_j \Rightarrow N_j$ for any $2 \leq j \leq n$ with $j \neq i$ and $N_1 \overset{int}{\Rightarrow} N_1$ by reflexivity of $\Rightarrow$ and $\overset{int}{\Rightarrow}$ (Lemma 3.7); hence $M \overset{int}{\Rightarrow} M'$ by applying the rule *right* for $\overset{int}{\Rightarrow}$.

The proof that $M \overset{int}{\Rightarrow} M'$ implies $M \xrightarrow{int}{}^{*}_{\mathsf{v}} M'$ is by straightforward induction on the derivation of $M \overset{int}{\Rightarrow} M'$, using that $\Rightarrow \subseteq \to_{\mathsf{v}}^*$ (Lemma 3.9.4).

(6) Since $\Rightarrow^* = \to_{\mathsf{v}}^*$ according to Lemma 3.9.4, Proposition 2.8 just says that $\Rightarrow$ is confluent. Anyway, we remark that $\Rightarrow$ does not enjoy the diamond property, see Section 6.

(7) Since $\overset{int}{\Rightarrow}{}^{*} = \xrightarrow{int}{}^{*}_{\mathsf{v}}$ (Lemma 3.9.5) and $\Rightarrow \subseteq \to_{\mathsf{v}}^*$ (Lemma 3.9.4), then Lemma 3.8.3 can be reformulated substituting $\xrightarrow{int}{}^{*}_{\mathsf{v}}$ for $\overset{int}{\Rightarrow}$, and $\to_{\mathsf{v}}^*$ for $\Rightarrow$.

(8) The proof is by induction on the derivation of $M \mathrel{\mathsf{R}} M'$, for any $\mathsf{R} \in \{\xrightarrow{h}_{\beta_v}, \xrightarrow{h}_{\sigma}\}$. ☐

Parallel reduction is closed under substitution, as stated by the following lemma.

Lemma 3.10 (Substitution vs. $\Rightarrow$)**.** *If* $M \Rightarrow M'$ *and* $V \Rightarrow V'$ *then* $M\{V/x\} \Rightarrow M'\{V'/x\}$.

Proof. By induction on the derivation of $M \Rightarrow M'$. Let us consider its last rule r.

- If $\mathsf{r} = \mathit{var}$ then $M = y\, M_1 \dots M_m$ and $M' = y\, M_1' \dots M_m'$ with $m \in \mathbb{N}$ and $M_i \Rightarrow M_i'$ for any $1 \leq i \leq m$. By induction hypothesis, $M_i\{V/x\} \Rightarrow M_i'\{V'/x\}$ for any $1 \leq i \leq m$. If $y \neq x$ then $M\{V/x\} = y\, M_1\{V/x\} \dots M_m\{V/x\}$ and $M'\{V'/x\} = y\, M_1'\{V'/x\} \dots M_m'\{V'/x\}$, so $M\{V/x\} \Rightarrow M'\{V'/x\}$ by applying the rule *var* for $\Rightarrow$. Otherwise $y = x$ and then $M\{V/x\} = VM_1\{V/x\} \dots M_m\{V/x\}$ and $M'\{V'/x\} = V'M_1'\{V'/x\} \dots M_m'\{V'/x\}$, hence $M\{V/x\} \Rightarrow M'\{V'/x\}$ by Lemma 3.9.1.
- If $\mathsf{r} = \lambda$ then $M = (\lambda y.M_0)M_1 \dots M_m$ and $M' = (\lambda y.M_0')M_1' \dots M_m'$ with $m \in \mathbb{N}$ and $M_i \Rightarrow M_i'$ for all $0 \leq i \leq m$; we can suppose without loss of generality that $y \notin \mathsf{fv}(V) \cup \{x\}$. By induction hypothesis, $M_i\{V/x\} \Rightarrow M_i'\{V'/x\}$ for all $0 \leq i \leq m$. By applying the rule λ for $\Rightarrow$, $M\{V/x\} = (\lambda y.M_0\{V/x\})M_1\{V/x\} \dots M_m\{V/x\} \Rightarrow (\lambda y.M_0'\{V'/x\})M_1'\{V'/x\} \dots M_m'\{V'/x\} = M'\{V'/x\}$.
- If $\mathsf{r} = \sigma_1$ then $M = (\lambda y.M_0)NLM_1 \dots M_m$ and $M' = (\lambda y.M_0'L')N'M_1' \dots M_m'$ with $m \in \mathbb{N}$, $L \Rightarrow L'$, $N \Rightarrow N'$ and $M_i \Rightarrow M_i'$ for any $0 \leq i \leq m$; we can suppose without loss of generality that $y \notin \mathsf{fv}(V) \cup \{x\}$. By induction hypothesis, $L\{V/x\} \Rightarrow L'\{V'/x\}$, $N\{V/x\} \Rightarrow N'\{V'/x\}$ and $M_i\{V/x\} \Rightarrow M_i'\{V'/x\}$ for any $0 \leq i \leq m$. Hence $M\{V/x\} \Rightarrow M'\{V'/x\}$ by applying the rule σ_1, since $M\{V/x\} = (\lambda y.M_0\{V/x\})N\{V/x\}L\{V/x\}M_1\{V/x\} \dots M_m\{V/x\}$ and $M'\{V'/x\} = (\lambda y.M_0'\{V'/x\}L'\{V'/x\})N'\{V'/x\}M_1'\{V'/x\} \dots M_m'\{V'/x\}$.

- If $\mathsf{r} = \sigma_3$ then $M = U((\lambda y.L)N)M_1 \dots M_m$ and $M' = (\lambda y.U'L')N'M_1' \dots M_m'$ with $m \in \mathbb{N}$ and $U, U' \in \Lambda_v$, $U \Rightarrow U'$, $L \Rightarrow L'$, $N \Rightarrow N'$ and $M_i \Rightarrow M_i'$ for any $1 \leq i \leq m$; we can suppose without loss of generality that $y \notin \mathsf{fv}(V) \cup \{x\}$. By induction hypothesis, $U\{V/x\} \Rightarrow U'\{V'/x\}$, $L\{V/x\} \Rightarrow L'\{V'/x\}$, $N\{V/x\} \Rightarrow N'\{V'/x\}$ and $M_i\{V/x\} \Rightarrow M_i'\{V'/x\}$ for any $1 \leq i \leq m$. So, $M\{V/x\} \Rightarrow M'\{V'/x\}$ by applying the rule σ_3, since $M\{V/x\} = U\{V/x\}((\lambda y.L\{V/x\})N\{V/x\})M_1\{V/x\} \dots M_m\{V/x\}$ and $M'\{V'/x\} = (\lambda y.U'\{V'/x\}L'\{V'/x\})N'\{V'/x\}M_1'\{V'/x\} \dots M_m'\{V'/x\}$.
- Finally, if $\mathsf{r} = \beta_v$, then $M = (\lambda y.M_0)V_0M_1 \dots M_m$ and $M' = M_0'\{V_0'/y\}M_1' \dots M_m'$ with $m \in \mathbb{N}$, $V_0 \Rightarrow V_0'$ and $M_i \Rightarrow M_i'$ for any $0 \leq i \leq m$; we can suppose without loss of generality that $y \notin \mathsf{fv}(V) \cup \{x\}$. By induction hypothesis, $V_0\{V/x\} \Rightarrow V_0'\{V'/x\}$ and $M_i\{V/x\} \Rightarrow M_i'\{V'/x\}$ for any $0 \leq i \leq m$. So, $M\{V/x\} \Rightarrow M'\{V'/x\}$ by applying the rule β_v, since $M\{V/x\} = (\lambda y.M_0\{V/x\})V_0\{V/x\}M_1\{V/x\} \dots M_m\{V/x\}$ and
$$\begin{aligned} M'\{V'/x\} &= M_0'\{V_0'/y\}\{V'/x\}M_1'\{V'/x\} \dots M_m'\{V'/x\} \\ &= M_0'\{V'/x\}\{V_0'\{V'/x\}/y\}M_1'\{V'/x\} \dots M_m'\{V'/x\}. \end{aligned}$$
□

The following lemma will play a crucial role in the proof of Lemmas 3.15-3.16 and shows that head σ-reduction $\xrightarrow{h}_{\sigma}$ can be postponed to head β_v-reduction $\xrightarrow{h}_{\beta_v}$.

Lemma 3.11 (Commutation of head reductions)**.**

(1) If $M \xrightarrow{h}_{\sigma} L \xrightarrow{h}_{\beta_v} N$ *then there exists* $L' \in \Lambda$ *such that* $M \xrightarrow{h}_{\beta_v} L' \xrightarrow{h}{}_{\sigma}^{=} N$*.*
(2) If $M \xrightarrow{h}{}_{\sigma}^{*} L \xrightarrow{h}{}_{\beta_v}^{*} N$ *then there exists* $L' \in \Lambda$ *such that* $M \xrightarrow{h}{}_{\beta_v}^{*} L' \xrightarrow{h}{}_{\sigma}^{*} N$*.*
(3) If $M \xrightarrow{h}{}_{\mathsf{v}}^{*} M'$ *then there exists* $N \in \Lambda$ *such that* $M \xrightarrow{h}{}_{\beta_v}^{*} N \xrightarrow{h}{}_{\sigma}^{*} M'$*.*

Proof. (1) By induction on the derivation of $M \xrightarrow{h}_{\sigma} L$. Let us consider its last rule r.

- If $\mathsf{r} = \sigma_1$ then $M = (\lambda x.M_0)N_0L_0M_1 \dots M_m$ and $L = (\lambda x.M_0L_0)N_0M_1 \dots M_m$ where $m \in \mathbb{N}$ and $x \notin \mathsf{fv}(L_0)$. Since $L \xrightarrow{h}_{\beta_v} N$, there are only two cases:
    - either $N_0 \xrightarrow{h}_{\beta_v} N_0'$ and $N = (\lambda x.M_0L_0)N_0'M_1 \dots M_m$ (according to the rule *right* for $\xrightarrow{h}_{\beta_v}$), then $M \xrightarrow{h}_{\beta_v} (\lambda x.M_0)N_0'L_0M_1 \dots M_m \xrightarrow{h}_{\sigma} N$;
    - or $N_0 \in \Lambda_v$ and $N = M_0\{N_0/x\}L_0M_1 \dots M_m$ (according to the rule β_v for $\xrightarrow{h}_{\beta_v}$, since $x \notin \mathsf{fv}(L_0)$), therefore $M \xrightarrow{h}_{\beta_v} N$.
- If $\mathsf{r} = \sigma_3$ then $M = V((\lambda x.L_0)N_0)M_1 \dots M_m$ and $L = (\lambda x.VL_0)N_0M_1 \dots M_m$ with $m \in \mathbb{N}$ and $x \notin \mathsf{fv}(V)$. Since $L \xrightarrow{h}_{\beta_v} N$, there are only two cases:
    - either $N_0 \xrightarrow{h}_{\beta_v} N_0'$ and $N = (\lambda x.VL_0)N_0'M_1 \dots M_m$ (according to the rule *right* for $\xrightarrow{h}_{\beta_v}$), then $M \xrightarrow{h}_{\beta_v} V((\lambda x.L_0)N_0')M_1 \dots M_m \xrightarrow{h}_{\sigma} N$;
    - or $N_0 \in \Lambda_v$ and $N = VL_0\{N_0/x\}M_1 \dots M_m$ (according to the rule β_v for $\xrightarrow{h}_{\beta_v}$, because $x \notin \mathsf{fv}(V)$), so $M \xrightarrow{h}_{\beta_v} N$.
- Finally, if $\mathsf{r} = \textit{right}$ then $M = VN_0M_1 \dots M_m$ and $L = VN_0'M_1 \dots M_m$ with $m \in \mathbb{N}$ and $N_0 \xrightarrow{h}_{\sigma} N_0'$. By Lemma 3.8.2, $N_0' \notin \Lambda_v$ and thus, since $L \xrightarrow{h}_{\beta_v} N$, the only possibility is that $N_0' \xrightarrow{h}_{\beta_v} N_0''$ and $N = VN_0''M_1 \dots M_m$ (according to the rule *right* for $\xrightarrow{h}_{\beta_v}$). By induction hypothesis, there exists $N_0''' \in \Lambda$ such that $N_0 \xrightarrow{h}_{\beta_v} N_0''' \xrightarrow{h}{}_{\sigma}^{=} N_0''$. Therefore, $M \xrightarrow{h}_{\beta_v} VN_0'''M_1 \dots M_m \xrightarrow{h}{}_{\sigma}^{=} N$.

(2) By hypothesis, there exist $m, n \in \mathbb{N}$ and $M_0, \dots, M_m, N_0, \dots, N_n \in \Lambda$ such that $M = M_0 \xrightarrow{h}_{\sigma} \dots \xrightarrow{h}_{\sigma} M_m = L = N_0 \xrightarrow{h}_{\beta_v} \dots \xrightarrow{h}_{\beta_v} N_n = N$. We prove by induction on $m \in \mathbb{N}$ that $M \xrightarrow{h}{}_{\beta_v}^{*} L' \xrightarrow{h}{}_{\sigma}^{*} N$ for some $L' \in \Lambda$.

- If $m = 0$ (resp. $n = 0$) then we conclude by taking $L' = N$ (resp. $L' = M$).
- Suppose $m, n > 0$: by applying Lemma 3.11.1 at most n times, there exist $N'_0, \dots, N'_{n-1} \in \Lambda$ such that $M_{m-1} \overset{h}{\to}_{\beta_v} N'_0 \overset{h}{\to}_{\beta_v}^{=} \dots \overset{h}{\to}_{\beta_v}^{=} N'_{n-1} \overset{h}{\to}_{\mathsf{v}}^{=} N$. By induction hypothesis (applied to $M = M_0 \overset{h}{\to}_{\sigma} \dots \overset{h}{\to}_{\sigma} M_{m-1} \overset{h}{\to}_{\beta_v}^{*} N$ or $M = M_0 \overset{h}{\to}_{\sigma} \dots \overset{h}{\to}_{\sigma} M_{m-1} \overset{h}{\to}_{\beta_v}^{*} N'_{n-1}$ depending on whether $N'_{n-1} \overset{h}{\to}_{\beta_v} N$ or $N'_{n-1} \overset{h}{\to}_{\sigma}^{=} N$, respectively), there exists $L' \in \Lambda$ such that $M \overset{h}{\to}_{\beta_v}^{*} L' \overset{h}{\to}_{\sigma}^{*} N$.

(3) By hypothesis, there exist $n \in \mathbb{N}$ and $L, M_1, N_1, \dots, M_n, N_n \in \Lambda$ such that $M \overset{h}{\to}_{\beta_v}^{*} L \overset{h}{\to}_{\sigma}^{+} M_1 \overset{h}{\to}_{\beta_v}^{+} N_1 \overset{h}{\to}_{\sigma}^{+} \dots \overset{h}{\to}_{\beta_v}^{+} N_{n-1} \overset{h}{\to}_{\sigma}^{+} M_n \overset{h}{\to}_{\beta_v}^{+} N_n \overset{h}{\to}_{\sigma}^{*} M'$ (i.e. n is the number of subsequences of the shape $\overset{h}{\to}_{\sigma}\overset{h}{\to}_{\beta_v}$ in the head v-reduction sequence from M to M'). We prove by induction on $n \in \mathbb{N}$ that $M \overset{h}{\to}_{\beta_v}^{*} N \overset{h}{\to}_{\sigma}^{*} M'$ for some $N \in \Lambda$.

- If $n = 0$ then $M \overset{h}{\to}_{\beta_v}^{*} L \overset{h}{\to}_{\sigma}^{*} M'$ and hence we conclude by taking $N = L$.
- Suppose $n > 0$. By applying the induction hypothesis to the head v-reduction sequence from M to M_n, $M \overset{h}{\to}_{\beta_v}^{*} N' \overset{h}{\to}_{\sigma}^{*} M_n \overset{h}{\to}_{\beta_v}^{+} N_n \overset{h}{\to}_{\sigma}^{*} M'$ for some $N' \in \Lambda$. By Lemma 3.11.2, $M \overset{h}{\to}_{\beta_v}^{*} N' \overset{h}{\to}_{\beta_v}^{*} N \overset{h}{\to}_{\sigma}^{*} N_n \overset{h}{\to}_{\sigma}^{*} M'$ for some $N \in \Lambda$. □

We are now ready to retrace Takahashi's method [Tak95] in our setting with β_v- and σ-reductions. The next four lemmas govern strong parallel reduction and will be used to prove Lemma 3.16, the key lemma stating that $\Rightarrow$ can be "sequentialized" according to $\Rrightarrow$.

Lemma 3.12. *If* $M \Rrightarrow M'$ *and* $N \Rightarrow N'$ *and* $M' \notin \Lambda_v$, *then* $MN \Rrightarrow M'N'$.

Proof. From the definition of $M \Rrightarrow M'$ it follows that $M \Rightarrow M'$ and $M \overset{h}{\to}_{\beta_v}^{*} L \overset{h}{\to}_{\sigma}^{*} L' \overset{\mathit{int}}{\Rightarrow} M'$ for some $L, L' \in \Lambda$. Hence, $MN \Rightarrow M'N'$ by Lemma 3.9.1, and $MN \overset{h}{\to}_{\beta_v}^{*} LN \overset{h}{\to}_{\sigma}^{*} L'N$ by Lemma 3.9.2. Since $M' \notin \Lambda_v$, $L'N \overset{\mathit{int}}{\Rightarrow} M'N'$ by Lemma 3.9.3. Therefore, $MN \Rrightarrow M'N'$. □

Lemma 3.13 (Applicative closure of $\Rrightarrow$)**.** *If* $M \Rrightarrow M'$ *and* $N \Rrightarrow N'$ *then* $MN \Rrightarrow M'N'$.

Proof. If $M' \notin \Lambda_v$ then $MN \Rrightarrow M'N'$ by Lemma 3.12, since $N \Rrightarrow N'$ implies $N \Rightarrow N'$.

Assume $M' \in \Lambda_v$: $MN \Rightarrow M'N'$ by Lemma 3.9.1, since $M \Rightarrow M'$ and $N \Rightarrow N'$. By hypothesis, there are $M_0, M'_0, N_0, N'_0 \in \Lambda$ such that $M \overset{h}{\to}_{\beta_v}^{*} M_0 \overset{h}{\to}_{\sigma}^{*} M'_0 \overset{\mathit{int}}{\Rightarrow} M'$ and $N \overset{h}{\to}_{\beta_v}^{*} N_0 \overset{h}{\to}_{\sigma}^{*} N'_0 \overset{\mathit{int}}{\Rightarrow} N'$. By Lemma 3.8.3, $M'_0 \in \Lambda_v$ since $M' \in \Lambda_v$, thus $M_0 = M'_0$ by Lemma 3.8.2 (and $M_0 \Rightarrow M'$ since $\overset{\mathit{int}}{\Rightarrow} \subseteq \Rightarrow$). Since $M_0 \in \Lambda_v$, using the rules *right* for $\overset{h}{\to}_{\beta_v}$ and $\overset{h}{\to}_{\sigma}$, we have $M_0 N \overset{h}{\to}_{\beta_v}^{*} M_0 N_0$ and $M_0 N_0 \overset{h}{\to}_{\sigma}^{*} M_0 N'_0$. By Lemma 3.9.2, $MN \overset{h}{\to}_{\beta_v}^{*} M_0 N$. By applying the rule *right* for $\overset{\mathit{int}}{\Rightarrow}$, we have $M_0 N'_0 \overset{\mathit{int}}{\Rightarrow} M'N'$. Therefore, $MN \overset{h}{\to}_{\beta_v}^{*} M_0 N \overset{h}{\to}_{\beta_v}^{*} M_0 N_0 \overset{h}{\to}_{\sigma}^{*} M_0 N'_0 \overset{\mathit{int}}{\Rightarrow} M'N'$ and hence $MN \Rrightarrow M'N'$. □

Lemma 3.14 (Substitution vs. $\overset{\mathit{int}}{\Rightarrow}$)**.** *If* $M \overset{\mathit{int}}{\Rightarrow} M'$ *and* $V \overset{\mathit{int}}{\Rightarrow} V'$ *then* $M\{V/x\} \overset{\mathit{int}}{\Rightarrow} M'\{V'/x\}$.

Proof. By induction on $M \in \Lambda$. Let us consider the last rule r of the derivation of $M \overset{\mathit{int}}{\Rightarrow} M'$.

- If $\mathsf{r} = \mathit{var}$ then $M = M'$ and there are only two cases: either $M = x$ and then $M\{V/x\} = V \overset{\mathit{int}}{\Rightarrow} V' = M'\{V'/x\}$; or $M = y \neq x$ and then $M\{V/x\} = y = M'\{V'/x\}$, therefore $M\{V/x\} \overset{\mathit{int}}{\Rightarrow} M'\{V'/x\}$ by reflexivity of $\overset{\mathit{int}}{\Rightarrow}$ (Lemma 3.7).
- If $\mathsf{r} = \lambda$ then $M = \lambda y.N$ and $M' = \lambda y.N'$ with $N \Rightarrow N'$; we can suppose without loss of generality that $y \notin \mathsf{fv}(V) \cup \{x\}$. We have $N\{V/x\} \Rightarrow N'\{V'/x\}$ according to Lemma 3.10, since $V \overset{\mathit{int}}{\Rightarrow} V'$ implies $V \Rightarrow V'$ (Lemma 3.8.5). By applying the rule λ for $\overset{\mathit{int}}{\Rightarrow}$, we have $M\{V/x\} = \lambda y.N\{V/x\} \overset{\mathit{int}}{\Rightarrow} \lambda y.N'\{V'/x\} = M'\{V'/x\}$.

- Finally, if $\mathsf{r} = \mathit{right}$ then $M = UNM_1 \dots M_m$ and $M' = U'N'M'_1 \dots M'_m$ for some $m \in \mathbb{N}$ with $U, U' \in \Lambda_v$ such that $U \Rightarrow U'$, $N \overset{\mathit{int}}{\Rrightarrow} N'$ and $M_i \Rightarrow M'_i$ for any $1 \leq i \leq m$. By induction hypothesis, $N\{V/x\} \overset{\mathit{int}}{\Rrightarrow} N'\{V'/x\}$. By Lemma 3.10, $U\{V/x\} \Rightarrow U'\{V'/x\}$ and $M_i\{V/x\} \Rightarrow M'_i\{V'/x\}$ for any $1 \leq i \leq m$, since $V \overset{\mathit{int}}{\Rrightarrow} V'$ implies $V \Rightarrow V'$ (Lemma 3.8.5). By applying the rule *right* for $\overset{\mathit{int}}{\Rrightarrow}$, we have
$$\begin{aligned} M\{V/x\} &= U\{V/x\}N\{V/x\}M_1\{V/x\} \dots M_m\{V/x\} \\ &\overset{\mathit{int}}{\Rrightarrow} U'\{V'/x\}N'\{V'/x\}M'_1\{V'/x\} \dots M'_m\{V'/x\} = M'\{V'/x\}. \end{aligned}$$
$\square$

Lemma 3.14 is used to prove the following substitution lemma for $\Rrightarrow$.

Lemma 3.15 (Substitution vs. $\Rrightarrow$)**.** *If $M \Rrightarrow M'$ and $V \Rrightarrow V'$ then $M\{V/x\} \Rrightarrow M'\{V'/x\}$.*

Proof. According to Lemma 3.10, $M\{V/x\} \Rightarrow M'\{V'/x\}$ since $M \Rightarrow M'$ and $V \Rightarrow V'$. By hypothesis, there exist $N, L \in \Lambda$ such that $M \xrightarrow{h}{}^{*}_{\beta_v} L \xrightarrow{h}{}^{*}_{\sigma} N \overset{\mathit{int}}{\Rrightarrow} M'$. By Lemma 3.9.8, $M\{V/x\} \xrightarrow{h}{}^{*}_{\beta_v} L\{V/x\}$ and $L\{V/x\} \xrightarrow{h}{}^{*}_{\sigma} N\{V/x\}$. By Lemma 3.14 (since $V \Rrightarrow V'$ implies $V \overset{\mathit{int}}{\Rrightarrow} V'$ according to Lemma 3.8.5), we have $N\{V/x\} \overset{\mathit{int}}{\Rrightarrow} M'\{V'/x\}$, thus $M\{V/x\} \xrightarrow{h}{}^{*}_{\beta_v} L\{V/x\} \xrightarrow{h}{}^{*}_{\sigma} N\{V/x\} \overset{\mathit{int}}{\Rrightarrow} M'\{V'/x\}$ and therefore $M\{V/x\} \Rrightarrow M'\{V'/x\}$. $\square$

Now we prove a key lemma, stating that parallel reduction $\Rightarrow$ coincides with strong parallel reduction $\Rrightarrow$ (the inclusion $\Rrightarrow \subseteq \Rightarrow$ holds trivially by definition of $\Rrightarrow$). In its proof, as well as in the proof of Corollary 3.18 and Theorem 3.4, our Lemma 3.11 plays a crucial role: indeed, since head σ-reduction well interacts with head β_v-reduction, Takahashi's method [Tak95] is still working when adding the reduction rules σ_1 and σ_3 to β_v-reduction.

Lemma 3.16 (Key Lemma)**.** *If $M \Rightarrow M'$ then $M \Rrightarrow M'$.*

Proof. By induction on the derivation of $M \Rightarrow M'$. Let us consider its last rule r.

- If $\mathsf{r} = \mathit{var}$ then $M = x\,M_1 \dots M_m$ and $M' = x\,M'_1 \dots M'_m$ where $m \in \mathbb{N}$ and $M_i \Rightarrow M'_i$ for all $1 \leq i \leq m$. By reflexivity of $\Rrightarrow$ (Lemma 3.7), $x \Rrightarrow x$. By induction hypothesis, $M_i \Rrightarrow M'_i$ for all $1 \leq i \leq m$. Therefore, $M \Rrightarrow M'$ by applying Lemma 3.13 m times.
- If $\mathsf{r} = \lambda$ then $M = (\lambda x.M_0)M_1 \dots M_m$ and $M' = (\lambda x.M'_0)M'_1 \dots M'_m$ where $m \in \mathbb{N}$ and $M_i \Rightarrow M'_i$ for all $0 \leq i \leq m$. By induction hypothesis, $M_i \Rrightarrow M'_i$ for all $1 \leq i \leq m$. According to Lemma 3.8.4, $\lambda x.M_0 \Rrightarrow \lambda x.M'_0$. So, $M \Rrightarrow M'$ by applying Lemma 3.13 m times.
- If $\mathsf{r} = \beta_v$ then $M = (\lambda x.M_0)VM_1 \dots M_m$ and $M' = M'_0\{V'/x\}M'_1 \dots M'_m$ where $m \in \mathbb{N}$, $V \Rightarrow V'$ and $M_i \Rightarrow M'_i$ for all $0 \leq i \leq m$. By induction hypothesis, $V \Rrightarrow V'$ and $M_i \Rrightarrow M'_i$ for all $0 \leq i \leq m$. Moreover, $M_0\{V/x\}M_1 \dots M_m \Rrightarrow M'$ by Lemma 3.15 and by applying Lemma 3.13 m times, thus $M_0\{V/x\}M_1 \dots M_m \xrightarrow{h}{}^{*}_{\beta_v} L \xrightarrow{h}{}^{*}_{\sigma} N \overset{\mathit{int}}{\Rrightarrow} M'$ for some $L, N \in \Lambda$. Therefore, $M \Rrightarrow M'$ since $M \xrightarrow{h}_{\beta_v} M_0\{V/x\}M_1 \dots M_m$.
- If $\mathsf{r} = \sigma_1$ then $M = (\lambda x.M_0)N_0L_0M_1 \dots M_m$ and $M' = (\lambda x.M'_0L'_0)N'_0M'_1 \dots M'_m$ where $m \in \mathbb{N}$, $L_0 \Rightarrow L'_0$, $N_0 \Rightarrow N'_0$ and $M_i \Rightarrow M'_i$ for any $0 \leq i \leq m$. By induction hypothesis, $N_0 \Rrightarrow N'_0$ and $M_i \Rrightarrow M'_i$ for any $1 \leq i \leq m$. By applying the rule σ_1 for $\xrightarrow{h}_{\sigma}$, we have $M \xrightarrow{h}_{\sigma} (\lambda x.M_0L_0)N_0M_1 \dots M_m$. By Lemma 3.9.1, $M_0L_0 \Rightarrow M'_0L'_0$ and thus $\lambda x.M_0L_0 \Rrightarrow \lambda x.M'_0L'_0$ according to Lemma 3.8.4. So $(\lambda x.M_0L_0)N_0M_1 \dots M_m \Rrightarrow M'$ by applying Lemma 3.13 $m+1$ times, hence there

are $L, N \in \Lambda$ such that $M \xrightarrow{h}_{\sigma} (\lambda x.M_0L_0)N_0M_1 \ldots M_m \xrightarrow{h}{}^{*}_{\beta_v} L \xrightarrow{h}{}^{*}_{\sigma} N \overset{\mathit{int}}{\Rrightarrow} M'$. By Lemma 3.11.2, there is $L' \in \Lambda$ such that $M \xrightarrow{h}{}^{*}_{\beta_v} L' \xrightarrow{h}{}^{*}_{\sigma} L \xrightarrow{h}{}^{*}_{\sigma} N \overset{\mathit{int}}{\Rrightarrow} M'$, so $M \Rrightarrow M'$.

- Finally, if $\mathsf{r} = \sigma_3$ then $M = V((\lambda x.L_0)N_0)N_1 \ldots N_n$ and $M' = (\lambda x.V'L_0')N_0'N_1' \ldots N_n'$ with $n \in \mathbb{N}$, $V \Rightarrow V'$, $L_0 \Rightarrow L_0'$ and $N_i \Rightarrow N_i'$ for any $0 \leq i \leq n$. By induction hypothesis, $N_i \Rrightarrow N_i'$ for any $0 \leq i \leq n$. By the rule σ_3 for $\xrightarrow{h}_{\sigma}$, we have $M \xrightarrow{h}_{\sigma} (\lambda x.VL_0)N_0N_1 \ldots N_n$. By Lemma 3.9.1, $VL_0 \Rightarrow V'L_0'$ and thus $\lambda x.VL_0 \Rrightarrow \lambda x.V'L_0'$ according to Lemma 3.8.4. So $(\lambda x.VL_0)N_0N_1 \ldots N_n \Rrightarrow M'$ by applying Lemma 3.13 $n+1$ times, hence there are $L, N \in \Lambda$ such that $M \xrightarrow{h}_{\sigma} (\lambda x.VL_0)N_0N_1 \ldots N_n \xrightarrow{h}{}^{*}_{\beta_v} L \xrightarrow{h}{}^{*}_{\sigma} N \overset{\mathit{int}}{\Rrightarrow} M'$. By Lemma 3.11.2, there is $L' \in \Lambda$ such that $M \xrightarrow{h}{}^{*}_{\beta_v} L' \xrightarrow{h}{}^{*}_{\sigma} L \xrightarrow{h}{}^{*}_{\sigma} N \overset{\mathit{int}}{\Rrightarrow} M'$, therefore $M \Rrightarrow M'$. □

Next Lemma 3.17 and Corollary 3.18 show that internal parallel reduction can be shifted after head v-reduction.

Lemma 3.17 (Postponement, version 1). *If $M \overset{\mathit{int}}{\Rrightarrow} L$ and $L \xrightarrow{h}_{\beta_v} N$ (resp. $L \xrightarrow{h}_{\sigma} N$) then there exists $L' \in \Lambda$ such that $M \xrightarrow{h}_{\beta_v} L'$ (resp. $M \xrightarrow{h}_{\sigma} L'$) and $L' \Rightarrow N$.*

Proof. By induction on the derivation of $M \overset{\mathit{int}}{\Rrightarrow} L$. Let us consider its last rule r.

- If $\mathsf{r} = \mathit{var}$, then $M = x = L$ which contradicts $L \xrightarrow{h}_{\beta_v} N$ and $L \xrightarrow{h}_{\sigma} N$ by Lemmas 3.8.1-2.
- If $\mathsf{r} = \lambda$ then $L = \lambda x.L'$ for some $L' \in \Lambda$, which contradicts $L \xrightarrow{h}_{\beta_v} N$ and $L \xrightarrow{h}_{\sigma} N$ by Lemmas 3.8.1-2.
- Finally, if $\mathsf{r} = \mathit{right}$ then $M = VM_0M_1 \ldots M_m$ and $L = V'L_0L_1 \ldots L_m$ where $m \in \mathbb{N}$, $V \Rightarrow V'$ (so $V \overset{\mathit{int}}{\Rrightarrow} V'$ by Lemma 3.8.5), $M_0 \overset{\mathit{int}}{\Rrightarrow} L_0$ (thus $M_0 \Rightarrow L_0$ since $\overset{\mathit{int}}{\Rrightarrow} \subseteq \Rightarrow$) and $M_i \Rightarrow L_i$ for any $1 \leq i \leq m$.
  - If $L \xrightarrow{h}_{\beta_v} N$ then there are only two cases, depending on the last rule r' of the derivation of $L \xrightarrow{h}_{\beta_v} N$.
    - If $\mathsf{r}' = \beta_v$ then $V' = \lambda x.N_0'$, $L_0 \in \Lambda_v$ and $N = N_0'\{L_0/x\}L_1 \ldots L_m$, thus $M_0 \in \Lambda_v$ and $V = \lambda x.N_0$ with $N_0 \Rightarrow N_0'$ by Lemma 3.8.3. By Lemma 3.10, $N_0\{M_0/x\} \Rightarrow N_0'\{L_0/x\}$. Let $L' = N_0\{M_0/x\}M_1 \ldots M_m$: so $M = (\lambda x.N_0)M_0M_1 \ldots M_m \xrightarrow{h}_{\beta_v} L'$ (apply the rule β_v for $\xrightarrow{h}_{\beta_v}$) and $L' \Rightarrow N$ by applying Lemma 3.9.1 m times.
    - If $\mathsf{r}' = \mathit{right}$ then $N = V'N_0L_1 \ldots L_m$ with $L_0 \xrightarrow{h}_{\beta_v} N_0$. By induction hypothesis, there exists $L_0' \in \Lambda$ such that $M_0 \xrightarrow{h}_{\beta_v} L_0' \Rightarrow N_0$. Let $L' = VL_0'M_1 \ldots M_m$: so $M \xrightarrow{h}_{\beta_v} L'$ (apply the rule *right* for $\xrightarrow{h}_{\beta_v}$) and $L' \Rightarrow N$ by applying Lemma 3.9.1 $m+1$ times.
  - If $L \xrightarrow{h}_{\sigma} N$ then there are only three cases, depending on the last rule r' of the derivation of $L \xrightarrow{h}_{\sigma} N$.
    - If $\mathsf{r}' = \sigma_1$ then $m > 0$, $V' = \lambda x.N_0'$ and $N = (\lambda x.N_0'L_1)L_0L_2 \ldots L_m$, thus $V = \lambda x.N_0$ with $N_0 \Rightarrow N_0'$ by Lemma 3.8.3. Using Lemmas 3.9.1 and 3.8.4, we have $\lambda x.N_0M_1 \Rightarrow \lambda x.N_0'L_1$. Let $L' = (\lambda x.N_0M_1)M_0M_2 \ldots M_m$: so $M = (\lambda x.N_0)M_0M_1 \ldots M_m \xrightarrow{h}_{\sigma} L'$ (apply the rule σ_1 for $\xrightarrow{h}_{\sigma}$) and $L' \Rightarrow N$ by applying Lemma 3.9.1 m times.

- $*$ If $\mathsf{r}' = \sigma_3$ then $L_0 = (\lambda x.L_{01})L_{02}$ and $N = (\lambda x.V'L_{01})L_{02}L_1 \dots L_m$. Since $M_0 \overset{int}{\Rightarrow} (\lambda x.L_{01})L_{02}$, then by simple inspection of the rules for $\overset{int}{\Rightarrow}$ (Definition 3.6) we infer that $M_0 = V_0 M_{02}$ with $V_0 \Rightarrow \lambda x.L_{01}$ and $M_{02} \overset{int}{\Rightarrow} L_{02}$ (so $M_{02} \Rightarrow L_{02}$ because $\overset{int}{\Rightarrow} \subseteq \Rightarrow$). By Lemmas 3.8.5 and 3.8.3, from $V_0 \Rightarrow \lambda x.L_{01}$ it follows that $V_0 = \lambda x.M_{01}$ with $M_{01} \Rightarrow L_{01}$. By Lemmas 3.9.1 and 3.8.4, $\lambda x.VM_{01} \Rightarrow \lambda x.V'L_{01}$. Let $L' = (\lambda x.VM_{01})M_{02}M_1 \dots M_m$: so $M = V((\lambda x.M_{01})M_{02})M_1 \dots M_m \xrightarrow{h}_{\sigma} L'$ (apply the rule σ_3 for $\xrightarrow{h}_{\sigma}$) and $L' \Rightarrow N$ by applying Lemma 3.9.1 $m+1$ times.
- $*$ If $\mathsf{r}' = \mathit{right}$ then $N = V'N_0L_1 \dots L_m$ with $L_0 \xrightarrow{h}_{\sigma} N_0$. By induction hypothesis, there exists $L_0' \in \Lambda$ such that $M_0 \xrightarrow{h}_{\sigma} L_0' \Rightarrow N_0$. Let $L' = VL_0'M_1 \dots M_m$: so $M \xrightarrow{h}_{\sigma} L'$ (apply the rule *right* for $\xrightarrow{h}_{\sigma}$) and $L' \Rightarrow N$ by applying Lemma 3.9.1 $m+1$ times. $\square$

Corollary 3.18 (Postponement, version 2)**.** *If* $M \overset{int}{\Rightarrow} L$ *and* $L \xrightarrow{h}_{\beta_v} N$ *(resp.* $L \xrightarrow{h}_{\sigma} N$*), then there exist* $L', L'' \in \Lambda$ *such that* $M \xrightarrow{h}{}^{+}_{\beta_v} L' \xrightarrow{h}{}^{*}_{\sigma} L'' \overset{int}{\Rightarrow} N$ *(resp.* $M \xrightarrow{h}{}^{*}_{\beta_v} L' \xrightarrow{h}{}^{*}_{\sigma} L'' \overset{int}{\Rightarrow} N$*).*

Proof. Immediate from Lemmas 3.17 and 3.16, applying Lemma 3.11.2 if $L \xrightarrow{h}_{\sigma} N$. $\square$

Proof of Sequentialization (Theorem 3.4 on page 8).
By Lemma 3.9.4, $M \Rightarrow^* M'$ and thus there are $m \in \mathbb{N}$ and $M_0, \dots, M_m \in \Lambda$ such that $M = M_0$, $M_m = M'$ and $M_i \Rightarrow M_{i+1}$ for any $0 \leq i < m$. We prove by induction on $m \in \mathbb{N}$ that there are $L, N \in \Lambda$ such that $M \xrightarrow{h}{}^{*}_{\beta_v} L \xrightarrow{h}{}^{*}_{\sigma} N \overset{int}{\Rightarrow}{}^{*} M'$, so $N \xrightarrow{int}{}^{*}_{\mathsf{v}} M'$ by Lemma 3.9.5.

- If $m = 0$ then $M = M_0 = M'$ and hence we conclude by taking $L = M' = N$.
- Suppose $m > 0$. By induction hypothesis applied to $M_1 \Rightarrow^* M'$, there are $L', N' \in \Lambda$ such that $M_1 \xrightarrow{h}{}^{*}_{\beta_v} L' \xrightarrow{h}{}^{*}_{\sigma} N' \overset{int}{\Rightarrow}{}^{*} M'$. By applying Lemma 3.16 to $M \Rightarrow M_1$, there exist $L_0, N_0 \in \Lambda$ such that $M \xrightarrow{h}{}^{*}_{\beta_v} L_0 \xrightarrow{h}{}^{*}_{\sigma} N_0 \overset{int}{\Rightarrow} M_1$. By applying Corollary 3.18 repeatedly, there is $N \in \Lambda$ such that $N_0 \xrightarrow{h}{}^{*}_{\mathsf{v}} N \overset{int}{\Rightarrow} N'$, and hence $M \xrightarrow{h}{}^{*}_{\mathsf{v}} N \overset{int}{\Rightarrow}{}^{*} M'$. According to Lemma 3.11.3, there is $L \in \Lambda$ such that $M \xrightarrow{h}{}^{*}_{\beta_v} L \xrightarrow{h}{}^{*}_{\sigma} N \overset{int}{\Rightarrow}{}^{*} M'$. $\square$

4. Standardization

This section is devoted to prove the standardization theorem for λ_v^σ, stating that if $N \to_{\mathsf{v}}^* L$ then there is a "standard" v-reduction sequence from N to L (Theorem 4.6). Roughly speaking, a reduction sequence is *standard* if the "positions" of the reduced redexes move from left to right.[1] Actually, in a call-by-value λ-calculus (such as λ_v^σ), this "left-to-right" order is more delicate to define, since β-redexes can be fired only after their arguments have been reduced to a value,[2] but the essence is the same: a standard reduction sequence begins with head reduction steps, and then continues with internal reduction steps selecting redexes according to a "left-to-right" order. Our choice to prioritize head reduction over internal

---

[1] In ordinary λ-calculus, standard sequences (for β-reduction) can be described as follows: "After each contraction of a redex R, index the λ's of redexes to the left of R. Redexes with indexed λ's are not allowed to be contracted anymore. Indexed λ's remain indexed after contractions of other redexes" [Bar84, p. 297].

[2] E.g., according to [Bar84, Cra09] (and us), in λ_v the β_v-reduction sequence $\Delta(II) \xrightarrow{h}_{\beta_v} \Delta I \xrightarrow{h}_{\beta_v} II$ is standard, even if the (only) β_v-redex in $\Delta(II)$ seems to be "on the right" of the β_v-redex ΔI reduced later. The subtlety is that in λ_v, unlike λ, new redexes can be created in the following way: a β_v-reduction step in the argument of a β-(not β_v-)redex R may turn the argument itself into a value, turning R into a β_v-redex.

reduction (followed also by [Plo75, Cra09] for λ_v) entails that in a standard sequence some changes of positions from right to left for the selected redexes may take place when passing from the head reduction phase to the internal reduction one: e.g. according to [Plo75, Cra09] (and us), in λ_v the sequence $(\lambda x.I\Delta)(y(II)) \overset{h}{\to}_{\beta_v} (\lambda x.I\Delta)(zI) \to_{\beta_v} (\lambda x.\Delta)(zI)$ is standard, even if the β_v-redex II is "on the right" of the β_v-redex $I\Delta$ (fired later as internal, since it is under the scope of a λ). Actually, in λ_v^σ another intricacy arises in defining a standard order: there are not only β-redexes but also σ-redexes and they may overlap. Our approach is to prioritize head β_v-redexes over head σ-redexes (this idea extends iteratively to subterms).

We define the notion of standard reduction sequence by closely following the approach used in [Plo75, Cra09], so the redex-order is defined by induction on the structure of terms, without involving any (tricky) notion of residual redex.

Definition 4.1 (Standard head sequence)**.** For any $k, m \in \mathbb{N}$ with $k \leq m$, a *standard head sequence*, denoted by $\lceil M_0, \dots, M_k, \dots, M_m \rceil^{head}$, is a finite sequence $(M_0, \dots, M_k, \dots, M_m)$ of terms such that $M_i \overset{h}{\to}_{\beta_v} M_{i+1}$ for any $0 \leq i < k$, and $M_i \overset{h}{\to}_\sigma M_{i+1}$ for any $k \leq i < m$.

In other words, a standard head sequence is a head v-reduction sequence where the head β_v-reduction steps precede all the head σ-reduction steps, without any order between head σ_1- and head σ_3-reduction steps. Note that when $k = m$ (resp. $k = 0$) then the standard head sequence consists only of head β_v-(resp. head σ-)reduction steps. It is easy to check that $\lceil M \rceil^{head}$ for any $M \in \Lambda$ (apply Definition 4.1 with $m = 0$).

Using the above definition of standard head sequence, we define by mutual induction the notions of standard sequence and standard inner sequence of terms (Definition 4.2).

Definition 4.2 (Standard and standard inner sequences)**.** *Standard sequences* and *standard inner sequences* of terms, denoted by $\lceil N_0, \dots, N_n \rceil^{std}$ and $\lceil N_0, \dots, N_n \rceil^{in}$ respectively (with $n \in \mathbb{N}$ and $N_0, \dots, N_n \in \Lambda$), are defined by mutual induction as follows:

1. if $\lceil M_0, \dots, M_m \rceil^{head}$ and $\lceil M_m, \dots, M_{m+n} \rceil^{in}$, then $\lceil M_0, \dots, M_m, \dots, M_{m+n} \rceil^{std}$;
2. $\lceil M \rceil^{in}$, for any $M \in \Lambda$;
3. if $\lceil M_0, \dots, M_m \rceil^{std}$ then $\lceil \lambda z.M_0, \dots, \lambda z.M_m \rceil^{in}$;
4. if $\lceil V_0, \dots, V_h \rceil^{std}$ and $\lceil N_0, \dots, N_n \rceil^{in}$, then $\lceil V_0 N_0, \dots, V_h N_0, \dots, V_h N_n \rceil^{in}$ (where $V_0, \dots, V_h \in \Lambda_v$);
5. if $\lceil N_0, \dots, N_n \rceil^{in}$, $\lceil L_0, \dots, L_l \rceil^{std}$ and $N_0 \notin \Lambda_v$, then $\lceil N_0 L_0, \dots, N_n L_0, \dots, N_n L_l \rceil^{in}$.

Remark 4.3. It is easy to show (by mutual induction on the definition of standard and standard inner sequences) that, given $n \in \mathbb{N}$ and $N_0, \dots, N_n \in \Lambda$, if $\lceil N_0, \dots, N_n \rceil^{in}$ (resp. $\lceil N_0, \dots, N_n \rceil^{std}$) then $N_i \overset{int}{\to}_{\mathsf{v}} N_{i+1}$ (resp. $N_i \to_{\mathsf{v}} N_{i+1}$) for any $0 \leq i < n$.

In fact, the presence of a standard or standard inner sequence means that not only there is a v-reduction sequence or an internal v-reduction sequence, respectively, but also that this v-reduction sequence is performed selecting v-redexes according to the aforementioned "left-to-right" order, up to some intricacies already pointed out on p. 16. Indeed, in Definition 4.2, the rule (1) — the only one yielding standard sequences — says that standard sequences start by reducing first head β_v-redexes, then head σ-redexes and then internal v-redexes, where the head β_v-redex in a term is its (unique, if any) *leftmost-outermost* β_v-redex not under the scope of λ's, and head σ-redexes in a term are its (possibly not unique) *leftmost-outermost* σ-redexes not under the scope of λ's. Rules (4)-(5) in Definition 4.2 intuitively mean that the positions of the v-redexes reduced in a standard inner sequence move from left to right.

In order to give informative examples about standard and standard inner sequences, for any $\mathsf{r} \in \{\beta_v, \sigma_1, \sigma_3, \sigma\}$ we set $\overset{int}{\to}_{\mathsf{r}} = \to_{\mathsf{r}} \cap \overset{int}{\to}_{\mathsf{v}}$.

Example 4.4. Let $L = (\lambda y.Ix)(z(\Delta I))(II)$: one has that $L \xrightarrow{int}_{\beta_v} (\lambda y.Ix)(z(\Delta I))I \xrightarrow{int}_{\beta_v} (\lambda y.x)(z(\Delta I))I$ and $L \xrightarrow{h}_{\sigma_1} (\lambda y.Ix(II))(z(\Delta I)) \xrightarrow{h}_{\beta_v} (\lambda y.Ix(II))(z(II))$ are not standard sequences; but $L \xrightarrow{int}_{\beta_v} (\lambda y.Ix)(z(\Delta I))I$ and $L \xrightarrow{h}_{\beta_v} (\lambda y.Ix)(z(II))(II) \xrightarrow{h}_{\beta_v} (\lambda y.Ix)(zI)(II) \xrightarrow{h}_{\sigma_1} (\lambda y.Ix(II))(zI) \xrightarrow{int}_{\beta_v} (\lambda y.x(II))(zI) \xrightarrow{int}_{\beta_v} (\lambda y.xI)(zI)$ are standard sequences.

The next lemma states that standard head and inner sequences are standard sequences.

Lemma 4.5. *Given $n \in \mathbb{N}$, if $\lceil N_0, \dots, N_n \rceil^{in}$ (resp. $\lceil N_0, \dots, N_n \rceil^{head}$) then $\lceil N_0, \dots, N_n \rceil^{std}$.*

Proof. $\lceil N_0 \rceil^{head}$ (resp. $\lceil N_n \rceil^{in}$ by Definition 4.2.2), so $\lceil N_0, \dots, N_n \rceil^{std}$ by Definition 4.2.1. □

In particular, $\lceil N \rceil^{std}$ for any $N \in \Lambda$: apply Definition 4.2.2 and Lemma 4.5 for $n = 0$.

Note that the concatenation of two standard sequences is not standard, in general: take for instance a standard inner sequence followed by a standard head sequence.

For all $V_0, \dots, V_h \in \Lambda_v$, $\lceil V_0, \dots, V_h \rceil^{std}$ iff $\lceil V_0, \dots, V_h \rceil^{in}$: the left-to-right implication follows from the rule (1) of Definition 4.2 (the only one yielding standard sequences) and Remarks 3.8.1-2 ($\lceil V_0, \dots, V_h \rceil^{head}$ is impossible for $h > 0$); the converse holds by Lemma 4.5.

We can now state and prove the standardization theorem for λ_v^σ, one of the main result of this paper: if M v-reduces to M' then there exists a standard sequence from M to M'. The idea to build this standard sequence is to sequentialize (as stated in Theorem 3.4) the v-reduction sequence from M to M' iteratively according to a "left-to-right" order.

Theorem 4.6 (Standardization). *Let M and M' be terms.*

1. *If $M \xrightarrow{h}_{\mathsf{v}}^{*} M'$ then there is a standard head sequence $\lceil M, \dots, M' \rceil^{head}$.*
2. *If $M \xrightarrow{int}_{\mathsf{v}}^{*} M'$ then there is a standard inner sequence $\lceil M, \dots, M' \rceil^{in}$.*
3. *If $M \to_{\mathsf{v}}^{*} M'$ then there is a standard sequence $\lceil M, \dots, M' \rceil^{std}$.*

Proof. Theorem 4.6.3 is an immediate consequence of Theorems 4.6.1-2 and Theorem 3.4: indeed, if $M \to_{\mathsf{v}}^{*} M'$ then there is a term M'' such that $M \xrightarrow{h}_{\mathsf{v}}^{*} M'' \xrightarrow{int}_{\mathsf{v}}^{*} M'$ by sequentialization (Theorem 3.4), moreover $M \xrightarrow{h}_{\mathsf{v}}^{*} M''$ implies that there is a sequence $\lceil M, \dots, M'' \rceil^{head}$ by Theorem 4.6.1, and $M'' \xrightarrow{int}_{\mathsf{v}}^{*} M'$ implies that there is a sequence $\lceil M'', \dots, M' \rceil^{in}$ by Theorem 4.6.2. According to the rule (1) of Definition 4.2, $\lceil M, \dots, M'', \dots, M' \rceil^{std}$.

It remains to prove Theorems 4.6.1-2.

Now, Theorem 4.6.1 is exactly our Lemma 3.11.3, already proved.

Theorem 4.6.2 is proved by induction on $M' \in \Lambda$, using Theorem 4.6.1.

- If $M' = z$ then $M = z$ by Lemma 3.9.7, thus $\lceil z \rceil^{in}$ by the rule (2) of Definition 4.2.
- If $M' = \lambda z.L'$ then there is $L \in \Lambda$ such that $M = \lambda z.L$ and $L \to_{\mathsf{v}}^{*} L'$, by Lemma 3.9.7. By sequentialization (Theorem 3.4), there exists a term N such that $L \xrightarrow{h}_{\mathsf{v}}^{*} N \xrightarrow{int}_{\mathsf{v}}^{*} L'$. By Theorem 4.6.1, from $L \xrightarrow{h}_{\mathsf{v}}^{*} N$ it follows that there is a sequence $\lceil L, \dots, N \rceil^{head}$. By induction hypothesis applied to $N \xrightarrow{int}_{\mathsf{v}}^{*} L'$, there is a sequence $\lceil N, \dots, L' \rceil^{in}$. According to the rule (1) of Definition 4.2, $\lceil L, \dots, N, \dots, L' \rceil^{std}$. By the rule (3) of Definition 4.2, $\lceil \lambda z.L, \dots, \lambda z.N, \dots, \lambda z.L' \rceil^{in}$, that is $\lceil M, \dots, M' \rceil^{in}$.
- If $M' = N'L'$ then $M = NL$ for some $N, L \in \Lambda$ by Remark 2.7, since $\xrightarrow{int}_{\mathsf{v}}^{*} \subseteq \to_{\mathsf{v}}^{*}$ and $M' \notin \Lambda_v$. By Lemma 3.9.5, $NL \overset{int}{\Rightarrow}^{*} N'L'$; clearly, for each step of $\overset{int}{\Rightarrow}$ in $NL \overset{int}{\Rightarrow}^{*} N'L'$, the last rule of its derivation is an instance of the rule *right* for $\overset{int}{\Rightarrow}$ (the other rules deal with values, see Definition 3.6). There are two sub-cases.
  - If $N \in \Lambda_v$ then $N \Rightarrow^{*} N'$ and $L \overset{int}{\Rightarrow}^{*} L'$, so $N \to_{\mathsf{v}}^{*} N'$ and $L \xrightarrow{int}_{\mathsf{v}}^{*} L'$ by Lemmas 3.9.4-5. By sequentialization (Theorem 3.4), there is a term N'' such that

$N \overset{h}{\to}{}^*_{\mathsf{v}} N'' \overset{int}{\to}{}^*_{\mathsf{v}} N'$, and actually $N = N''$ by Lemmas 3.8.1-2 since N is a value; thus, $N \overset{int}{\to}{}^*_{\mathsf{v}} N'$. By induction hypothesis applied to $N \overset{int}{\to}{}^*_{\mathsf{v}} N'$ and $L \overset{int}{\to}{}^*_{\mathsf{v}} L'$, there are sequences $\lceil N, \dots, N' \rceil^{in}$ (hence $\lceil N, \dots, N' \rceil^{std}$ by Lemma 4.5) and $\lceil L, \dots, L' \rceil^{in}$. In particular, according to Remark 4.3, if $\lceil N, \dots, N' \rceil^{std} = (N_0, \dots, N_n)$ for some $n \in \mathbb{N}$ and $N_0, \dots, N_n \in \Lambda$ (with $N_0 = N$ and $N_n = N'$), then $N_i \to_{\mathsf{v}} N_{i+1}$ for all $0 \leq i < n$, and hence $N_0, \dots, N_n$ (that is, all the terms in $\lceil N, \dots, N' \rceil^{std}$) are values by Remark 2.7, since N_0 is a value. By applying the rule (4) of Definition 4.2, $\lceil NL, \dots, N'L, \dots, N'L' \rceil^{in}$.
- If $N \notin \Lambda_v$ (i.e. $N = VM_1 \dots M_m$ with $m > 0$, by Remark 2.2) then $N \overset{int}{\Rightarrow}{}^* N'$ and $L \Rightarrow^* L'$, so $N \overset{int}{\to}{}^*_{\mathsf{v}} N'$ and $L \to^*_{\mathsf{v}} L'$ by Lemmas 3.9.4-5. By sequentialization (Theorem 3.4), $L \overset{h}{\to}{}^*_{\mathsf{v}} L'' \overset{int}{\to}{}^*_{\mathsf{v}} L'$ for some term L''. By Theorem 4.6.1, there is a sequence $\lceil L, \dots, L'' \rceil^{head}$. By induction hypothesis applied to $N \overset{int}{\to}{}^*_{\mathsf{v}} N'$ and $L'' \overset{int}{\to}{}^*_{\mathsf{v}} L'$, there are sequences $\lceil N, \dots, N' \rceil^{in}$ and $\lceil L'', \dots, L' \rceil^{in}$. According to the rule (1) of Definition 4.2, $\lceil L, \dots, L'', \dots, L' \rceil^{std}$. By applying the rule (5) of Definition 4.2, $\lceil NL, \dots, N'L, \dots, N'L' \rceil^{in}$, that is $\lceil M, \dots, M' \rceil^{in}$. ☐

Theorem 4.6 gives only a *weak standardization*: it rearranges a v-reduction sequence from M to M' so as to obtain a standard sequence from M to M', but a standard sequence selects v-redexes following a *partial* (and not total, in general) *order* on v-redexes. Indeed, a standard sequence is not uniquely determined by its starting and end terms, and this is essentially due to two facts (exemplified by Examples 4.7-4.8, respectively):

(1) as already remarked on pp. 7-8, head σ-redexes may overlap and be incomparable;
(2) in a standard (head) sequence, there is no restriction on when ending a head β_v-reduction phase and beginning a head σ-reduction phase.

Example 4.7. The following σ-reduction sequences (fired σ-redexes are underlined)

$$\underline{I(\Delta I)}I \overset{h}{\to}_{\sigma_1} \underline{(\lambda x.xI)(\Delta I)} \overset{h}{\to}_{\sigma_3} (\lambda z.(\lambda x.xI)(zz))I \qquad \text{and}$$

$$\underline{I(\Delta I)}I \overset{h}{\to}_{\sigma_3} \underline{(\lambda z.I(zz))II} \overset{h}{\to}_{\sigma_1} (\lambda z.\underline{I(zz)I})I \overset{int}{\to}_{\sigma_1} (\lambda z.(\lambda x.xI)(zz))I$$

are both — different — standard sequences from $I(\Delta I)I$ to $(\lambda z.(\lambda x.xI)(zz))I$.

Example 4.8. The following head v-reduction sequences (fired v-redexes are underlined)

$$I(\underline{\Delta\Delta})I \overset{h}{\to}_{\beta_v} \underline{I(\Delta\Delta)I} \overset{h}{\to}_{\sigma_1} (\lambda x.xI)(\Delta\Delta) \qquad \text{and} \qquad \underline{I(\Delta\Delta)I} \overset{h}{\to}_{\sigma_1} (\lambda x.xI)(\Delta\Delta)$$

are both — different — standard (head) sequences from $I(\Delta\Delta)I$ to $(\lambda x.xI)(\Delta\Delta)$.

Finally, we compare our notion of standardization with that for Plotkin's λ_v given in [Plo75, p. 137] and [Cra09]. To make the comparison possible we neglect σ-reduction and we recall that $\overset{h}{\to}_{\beta_v}$ is exactly Plotkin's left-reduction [Plo75, p. 136]. As remarked in [HZ09, p. 149], both $(\lambda z.II)(II) \overset{int}{\to}_{\beta_v} (\lambda z.I)(II) \overset{h}{\to}_{\beta_v} (\lambda z.I)I$ and $(\lambda z.II)(II) \overset{h}{\to}_{\beta_v} (\lambda z.II)I \overset{int}{\to}_{\beta_v} (\lambda z.I)I$ are standard sequences from $(\lambda z.II)(II)$ to $(\lambda z.I)I$ according to [Plo75, Cra09]. However, only the second sequence is standard in our sense (our standardization restricted to $\to_{\beta_v}$ is exactly the parametric standardization of [PR04] for λ_v, which imposes a total order on β_v-redexes). Without the distinction in Definition 4.2 between standard and standard inner sequences, both the above sequences would be standard; indeed, [Plo75, Cra09] do not make this distinction and their standardization imposes only a partial order on β_v-redexes.

5. Conservativity

We now present our main contribution: the shuffling calculus λ_v^σ is a *conservative* extension of λ_v. To be precise, we will prove that λ_v^σ is sound with respect to the observational equivalence introduced by Plotkin in [Plo75] for λ_v (Corollary 5.4), and that the notions of potential valuability and solvability for λ_v, introduced in [PR99], coincide with the respective notions for λ_v^σ (Theorem 5.7). This justifies the idea that λ_v^σ is a useful tool for studying properties of λ_v, as stated in [CG14]. All these results can be proved using standardization for λ_v^σ. Actually, the following corollary of sequentialization (Theorem 3.4) is enough.

Corollary 5.1 (Reduction to a value)**.** *Let $M \in \Lambda$ and $V \in \Lambda_v$.*

1. *If $M \to_{\mathsf{v}}^* V$ then there exists $V' \in \Lambda_v$ such that $M \xrightarrow{h}{}^{*}_{\beta_v} V' \xrightarrow{int}{}^{*}_{\mathsf{v}} V$.*
2. *$M \xrightarrow{h}{}^{*}_{\beta_v} V$ if and only if $M \xrightarrow{h}{}^{*}_{\mathsf{v}} V$.*

Proof. (1) By sequentialization (Theorem 3.4), $M \xrightarrow{h}{}^{*}_{\beta_v} L \xrightarrow{h}{}^{*}_{\sigma} N \xrightarrow{int}{}^{*}_{\mathsf{v}} V$ for some $N, L \in \Lambda$. By Lemma 3.9.7, $N \in \Lambda_v$ and thus $L = N$ according to Lemma 3.8.2.

(2) $\Leftarrow$: By Lemma 3.11.3, $M \xrightarrow{h}{}^{*}_{\beta_v} N \xrightarrow{h}{}^{*}_{\sigma} V$ for some N, and $N = V$ by Lemma 3.8.2. $\Rightarrow$: Trivial, since $\xrightarrow{h}_{\beta_v} \subseteq \xrightarrow{h}_{\mathsf{v}}$. □

Let us recall the notion of observational equivalence introduced by Plotkin [Plo75] for λ_v. Informally, two terms are observationally equivalent if they can be substituted for each other in all contexts without observing any difference in their behaviour.

Definition 5.2 (Halting, observational equivalence)**.** Let $M \in \Lambda$.

- We say that (*the evaluation of*) M *halts* if there exists $V \in \Lambda_v$ such that $M \xrightarrow{h}{}^{*}_{\beta_v} V$.
- The (*call-by-value*) *observational equivalence* is an equivalence relation $\cong$ on Λ defined by: $M \cong N$ if, for every context $\mathtt{C}$, one has that $\mathtt{C}(\!|M|\!)$ halts iff $\mathtt{C}(\!|N|\!)$ halts.

Original Plotkin's definition of call-by-value observational equivalence [Plo75, p. 144] also requires that $\mathtt{C}(\!|M|\!)$ and $\mathtt{C}(\!|N|\!)$ are closed terms, according to the tradition identifying programs with closed terms. However, the two equivalences coincide.

Clearly, the notions of halting and observational equivalence can be defined also for λ_v^σ, using $\xrightarrow{h}_{\mathsf{v}}$ instead of $\xrightarrow{h}_{\beta_v}$ in Definition 5.2. But head σ-reduction plays no role neither in deciding the halting problem for evaluation (Corollary 5.1.1), nor in reaching a particular value (Corollary 5.1.2). Therefore, we can conclude that the notions of halting and observational equivalence in λ_v^σ *coincide* with those in λ_v, respectively.

Now we compare the equational theory of λ_v^σ with Plotkin's observational equivalence.

Theorem 5.3 (Adequacy of v-reduction)**.** *If $M \to_{\mathsf{v}}^* M'$ then: M halts iff M' halts.*

Proof. If M' halts then $M' \xrightarrow{h}{}^{*}_{\beta_v} V \in \Lambda_v$ and hence $M \to_{\mathsf{v}}^* M' \to_{\mathsf{v}}^* V$ since $\xrightarrow{h}_{\beta_v} \subseteq \to_{\mathsf{v}}$. By Corollary 5.1.1, there exists $V' \in \Lambda_v$ such that $M \xrightarrow{h}{}^{*}_{\beta_v} V'$. Thus, M halts.

Conversely, if M halts then $M \xrightarrow{h}{}^{*}_{\beta_v} V \in \Lambda_v$, so $M \to_{\mathsf{v}}^* V$ since $\xrightarrow{h}_{\beta_v} \subseteq \to_{\mathsf{v}}$. By confluence of $\to_{\mathsf{v}}$ (Proposition 2.8, since $M \to_{\mathsf{v}}^* M'$) and Remark 2.7 (as $V \in \Lambda_v$), $V \to_{\mathsf{v}}^* V'$ and $M' \to_{\mathsf{v}}^* V'$ for some $V' \in \Lambda_v$. By Corollary 5.1.1, $M' \xrightarrow{h}{}^{*}_{\beta_v} V''$ for some $V'' \in \Lambda_v$. Therefore, M' halts. □

Corollary 5.4 (Soundness with respect to λ_v)**.** *If $M =_{\mathsf{v}} N$ then $M \cong N$.*

Proof. Let $\mathtt{C}$ be a context. By confluence of $\to_{\mathsf{v}}$ (Proposition 2.8), $M =_{\mathsf{v}} N$ implies that there exists $L \in \Lambda$ such that $M \to_{\mathsf{v}}^* L$ and $N \to_{\mathsf{v}}^* L$, hence $\mathtt{C}(\!|M|\!) \to_{\mathsf{v}}^* \mathtt{C}(\!|L|\!)$ and $\mathtt{C}(\!|N|\!) \to_{\mathsf{v}}^* \mathtt{C}(\!|L|\!)$. By Theorem 5.3, $\mathtt{C}(\!|M|\!)$ halts iff $\mathtt{C}(\!|L|\!)$ halts iff $\mathtt{C}(\!|N|\!)$ halts. Therefore, $M \cong N$. □

Plotkin [Plo75, p. 144] has already proved that $M =_{\beta_v} N$ implies $M \cong N$: we point out that our Corollary 5.4 is not obvious since λ_v^σ equates more than Plotkin's λ_v (indeed, $=_{\beta_v} \subseteq =_{\mathsf{v}}$ since $\to_{\beta_v} \subseteq \to_{\mathsf{v}}$, and Example 2.9 shows that this inclusion is strict). Corollary 5.4 means that λ_v^σ is sound with respect to the operational semantics of λ_v. In a way, adding σ-reduction rules to β_v-reduction is harmless with respect to Plotkin's notion of observational equivalence for λ_v: λ_v^σ does not equate too much.

The converse of Corollary 5.4 does not hold since $\lambda x.x(\lambda y.xy) \cong \Delta$ but $\lambda x.x(\lambda y.xy)$ and Δ are different v-normal forms, so $\lambda x.x(\lambda y.xy) \neq_{\mathsf{v}} \Delta$ by confluence of $\to_{\mathsf{v}}$ (Proposition 2.8).

Another remarkable consequence of Corollary 5.1.1 is Theorem 5.7 below: the notions of potential valuability and solvability for the shuffling calculus λ_v^σ (studied in [CG14]) coincide with the corresponding ones for Plotkin's λ_v (studied in [PR99, RP04, PPR05, PPR11]).

Definition 5.5 (Potential valuability, solvability)**.** Let N be a term and $x_1, \dots, x_k$ be pairwise distinct variables (with $k \in \mathbb{N}$) such that $\mathsf{fv}(N) = \{x_1, \dots, x_k\}$:

- N is v-*potentially valuable* (resp. β_v-*potentially valuable*) if there are values $V_1, \dots, V_k$, V such that $N\{V_1/x_1, \dots, V_k/x_k\} \to_{\mathsf{v}}^* V$ (resp. $N\{V_1/x_1, \dots, V_k/x_k\} \to_{\beta_v}^* V$);
- N is v-*solvable* (resp. β_v-*solvable*) whenever there are $n \in \mathbb{N}$ and terms $M_1, \dots, M_n$ such that $(\lambda x_1 \dots x_k.N)M_1 \cdots M_n \to_{\mathsf{v}}^* I$ (resp. $(\lambda x_1 \dots x_k.N)M_1 \cdots M_n \to_{\beta_v}^* I$).

The notions of potential valuability and solvability are parametric with respect to the reduction rules, so any variant of the λ-calculus has its own notions of potential valuability and solvability: Definition 5.5 introduces them for λ_v and λ_v^σ. Clearly, potential valuability is interesting only in a call-by-value setting, where a β-redex can be reduced only when its argument is a value: potentially valuable terms are those that, *up to a suitable substitution*, can be evaluated or placed in argument position without yielding a stuck β-redex.

The relevance of β-solvability for ordinary (call-by-name) λ-calculus is clearly presented in [Bar84], where this notion has been proved to grasp the idea of "meaningful program", i.e., a program that can produce any given output when supplied by suitable arguments. It is well known that, in λ, β-solvability is operationally characterized by head β-reduction: a term is β-solvable iff it is head β-normalizable. In a call-by-value setting, β_v-solvability and v-solvability are just the corresponding notions of solvability for λ_v and λ_v^σ, respectively.

In [PR99, RP04, PPR11] it has been proved that β_v-solvable terms are a proper subset of the β_v-potentially valuable terms, and it has been pointed out that β_v-reduction is too weak in order to characterize both these properties: an operational characterization of β_v-potential valuability and β_v-solvability cannot be given inside λ_v because of the problem of "premature" β_v-normal forms described in Section 1, e.g. the terms M and N in Eq. 1.1 are β_v-normal but neither β_v-solvable nor β_v-potentially valuable. In fact, β_v-solvability and β_v-potential valuability have been operationally characterized using two lazy strategies on — *call-by-name* — β-reduction (see [RP04, Theorems 3.1.9 and 3.1.14]), which is disappointing and unsound for λ_v: according to these lazy strategies, stuck β-redexes can be fired (even if the argument is not a value), for instance $(\lambda y.M)(xI)$ reduces to $M\{xI/y\}$.

On the other hand, concerning λ_v^σ, Theorems 24-25 in [CG14] give semantic and operational characterizations of v-potentially valuability and v-solvability. Interestingly, the operational characterizations rest on v-reduction strategies and then are *internal* to λ_v^σ. Let us recall these theorems (see Proposition 5.6 below) and, firstly, the notions involved in it.

For every term M with $\mathsf{fv}(M) \subseteq \{x_1, \dots, x_n\}$ and $\vec{x} = (x_1, \dots, x_n)$, we denote by $\llbracket M \rrbracket_{\vec{x}}$ (resp. $\llbracket M \rrbracket_{\vec{x}}^{\mathsf{s}}$) its *semantics* (resp. *stratified semantics*) in a relational model for λ_v^σ and λ_v.

All the details about this denotational model are in [CG14], for our purpose it is enough to recall that $\llbracket M \rrbracket_{\vec{x}}$ is a set such that $\llbracket M \rrbracket^{\mathsf{s}}_{\vec{x}} \subseteq \llbracket M \rrbracket_{\vec{x}}$, and if $M \to_{\mathsf{v}} N$ then $\llbracket M \rrbracket_{\vec{x}} = \llbracket N \rrbracket_{\vec{x}}$.

The reductions $\to_{\mathsf{w}}$ and $\to_{\mathsf{s}}$ are the closures of $\mapsto_{\beta_v} \cup \mapsto_{\sigma_1} \cup \mapsto_{\sigma_3}$ under weak and stratified contexts, respectively, where weak contexts (denoted by $\mathtt{W}$) and stratified contexts (denoted by $\mathtt{S}$) are special kinds of contexts defined as follows (see [CG14] for more details):

$$\mathtt{W} ::= (\!|\cdot|\!) \mid \mathtt{W}M \mid M\mathtt{W} \mid (\lambda x.\mathtt{W})M \qquad\qquad \mathtt{S} ::= \mathtt{W} \mid \lambda x.\mathtt{S} \mid \mathtt{S}M\,.$$

Note that $\to_{\mathsf{w}}$ and $\to_{\mathsf{s}}$ are two (non-deterministic but confluent) sub-reductions of $\to_{\mathsf{v}}$.

Proposition 5.6 (Semantic and operational characterization of v-potential valuability and v-solvability, [CG14])**.** *Let M be a term with $\mathsf{fv}(M) \subseteq \{x_1, \dots, x_n\}$ and $\vec{x} = (x_1, \dots, x_n)$.*

(1) Semantic and operational characterization of v-potential valuability ([CG14, Theorem 24])*: M is v-potentially valuable iff $\llbracket M \rrbracket_{\vec{x}} \neq \emptyset$ iff M is w-normalizable.*

(2) Semantic and operational characterization of v-solvability ([CG14, Theorem 25])*: M is v-solvable iff $\llbracket M \rrbracket^{\mathsf{s}}_{\vec{x}} \neq \emptyset$ iff M is s-normalizable.*

Thanks to standardization for λ_v^σ (actually, Corollary 5.1.1), we can prove Theorem 5.7 below, which reconciles the results about solvability and potential valuability for λ_v^σ and λ_v.

Theorem 5.7 (Potential valuability and solvability for λ_v^σ and λ_v)**.** *Let M be a term:*

(1) M is v-potentially valuable if and only if M is β_v-potentially valuable;

(2) M is v-solvable if and only if M is β_v-solvable.

Proof. In both points, the implication from right to left is trivial since $\to_{\beta_v} \subseteq \to_{\mathsf{v}}$. Let us prove the other direction. Let $\mathsf{fv}(M) = \{x_1, \dots, x_m\}$ for some $m \in \mathbb{N}$.

(1) Since M is v-potentially valuable, there exist some values $V, V_1, \dots, V_m$ such that $M\{V_1/x_1, \dots, V_m/x_m\} \to_{\mathsf{v}}^{*} V$; then, by Corollary 5.1.1 and because $\xrightarrow{h}_{\beta_v} \subseteq \to_{\beta_v}$, $M\{V_1/x_1, \dots, V_m/x_m\} \to_{\beta_v}^{*} V'$ for some $V' \in \Lambda_v$. So, M is β_v-potentially valuable.

(2) Since M is v-solvable, there exist terms $N_1, \dots, N_n$ (for some $n \geq 0$) such that $(\lambda x_1 \dots x_m.M)N_1 \cdots N_n \to_{\mathsf{v}}^{*} I$; then, by Corollary 5.1.1 and because $\xrightarrow{h}_{\beta_v} \subseteq \to_{\beta_v}$, there exists $V \in \Lambda_v$ such that $(\lambda x_1 \dots x_m.M)N_1 \cdots N_n \to_{\beta_v}^{*} V \xrightarrow{int}{}_{\mathsf{v}}^{*} I$. According to Lemma 3.9.7 , $V = \lambda x.N$ for some $N \in \Lambda$ such that $N \to_{\mathsf{v}}^{*} x$. By Corollary 5.1.1, there is $V' \in \Lambda_v$ such that $N \xrightarrow{h}{}_{\beta_v}^{*} V' \xrightarrow{int}{}_{\mathsf{v}}^{*} x$, hence $V' = x$ by Lemma 3.9.7 again. Since $\xrightarrow{h}_{\beta_v} \subseteq \to_{\beta_v}$, $N \to_{\beta_v}^{*} x$ and thus $V = \lambda x.N \to_{\beta_v}^{*} I$, so M is β_v-solvable. □

According to Theorem 5.7, the notions of potential valuability and solvability for λ_v^σ *coincide* with the respective ones for Plotkin's λ_v. So, the semantic (via a relational model) and operational (via two sub-reductions of $\to_{\mathsf{v}}$) characterizations of v-potential valuability and v-solvability given in Proposition 5.6 are also semantic and operational characterizations of β_v-potential valuability and β_v-solvability. The difference is that these notions are characterized operationally *inside* λ_v^σ (using call-by-value reductions), while it is impossible to characterize them operationally inside λ_v. This shows how λ_v^σ is a useful, conservative and "complete" tool for studying semantic and operational properties of Plotkin's λ_v.

For the sake of completeness, we mention another conservativity result of λ_v^σ with respect to λ_v, proved in [Gue15, Theorem 21] thanks to our sequentialization: it shows that the notions of head reduction for λ_v^σ and λ_v are equivalent from the termination viewpoint.

Proposition 5.8 (Head normalization, [Gue15])**.** *Let $N \in \Lambda$. The following are equivalent:*

(1) N is head v*-normalizable;*
(2) N is head β_v-normalizable;
(3) $N =_{\mathsf{v}} L$ for some head v*-normal L;*
(4) N is strongly head v*-normalizing.*

The equivalence (1)⇔(4) means that normalization and strong normalization are equivalent for head v-reduction (for head β_v-reduction they are trivially equivalent since head β_v-reduction is deterministic), therefore if one is interested in studying the termination of head v-reduction, no difficulty arises from its non-determinism. The equivalence (4)⇔(2) or (1)⇔(2) says that the evaluation defined for Plotkin's λ_v (head β_v-reduction) terminates if and only if the evaluation defined for λ_v^σ (head v-reduction) terminates: σ-rules play no role in deciding the termination of a head v-reduction sequence (in a way, this generalizes Corollary 5.1.2), they can only activate hidden β_v-redexes that are not in head position.

Standardization is related to normalization. In [Gue15, Theorem 24] a family of normalizing strategies for λ_v^σ has been introduced: a term M is v-normalizable iff M v-reduces to its v-normal form selecting v-redexes in a special order defined in [Gue15, Definition 22]. Actually, these normalizing strategies are a special case of standard sequences.

Definition 5.9 (Strict standard head sequence)**.** A *strict standard head sequence* is a finite sequence $(M_0, \dots, M_k, \dots, M_m)$ of terms (with $k \leq m$) such that M_k is head β_v-normal, M_m is head v-normal, $M_i \xrightarrow{h}_{\beta_v} M_{i+1}$ for any $0 \leq i < k$, and $M_i \xrightarrow{h}_{\sigma} M_{i+1}$ for any $k \leq i < m$.

A *strict standard sequence* is then defined by replacing the notion of standard head sequence with the notion of strict standard head sequence in Definition 4.2. So, normalization theorem proved in [Gue15, Theorem 24] can be reformulated as follows:

Proposition 5.10 (Normalization, [Gue15])**.** *Let M be a term: M is* v*-normalizable iff there exists a strict standard sequence from M to its* v*-normal form.*

The proof of the left-to-right direction of Proposition 5.10 (the right-to-left one is trivial) relies on Proposition 5.8, see [Gue15] for details: the idea is that, given a v-normalizable (and then head v-normalizable) term M, one performs — deterministically — head β_v-reduction steps from M as long as a head β_v-normal form N is reached (according to Proposition 5.8, a term is head v-normalizable iff it is head β_v-normalizable); then, one performs head σ-reduction steps from N (where head σ_1- and head σ_3-reduction steps can be performed in whatever order) as long as a head v-normal form L is reached (such a L always exists because $\xrightarrow{h}_{\sigma}$ is strongly normalizing and preserves β_v-normal forms); finally, one performs internal v-reduction steps starting from L by iterating this strategy on the subterms of L, according to the standard left-to-right order, as long as the v-normal form of M is reached.

Clearly, Theorem 4.6 fails if in its statement "standard sequence" is replaced by "strict standard sequence": $I\Delta I \xrightarrow{h}_{\sigma_1} (\lambda.xI)\Delta$ is a standard sequence but there is no strict standard sequence from $I\Delta I$ to $(\lambda.xI)\Delta$, since $I\Delta I \xrightarrow{h}_{\beta_v} \Delta I \xrightarrow{h}_{\beta_v} II \xrightarrow{h}_{\beta_v} I$ and I is (head) v-normal. Similarly, $(\Delta\Delta)(II) \xrightarrow{int}_{\beta_v} (\Delta\Delta)I$ is a standard sequence but there is no strict standard sequence from $(\Delta\Delta)(II)$ to $(\Delta\Delta)I$, since $(\Delta\Delta)(II)$ is not head β_v-normalizable.

6. Conclusions

It has been proved in [PR99, Pao02, RP04, PPR11] that β_v-reduction is too weak to characterize operationally some semantical properties of λ_v, such as separability, potentially valuability and solvability. The main motivation behind the introduction of λ_v^σ in [CG14] was to achieve a call-by-value language where potential valuability and solvability can be characterized operationally without resorting to reductions external the call-by-value paradigm: λ_v^σ allows an internal operational characterization of such notions [CG14, Theorems 24-25]. In this paper we close the game, by proving that λ_v^σ is a conservative extension of λ_v: in particular, λ_v^σ is sound with respect to the opeartional semantics of λ_v (Corollary 5.4), and potential valuability and solvability for λ_v^σ coincide with the respective notions for λ_v (Theorem 5.7). So, λ_v^σ is a useful framework for studying semantic and operational properties of λ_v. The technical tool on which the proofs of these conservativity properties are based is an interesting result in its own, namely standardization for λ_v^σ (Theorem 4.6).

Standardization for λ_v^σ has been proved using parallel reduction. Let us recall that parallel reduction in λ-calculus has been defined by Tait and Martin-Löf in order to prove confluence of β-reduction, without referring to the tricky notion of residuals. Takahashi in [Tak89, Tak95] has simplified this technique and showed that it can be successfully applied also to prove standardization for λ. However, in λ_v^σ our parallel reduction $\Rightarrow$ cannot be used to prove confluence of $\to_{\mathsf{v}}$, since $\Rightarrow$ does not enjoy the diamond property. Indeed, consider

$$\begin{tikzcd}
 & (\lambda x.M)\big((\lambda y.N)(zz)\big)L \arrow[dl, Rightarrow, "\text{(by applying the rule } \sigma_1\text{)}"'] \arrow[dr, Rightarrow, "\text{(by applying the rule } \sigma_3\text{)}"] & \\
M_1 = (\lambda x.ML)\big((\lambda y.N)(zz)\big) & & \big(\lambda y.(\lambda x.M)N\big)(zz)L = M_2
\end{tikzcd}$$

It is easy to check that there is no term M' such that $M_1 \Rightarrow M'$ and $M_2 \Rightarrow M'$.

The proof of the standardization theorem is based on a sequentialization property, imposing a total order between β_v-redexes, but a partial one between σ-redexes. We conjecture that a total order between all v-redexes can be provided by defining a suitable notion of head σ-reduction that properly interleaves head σ_1- and head σ_3-reduction steps. Anyway, we do not fully explored this possibility because we are unaware of interesting applications.

Postponements of head σ-reduction to head β_v-reduction (Lemma 3.11) and of internal v-reduction to head v-reduction (Corollary 3.18) suggest the idea that, in order to avoid the issues affecting λ_v when dealing with open terms and stuck β-redexes, it is enough to restrict our shuffling calculus λ_v^σ by allowing (local head) σ-reduction steps only when a (local head) β_v-normal form is reached. This approach generalizes the idea behind strict standard sequences defined in Section 5. In fact, this restricted shuffling calculus is a "minimalistic" extension of Plotkin's λ_v solving the problem of premature β_v-normal forms. Since values are head β_v-normal and I is v-normal, Corollary 5.1.1 and Proposition 5.10 ensure that the conservativity result given by Theorem 5.7 (as well as Corollary 5.4) would still hold in this restricted shuffling calculus. But solving the problem of premature β_v-normal forms is only the first step in the direction of a deep analysis of λ_v and, more generally, of call-by-value settings: the *whole* shuffling calculus λ_v^σ seems to be an adequate framework for this task (Corollary 5.4 and Theorem 5.7 exemplify how call-by-value properties can be correctly studied inside the whole λ_v^σ) and its study is more elegant and simpler without imposing any "clumsy" syntactic restrictions on the definition of shuffling calculus reduction rules.

Future work. We plan to continue to explore the call-by-value setting, using the shuffling calculus λ_v^σ. As a first step, we would like to revisit and improve the Separability Theorem given in [Pao02] for λ_v. Still the issue is more complex than in the call-by-name, indeed in ordinary λ-calculus different $\beta\eta$-normal forms can be separated (by the Böhm Theorem), while in λ_v there are different normal forms that cannot be separated, but which are only semi-separable (e.g. I and $\lambda z.(\lambda u.z)(zz)$). We hope to completely characterize separable and semi-separable normal forms in λ_v^σ. This should be a first step aimed to define a semantically meaningful notion of approximants. Then, we should be able to provide a new insight on the denotational analysis of the call-by-value, maybe overcoming limitations as that of the absence of fully abstract filter models [RP04, Theorem 12.1.25]. Last but not least, an unexplored but challenging research direction is the use of our commutation σ-rules to improve and speed up the call-by-value evaluation. We do not have any concrete evidence supporting such possibility, but since λ_v^σ is strongly related to the calculi presented in [HZ09, AP12] (see [AG16] for a comparison), which are endowed with explicit substitutions, we believe that a sharp use of commutations could have a relevant impact on the evaluation.

Acknowledgements. This work has been supported by LINTEL TO_Call1_2012_0085, a Research Project funded by the "Compagnia di San Paolo", and by the A*MIDEX project (ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French Government program, managed by the French National Research Agency (ANR).

The authors wish to thank the anonymous referees for their insightful comments.

References

[Acc15] Beniamino Accattoli. Proof nets and the call-by-value λ-calculus. *Theoretical Compuer Science*, 606:2–24, 2015.

[AG16] Beniamino Accattoli and Giulio Guerrieri. Open Call-by-Value. In *Programming Languages and Systems – 14th Asian Symposium (APLAS 2016)*, volume 10017 of *Lecture Notes in Computer Science*, pages 206–226. Springer-Verlag, 2016.

[AP12] Beniamino Accattoli and Luca Paolini. Call-by-Value Solvability, Revisited. In *Functional and Logic Programming*, volume 7294 of *Lecture Notes in Computer Science*, pages 4–16. Springer-Verlag, 2012.

[AS15] Beniamino Accattoli and Claudio Sacerdoti Coen. On the Relative Usefulness of Fireballs. In *30th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2015*, pages 141–155. IEEE Computer Society, 2015.

[Bar84] Henk Barendregt. *The Lambda Calculus: Its Syntax and Semantics*, volume 103 of *Studies in logic and the foundation of mathematics*. North Holland, 1984.

[CF58] Haskell B. Curry and Robert Feys. *Combinatory Logic*, volume 1. North Holland, 1958.

[CG14] Alberto Carraro and Giulio Guerrieri. A Semantical and Operational Account of Call-by-Value Solvability. In *Foundations of Software Science and Computation Structures*, volume 8412 of *Lecture Notes in Computer Science*, pages 103–118. Springer-Verlag, 2014.

[CH00] Pierre-Louis Curien and Hugo Herbelin. The duality of computation. In *Proceedings of the Fifth ACM SIGPLAN International Conference on Functional Programming (ICFP '00)*, pages 233–243. ACM, 2000.

[Cra09] Karl Crary. A Simple Proof of Call-by-Value Standardization. Technical Report CMU-CS-09-137, Carnegie Mellon University, 2009.

[DL07] Roy Dyckhoff and Stéphane Lengrand. Call-by-Value lambda-calculus and LJQ. *Journal of Logic and Computation*, 17(6):1109–1134, 2007.

[EG16] Thomas Ehrhard and Giulio Guerrieri. The Bang Calculus: an untyped lambda-calculus generalizing call-by-name and call-by-value. In *Proceedings of the 18th International Symposium on Principles and Practice of Declarative Programming (PPDP 2016)*, pages 174–187. ACM, 2016.

[EHR92] Lavinia Egidi, Furio Honsell, and Simona Ronchi Della Rocca. Operational, Denotational and Logical Descriptions: A Case Study. *Fundamenta Informaticae*, 16(2):149–169, 1992.

[Gir87] Jean-Yves Girard. Linear logic. *Theoretical Computer Science*, 50(1):1–102, 1987.

[GL02] Benjamin Grégoire and Xavier Leroy. A compiled implementation of strong reduction. In *Proceedings of the Seventh ACM SIGPLAN International Conference on Functional Programming (ICFP '02)*, pages 235–246. ACM, 2002.

[GPR15] Giulio Guerrieri, Luca Paolini, and Simona Ronchi Della Rocca. Standardization of a Call-By-Value Lambda-Calculus. In *13th International Conference on Typed Lambda Calculi and Applications (TLCA 2015)*, volume 38 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 211–225, 2015.

[Gue15] Giulio Guerrieri. Head reduction and normalization in a call-by-value lambda-calculus. In *2nd International Workshop on Rewriting Techniques for Program Transformations and Evaluation (WPTE 2015)*, volume 46 of *OpenAccess Series in Informatics (OASIcs)*, pages 3–17, 2015.

[Hin78] Roger Hindley. Standard and normal reductions. *Transactions of the American Mathematical Society*, pages 253–271, 1978.

[Hue80] Gérard Huet. Confluent Reductions: Abstract Properties and Applications to Term Rewriting Systems. *Journal of ACM*, 27(4):797–821, 1980.

[HZ09] Hugo Herbelin and Stéphane Zimmermann. An Operational Account of Call-by-Value Minimal and Classical lambda-Calculus in "Natural Deduction" Form. In *Typed Lambda Calculi and Applications, 9th International Conference, TLCA 2009*, volume 5608 of *Lecture Notes in Computer Science*, pages 142–156. Springer-Verlag, 2009.

[JGS93] Neil D. Jones, Carsten K. Gomard, and Peter Sestoft. *Partial Evaluation and Automatic Program Generation.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.

[Klo80] Jan Willem Klop. Combinatory Reduction Systems. *Mathematical Centre Tracts*, 127, 1980.

[Kri90] Jean-Louis Krivine. *Lambda-Calcul : Types et Modèles.* Études et Recherches en Informatique. Masson, 1990.

[Las05] Søren B. Lassen. Eager Normal Form Bisimulation. In *20th IEEE Symposium on Logic in Computer Science (LICS 2005), Proceedings*, pages 345–354. IEEE Computer Society, 2005.

[Mit79] Gerd Mitschke. The Standardization Theorem for λ-Calculus. *Mathematical Logic Quarterly*, 25(1-2):29–31, 1979.

[Mog88] Eugenio Moggi. Computational lambda-calculus and monads. Technical report, Edinburgh University, 1988. Tech. Report ECS-LFCS-88-66.

[Mog89] Eugenio Moggi. Computational Lambda-Calculus and Monads. In *Proceedings of the 4th Symposium on Logic in Computer Science (LICS'89)*, pages 14–23. IEEE Computer Society, 1989.

[MOTW95] John Maraist, Martin Odersky, David N. Turner, and Philip Wadler. Call-by-name, call-by-value, call-by-need and the linear lambda calculus. *Electronic Notes in Theoretical Computer Science*, 1:370–392, 1995.

[MOTW99] John Maraist, Martin Odersky, David N. Turner, and Philip Wadler. Call-by-name, call-by-value, call-by-need and the linear lambda calculus. *Theoretical Computer Science*, 228(1–2):175–210, 1999.

[Pao02] Luca Paolini. Call-by-Value Separability and Computability. In *Italian Conference in Theoretical Computer Science*, volume 2202 of *Lecture Notes in Computer Science*, pages 74–89. Springer-Verlag, 2002.

[Plo75] Gordon D. Plotkin. Call-by-name, call-by-value and the lambda-calculus. *Theoretical Computer Science*, 1(2):125–159, 1975.

[PPR05] Luca Paolini, Elaine Pimentel, and Simona Ronchi Della Rocca. Lazy strong normalization. In *Proceedings of Intersection Types and Related Systems (ITRS'04)*, volume 136C of *Electronic Notes in Theoretical Computer Science*, pages 103–116, 2005.

[PPR11] Luca Paolini, Elaine Pimentel, and Simona Ronchi Della Rocca. Strong Normalization from an unusual point of view. *Theoretical Computer Science*, 412(20):1903–1915, 2011.

[PR99] Luca Paolini and Simona Ronchi Della Rocca. Call-by-value Solvability. *Theoretical Informatics and Applications*, 33(6):507–534, 1999. RAIRO Series, EDP-Sciences.

[PR04] Luca Paolini and Simona Ronchi Della Rocca. Parametric parameter passing lambda-calculus. *Information and Computation*, 189(1):87–106, 2004.

[Reg92] Laurent Regnier. *Lambda calcul et réseaux.* PhD thesis, Université Paris 7, 1992.

[Reg94] Laurent Regnier. Une équivalence sur les lambda-termes. *Theoretical Computer Science*, 126(2):281–292, 1994.

[RP04] Simona Ronchi Della Rocca and Luca Paolini. *The Parametric λ-Calculus: a Metamodel for Computation.* Texts in Theoretical Computer Science: An EATCS Series. Springer-Verlag, 2004.

[SF92] Amr Sabry and Matthias Felleisen. Reasoning About Programs in Continuation-passing Style. *SIGPLAN Lisp Pointers*, V(1):288–298, 1992.

[SF93] Amr Sabry and Matthias Felleisen. Reasoning about programs in continuation-passing style. *Lisp and Symbolic Computation*, 6(3-4):289–360, 1993.

[SW97] Amr Sabry and Philip Wadler. A Reflection on Call-by-Value. *ACM Transactions on Programming Languages and Systems*, 19(6):916–941, 1997.

[Tak89] Masako Takahashi. Parallel Reduction in λ-Calculus. *Journal of Symbolic Computation*, 7(2):113–123, 1989.

[Tak95] Masako Takahashi. Parallel Reductions in lambda-Calculus. *Information and Computation*, 118(1):120–127, 1995.

[Ter03] Terese. *Term Rewriting Systems*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2003.